ИЗКУСТВЕНИЯТ ИНТЕЛЕКТ: тенденции, регулации и сценарии за развитие

45 години Българска стопанска камара
в подкрепа на българския бизнес

Април 2025 г.

„ИЗКУСТВЕНИЯТ ИНТЕЛЕКТ: тенденции, регулации и сценарии за развитие“

Автори: Христина Каспарян и Миряна Недева, Българска стопанска камара

Анализът е актуален към април 2025 г.

При изготвянето на анализа са използвани различни модели на изкуствен интелект за откриване на източници и оптимизация на текста.

Съдържание:

Списък със съкращения

AI	Artificial Intelligence (Изкуствен интелект)
API	Application programming interface (Приложен програмен интерфейс)
APNSA	Assistant to the President for National Security Affairs (Помощник на президента по въпросите на националната сигурност)
APST	Assistant to the President for Science and Technology (Помощник на президента по наука и технологии)
ASIC	accelerator chips, вид чипове, използвани за ИИ
CEN	European Committee for Standardization (Европейски комитет по стандартизация)
CENELEC	European Committee for Electrotechnical Standardization (Европейски комитет по електротехническа стандартизация)
CERN	European Organization for Nuclear Research (Европейска организация за ядрени изследвания)
CRA	Cyber Resilience Act (Акт за киберустойчивост)
CSET	Center for Security and Emerging Technologies (Център за сигурност и нововъзникващи технологии)
CSIS	Center for Strategic and International Studies (Център за стратегически и международни изследвания)
DARPA	Defense Advanced Research Projects Agency (Агенцията за перспективни изследователски проекти за отбраната)
DEFT	Data Free Flow with Trust (Свободен поток от данни с доверие)
DMA	Digital Markets Act (Акт за цифровите пазари)
DSA	Digital Services Act (Акт за цифровите услуги)
EBNA	European Business Nuclear Alliance (Европейски бизнес ядрен алианс)
FLAP	Франкфурт, Лондон, Амстердам, Париж
FLOPs	Floating point operations per second (мярка за изчислителната мощност на компютрите)
FSB	Financial Stability Board
GAN	generative adversarial network (Генеративна противникова мрежа)
GDPR	General Data Protection Regulation (Общ регламент за защита на данните)
GPAI	Global Partnership on AI (Глобално партньорство за изкуствен интелект)
GPU	Graphics processing unit (графичен процесор)
HAIP	Hiroshima AI Process (Хирошимски процес за ИИ)
IDS	intrusion detection system (система за откриване на проникване)
IEA	International Energy Agency (Международна енергийна агенция)
IoT	Internet of Things (Интернет на нещата)
IP	Internet Protocol (интернет протокол)
IPS	intrusion prevention system (Система за предотвратяване на проникване)
ISO	Международна организация по стандартизация
ISO	Международна организация по стандартизация

ITU	Международен телекомуникационен съюз
KPI	Key performance indicators
KYC	Know your customer
LLM	Large language models (Големи езикови модели)
ML	machine learning
MW	мегават
MWh	мегаватчас
NLP	natural language processing
PCAST	President's Council of Advisors on Science and Technology (Президентски съвет на съветниците по наука и технологии)
PUE	power usage effectiveness (ефективност на използване на енергията)
R&D	research and development
ROI	Return on investment (Възвращаемост на инвестиции)
SOC	security operations center (център за операции по сигурност)
STEM	Science, technology, engineering and mathematics
TPU	Tensor Processing Units (специализирани чипове за ИИ)
UEBA	user entity and behaviour analytics
URL	Uniform Resource Locator (универсален указател на ресурс)
VAE	вариационен автоенкодер
VIP	very important person
VLOPs	Very Large Online Platforms (Много големи онлайн платформи)
WIPO	Световна организация за интелектуална собственост
БВП	брутен вътрешен продукт
БСК	Българска стопанска камара
ЕИСК	Европейски икономически и социален комитет
ЕК	Европейска комисия
ЕС	Европейски съюз
ИИ	Изкуствен интелект
ИКТ	Информационни и комуникационни технологии
МВФ	Международен валутен фонд
МОТ	Международна организация на труда
МСП	Малки и средни предприятия
НИРД	Научно-изследователска и развойна дейност
ОИСР	Организация за икономическо сътрудничество и развитие
ООН	Организация на обединените нации
САЩ	Съединени американски щати
СИФ	Световен икономически форум
ЮНЕСКО	Организацията на обединените нации за образование, наука и култура

1. Въведение

Бъдещето е тук и ще преобърне всичко. Готови ли сме?

Изкуственият интелект (ИИ) вече не е научна фантастика, а двигателят на чисто нова революция. Подобно на електричеството преди век или на интернет преди две десетилетия, ИИ е на път да промени тотално начина, по който работим, живеем и дори мислим. Той вече е част от живота ни, а влиянието му тепърва ще набира умопомрачителна скорост. Настоящият анализ ще ви потопи в най-горещите тенденции около ИИ и икономиката, като хвърли светлина върху шансовете и предизвикателствата пред България в тази нова ера. **Трябва да се вземе предвид, че анализът не претендира за изчерпателност, поради динамиката, с която се развива тази технология. Той по-скоро поставя основа за по-нататъшни разработки и допълнителен преглед.** Разработените сценарии за развитие на регулациите, киберсигурността, производството и финансите имат за цел да провокират стратегическото мислене и да предизвикат обществен дебат доколко сме готови за бъдещето, към което се е устремил светът.

Парите следват ИИ

Очаква се до 2030 година ИИ да налее над 15,7 трилиона долара в световната икономика. Инвестициите в ИИ се увеличават с шеметна скорост, като само през 2023 г. компаниите са похарчили над 166 милиарда долара за ИИ технологии. А появата на "умни" системи като генеративния ИИ отваря врати към невиждани досега възможности за автоматизация, индивидуално отношение към клиентите и куп иновации във всички сектори на икономиката.

ЕС и България на кръстопът

За съжаление, България все още изостава от средното ниво за ЕС по отношение на готовността за ИИ и сме далеч от челните места по научни пробиви в тази сфера. От своя страна, ЕС също значително изостава от водещите държави в тази област САЩ и Китай. Големите спънки са липсата на топ специалисти, слабото развитие на науката и малкото инвестиции в проучвания и нови технологии. С правилните стъпки и политики и съответните стратегически решения изоставането може да бъде намалено и страната да се насочи към ниши, които да доведат до ползи както за икономиката, така и за обществото.

ИИ – суперсила за бизнеса

ИИ не е просто поредната модна дума. Той идва на помощ за повишаване на ефективността и намаляване на разходите във всички сектори на икономиката и обществения живот.

В здравеопазването помага за по-бърза и точна диагностика, спестявайки до 50% от разходите. В заводите прогнозира кога една машина ще се счупи и автоматизира процеси, намалявайки загубите. В енергетиката прави мрежите по-умни и помага за по-добро управление на електропреносната мрежа. В земеделието оптимизира добивите и прави селското стопанство по-ефективно. Във финансите автоматизира анализи, оценява кредити и открива измами за секунди. В науката ускорява проучванията и помага на учените да правят открития по-бързо.

Настоящият анализ маркира ползите и предизвикателствата пред отделните сектори, като предоставя начална информация за тях.

Не без пречки

Внедряването на ИИ си има своите предизвикателства. Необходими са сериозни инвестиции в хардуер, софтуери и развойна дейност. Събирането и обработката на качествени данни, които да захранят ИИ, изисква сериозни разходи. Самото обучение на

моделите изисква средства, а и битката за привличане и задържане на специалисти става все по-ожесточена. Що се отнася до регулациите, те задават правилата, но ако при тях не се намери правилния баланс, те могат да забавят и оскъпят процеса, особено за малкия бизнес. Чрез Акта за ИИ Европейският съюз създава рамка от правила, която цели да направи ИИ по-безопасен и прозрачен. Тези правила могат да са както трамплин, така и спирачка за новите технологии – всичко зависи от това как се прилагат регулациите.

Нашата пътна карта

Анализът предполага предприемането на конкретни действия. На първо място, това е наличието на ясен план и преразглеждане на всички стратегически документи. Насърчаването на новите бизнес модели и внедряването на ИИ в различните сектори е от стратегическо значение за икономическия просперитет на страната. Инвестициите в наука следва да бъдат ориентирани към изследвания и разработки с висока международна стойност и увеличена патентна дейност. ИИ предполага дълбока и мащабна цялостна реформа на образователната система, при която критичното мислене и аналитичните способности са базови умения, които трябва да се възпитават и надграждат от ранна детска възраст.

ИИ е не просто технология, а огромна възможност за България да расте, да е по-ефективна и да създава продукти и услуги с висока добавена стойност. Предизвикателствата са налице, но с правилните ходове, инвестиции и общи усилия те могат да се превърнат в конкурентни предимства. Въпросът не е дали ще приемем тази нова реалност, а как ще я управляваме – активно, умно и с поглед към бъдещето.

ИИ е вече тук – време е да решим дали ще бъдем пасивни наблюдатели, или активни участници в новата икономическа ера!

2. Изводи

1. Изкуственият интелект (ИИ) се утвърждава като технология с общо предназначение (General Purpose Technology) подобно на електричеството, компютрите или интернет, която навлиза във всички сфери на обществения живот и променя фундаментално световната икономика. Успешната цифрова трансформация е процес на стратегическо преосмисляне на всички икономически и социални политики, а не само проблем на използване на ново технологично решение. Европейската и световната икономика се намират на кръстопът – да възприемат широко и отговорно ИИ, което би означавало нова епоха на растеж, иновации и благосъстояние, или да подценят промяната, рискувайки да загубят конкурентоспособност и благоденствие.
2. ИИ се развива с ускорени темпове, като в някои задачи многократно превъзхожда човешките възможности, въпреки че все още изостава в по-сложни области. Проучванията демонстрират, че ИИ повишава производителността на работниците и служителите и води до по-качествена работа, като същевременно ускорява научния прогрес. Въпреки тези позитивни тенденции, обществото е все по-информирано с потенциалното му въздействие и изразява нарастваща тревожност по отношение на неговото развитие.
3. ЕС има историческия шанс да създаде „европейски модел" на изкуствен интелект – сигурен, прозрачен, социално отговорен и иновативен. Създаването на конкурентоспособен модел следва да се изгражда в постоянен диалог с индустрията и заетите, с мащабни инвестиции и гъвкави, адаптивни механизми на управление, които подкрепят и насочват визионерския дух на европейските иноватори, без да ги възпрепятстват чрез тежки регулации.
4. Внедряването на изкуствен интелект е в пряка корелация с технологичния суверенитет. ЕС не притежава технологичен и, по-конкретно – цифров суверенитет, като това се отнася в още по-висока степен за България. Необходимо е преосмисляне на подхода при регулациите и развитието на регулаторна рамка, която подкрепя технологичния напредък и привлича технологични таланти.
5. Необходим е по-голям фокус върху разработване на стандартизирани оценки за отговорност, справяне със заплахата от дезинформация и други форми на злоупотреба, повишаване на прозрачността при разработването на ИИ и задълбочено разбиране на потенциалните краткосрочни и дългосрочни рискове, за да се гарантира безопасното и етично развитие на изкуствения интелект.
6. Формирането на целенасочена национална политика в областта на изкуствения интелект в България е от изключително голямо значение, като тя трябва да се осъществи с активното участие на индустрията, фокусирана върху инвестиции, развитие на човешкия капитал, индустриална трансформация и балансирана регулация за икономически растеж и дигитална независимост, в т.ч:
    - Целенасочени инвестиции в качествени научни проекти чрез конкурентни програми с акцент върху оригиналност, приложимост и международно сътрудничество.
    - Стимулиране на международни публикации в престижни научни списания чрез програми за научна мобилност, съвместни проекти с водещи университети и допълнително финансиране за реномирани публикации;
    - Национална стратегия за ИИ и трудова заетост, която очертава пътя за адаптация на различните сектори;
    - Стимулиране на партньорства между бизнес, академични институции и публичния сектор за създаване на иновационни модели на заетост;

- Развитие на научната инфраструктура и подкрепа за изследователски групи особено в университетите и институтите, които работят в областта на машинното обучение, обработка на естествен език и роботика;
- Публично-частно партньорство чрез интегриране на научната дейност с индустрията, създаване на технологични паркове и инкубатори около водещи университети и научни институти.

7. Инвестициите в центрове за данни са необходими с оглед на стратегическата сигурност на страната. БСК отчита нарастващото значение на инфраструктурата, необходима за поддържането на големи данни. Ефективността на системите за изкуствен интелект пряко зависи от качеството и количеството на данните. Първоначалните инвестиции в адекватна инфраструктура за съхранение и обработка на големи данни са високи, но дългосрочната възвръщаемост на инвестицията се очаква да надхвърли значително първоначалните разходи, особено с намаляване на цената за съхранение и обработка на данни с течение на времето.
8. В дългосрочен план трябва да се укрепва и защитава ролята на България като европейски доставчик на критичните и стратегическите материали, необходими за преноса на данни, особено за онези от тях, на които България е сред водещите европейски доставчици. Това създава допълнителни възможности за икономическо развитие на България.
9. Енергийният отпечатък на ИИ е значителен и тенденцията е за неговото нарастване едновременно с нарастването на използването на ИИ. През следващото десетилетие енергийната ефективност ще се превърне в ключов критерий при проектирането на модели с ИИ и хардуер. Инвестициите в енергийна инфраструктура са необходими наред с подкрепа за иновации в областта на енергетиката.
10. ИИ предполага и фокусиране на усилия върху обезпечаване на националната сигурност чрез интегриране на ИИ в киберсигурността, с цел по-добро покритие на заплахи, по-бързо откриване на атаки и по-ефективна превенция.
11. Инвестициите в иновативни технологични решения и цялостна система за управление на електронните отпадъци са необходими, доколкото е-отпадъците съдържат ценни ресурси, които са ключови за развитието на ИИ. Икономическата ефективност от по-широкото използване на ИИ и съпровождащата я ускорена употреба на все по-нови електронни устройства е в пряка зависимост от иновациите и намирането на устойчиви екологосъобразни решения и технологии за рециклиране на излезлите от употреба устройства, особено по отношение извличането на ценните елементи, вложени в тяхното производство.
12. Задължително е и отразяването на нововъзникващите индустриални сектори в резултат от развитието на ИИ в стратегическите документи на страната, в т.ч.: в подготовката и изпълнението на Иновационната стратегия за интелигентна специализация, съпроводена със съответния анализ в кои от тези нови индустрии България може да има конкурентно предимство и силни международни позиции. Обмисляне на политики и мерки за подкрепа на иновативни бизнес модели и трансформация на съществуващите бизнес модели е ключово за запазване на конкурентоспособността на българските предприятия в дългосрочен план.
13. Използването на ИИ налага преструктуриране и пренастройване на съществуващите международни правила за защита на интелектуалната собственост. Управлението на интелектуалната собственост е един от най-важните инструменти на бизнеса за придобиване на дългосрочни конкурентни предимства, възвръщаемост на инвестициите и повишаване на икономическата ефективност. БСК продължава да следи внимателно световния дебат по отношение правата на интелектуална собственост и ИИ.

14. Политиките в областта на образованието и пазара на труда могат да позиционират страната ни сред регионалните лидери и този шанс не трябва да бъде пропускан. При разработката на водещи модели човешката експертиза ще остане незаменима. Предприятията ще трябва да планират значителни инвестиции в човешки капитал. Най-големи икономически ползи ще имат държавите, които трансформират своите образователни системи по начин, който отговаря на търсенето на конкретни умения на пазара на труда.

3. Дефиниции и категоризация

Общо възприета дефиниция за "изкуствен интелект" на международно равнище към този момент не съществува. Съществуват различни дефиниции, дадени от международни организации, корпорации, водещи разработчици на ИИ модели, както и ЕС, и различни държави. Технологията може да бъде категоризирана и, съответно - дефинирана по различен начин според нейните способности, функционалности, тип алгоритъм и начин на трениране. В настоящия анализ се представят изчерпателен брой дефиниции, за да се очертае по-ясна и точна картина за тази така ключова за развитието на света технология. Изброените категории очертават рамка за разбиране на текущите и бъдещите възможности на изкуствения интелект. В настоящия анализ в Приложение №1 са включени дефинициите на Европейския съюз, Организацията за икономическо сътрудничество и развитие (ОИСР), Международната организация по стандартизация (International Organization for Standardization, ISO), университета "Станфорд", колежа "Итън", Google, NVIDIA, IBM, Microsoft, Avast и Oracle.

Настоящият документ съдържа прогнози и очаквания, които се основават на анализи и данни от различни източници. Тези прогнози са изготвени с използването на различни методологии и се базират на информация, която е била актуална към момента на тяхното съставяне. Предвид динамиката на пазарните условия и други фактори, е възможно бъдещите събития да се развият по начин, който съществено се различава от представените прогнози. Ето защо, предоставената информация следва да се възприема като източник на информация и ориентация, а не като окончателно предвиждане на бъдещето.

4. Въведение и тенденции в развитието на ИИ

4.1. Глобална технологична надпревара в областта на ИИ

Настоящата точка от анализа се фокусира върху изследване на глобалната технологична надпревара в областта на ИИ, през призмата на данни и статистически показатели. Разбирането на тази динамична конкуренция изисква анализ на международни индекси и авторитетни изследвания от водещи организации, които следят отблизо развитието на ИИ. Чрез синтез и интерпретация на информацията, предоставена от тези организации, ще бъде представена подробна картина на глобалната конкуренция и ще бъдат очертани водещите държави и региони, ключовите тенденции в инвестициите, развитието на таланти и научните публикации в областта на ИИ, както и използването на ИИ и ефектите от това.

4.1.1 Stanford AI Index 2024

За седма поредна година университетът Станфорд представи своя AI Index. В тазгодишното издание са направени няколко основни извода, които очертават една ясна картина на продължаващо бързо развитие на технологията, съпътствана от несигурност в обществеността относно потенциалните ефекти от ИИ.[1] **AI Index 2024** показва, че изкуственият интелект продължава да напредва, като в някои задачи вече превъзхожда човешките възможности, въпреки че все още изостава в по-сложни области. Индустрията остава основен двигател на авангардните изследвания в областта на ИИ, като инвестициите в генеративен ИИ отбелязват значителен скок. САЩ запазва водещата си позиция като източник на най-добрите ИИ модели, докато се наблюдава нарастване на регулациите, свързани с ИИ. Проучванията демонстрират, че ИИ повишава производителността на работниците и води до по-качествена работа, като същевременно ускорява научния прогрес. Въпреки тези позитивни тенденции, обществеността става все по-наясно с потенциалното въздействие на ИИ и изразява нарастваща нервност по отношение на неговото развитие[2].

За целите на настоящия анализ ще се обърне внимание на **Глава 3** и **Глава 4** от доклада "AI Index 2024", които разглеждат развитието на отговорен ИИ и икономическото въздействие на технологията, като се фокусира върху конкретни механизми, чрез които ИИ трансформира работните процеси и индустриите.

Глава 3 на доклада AI Index разглеждаща отговорния ИИ, разкрива няколко ключови и взаимосвързани предизвикателства в областта. Липсата на стандартизирани оценки за отговорността на големите езикови модели (LLM) сериозно затруднява обективното сравнение на рисковете и ограниченията между различните модели. Лесното генериране и трудното откриване на политически дълбоки фалшификати (deepfakes) представлява нарастваща заплаха за изборните процеси и общественото мнение[3]. Друго предизвикателство, което изследванията разкриват е, че уязвимостите в LLM са по-комплексни от очакваното, което изисква по-задълбочени методи за тестване и оценка на сигурността. Това води до нарастващи притеснения сред бизнесаь породени от рисковете, свързани с ИИ, като основните опасения са свързани с поверителността, сигурността на данните и надеждността. Докладът установява, както че LLM могат да генерират съдържание, защитено от авторско право, което повдига важни правни

[1] Maslej, N. et al., 2024. *Artificial Intelligence Index Report 2024*, Human-Centered Artificial Intelligence. United States of America. Retrieved from https://hai.stanford.edu/ai-index/2024-ai-index-report
[2] Maslej, N. et al., 2024. *Artificial Intelligence Index Report 2024*, Human-Centered Artificial Intelligence. United States of America. Retrieved from https://hai.stanford.edu/ai-index/2024-ai-index-report
[3] Ibid

въпроси, така и ниска степен на прозрачност от страна на разработчиците на ИИ, особено по отношение на данните за обучение и методологиите, което възпрепятства по-нататъшното разбиране на устойчивостта и безопасността на ИИ системите[4]. Дебатът между краткосрочни и дългосрочни заплахи, породени от ИИ, затруднява анализирането на екстремните рискове от ИИ. Всички тези данни подчертават необходимостта от по-голям фокус върху разработването на стандартизирани оценки за отговорност, справянето с рисковете от дълбоки фалшификати и други форми на злоупотреба, повишаване на прозрачността при разработването на ИИ и задълбочено разбиране на потенциалните краткосрочни и дългосрочни рискове, за да се гарантира безопасното и етично развитие на изкуствения интелект.

Глава 4 от доклада AI Index разглеждаща икономическите ефекти на ИИ, като подчертава, че внедряването на генеративни ИИ инструменти води до значителни подобрения в производителността. Особено важно е, че най-голям ефект се наблюдава при служители с по-малък опит, което индикира потенциала на ИИ да демократизира производителността и да намали разликата в уменията.[5] Освен количественото увеличение на производителността, докладът отчита приноса на ИИ за подобряване на качеството на работа, осигурявайки бърз достъп до важна информация от практическо значение и така подпомага вземането на по-информирани решения. Докладът потвърждава, че ИИ вече не е нишов технологичен тренд, а фундаментална сила, която преобразява различни сектори на икономиката. Компаниите, които успешно интегрират ИИ в своите операции, отчитат конкретни икономически ползи, включително увеличени приходи (59% от анкетираните организации) и оптимизирани разходи (42% от анкетираните организации)[6]. Това свидетелства за реалната възвръщаемост на инвестициите в ИИ технологии и стимулира по-нататъшното им внедряване. Въпреки общото намаление на обявите за работа (основно от водещите технологични компании в САЩ), докладът констатира устойчиво търсене на специфични ИИ умения, като машинно обучение и изкуствен интелект[7]. Това е ключов индикатор за продължаващата нужда от квалифицирани специалисти в тази област. Въпреки спада в общите корпоративни инвестиции, докладът отбелязва стабилност в инвестициите в ИИ стартъпи, особено в областта на генеративния ИИ. Това показва, че инвеститорският интерес към иновациите в тази сфера остава силен. Нарастващият процент на компании, внедрили ИИ в поне една бизнес функция (55% през 2023 г., значителен ръст в сравнение с 20% през 2017 г.)[8], е ясен знак за ускореното приемане на ИИ в корпоративния свят. Докладът обръща внимание и на нарастващата индустриална автоматизация, където Китай е безспорен лидер по внедрени индустриални роботи. Ръстът на колаборативните и сервизните роботи допълнително илюстрира разширяващото се приложение на ИИ в производствените процеси.[9]

4.1.2 SCImago Journal & Country Rank

В условията на нарастваща глобална конкуренция и ускорено развитие на технологии, ИИ заема централно място не само в икономическото развитие и трансформацията на публичната администрация, но и в научноизследователската дейност. Научните изследвания и международното им влияние в областта на ИИ се превръщат в стратегически индикатори за иновационния капацитет на една държава. Анализът на данните от SCImago Journal & Country Rank, конкретно за категорията

[4] Ibid
[5] Maslej, N. et al., 2024. *Artificial Intelligence Index Report 2024*, Human-Centered Artificial Intelligence. United States of America. Retrieved from https://hai.stanford.edu/ai-index/2024-ai-index-report
[6] Ibid
[7] Ibid
[8] Ibid
[9] Maslej, N. et al., 2024. *Artificial Intelligence Index Report 2024*, Human-Centered Artificial Intelligence. United States of America. Retrieved from https://hai.stanford.edu/ai-index/2024-ai-index-report

"Изкуствен интелект" и фокус върху страните от Източна Европа, предлага ясна картина за разпределението на научната активност и влиянието на научните общности в региона.

Регионът на Източна Европа демонстрира широко разнообразие както в количеството, така и в качеството на научните публикации в областта на ИИ. Водещи по обем на продукцията са Русия и Полша, съответно с 15,909 и 14,237 научни публикации. [10] Това потвърждава съществуването на сериозни научни школи и инфраструктура в тези държави, които инвестират в дългосрочен план в развитието на ИИ като стратегическа научна област. По отношение на научното влияние, измервано чрез среден брой цитирания на публикация и h-индекс, се откроява Словения. Страната има най-високо средно цитиране от 23.25 на документ и h-индекс от 95, което я позиционира като държава с изключително силна академична репутация в сферата на ИИ.[11] Този резултат подчертава значението на фокусирани усилия върху качеството на научната работа, международното сътрудничество и публикуване в престижни издания.

България заема осмо място в Източна Европа по брой научни публикации в областта на изкуствения интелект, с общо 3,045 документа.[12] Това количество поставя страната в средната група по обем, което показва наличие на активна изследователска дейност. Въпреки това, качествените показатели остават по-ниски спрямо водещите страни. Средният брой цитирания на документ е едва 4.20 – значително под нивото на Словения и дори под регионалната средна стойност. H-индексът е 42, което показва умерено влияние на българските изследвания в международен контекст [13] . Тези стойности говорят за ограничено въздействие и видимост в глобалната академична екосистема.

Макар България да демонстрира стабилно присъствие в количествено отношение, данните от SCImago ясно показват необходимостта от повишаване на качеството, въздействието и международната разпознаваемост на българските научни изследвания в сферата на изкуствения интелект.

4.1.3 Anthropic AI index

ИИ бързо се превръща в съществен фактор, променящ структурата на съвременния трудов пазар. Анализът на данните от **Anthropic Economic Index**, публикуван през февруари 2025 г., предоставя ключова информация за начина, по който ИИ се използва в различни професионални сфери, както и за неговото потенциално въздействие върху заетостта, производителността и професионалното развитие. Anthropic е един от световните лидери в разработването на ИИ модели - Claude, основан през 2021 г., който се занимава с изследвания в областта на безопасността при разработването и използването на ИИ модели. При изготвянето на индекса Anthropic анализира милиони анонимни разговори в Claude.ai., като набора от данни е open-source.

Данните от Anthropic Economic Index сочат, че към момента ИИ се използва най-интензивно в дейности, свързани със софтуерна разработка и техническо писане, където задачите често са структурирано формализирани и податливи на автоматизация или интелигентно подпомагане[14]. Около 36% от професиите използват ИИ в поне една четвърт от своите задачи, а приблизително 4% – в над три четвърти от ежедневната си работа. Особено важно разграничение се прави между допълване (augementation) и

[10] SCImago Journal & Country Rank, "Artificial Intelligence; Eastern Europe, 1996-2023" 2024 https://www.scimagojr.com/countryrank.php?category=1702&area=1700&region=Eastern%20Europe 13.03.2025
[11] Ibid
[12] SCImago Journal & Country Rank, "Artificial Intelligence; Eastern Europe, 1996-2023" 2024 https://www.scimagojr.com/countryrank.php?category=1702&area=1700&region=Eastern%20Europe 13.03.2025
[13] Ibid
[14] Anthropic, "*Anthropic Economic Index*" 2025, Retrieved from https://www.anthropic.com/news/the-anthropic-economic-index , 13.03.2025

автоматизация (automation): данните показват, че ИИ се използва в 57% от случаите за подпомагане на човешкия труд и само в 43% за пълна автоматизация на задачи[15]. Това поставя основата за нов модел на сътрудничество между човека и машината, при който технологиите не заместват, а разширяват възможностите на работещите.

Анализът на Anthropic показва, че ИИ е най-разпространен в професии със средни до високи доходи, като компютърни програмисти, инженери, анализатори и специалисти по данни.[16] Същевременно, използването на ИИ е по-слабо в най-ниско и най-високо платените професии, което подчертава потенциален риск от засилване на социалното и икономическото неравенство при липса на адекватни образователни и социални политики. От друга страна, технологиите предлагат възможност за подобряване на баланса между работа и личен живот, особено чрез автоматизация на рутинни дейности. Това позволява на служителите да се фокусират върху по-креативни, стратегически и удовлетворяващи аспекти на своята работа, с потенциални ползи за психичното здраве и работната мотивация.

Въпреки широко разпространените опасения за масова загуба на работни места, доказателствата сочат, че ИИ се използва основно за допълване, а не за заместване на работната ръка. Това дава основание да се прогнозира, че много от съществуващите роли ще бъдат трансформирани, а не елиминирани[17]. Освен това, ще възникнат нови професии и специализации, които ще изискват съвместна работа между хора и ИИ системи, което се потвърждава и в AI Index 2024 на Станфорд[18]. Развитието на такива хибридни професии ще изисква динамично и гъвкаво професионално образование, ориентирано към дигитални умения, етика в технологиите, междуфункционални компетентности и критично мислене.

На глобално ниво, както и в рамките на Европейския съюз, интеграцията на ИИ в индустрията и публичната администрация има потенциал да повиши производителността, да съкрати оперативните разходи и да ускори дигиталната трансформация. В същото време, обаче, съществува риск от задълбочаване на неравенствата, поради различната степен на достъп до технологии и знания между отделни социални групи и региони. Този риск може да бъде смекчен чрез целенасочени образователни програми, подкрепа за преквалификация и насърчаване на дигитализация със социално включване (inclusivity). Особено важно е в държавната политика да се заложи подкрепа за най-засегнатите сектори – не чрез защита от промяната, а чрез инвестиции в адаптация и развитие.

ИИ няма да замени хората – ще замени конкретни задачи и ще създаде нови възможности. Подготовката на хората за съвместна работа с ИИ не е просто икономическа необходимост – тя е въпрос на социална устойчивост, демократична стабилност и дългосрочен икономически растеж и конкурентоспособност. Политиките, които ще се приемат днес, ще определят дали ИИ ще бъде инструмент за обогатяване на човешкия труд – или за задълбочаване на социалните предизвикателства.

4.1.4 Oxford Insights – Government AI Readiness Index 2024

С цел да се очертае ясна картина за предизвикателства и възможностите, които изкуственият интелект носи със себе си, ще разгледаме и доклада **Government AI Readiness Index 2024** на Oxford Insights, който разкрива ключовите глобални и регионални тенденции.

През 2024 г. се наблюдава значително активизиране по отношение на националните ИИ стратегии. Общо 12 нови стратегии са били публикувани или обявени,

[15] Ibid
[16] Ibid
[17] Ibid
[18] Maslej, N. et al., 2024. *Artificial Intelligence Index Report 2024*, Human-Centered Artificial Intelligence. United States of America. Retrieved from https://hai.stanford.edu/ai-index/2024-ai-index-report

което представлява тройно увеличение спрямо предходната година, като повече от половината от тях произлизат от страни с ниски и средни доходи като Етиопия, Нигерия, и Узбекистан.[19] Това ясно показва, че ИИ вече не е привилегия на богатите нации, а глобално разпознат инструмент за развитие, управление и подобряване на обществените услуги. Същевременно, САЩ продължават да заемат лидерска позиция в индекса с резултат от 87.03, значително над глобалното средно ниво от 47.59 точки, а държави с по-нисък доход постигат реален напредък чрез изграждане на основите на ИИ управлението – чрез ясно дефинирани стратегии, етични рамки и подобрен достъп до данни.[20] Изводът е, че глобалното разделение в ИИ готовността вече не се основава толкова на политическа воля или стратегическа визия, а главно на разликите в технологичния капацитет, иновационната екосистема и човешкия капитал.

Анализът на ИИ готовността сред водещите глобални сили показва отчетливи различия във фокуса и силните страни. САЩ продължава да доминира, благодарение на зрелостта на своя технологичен сектор, силната си иновационна култура и балансирана регулаторна рамка. В същото време, Европейският съюз, със среден резултат от около 70 точки, демонстрира стабилност и висока ефективност по отношение на етичните стандарти, управление и правна регулация.[21] Китай е на сравнително близко ниво (72.01), с добре балансирано представяне, особено в инфраструктурата (80.18).[22] Въпреки това, ограничената публичност и прозрачност на данните за Китай представлява предизвикателство за пълната оценка на реалния му капацитет. Изводът е, че ЕС има силен институционален капацитет и регулаторна рамка (Акт за ИИ), но остава в сянката на САЩ, що се отнася до иновационен и технологичен капацитет. Необходима е целенасочена политика за стимулиране на стартиращи екосистеми и частно-обществено партньорство в ИИ сектора.

Регионът на Източна Европа отчита средна стойност от 57.88, което го позиционира над глобалната средна стойност. Най-силно представящата се страна е Естония (72.62) – единствена в региона в топ 25 в света, следвана от Чехия (70.23), Полша (67.51) и Литва (67.80).[23] Технологичният сектор остава слабо звено на региона – с резултат от 39.03, което го поставя сред най-ниските в световен мащаб. Това ясно посочва нуждата от инвестиции в иновации, стимули за стартиращи компании и развитие на човешкия капитал в STEM области. Важно е да се отбележи, че все повече държави от региона формулират или обновяват своите ИИ стратегии – примери за това са Молдова, Румъния и Полша, които акцентират както върху икономическата конкурентоспособност, така и върху етичната употреба на ИИ. С резултат от 60.64 точки България се позиционира над регионалната средна стойност (57.88) и над глобалната (47.59). По отношение на трите стълба, страната показва силно представяне в: "Правителство"- 65.19, "Данни и инфраструктура" - 78.85, но сериозно изостава в "Технологичен сектор" - едва 37.88.[24]

Докладът ясно показва, че България е на кръстопът: разполага с институционални и инфраструктурни предимства, но се нуждае от дълбока и бърза технологична трансформация, за да навакса сериозното си изоставане. Въвеждането и следването на национална ИИ стратегия с измерими цели, развитие на предприемаческа екосистема и инвестиции в човешкия капитал са наложителни за издигането на страната сред дигитално напредналите нации.

[19] Oxford Insights, "Government AI Readiness Index 2024", 2024, Retrieved from https://oxfordinsights.com/ai-readiness/ai-readiness-index/
[20] Oxford Insights, "Government AI Readiness Index 2024", 2024, Retrieved from https://oxfordinsights.com/ai-readiness/ai-readiness-index/
[21] Ibid
[22] Ibid
[23] Ibid
[24] Oxford Insights, " Government AI Readiness Index 2024", 2024, Retrieved from https://oxfordinsights.com/ai-readiness/ai-readiness-index/

4.2. Тенденции при развитието на ИИ

MIT Sloan Management Review очертава пет ключови тенденции в областта на ИИ и науката за данни за 2025 г., които ще оформят бъдещата стратегия на организациите. Една от водещите тенденции е възходът на **агентния ИИ**, характеризиращ се със способността на системите да изпълняват задачи автономно.[25] Въпреки значителния потенциал, който тази технология носи, нейното ефективно внедряване и управление представляват съществени предизвикателства, които изискват внимателно планиране и експертиза.

Нарастващото осъзнаване на необходимостта от доказване на икономическата ефективност на ИИ инициативите води до засилен фокус върху **измерването на възвръщаемостта на инвестициите (ROI)**.[26] Организациите все повече търсят начини за количествено оценяване на стойността, която проектите с ИИ генерират, което е критично за вземане на информирани решения относно бъдещи инвестиции и стратегически насоки.

Въпреки технологичния напредък, много компании все още се сблъскват с предизвикателството за изграждане на **култура, базирана на данни**.[27] Създаването на среда, в която решенията се основават на задълбочен анализ на данни, изисква не само подходящи инструменти и технологии, но и силна ангажираност от страна на ръководството и инвестиции в обучението на персонала за ефективно използване на данните в процеса на вземане на решения. Друга важна тенденция е свързана с **управлението на неструктурирани данни**.[28] С експоненциалното нарастване на обема на данни под формата на текст, изображения, видео и аудио, организациите трябва да разработят ефективни стратегии за съхранение, обработка и анализ на тези данни, за да извлекат ценни инсайти и да получат конкурентно предимство. Накрая, ролите и отговорностите на **лидерските позиции в областта на ИИ и науката за данни** продължават да еволюират.[29] Организационните структури се адаптират, за да отразят нарастващото значение на тези области, което води до промени в йерархиите, отговорностите и начините за отчитане на резултатите. Тези очертаващи се тенденции подчертават необходимостта организациите да бъдат гъвкави и да възприемат проактивен подход към ИИ и управлението на данни. Инвестирайки стратегически в подходящи технологии, изграждайки култура, ориентирана към данните, и развивайки адекватно лидерство, компаниите могат да се възползват максимално от възможностите, които ИИ предлага, и да поддържат своята конкурентоспособност в бързо променящата се бизнес среда.

На 18 март 2025 г., по време на GTC AI Conference, Дженсен Хуанг, изпълнителният директор и основател на NVIDIA, водеща световна компания в областта на високопроизводителния компютърен хардуер и ключов участник в глобалната ИИ екосистема, представи визията на компанията за бъдещето на технологиите в ерата на изкуствения интелект и очерта технологичните тенденции. В своята реч Хуанг акцентира върху фундаменталната роля на ускорените изчислителни мощности като движеща сила за по-нататъшния прогрес в ИИ, тъй като традиционните методи за общо изчисление вече не могат да отговорят на нарастващите нужди на съвременните, все по-сложни ИИ модели.[30]

---

[25] MIT Sloan, "Five Trends in AI and data science for 2025", 2025, https://sloanreview.mit.edu/article/five-trends-in-ai-and-data-science-for-2025/ 12.03.2025
[26] Ibid., MIT Sloan
[27] MIT Sloan, "*Five Trends in AI and data science for 2025*", 2025, https://sloanreview.mit.edu/article/five-trends-in-ai-and-data-science-for-2025/ 12.03.2025
[28] Ibid
[29] Ibid
[30] "*GTC March 2025 Keynote with NVIDIA CEO Jensen Huang*", 2025, https://www.nvidia.com/gtc/keynote/ 24.03.2025

Основен фокус се поставя върху практическите приложения на ИИ в широк спектър от сектори, включително медицина, роботика, автомобилна индустрия и научни изследвания. Внедряването на ИИ технологии в тези области вече създават нови възможности и решават комплексни проблеми. Ключов елемент от визията на NVIDIA, е значението на софтуерните инструменти и библиотеки, за улесняване на процеса на разработка и внедряване на ИИ решения, като по този начин тези софтуерни екосистеми се превръщат в катализатор за трансформацията на всички индустрии чрез ИИ.[31]

Темата за бъдещето на взаимодействието „човек-компютър" е водеща, защото ИИ ще позволи по-интуитивни и естествени форми на комуникация чрез глас, зрение и други сензорни модалности. Тази еволюция ще доведе до появата на нови професионални направления, свързани с разработването, внедряването и поддръжката на ИИ системи, което налага необходимостта от мащабна преквалификация на работната сила за адаптиране към тези променящи се изисквания.[32]

В допълнение, нараства значението на платформите за създаване на симулационни среди и дигитални близнаци, които ще играят ключова роля в развитието на автономни системи и роботика, предоставяйки инструменти за симулация, обучение и дизайн. Според Хуанг, тези процеси на адаптация и трансформация вече са в ход и не представляват бъдещи предизвикателства, пред които бизнесът тепърва ще се изправя. Напротив, успехът ще бъде на страната на организациите, които вече активно се адаптират към новата ера на ИИ. Дженсен Хуанг представи ясна траектория за развитието на ИИ, подчертавайки значението на ускореното изчисление, софтуерната екосистема и реалните приложения в различни индустрии, като акцентира върху необходимостта от непрекъсната адаптация и инвестиции в нови умения, за да се възползва човечеството от пълния потенциал на ИИ.[33]

Статистическите данни за изкуствения интелект към 2025 г., предоставени от Hostinger, очертават ясна картина на бърз и мащабен растеж. Прогнозите сочат, че глобалният пазар на ИИ ще продължи да нараства с впечатляващ годишен темп, достигайки 305,9 милиарда долара в края на 2024 г. и се очаква да се разширява със средногодишен темп на растеж от 28,46% в периода 2024-2030 г.[34] Тези темпове на растеж потвърждават динамичното развитие на сектора и нарастващото му значение в световната икономика. Очаква се ИИ да окаже съществено икономическо въздействие, като до 2030 г. се прогнозира принос от над 15,7 трилиона долара към световната икономика. Тази мащабна проекция подчертава трансформиращия потенциал на ИИ в различни индустрии. Очаква се, например, в здравеопазването ИИ да намали разходите за лечение с до 50%.[35]

От гледна точка на пазара на труда, ИИ се очаква да има двоен ефект. Прогнозите сочат създаването на значителен брой нови работни места, достигайки 97 милиона до 2025 г. и 133 милиона до 2030 г. Въпреки този оптимизъм, съществуват и опасения относно потенциалната замяна на работни места, като се цитира цифра от до 85 милиона работни места по света.[36] Тези противоречиви прогнози подчертават необходимостта от проактивни мерки за преквалификация и адаптация на работната сила.

Инвестициите в ИИ остават ключов двигател на растежа. Бизнесите вече отделят до 20% от технологичния си бюджет за ИИ, а значителен процент от компаниите (58%) планират да увеличат инвестициите си през 2025 г. Генеративният ИИ се очертава като водеща технология, използвана от 51% от компаниите за създаване на съдържание,

[31] Ibid
[32] Ibid
[33] Ibid
[34] "*AI Statistics and trends: New Research for 2025*", 2025, https://www.hostinger.com/tutorials/ai-statistics 11.03.2025
[35] Ibid
[36] Ibid

клиентска поддръжка и автоматизация. Прогнозите за пазара на генеративен ИИ са изключително високи, достигайки 1,3 трилиона долара до 2032 г. Приемането на ИИ в бизнеса също е във възходяща фаза. Високият процент изпълнителни директори (86%) вярват в значителното влияние на ИИ върху техния бизнес. В допълнение, ИИ има потенциал да увеличи производителността на бизнеса с до 40%, а 37% от бизнес лидерите посочват ИИ като стимул да повишат квалификацията на своите служители през следващите две или три години.[37]

Данните на Hostinger ясно показват, че тази технология е на път да претърпи експоненциален растеж и да окаже дълбоко въздействие върху икономиката и пазара на труда. Въпреки потенциалните предизвикателства, свързани със замяната на работни места и опасенията за сигурност и пристрастия, инвестициите и внедряването на ИИ продължават да се ускоряват, което затвърждава ключовата роля на тази технология в бъдещето. Очертава се ясна картина на динамично развитие в областта на изкуствения интелект. Тенденциите сочат към все по-автономни ИИ системи, нужда от измерима икономическа ефективност на ИИ инициативите, критична важност на култура, базирана на данни, и ефективно управление на нарастващите обеми неструктурирани данни.

Технологичният напредък се движи от ускорени изчислителни мощности, което ще е ключа към успешната дигитална трансформация, която ще засяга всички индустрии.

Очаква се ИИ да доведе до значителен ръст на глобалния пазар, генерирайки трилиони долари и създавайки милиони нови работни места, въпреки потенциалното изместване на други. [38] Инвестициите в ИИ продължават да нарастват, като генеративният ИИ се очертава като доминираща технология, а бизнесите все повече осъзнават необходимостта от адаптация и преквалификация на персонала, за да се възползват от пълния потенциал на ИИ в ерата на дигитална трансформация.[39]

[37] "*AI Statistics and trends: New Research for 2025*", 2025, https://www.hostinger.com/tutorials/ai-statistics 11.03.2025
[38] Ibid
[39] Ibid

5. Регулации

ИИ е трансформираща технология, която навлиза дълбоко в живота ни, което налага внимателно разглеждане на регулаторните рамки в световен мащаб, с цел намиране на оптималния регулаторен баланс. Тези мерки са стратегически важни за бъдещото развитие на ИИ, балансирането на иновациите с обществената безопасност и справянето с етичните и социално-икономическите последици. В епоха на глобална взаимосвързаност, познаването и сравнителният анализ на регулаторните подходи на различните страни са ключови за международното сътрудничество и предотвратяването на регулаторна фрагментация. Чрез анализ на тези рамки ще се идентифицират тенденции и потенциални модели за ефективно управление на ИИ, което да максимизира ползите и минимизира вредите, като дава възможност за формиране на балансирани политики, които изграждат доверие в технологиите, същевременно позволявайки тяхното развитие.

Настоящият анализ ще изследва регулаторните рамки, включително предстоящи законодателни инициативи, на държавите-членки на **Международната мрежа на институтите по безопасен ИИ**, създадена на 20 ноември 2024 г. в Сан Франциско, САЩ.[40] Мрежата обединява представители на Австралия, Великобритания, Европейския съюз, Канада, Кения, Република Корея, САЩ, Сингапур, Франция и Япония. Независимо че Китай не е част от тази мрежа, регулаторните рамки в страната също ще бъдат обект на анализ, предвид мащаба и влиянието на китайската икономика и технологичен прогрес върху глобалните процеси. Ще бъдат разгледани и международните инициативи на Г-7 и ОИСР, които са основополагащи за развитието на регулаторните рамки в голям брой държави по света. Поради ограничението върху обема на настоящия анализ, няма да бъдат разгледани всички регулаторни мерки и инициативи на съответните държави, а само водещите регулации и предстоящи такива, и стратегиите, в чиито рамки се разработват тези регулации. Допълнително ограничение на анализа е, че няма да бъдат изследвани регулациите на ниво щат, регион и/или област в изброените държави. Анализът се фокусира върху регулаторните рамки на ЕС, САЩ и Китай, като в *Приложение 3* са анализирани регулаторните рамки на останалите държави-членки на Международната мрежа на институтите за безопасен ИИ. В допълнение, *Приложение 4* включва прогнозни сценарии на базата на текущото развитие на регулаторните рамки на съответните държави.

5.1. G7 Hiroshima AI Process

В рамките на последователните срещи на страните от Г-7 се наблюдава значителна еволюция в подхода към изкуствения интелект, започвайки с принципите на Хирошимския процес за ИИ (G7 Hiroshima AI Process, (HAIP)) през 2023 г., които послужиха като основополагаща рамка от политически ангажименти.[41] [42] Тези принципи очертават ключови насоки за разработването и управлението на усъвършенствани ИИ системи, като се акцентира върху управлението на риска през целия жизнен цикъл на ИИ. Освен това, принципите подчертават важността на мерките за сигурност и защитата на личните данни, като признават чувствителния характер на информацията, използвана и генерирана от ИИ.[43] Отговорността, прозрачността и споделянето на информация между всички заинтересовани страни са определени като съществени за изграждане на

[40] "First meeting of the International Network of AI Safety Institutes" , 2024 https://digital-strategy.ec.europa.eu/en/news/first-meeting-international-network-ai-safety-institutes 10.03.2025
[41] "G7 Leaders' Statement on the Hiroshima AI Process" , 2023 https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document01_en.pdf 10.03.2025
[42] "Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI System ", 2023 https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document04_en.pdf 10.03.2025
[43] Ibid., Водещи Принципи

доверие и осигуряване на отчетност.[44] Тези принципи са строго залегнати в разработения Международен кодекс за поведение публикуван заедно с обявяването на HAIP.[45] Документът от Хирошима също така призова за инвестиции в сигурност, етика и социални изследвания, за да се адресират по-ефективно потенциалните въздействия на ИИ, и да се подкрепи разработването на международни технически стандарти и етични рамки, което цели да осигури съгласуван глобален подход.[46]

Последвалата Декларация от министрите на индустрията и цифровите технологии във Верона и Тренто през 2024 г. надгражда тези фундаментални принципи, като преминава към по-конкретни действия и оперативни механизми, насочени към подпомагане на безопасното внедряване и използване на ИИ.[47] В инициативата се включват и нови партньори като Бразилия, Украйна и Република Корея, както и представители на ООН, което сигнализира за разширяване на глобалния диалог и ангажираност по темата. Декларацията задълбочава подхода към икономическото въздействие на ИИ, като поставя специален акцент върху индустриалната конкурентоспособност и потенциала на ИИ за стимулиране на растежа.[48] Важен елемент е управлението на данни чрез принципа „Свободен поток от данни с доверие" (DFFT), който има за цел да улесни трансграничния обмен на данни при спазване на стандартите за защита.[49] Документът предлага и конкретни инструменти за изпълнение на целите, включително създаване на доклади, анализи, инструментариуми, сборници, както и организиране на срещи и контактни групи за улесняване на сътрудничеството, като също така признава нуждата от „институционализиране" на ИИ управлението чрез създаване на механизми за наблюдение, координация и евентуално регулиране на ИИ технологиите.[50]

И двата документа се фокусират върху намирането на деликатен баланс между насърчаването на иновациите и необходимостта от ефективна регулация. От една страна, Г-7 се стреми да подкрепи иновациите чрез насърчаване на отвореност, международно сътрудничество, фундаментални и приложни научни изследвания, осигуряване на достъп до необходимата инфраструктура и изграждане на сигурни мрежи за обмен на данни и знания. От друга страна, съществува ясна ангажираност за ограничаване на потенциалните рискове, свързани с ИИ, чрез прилагане на мерки за управление на сигурността, осигуряване на прозрачност на ИИ системите, установяване на ясна отговорност за тяхното използване, провеждане на оценки на въздействието и гарантиране на защитата на основните права и свободи на гражданите. Това двустълбово развитие се очертава като централен елемент на новия международен ред в областта на ИИ, където икономическата и технологичната конкурентоспособност се разглеждат като неразривно свързани с етичните и регулаторни рамки.

Нещо повече, страните от Г-7 демонстрират ясно изразено желание да наложат глобална управленска рамка за ИИ, която да се основава на споделени демократични ценности, принципите на върховенството на закона и безусловното зачитане на човешките права. Те активно насърчават оперативната съвместимост между националните регулации и международните стандарти, за да се избегне създаването на фрагментиран и несъгласуван глобален ИИ пейзаж. Същевременно се подкрепя

[44] "Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI System ", 2023 https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document04_en.pdf 10.03.2025
[45] "Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems", 2023, https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document05_en.pdf 1.03.2025
[46] Ibid - Кодекс
[47] "Ministerial declaration", 2024, https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document07_en.pdf , 14.03.2025
[48] Ibid - Декларация
[49] Ibid - Декларация
[50] "Ministerial declaration", 2024, https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document07_en.pdf , 14.03.2025

създаването на отворена и сигурна интернет среда, като се отхвърлят тенденциите към технологична изолация и фрагментация на глобалната мрежа. В този контекст, Г-7 се превръща де факто в архитект на международната ИИ политика, като други ключови международни организации като ОИСР, Програмата на ООН за развитие, ЮНЕСКО и Международния телекомуникационен съюз (ITU) са привлечени като важни партньори и координатори за постигане на тези глобални цели, заедно с Република Корея, Украйна, Бразилия и ОАЕ. Страните от Г-7 не само формулират свои собствени национални политики, но и се стремят да наложат стандарти, които да станат универсални и приложими в глобален мащаб, оказвайки значително влияние върху бъдещото развитие и регулиране на ИИ по целия свят.

5.2. Обсерватория за политики в областта на изкуствения интелект (ОИСР)

В началото на 2020 г. беше създадена Обсерваторията за политики за изкуствен интелект (ИИ) на Организацията за икономическо сътрудничество и развитие (ОИСР), която представлява водеща международна платформа за сътрудничество и стандартизация в областта на ИИ.[51] Чрез ангажиране на най-развитите икономики в света, успоредно с развиващите се страни, които полагат значителни усилия за икономически и обществен прогрес, Обсерваторията играе ключова роля в глобалното оформяне на политиките в сферата на ИИ, като разработва и каталог с инструменти и показатели за надежден ИИ, който всеки може да използва при разработването, внедряването и използването на ИИ, с цел спазване на етичните принципи за безопасен и надежден ИИ. Следва да се отбележи, че България участва в този важен форум не като самостоятелен субект, а като интегрална част от Европейския съюз, въпреки че други страни-членки на ЕС участват и самостоятелно. В настоящия момент Обсерваторията се председателства съвместно от Словакия и Сърбия, което подчертава многостранния и приобщаващ характер на нейната дейност.[52]

Принципите на ОИСР за ИИ са насочени към насърчаване на иновативен и надежден ИИ в полза на хората и планетата. Те се основават на ценности като приобщаващ растеж, устойчиво развитие, благосъстояние и зачитане на човешките права и демократичните ценности.[53]

Основните принципи за отговорно управление на надежден ИИ, препоръчани от ОИСР, включват отчетност на ИИ участниците за правилното функциониране на ИИ системите и осигуряване на сигурност и безопасност през целия им жизнен цикъл.[54] Те, също така, подчертават необходимостта от справедливост и прозрачност при разработването и използването на ИИ, както и зачитане на върховенството на закона, човешките права и демократичните ценности. ОИСР отправя и препоръки към правителствата и заинтересованите страни, които са свързани с инвестиране в ИИ изследвания и развитие, насърчаване на дигитална екосистема за ИИ, оформяне на гъвкава политическа и регулаторна среда, укрепване на уменията за бъдеще, основано на ИИ, и насърчаване на международното сътрудничество в областта.[55]

Принципите на ОИСР за ИИ играят ключова роля в оформянето на глобалния ИИ пейзаж, като предлагат ценни насоки за правителствата, организациите и отделните лица, ангажирани с разработването и използването на ИИ, с цел да се гарантира, че тази технология служи на общото благо и се развива по надежден и етичен начин, благодарение на създаването на по-хармоничен и съгласуван международен подход към управлението на ИИ.

[51] "The OECD Artificial Intelligence Policy Observatory" https://oecd.ai/en/ , 11.03.2025
[52] "About the Global Partnership on AI (GPAI)" , 2024, https://oecd.ai/en/about/what-we-do , 11.03.2025
[53] "AI Principles", 2019, https://www.oecd.org/en/topics/sub-issues/ai-principles.html 11.03.2025
[54] Ibid - "AI Principles"
[55] "AI Principles", 2019, https://www.oecd.org/en/topics/sub-issues/ai-principles.html 11.03.2025

5.2.1 Рамка за отчитане на Г7 – Международен кодекс за поведение

Като част от HAIP, Г-7 стартира доброволна рамка за докладване, за да насърчи прозрачността и отчетността сред организациите, разработващи модерни системи за ИИ. Резултатите ще улеснят прозрачността и сравнимостта на мерките за намаляване на риска и ще допринесат за идентифициране и разпространение на добри практики.[56] Тази инициатива за сътрудничество с Г-7 очертава ОИСР като ключов международен орган за стандартизация и сътрудничество в областта на ИИ, чиито принципи са тясно свързани с инициативите на Г-7. Може да се заключи, че политиките за ИИ, разработвани в рамките на HAIP, ще окажат значително глобално влияние, като ОИСР ще играе централна роля в тяхното разпространение и прилагане за насърчаване на отговорно и етично развитие на ИИ в световен мащаб.

5.2.2 Форуми на високо равнище за безопасен ИИ

Декларацията от Блечли, подписана на първия в историята Международен форум на високо равнище за безопасност на изкуствения интелект, проведен на 1-2 ноември 2023 г., във Великобритания, постави началото на множество политики за етичен ИИ, разработвани оттогава досега. Тази декларация, подписана от държави, водещи в развитието на ИИ технологии, както и от Европейския съюз, изразява общо разбиране за спешната необходимост от идентифициране и съвместно управление на потенциалните рискове, свързани с ИИ, особено с т.нар. "граничен/авангарден ИИ" (frontier AI).[57]

Декларацията подчертава както безпрецедентните възможности, които ИИ предлага за прогрес и ползи за човечеството, така и сериозните рискове, свързани с неговото развитие и използване, включително по отношение на човешките права, справедливостта, прозрачността, безопасността, отчетността, етиката и смекчаването на пристрастията. Декларацията отразява общото убеждение, че нито една част от обществото не може да се справи сама с въздействието на граничния ИИ и че реализирането на неговия потенциал изисква устойчивото внимание и сътрудничество на правителствата, бизнеса, академичните среди и гражданското общество.[58]

Сеулската декларация, приета на срещата на върха за ИИ в Република Корея на 21 май 2024 г., потвърждава общия ангажимент за международно сътрудничество и диалог в областта на изкуствения интелект. Декларацията надгражда работата от срещата на върха в Блечли Парк, като признава взаимосвързаността на безопасността, иновациите и приобщаването на ИИ като ключови приоритети в международното управление на ИИ.[59] Подчертава се важността на оперативната съвместимост между ИИ рамките за управление, базирани на подход, ориентиран към риска, и се потвърждава подкрепата за прилагането на Международния кодекс за поведение от HAIP за организации, разработващи усъвършенствани ИИ системи, като отбелязват особената отговорност на тези, които разработват и внедряват граничен/ авангарден ИИ.[60] Също така се обръща внимание на значението на активната многостранна колаборация и ангажираността с други международни инициативи като ООН, Г7, Г20 и ОИСР.

[56] "G7 reporting framework – Hiroshima AI Process (HAIP) international code of conduct for organizations developing advanced AI systems", 2024, https://transparency.oecd.ai/ , 11.03.2025
[57] "The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023", 2023, https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023 , 11.03.2025
[58] Ibid
[59] "Seoul Declaration for Safe, Innovative and Inclusive AI by Participants Attending the Leaders' Session of the AI Seoul Summit, 21st May, 2024", 2024, https://www.mofa.go.kr/eng/brd/m_5674/view.do?page=1&seq=321007 , 11.03.2025
[60] "Seoul Declaration for Safe, Innovative and Inclusive AI by Participants Attending the Leaders' Session of the AI Seoul Summit, 21st May, 2024", 2024, https://www.mofa.go.kr/eng/brd/m_5674/view.do?page=1&seq=321007 , 11.03.2025

Декларацията приветства създаването и разширяването на институции за безопасност на ИИ и насърчава сътрудничеството в научните изследвания в тази област.[61]

На 11 февруари 2025 г., в рамките на **AI Action Summit в Париж**, беше приета декларация, озаглавена "Ангажимент за надежден изкуствен интелект в света на труда", която подчертава ангажимента на подписалите я страни към отговорното внедряване на ИИ технологии в работната среда.[62] Декларацията изрично се позовава на Декларацията на срещата на върха за действие по изкуствен интелект, Парижкия пакт за хората и планетата, резолюции на Общото събрание на ООН, Глобалния дигитален пакт, както и инициативи на Г7 и Г20. Тя, също така, взема предвид предстоящата препоръка на ОИСР относно ИИ на пазара на труда и съвместното изявление на ангажираните групи на Г7 – Labour 7 (L7) и Business 7 (B7).[63]

Компаниите, присъединили се към ангажимента, се задължават да насърчават социалния диалог с представители на работниците относно внедряването на ИИ, да инвестират в човешки капитал чрез обучение и преквалификация, да гарантират безопасност, здраве, автономност и достойнство на работното място, да предотвратят дискриминацията на пазара на труда, да защитят поверителността на работниците и да насърчават производителността и приобщаването във веригите на стойността. Важен акцент е поставен върху спазването на трудовите стандарти във веригата за доставки на ИИ, включително достойни заплати и безопасни условия на труд. Компаниите се ангажират да следят напредъка си чрез рамка за измерване и да си сътрудничат с правителствата чрез публично-частни партньорства.[64]

За разлика от предишните форуми за безопасност на ИИ (AI Safety Summit), този AI Action Summit има за цел да стимулира конкретни действия. Въпреки това, САЩ не подписва тази декларация. Забелязва се съществено разминаване в подхода към регулирането на ИИ между Европа (водена от инициативата на Франция) и САЩ. Европейският подход, илюстриран от тази декларация, силно набляга на социалните аспекти и защитата на работниците при внедряването на ИИ. За разлика от това, американският подход е по-скоро ориентиран към стимулиране на иновациите чрез дерегулация и запазване на лидерската позиция, с по-малко изявен акцент върху специфични мерки за защита на работниците в контекста на ИИ.

Тази декларация затвърждава тенденцията за по-регулирано и социално ориентирано развитие на ИИ в ЕС, като се стреми да гарантира, че технологичният прогрес не става за сметка на основните трудови права. В световен мащаб, отказът на САЩ да се присъедини към тази инициатива очертава съществуващите различия във вижданията за бъдещето на ИИ и неговото въздействие върху пазара на труда. Въпреки това, декларацията представлява важна стъпка към създаване на глобални норми и стандарти за надежден ИИ в света на труда и може да послужи като основа за бъдещи международни споразумения.

5.3. Съпоставка на регулаторните рамки и проблеми на регулациите

5.3.1 Европейски съюз

В рамките на настоящия анализ регулациите в ЕС относно ИИ ще бъдат разгледани само на ниво ЕС, не и на ниво държава-членка, с изключение на България и Франция.

[61] Ibid
[62] "Pledge for a Trustworthy AI in the World of Work", 2025, https://www.elysee.fr/emmanuel-macron/2025/02/11/pledge-for-a-trustworthy-ai-in-the-world-of-work 11.03.2025
[63] Ibid
[64] "Pledge for a Trustworthy AI in the World of Work", 2025, https://www.elysee.fr/emmanuel-macron/2025/02/11/pledge-for-a-trustworthy-ai-in-the-world-of-work 11.03.2025

Франция ще бъде разгледана, поради факта, че беше страната домакин на тазгодишното издание на международния форум на високо равнище за ИИ (AI Action Summit) след Великобритания през 2023 г. и Република Корея през 2024 г. Важно е да се отбележи, че от всички институти от Международната мрежа от институти за безопасен ИИ в света, единствено този на ЕС т.нар Служба по ИИ (AI Office) има регулаторни функции.

Регламент (ЕС) 2024/1689 – Акт за изкуствения интелект (AI Act)

Актът за изкуствения интелект[65] цели създаване на правна основа за сигурност, прозрачно и етично използване на системите с изкуствен интелект в ЕС. Основната му цел е да гарантира, че ИИ системите, внедрени в ЕС, отговарят на основните права, безопасността и демократичните ценности (чл. 1).

Актът забранява използването на определени ИИ практики, които се считат за неприемливо високорискови. Сред тях са подсъзнателно повлияване на човешкото поведение с цел манипулация (чл. 5, параграф 1, буква а); експлоатиране на уязвимости на специфични групи (напр., деца, хора с увреждания) (чл. 5, параграф 1, буква б); социално оценяване от публични органи (social scoring) (чл. 5, параграф 1, буква в); разпознаване на емоции на работното място и в учебни заведения (чл. 5, параграф 1, буква е), биометрична категоризация по расов, етнически, политически или религиозен признак (чл. 5, параграф 1, буква ж).

Регламентът въвежда хармонизиран подход, основан на риска, при който системите се класифицират според потенциалното им въздействие върху обществото. ИИ системите се класифицират в четири категории според риска, който представляват: неприемлив, висок, ограничен и минимален риск (чл. 6). Най-строги изисквания се прилагат към високорисковите системи, включително тези, използвани в критични инфраструктури, образованието, здравеопазването, правоприлагането и заетостта (Приложение III). Всяка високорискова ИИ система трябва да премине процедура по оценка на съответствието (чл. 43), да има система за управление на риска (чл. 9), да осигурява прозрачност (чл. 13), човешки надзор (чл. 14) и надеждност (чл. 15).

За прилагането на Акта се предвиждат: назначаване на национални компетентни органи (чл. 70); създаване на Европейски съвет по изкуствен интелект (AI Board) (чл. 65); въвеждане на механизъм за мониторинг от доставчиците след пускането на пазара (чл. 72); възможност за създаване на регулаторни лаборатории (regulatory sandboxes) за иновативни решения (чл. 57); строги санкции при неспазване – до 35 млн. евро или 7% от годишния оборот (чл. 99).

Актът за ИИ представлява регулаторна рамка, която поставя в центъра си защитата на основни човешки права, сигурността и социалната справедливост, прилагана върху технологии с изкуствен интелект. Социалният ефект на регламента се изразява в няколко ключови направления:

1. Защита на човешките права и предотвратяване на злоупотреби

Регламентът забранява използването на ИИ системи, които могат да нарушат човешкото достойнство и личната автономия – например, системи за социален рейтинг, манипулиране на поведение, биометрична категоризация по расов, етнически или политически признак, както и разпознаване на емоции в образователен или трудов контекст (чл. 5, параграф 1). Това е директен отговор на обществените притеснения относно навлизането на ИИ в личната сфера.

2. Гарантиране на човешки надзор

Актът за ИИ задължава разработчиците и потребителите на високорискови ИИ системи да гарантират, че решенията, взети от ИИ, подлежат на **ефективен човешки**

[65] OB L, 2024/1689, 12.7.2024, ELI: http://data.europa.eu/eli/reg/2024/1689/oj

надзор, с цел предотвратяване на вреди (чл. 14). Това е от критично значение за контексти като подбор на персонал, достъп до образование, социални услуги или кредитиране.

3. Прозрачност и обяснимост

Регламентът въвежда изисквания за **прозрачност на алгоритмите** (чл. 13), включително задължение за информация дали взаимодействието е с ИИ система, особено когато това може да повлияе на вземането на решения или оценката на човек. Това подкрепя правото на потребителите и гражданите на информираност, особено в чувствителни области.

4. Предпазване на уязвими групи

Особен акцент е поставен върху предотвратяването на дискриминация и неравен достъп до услуги. Чрез ограничаване на ИИ практики, които експлоатират **уязвимостите на лица, поради възраст, увреждане или социален статус** (чл. 5, параграф 1, б), актът служи като щит за социална справедливост в цифровата трансформация.

5. Укрепване на общественото доверие

Чрез сертификация, мониторинг след пускането на пазара (чл. 72) и санкционен режим (чл. 99) регламентът цели да засили **доверието в технологията**, което е предпоставка за нейното масово приемане. Това е особено важно в държави и региони с по-ниска цифрова зрялост.

Икономическият ефект на Акта за ИИ е дълбоко амбивалентен – от една страна, той поставя стабилна правна основа за доверени ИИ технологии, от друга страна – въвежда тежка административна и оперативна рамка, която може да възпрепятства иновациите и конкуренцията, особено сред МСП и стартиращите компании.

1. Разходи за съответствие

Регламентът изисква от доставчиците на високорискови системи да внедрят цялостна **система за управление на риска** (чл. 9), подробна техническа документация (чл. 11), система за проследимост (чл. 12), план за мониторинг след пускането на пазара (чл. 72) и да подлежат на оценка на съответствие (чл. 43–55). Това води до: увеличаване на **фиксираните разходи** за вход на пазара; необходимост от вътрешни или външни **екипи за правно и техническо съответствие**; повишена **правна несигурност**, свързана с бъдещи делегирани актове и стандарти. Това създава дисбаланс между големите технологични компании и МСП и стартиращите предприятия, като последните понасят по-тежко тежестта, която оказват новите задължения.

2. Инвестиционна атрактивност и глобална конкурентоспособност

Въпреки че Актът за ИИ предлага правна сигурност, **сложността и обхватът на регулацията** може да направи ЕС по-малко привлекателен за инвестиции в ИИ в сравнение със САЩ, Великобритания или Азия. Това е особено валидно в период на нарастваща конкуренция в генеративен ИИ и модели с общо предназначение.

Забавяне в приемането на **делегираните актове** би могло да създаде „регулаторна несигурност", което възпира инвеститорите и затруднява бизнес планирането.

3. Възможности за нови индустрии и етична стойност

Актът за ИИ същевременно отваря нови възможности в развитие на ИИ технологии, свързани с проверка на съответствието ("**compliance tech**") с всеобхватната и сложна регулаторна рамка на ЕС, правен ИИ, обясними и интерпретируеми ИИ системи

(Explainable AI); разрастване на пазара за **сертификационни и етични консултантски услуги**; позициониране на ЕС като **глобален еталон** за отговорен и сигурен ИИ, но това са технологии, които не се разработват с цел скалиране и оптимизиране на производствени процеси, което крие риска от още по-сериозно изоставане на ЕС в технологичен и индустриален суверенитет и да го направи зависим от чужди технологии. При добра реализация ЕС може да се превърне в износител на доверени технологии, рамки и практики за отговорен ИИ, което носи икономически ползи и глобално влияние, но само при условие че работи в изграждане и развитие на индустриален капацитет и балансиран регулаторен подход.

Актът за ИИ поставя солидна основа за изграждане на **социално отговорна и технологично устойчива икономика**, в която ИИ се използва в съответствие с човешките права и европейските ценности. От социална гледна точка, той представлява авангарден инструмент за защита на гражданите в цифровата ера. От икономическа гледна точка, обаче, той въвежда регулаторна тежест, която може да ограничи растежа и гъвкавостта на европейските ИИ иноватори – особено ако не бъде съпътстван от адекватна подкрепа, стандартизация и координирано прилагане и в дългосрочен план да намали допълнително конкурентоспособността на европейската икономика, превръщайки я в регионален играч.

Делегирани и изпълнителни актове по Акт за ИИ на ЕС

Регламентът за ИИ е подробна рамка, но той предвижда и значителен обем **вторично законодателство** – делегирани и изпълнителни актове, чрез които Комисията ще доразвива или актуализира някои технически аспекти. Фокусът занапред трябва да бъде поставен не върху нови регулации, а върху прилагането и осигуряването на необходимите ресурси, така че Актът за ИИ реално да функционира. Въпреки това, обемът от вторично законодателство, което трябва да бъде разработено, представлява сериозно предизвикателство, защото Актът за ИИ предвижда създаването на минимум още 26–27 различни акта. Някои от тези актове са по-скоро процедурни или стандартни, но има определени елементи, които са от особена важност и с висока степен на въздействие върху бизнеса и способността му да се развива.

Европейската комисия разполага с правомощия съгласно член 97 от Акта за ИИ, включително да работи съвместно с Борда по изкуствен интелект (AI Board) върху подготовката на изпълнителни и делегирани актове. Сред ключовите теми за изпълнение са: оспорването на правомощията на нотифицирания орган (чл. 37), създаването на общи технически спецификации (чл. 41), действия в случай на провал на кодексите на практика, за определяне на модел, пораждащ системни рискове по инициатива на ЕК (чл. 52 параграф 4), както и механизми за одобрение на успешни кодекси чрез изпълнителен акт, където Комисията, ако счете кодекса за неподходящ, може да установи общите правила за изпълнението на задълженията (чл. 50 параграф 7). Особено важни са и правилата за налагане на глоби спрямо доставчици на основни ИИ модели (чл. 101 параграф 6), както и детайлизирането на режима за регулаторни лаборатории (sandbox – чл. 58), тестване в реална среда (чл. 60) и процедурите по алармиране и подбор на експерти (чл. 68). От особена важност е и предстоящият изпълнителен акт за шаблона на плановете за последващ мониторинг на пазара (чл. 72).

В допълнение към изпълнителните актове, Актът за ИИ предвижда и значителен брой делегирани актове, чрез които Комисията ще може да изменя съдържанието на основни приложения и регулаторни изисквания. Сред тях са: промени в класификационните правила за високорискови ИИ системи (член 6), актуализация на Приложение III (член 7), техническата документация в Приложение IV (чл. 11) и методиката за оценка на съответствието по Приложения VI и VII (чл. 43), нови технически елементи към ЕС декларацията за съответствие (чл. 47). Предвиждат се, също така, делегирани актове за нова класификация на основните модели носещи

системен риск (чл. 51). **Чл. 52** упълномощава Комисията да изменя прага за "системен риск" при ИИ модели с общо предназначение (10^25 FLOPs) и да добавя показатели и бенчмаркове за оценка на високото въздействие, както и допълнения към технически приложения, свързани с измерване, изчисляване и доказване на съответствие (Приложение XIII). **Чл. 53** упълномощава ЕК да прави изменения в Приложения XI и XII, с оглед на технологичното развитие.

По отношение на насоките и ръководствата (guidelines), предвидени в Акта за ИИ, остава въпросът защо съзаконодателите са одобрили регулаторни изисквания, за които все още не е ясно как на практика ще бъдат приложени. Темите, за които ще се изискват насоки, включват членове 8–15 и 28, чл. 5 (отнасящ се до забранените практики), практически аспекти при съществени модификации на ИИ системи, прозрачност по чл. 52, взаимодействието на ИИ акта с Приложение II и секторното законодателство, както и тълкуването и прилагането на дефиницията на изкуствен интелект.

Кодекс за поведение - трета чернова и очаквано приемане през м. май 2025 г.

Като част от усилията за прилагане на Акта за изкуствения интелект, Кодексът за поведение за ИИ с общо предназначение представлява важен, макар и доброволен, инструмент за операционализиране на изискванията на Акта, особено за доставчиците на усъвършенствани ИИ модели.[66] Въпреки значителния напредък, отразен в третата редакция на документа, съществуват сериозни предизвикателства, породени от дисбаланса в представителството при разработването му, където едва 20% от участниците са представители на бизнеса. Тази ситуация крие риск от въвеждане на прекомерно строги или непрактични регулации, които могат да имат неблагоприятни последици освен за иновациите и конкурентоспособността на европейския бизнес, също и социални такива.

Предложеният Кодекс за поведение за ИИ с общо предназначение, неговата връзка с Акта за ИИ и потенциалните последици от предвидените мерки и изисквания разкриват сложна картина на усилия за управление на бързо развиващата се област на изкуствения интелект. Третата редакция на Кодекса бележи значителен напредък спрямо втората,[67] като акцентът е поставен върху структурно опростяване, добавяне на детайли и повишаване на разбираемостта. Това включва по-ясно разграничаване на общите ангажименти от специфичните мерки в ключови области като прозрачност, авторско право и безопасност,[68] както и премахване на отделни KPI-та в секцията за авторско право в полза на интегрирането им в самите мерки.[69]

Кодексът за поведение надгражда правната рамка, установена от Акта за ИИ, като предоставя по-детайлни и специфични насоки за изпълнение на задълженията, особено за доставчиците на ИИ модели с общо предназначение.[70] Той съдържа конкретни мерки

[66] European Commission, "Third Draft General-Purpose AI Code of Practice", 2025, Retrieved from https://digital-strategy.ec.europa.eu/en/library/third-draft-general-purpose-ai-code-practice-published-written-independent-experts

[67] European Commission, "Second Draft General-Purpose AI Code of Practice " 2024, Retrieved from https://digital-strategy.ec.europa.eu/en/library/second-draft-general-purpose-ai-code-practice-published-written-independent-experts

[68] European Commission, "1- Commitments -Third Draft General-Purpose AI Code of Practice", 2025, Retrieved from https://digital-strategy.ec.europa.eu/en/library/third-draft-general-purpose-ai-code-practice-published-written-independent-experts

[69] European Commission, "3- Copyright -Third Draft General-Purpose AI Code of Practice", 2025, Retrieved from https://digital-strategy.ec.europa.eu/en/library/third-draft-general-purpose-ai-code-practice-published-written-independent-experts

[70] European Commission, "1- Commitments -Third Draft General-Purpose AI Code of Practice", 2025, Retrieved from https://digital-strategy.ec.europa.eu/en/library/third-draft-general-purpose-ai-code-practice-published-written-independent-experts

за прозрачност, вкл. формуляр за документация на модела,[71] детайлни насоки за спазване на авторското право[72] и разширени мерки за безопасност и сигурност,[73] които надхвърлят общите изисквания на Акта, което следва да се счита за допълнителна регулация, наложена чрез документ, който е определян като правно необвързващ, но дефакто задължава бизнеса да спазва всички изисквания, които не са заложени в Акта за ИИ. Въпреки доброволния си характер, Кодексът включва по-строги изисквания в определени области, като по-подробна документация, ангажименти за смекчаване на риска от нарушаване на авторски права и изискване за контактно лице за авторски въпроси.[74]

Съществува, обаче, загриженост относно факта, че само 20% от разработчиците на Кодекса са представители на бизнеса, което крие риск от неадекватно отразяване на бизнес перспективите и опасенията. Това може да доведе до създаването на непрактични или прекомерно строги изисквания, които да задушат иновациите и да натоварят непропорционално бизнеса, особено МСП. Главният директор на Meta по глобалните въпроси открито заяви, че Meta няма да подпише Кодекса с настоящото му съдържание (третата редакция), защото "правилата, надхвърлят изискванията на Акта за изкуствения интелект и налагат неизпълними и технически неосъществими изисквания". [75] Същите притеснения бяха изразени и от президента на Google по глобалните въпроси, който определи Кодекса като "стъпка в грешната посока".[76] Имайки предвид липсата на технологичен суверенитет на ЕС и сериозната зависимост от американски дигитални технологии, потенциален отказ от големите технологични фирми да прекратят или силно да ограничат предлагането на своите технологии и услуги на Единния пазар би имало катастрофални последици за ЕС, както в икономически, така и в политически и социален аспект.

В контекста на предстоящото приемане на Кодекса и възможността Комисията самостоятелно да наложи изискванията, заложени в него, е налице потенциал за значителни последици за бизнеса (член 50). Задължителното спазване на детайлните изисквания ще доведе до увеличаване на разходите за съответствие и регулаторната тежест, което може да намали конкурентоспособността на европейските компании на световния пазар. Освен ясните негативни последици, които това би имало върху бизнеса и икономиката на ЕС, следва да се разгледат подробно и негативните социалните последици, които могат да възникнат в резултат на предвидената тежка регулация за бизнеса. Някои от ключовите социални последици включват:

- **Увеличаване на безработицата и социално напрежение:** Ако европейските предприятия, ангажирани с разработката и внедряването на ИИ, изпитват затруднения в конкуренцията с компании от юрисдикции с по-либерално законодателство, това може да доведе до намаляване на тяхната

[71] European Commission, "2- Transparency -Third Draft General-Purpose AI Code of Practice", 2025, Retrieved from https://digital-strategy.ec.europa.eu/en/library/third-draft-general-purpose-ai-code-practice-published-written-independent-experts

[72] European Commission, "3- Copyright -Third Draft General-Purpose AI Code of Practice", 2025, Retrieved from https://digital-strategy.ec.europa.eu/en/library/third-draft-general-purpose-ai-code-practice-published-written-independent-experts

[73] European Commission, "4- Safety and Security -Third Draft General-Purpose AI Code of Practice", 2025, Retrieved from https://digital-strategy.ec.europa.eu/en/library/third-draft-general-purpose-ai-code-practice-published-written-independent-experts

[74] European Commission, "3- Copyright -Third Draft General-Purpose AI Code of Practice", 2025, Retrieved from https://digital-strategy.ec.europa.eu/en/library/third-draft-general-purpose-ai-code-practice-published-written-independent-experts

[75] Browne, R. "Google, Meta execs blast Europe over strict AI regulation as Big Tech ups the ante", 2025 https://www.cnbc.com/2025/02/21/google-meta-execs-blast-europe-over-strict-ai-regulation.html 19.03.2025

[76] Haeck, P. "EU rules for advanced AI are step in wrong direction, Google says", 2025, https://www.politico.eu/article/google-eu-rules-advanced-ai-artificial-intelligence-step-in-wrong-direction/ 19.03.2025

рентабилност и необходимост от оптимизиране на разходите, често чрез съкращаване на работна ръка. Загубата на работни места не само ще засегне пряко заетите в ИИ сектора, но може да има верижен ефект върху свързани индустрии и да повиши общото ниво на безработица, което води до социално напрежение и нестабилност. Особено уязвими могат да бъдат региони, силно зависими от технологичния сектор и иновациите.

- **Задълбочаване на дигиталното разделение:** Тежката регулация може да даде предимство на големите технологични компании, които разполагат с ресурсите да се справят с административната и финансова тежест, докато по-малките предприятия и стартиращи компании могат да бъдат изтласкани от пазара или да бъдат възпрепятствани в своето развитие, което противоречи на всички уверения, дадени от ЕК, че иска да подпомогне МСП да развиват иновации. Това може да доведе до концентрация на ИИ технологиите в ръцете на ограничен брой играчи, което да ограничи достъпа до тези технологии за по-широк кръг от бизнеси и потребители, задълбочавайки дигиталното разделение и неравенството в обществото.
- **Ограничаване на достъпа до социални услуги и иновации:** Ако разходите за съответствие направят ИИ базираните решения по-скъпи, това може да ограничи внедряването им в ключови социални сектори като здравеопазване, образование и обществени услуги. Например, по-малки болници или образователни институции може да не могат да си позволят ИИ инструменти за диагностика или персонализирано обучение, което да забави подобряването на качеството и достъпността на тези услуги. Забавянето на технологичния прогрес в тези области може да има преки негативни последици за благосъстоянието на гражданите.
- **Нарастване на социалното недоволство и недоверие:** Ако регулациите доведат до икономическа стагнация или загуба на работни места, това може да доведе до нарастване на социалното недоволство и недоверие към управляващите и регулаторните органи както на национално, така и на европейско ниво. Също така, ако европейските ИИ решения изостават от тези, разработени в други региони, може да засили чувството на технологична зависимост и загуба на суверенитет, което също може да подхрани обществено недоволство.
- **Възпрепятстване на решаването на социални проблеми:** ИИ има потенциала да помогне за решаването на множество социални проблеми, като климатичните промени, градското планиране, управлението на бедствия и други. Ако тежката регулация забави развитието и внедряването на тези технологии, това може да възпрепятства намирането на ефективни решения на тези належащи социални предизвикателства.
- **Етични дилеми и социална справедливост:** Въпреки че регулацията има за цел да гарантира етичното използване на ИИ, прекомерната тежест може да отклони фокуса от същинските етични въпроси и да създаде нови форми на социална несправедливост, свързани с ограничени възможности и достъп до технологии.

Макар че регулирането на ИИ е необходимо за минимизиране на рисковете и насърчаване на етичното му използване, **небалансираната и прекомерно тежка регулация, водеща до липса на конкурентоспособност на бизнеса, може да има сериозни и широкообхватни негативни социални последици**, включващи

загуба на работни места, задълбочаване на неравенствата, забавяне на социалния прогрес и нарастване на общественото недоволство. Ето защо е изключително важно регулаторните органи, включително и на национално ниво, да подходят с **внимание и гъвкавост**, като се стремят към намиране на оптимален баланс между регулиране и стимулиране на иновациите, за да се избегнат нежелани социални ефекти.

Вторичното законодателство е *нож с две остриета* за икономиката на ЕС и всяка от страните-членки. От една страна, то е необходимо за **гъвкавостта на регулацията** – ИИ технологиите се развиват бързо и вграждането на механизъм за бързо адаптиране (чрез делегирани актове) гарантира, че регламентът няма да остарее или да остави вратички. Например, ако утре се появи нов високорисков сценарий, Комисията може да го добави в приложение III, без да чака години. Това е добре и за инвестициите, защото създава **по-стабилна дългосрочна рамка** – инвеститорите знаят, че правилата могат да се нагодят към бъдещи рискове, вместо да бъдат отменяни изцяло. От друга страна, обаче, прекомерната зависимост от вторични актове поражда **несигурност в краткосрочен план**. Днес бизнесът знае главните акценти, поставени в Акта за ИИ, но много детайли още не са изяснени: например, как точно ще изглеждат задължителните формуляри за оценка на правата на човека или какви конкретни **бенчмаркове** ще се изискват за тестване на "високо въздействие" при един генеративен модел. Компаниите може да са предпазливи да инвестират в определена технология, ако подозират, че след година чрез делегиран акт тя може да попадне в забранен списък или да бъде обложена с нови изисквания. Това е особено валидно за определени случаи – например, един чатбот, който утре може да бъде квалифициран като "високорисков", ако излезе анализ, че масово подвежда потребители. Такава несигурност може да **забави решения за инвестиции или да ги пренасочи към страни извън ЕС**, докато не се издадат всички ключови актове и насоки, забавяйки технологичното развитие на европейската икономика.

Делегираните актове потенциално могат да се използват и за **повишаване/понижаване на праговете**, което директно влияе на броя засегнати фирми. Например, ако технологията напредне и 10^25 FLOPs мощност стане лесно достижим ресурс, Комисията може да повиши прага, за да не маркира твърде много модели като "системно рискови". Това би *облекчило* достъпа на нови модели до пазара, като ги остави извън най-строгия режим, докато са още експериментални. Обратно, ако определени практики заобикалят закона (напр., чрез леко модифициране на ИИ система, за да не попада формално в дефиницията на високорискова), Комисията може да **затегне дефинициите** чрез делегиран акт. От гледна точка на иновациите, е важно тези промени да стават **прозрачно и с участие на експерти и заинтересовани страни** – т.е. Комисията да провежда консултации, преди да приеме вторичните актове. Ако вторичното законодателство се прави в диалог с индустрията, то може да *помогне* на иновациите – например, чрез **изпълнителни актове за стандартизация**, които създават ясни тестови процедури, достъпни и за стартиращи компании.

За сравнение, САЩ считат, че под регулаторен надзор трябва да попадат тези модели, които страната определя като "авангардни модели" (frontier models) – модели, достигнали ниво 10^26 FLOPs на изчислителна мощност,[77] което означава, че ако САЩ въведе регулации за такива модели сега, ще бъде засегнат единствено моделът Grok 3, който е захранван от суперкомпютърния клъстър Colossus на компанията xAI, което осигурява поле за развитие на всички останали модели, разработени в САЩ, които са и водещите в световен мащаб, осигурявайки им допълнително стратегическо предимство. Това означава, че ЕК трябва да регулира, както Grok 3, така и всички останали

[77] Federal Register, "Framework for Artificial Intelligence Diffusion", 2025, https://www.federalregister.gov/documents/2025/01/15/2025-00636/framework-for-artificial-intelligence-diffusion 17.03.2025

съществуващи модели,[78] което би се оказало непосилна задача, имайки предвид че капацитетът за изпълнение на заложените мерки тепърва се изгражда.

Нереализирането навреме на вторичните актове би било проблем – Актът за ИИ се предвижда да влезе в сила през 2026 г., но дотогава Комисията трябва да изработи десетки актове и **да изгради капацитета на т.нар. Служба за ИИ.** Съществува сериозен риск от **бавно административно изпълнение,** което би довело до вакуум или объркване през преходния период. Например, ако хармонизираните стандарти за определени изисквания не са готови до 2026 г., бизнесът ще се чуди как да покаже съответствие. Потенциалната правна, оперативна и финансова несигурност за всички разработчици и внедрители на ИИ технологии, опериращи на пазара на ЕС, би навредило на пазарния достъп, защото несигурността може да спре пускането на продукти на пазара и би отблъснало потенциалните разработчици и внедрители от инвестиции в ИИ технологии, което ще окаже дълготраен негативен ефект върху всички страни-членки. Затова от стратегическа важност е **ускоряването на стандартите** (вкл. чрез работа с CEN/CENELEC, ISO и др.) и наемане на достатъчно експерти в Службата за ИИ.

Делегираните и изпълнителни актове дават гъвкавост, което е позитив за дългосрочното развитие на етичния ИИ, но в краткосрочен аспект представляват още един слой на неизвестност за бизнеса. За да се минимизират негативите, Комисията следва да комуникира рано намеренията си (напр., публикуване на пътна карта кои области ще са обект на делегирани актове през първата година); да привлича индустрията в изготвянето (публични консултации, пилотни проекти); да използва делегираните актове, **само когато е нужно** – т.е. да не променя прагове и списъци без ясна обосновка, тъй като всяка промяна разстройва инвестиционните планове, а ако ще бъдат налагани изменения, то те да се случват в консултация с всички засегнати страни и бизнесът да не бъде с ограничено право на глас, което е видимо при разработването на Кодекса за поведение към Акта за ИИ; да гарантира, че изискванията не стават по-рестриктивни от необходимото, за да не се въведе "пълзящо" затягане на режима, което не е било предвидено първоначално от законодателя, което се случва в момента при разработването на Кодекса за поведение и поражда риск от неприемане и подкопаване на кредибилността на самия Акт за ИИ, а оттам и на европейските институции.

Директива (ЕС) 2024/2831 относно подобряването на условията на труд при работа през платформа

Директивата за подобряване на условията на труд при работа през платформи представлява фундаментална реформа в трудовите отношения в ЕС, насочена към преодоляване на правната несигурност за милиони платформени работници. Основният механизъм, заложен в директивата, е презумпцията за трудово правоотношение при наличие на контрол от страна на платформата (чл. 4), която трансформира досегашния модел на самонаемане. Осъществяването на тази презумпция се обвързва с определени конкретни критерии, свързани с начина на възнаграждение, контрол на автономността и алгоритмичен надзор върху производителността.[79]

Един от най-иновативните и значими елементи на директивата е въвеждането на строги правила за алгоритмичен мениджмънт – автоматизирани системи, които разпределят задачи, оценяват представянето на работниците, изчисляват възнаграждение или вземат решения за прекратяване на трудовото отношение.

Членове 6 до 10 от директива установяват конкретни права за работниците:

[78] "Notable AI Models", updated 2025, https://epoch.ai/data/notable-ai-models, 25.03.2025.
[79] *OJ L, 2024/2831, 11.11.2024, ELI: http://data.europa.eu/eli/dir/2024/2831/oj*

- **Право на човешки надзор:** Всички автоматизирани решения, които имат съществено въздействие върху условията на труд, трябва да бъдат наблюдавани и одобрени от човек.
- **Право на информация и обяснение:** Работниците имат право да знаят как алгоритмите влияят върху тяхната заетост – напр., как се изчислява производителността или възнаграждението.
- **Забрана на обработка на лични и чувствителни данни:** Алгоритмичните системи не могат да използват биометрични или други чувствителни данни като критерий за вземане на решения.

Правата на работниците, заложени в директивата, се очаква да доведат до:

- **Повишена защита на работниците** – платформените работници ще имат по-добра социална защита и трудови права;
- **Намаляване на злоупотребите** – предотвратява експлоатацията чрез фалшиво самонаемане;
- **Подобрена прозрачност** – дава на регулаторите и работниците повече информация за управлението на алгоритмите, които платформата използва.

Тези разпоредби целят да адресират нарастващите социални притеснения относно дигиталната експлоатация, липсата на отчетност при автоматизирани уволнения и намаляването на човешкия елемент в трудовите отношения, и да поставят основата за нова форма на „алгоритмична справедливост" в цифровата икономика на труда. Директивата дава право на колективни действия, като подсилва правото на платформените работници да организират синдикати и да участват в колективни преговори. Друг важен елемент на директивата е и трансграничната ѝ приложимост – тя ще се прилага за всички платформи, независимо от юрисдикцията им, стига дейността им да засяга работници в ЕС (чл. 1). Това обвързва и платформи, регистрирани извън Съюза, със стандартите за защита на трудовите и лични права в ЕС, което крие и своите рискове, поради факта, че едни от най-големите платформи на пазара на ЕС са базирани в трети страни.

От гледна точка на изкуствения интелект, директивата налага нови изисквания върху използването на ИИ за вземане на решения, свързани с труда, като човешки надзор, прозрачност на логиката и обяснимост на алгоритмите, което поставя сериозни изисквания пред архитектурите на ИИ системите, използвани от платформите.

Освен социални ефекти, директивата има и съответните икономически такива, които могат да бъдат негативни:

- **Рискове за иновациите в ИИ** – платформите може да ограничат използването на ИИ, за да избегнат правни усложнения.
- **По-малко гъвкавост за работещите през платформи** – много хора предпочитат гъвкавостта на самонаемането. Принудителното преминаване към трудови договори може да намали възможностите за гъвкавост на заетостта, която обикновено се търси от предлагащите труда си през платформи.
- **Разходи за съответствие и адаптация:** Компаниите ще трябва да извършат цялостен преглед на алгоритмичните си системи, да правят оценка на въздействието върху личните данни (чл. 7) и да предоставят безплатно детайлна информация на работниците, техните представители и националните компетентни органи за използваните от тях алгоритмични системи, касаещи наблюдение и вземане на решения по всяко време, когато такава информация се поиска (чл. 9). Това създава значителни разходи и изисква много време, което ще засегне по-сериозно МСП и стартиращи платформи.

- **Правна несигурност и повишен риск от спорове:** Изискванията за прозрачност и презумпцията за трудово правоотношение (чл. 4) ще отворят вратата за множество трудови спорове, което ще натовари платформите с правна несигурност.
- **Правна несигурност при трансгранични операции:** Приложимостта на директивата към всички платформи, действащи на територията на ЕС, независимо от седалището им (чл. 1), създава правен риск за глобалните компании, чиито модели трябва да бъдат адаптирани спрямо европейските стандарти.
- **Потенциален отлив на инвестиции и риск от пазарна фрагментация:** Платформи, които разчитат на висока степен на автоматизация (вкл. ИИ-базирани платформи), могат да пренасочат дейността си към по-гъвкави юрисдикции, като САЩ или Азия, поради натрупващата се регулаторна тежест.
- **Стратегическа несигурност** по отношение на конкурентоспособността и дигиталното лидерство на ЕС.

При адекватно прилагане на директивата, дългосрочните ефекти могат да бъдат и положителни:

- **Стимул за по-етично и отговорно ИИ:** Директивата създава стимул за разработване на „обясним" и прозрачен ИИ (explainable AI), който е по-съвместим със социалните очаквания и човешките права – това може да се превърне в конкурентно предимство за европейските иноватори, стига МСП и стартиращите фирми да имат осигурен нужния финансов ресурс да разработят такива платформи.
- **Устойчиви бизнес модели:** Директивата укрепва социалната роля на платформите, включително чрез гарантиране и насърчаване на **правото на колективни действия** и синдикална организация (чл. 25), което може да доведе до по-устойчиви и социално отговорни бизнес модели.
- **Възможности за нови пазарни ниши:** Регулацията отваря място за нови услуги и продукти в сферата на ИИ етика, правни технологии (legal tech), създавайки допълнителна иновационна стойност.

В следствие на приемането на директивата платформените компании ще трябва да преразгледат значителна част от своите алгоритмични архитектури и бизнес модели, което може да доведе до високи разходи за адаптация и съответствие. Това ще доведе до повишение на правните рискове, свързани с оспорване на автоматизирани решения и с трудовото законодателство в различни държави членки. В дългосрочен план, директивата може да насърчи преминаването към по-устойчиви модели на „алгоритмично сътрудничество", но в краткосрочен аспект ще доведе до регулаторно напрежение, особено при стартиращи платформи, които имат с ограничени ресурси.

Регулацията на алгоритмичния мениджмънт чрез Директивата за работа през платформи, от една страна, гарантира основни трудови права, предотвратява дигитална експлоатация и подкрепя социалната справедливост в новите форми на труд, от друга страна, увеличава регулаторната тежест, рисковете и административните разходи за бизнеса. За да се избегне загубата на конкурентоспособност, е необходимо ЕС да прилага тези разпоредби по начин, който позволява гъвкавост, поетапно прилагане и подкрепа за стартиращите платформи. Съчетаването на високи социални стандарти с реалистична оценка на икономическите условия и адекватни политики за развитие е ключът към успеха на европейския модел за дигитална икономика.

Регламент (ЕС) 2022/2065 – Акт за цифровите услуги (DSA)

Актът за цифровите услуги има за цел да създаде по-безопасна и отговорна онлайн среда, включително чрез засилени задължения за много големи онлайн платформи и посредници. Макар че не регулира ИИ пряко, DSA оказва съществено влияние върху свободата за използване и внедряване на ИИ технологии, особено в области като автоматизирана модерация, препоръчващи системи и персонализирано съдържание.[80]

Чл. 34 от DSA изисква от много големите платформи (VLOPs) да идентифицират, анализират и смекчават системни рискове, включително такива, свързани с използването на ИИ алгоритми. Изискванията за прозрачност и достъп до обяснения на алгоритмични процеси (чл. 27) могат да ограничат търговската тайна и иновациите, особено за компании, предлагащи нови ИИ услуги чрез платформени бизнес модели. Освен това, задълженията за човешки надзор върху автоматизирани системи могат да ограничат автоматизацията и да увеличат оперативните разходи, което прави пазарния достъп по-труден за стартиращи компании и иновативни МСП.

Регламент (ЕС) 2022/1925 – Акт за цифровите пазари (DMA)

Актът за цифровите пазари цели да ограничи монополното поведение на "пазачи" ("gatekeepers"), които контролират достъпа, и да стимулира конкуренцията в цифровия пазар.[81] От гледна точка на ИИ, DMA има двойствен ефект. От една страна, чрез задължението за оперативна съвместимост и достъп до данни DMA може да отвори възможности за нови ИИ компании да разработват услуги, използвайки данни и интерфейси на големите играчи (чл. 6 параграф10). От друга страна, задълженията за неоткриване на потребителска активност и ограничената интеграция между платформи могат да възпрепятстват бизнес модели, базирани на комбиниране на ИИ решения в различни услуги. DMA, също така, задължава "пазачите" да не използват привилегировани данни за конкурентни предимства – нещо, което има значение за ИИ модели, обучавани върху данни от платформата. Това създава нови юридически и технически предизвикателства пред разработката на мощни ИИ системи в контекста на големи екосистеми и създава риск от загуба на конкурентоспособност на европейския бизнес, при разработването на ИИ технологии.

Регламент (ЕС) 2024/2847 – Акт за киберустойчивост (CRA)

Актът за киберсутойчивост налага изисквания за киберсигурност върху всички свързани цифрови продукти, включително такива, в които е вграден ИИ.[82] Регламентът предвижда задължителна оценка на риска, нотификации при уязвимости и изисквания за поддръжка по време на целия жизнен цикъл (чл. 8, 12-14). От гледна точка на ИИ, това означава, че всеки разработчик на ИИ софтуер или устройство, използващо ИИ функционалност, трябва да съобрази архитектурата си не само с етичните изисквания на Акта за ИИ, но и с високите стандарти за киберустойчивост на CRA. Това може да увеличи разходите за разработка и тестване, особено при продукти, насочени към пазари с по-кратки цикли на иновация. Нещо повече, за високорискови цифрови продукти (вкл. определени ИИ решения) CRA изисква сертифициране от трета страна (чл. 39), което може да забави навлизането на нови технологии на пазара и да повлияе негативно върху времето до реализация (time-to-market) за ИИ иновации, което да даде стратегическо предимство на продукти, разработени в трети страни.

[80] *OJ L 277, 27.10.2022,* ELI: http://data.europa.eu/eli/reg/2022/2065/oj
[81] OJ L 265, 12.10.2022, ELI: http://data.europa.eu/eli/reg/2022/1925/oj
[82] OJ L, 2024/2847, 20.11.2024, ELI: http://data.europa.eu/eli/reg/2024/2847/oj

Регламент (ЕС) 2016/679 – Общ регламент за защита на данните (GDPR)

Общият регламент за защита на данните (GDPR), в сила от 2018 г., установява строги правила за обработка на лични данни в ЕС.[83] Той изисква данните да се обработват законосъобразно и прозрачно, само при наличие на правно основание (чл. 5–6) и предоставя ключови права на субектите – достъп, коригиране, изтриване („правото да бъдеш забравен"), преносимост, възражение и ограничаване (чл. 15–21). Въвежда се принципът „защита по дизайн и по подразбиране" (чл. 25) и задължение за уведомление при пробиви в сигурността (чл. 33–34). Особено важен е чл. 22, който забранява изцяло автоматизирани решения с правен ефект без човешка намеса, освен при изключения – ключово изискване за ИИ системи.

GDPR е емблематичен пример за стремежа на ЕС да защити фундаментални права – личната неприкосновеност – дори ако това изисква налагане на значителни ограничения върху бизнеса. Регулацията несъмнено създаде регулаторна тежест, с която фирмите продължават да се справят. Видяхме, че тя доведе до леко понижение на приходи и пазарна оценка в чувствителни сектори,[84] но в същото време, GDPR постигна основната си цел – засилена защита на данните – и дори доведе до непредвидени положителни ефекти като намаляване на пробивите.[85] Иронично обаче, GDPR консолидира позициите на технологични гиганти, които разполагат с директен достъп до данни чрез собствени услуги, докато малките играчи са ограничени.[86]

От гледна точка на иновациите, регламентът има двоен ефект. Макар да засилва доверието и правата на гражданите, той затруднява достъпа до данни – критичен ресурс за ИИ. След влизането му в сила, рисковите инвестиции в европейски технологични стартъпи намаляват, а много ИИ компании преместиха дейности извън ЕС, за да избегнат сложността на съответствие. Това поражда притеснения, че ЕС изостава иновационно и губи конкурентоспособност в глобалния ИИ и цялостния технологичен пейзаж.

Пресечни точки в съществуващата регулаторна рамка на ЕС върху изкуствения интелект и икономическото развитие

Регулаторната архитектура на Европейския съюз, включваща Акта за изкуствения интелект (AI Act), Общия регламент за защита на личните данни (GDPR), Акта за цифровите услуги (DSA), Акта за цифровите пазари (DMA), Акта за киберустойчивост (CRA) и Директивата за работа чрез платформи (Platform Work Directive), формира изключително комплексна среда, която оказва дълбоко и взаимносвързано въздействие върху иновационната екосистема и ИИ сектора в Европа. Целта на тази рамка е да наложи модел на етично регулирана иновация, но в същото време изисква значителни човешки, административни и правни ресурси, което води до тежест, особено за стартиращите компании и малките и средните предприятия.

Съществуват няколко ключови пресечни точки между нормативните актове:

Актът за ИИ и GDPR създават паралелен и взаимнодопълващ се режим за системи, които обработват лични данни чрез автоматизирани решения. Алгоритми, особено тези с общо предназначение (GPAI), подлежат на едновременно спазване на изискванията за защита на данните и изискванията за прозрачност, одит и човешки

[83] OJ L 119, 4.5.2016, ELI: http://data.europa.eu/eli/reg/2016/679/oj
[84]Motoki, F., & Pinto, J."Regulating data: Evidence from corporate America" . Journal of Business Finance & Accounting, 52(1), 541-568. https://doi.org/10.1111/jbfa.12820
[85] Ibid
[86] Moazed, Alex, "How GDPR is Helping Big Tech and Hurting the Competition", https://www.applicoinc.com/blog/how-gdpr-is-helping-big-tech-and-hurting-the-competition/#:~:text=Giants%20can%20thrive%20in%20the%20face%20of%20GDPR%20and%20other%20privacy%20laws&text=Google%2C%20for%20example%2C%20has%20seen,America%20and%20Europe%20lost%20ground 23.03.2025

надзор от Акта за ИИ. Това налага двуслойни механизми за съответствие, което води до административно дублиране и нужда от интегрирани екипи по съответствие.

Актът за ИИ и Актът за цифровите услуги взаимодействат при регулирането на големите онлайн платформи. DSA задължава много големите онлайн платформи да оценяват системните рискове, включително такива, произтичащи от алгоритми с ИИ, докато Акта за ИИ ги класифицира като високорискови при използване за модериране на съдържание или персонализация. Потенциалният дублиращ ефект изисква координирана методология за одит и отчетност.

Актът за ИИ и Актът за цифровите пазари поставят напрежение между изискванията за оперативна съвместимост и достъп до данни (DMA) и изискванията за защита на сигурността, бизнес тайната и устойчивостта на системите от Акта за ИИ. Този баланс ще бъде определящ за модела на ИИ иновации в платформената икономика и може да постави ЕС в конкурентен риск спрямо юрисдикции с по-гъвкави подходи към споделената инфраструктура.

Актът за ИИ и Актът за киберустойчивост изискват ИИ системите да бъдат подложени както на техническа, така и на оценка за киберсигурността още на етапа на проектиране (CRA). Системите ще трябва да изпълняват критерии едновременно по линия на надеждност, сигурност и правна съвместимост, което предполага необходимост от интердисциплинарни екипи и нови роли в ИТ и правния сектор.

Актът за ИИ и Директива за работа през платформи съвместно регулират алгоритмичния мениджмънт в платформената заетост. Изискванията за прозрачност, обяснимост и право на обжалване на автоматизирани решения се удвояват, което създава пречки за прилагане на ИИ решения в сектори като мобилни доставки, споделено пътуване и дигитални микрозадания.

„Компас за конкурентоспособността "

На 29 януари 2025 г. Европейската комисия представи „Компаса за конкурентоспособността" – стратегическа рамка за укрепване на икономическите позиции на ЕС чрез три стълба: иновации, декарбонизация и икономическа сигурност.[87] Рамката поставя ИИ като ключов елемент във всяко от направленията.

Първият стълб е насочен към преодоляване на изоставането в иновациите. В него се предвиждат редица инициативи с пряко значение за ИИ. Подготвяният Европейски закон за иновациите (очакван в края на 2025 или началото на 2026 г.) цели да опрости регулациите и да улесни достъпа до рисков капитал чрез програми като TechEU, за да се повиши комерсиализацията на иновации. ИИ Фабриките, обявени в началото на 2025 г., ще изградят свързани хъбове между индустрия, академични среди и суперкомпютърни центрове. Очакваната през третото тримесечие на 2025 г. Стратегия за приложен ИИ ще се фокусира върху внедряването на ИИ в приоритетни сектори чрез насочени инвестиции в НИРД. Законът за облачните пространства и ИИ (планиран за края на 2025 или началото на 2026 г.) ще насърчава енергийно ефективна инфраструктура за ИИ и мобилизация на инвестиции. Законът за съвременните материали, предвиден за 2026 г., ще осигури достъп до технологии и материали, необходими за скалируемостта на ИИ решения. Допълнително, стратегията за природните науки ще подкрепи реиндустриализацията чрез биотехнологии, биоикономика и ИИ, чрез регулации, инвестиции и изграждане на умения.

Вторият стълб се фокусира върху декарбонизацията и конкурентоспособността, водени от Чистата индустриална сделка и Плана за достъпна енергия. Последният е особено важен за ИИ, тъй като високите енергийни разходи и несигурният енергиен

[87] European Commission, "Competitiveness Compass", 2025, retrieved from https://commission.europa.eu/document/download/10017eb1-4722-4333-add2-e0ed18105a34_en

достъп са пречка за развитието на изчислително интензивни технологии и забавят реиндустриализацията на ЕС, намалявайки неговата конкурентоспособност.

Третият стълб – икономическа сигурност – включва Стратегията за готовност на ЕС, която цели устойчивост на икономиката при кризи. Това включва стабилна ИИ екосистема, която може да подкрепи реиндустриализацията и икономическата адаптивност.

Към тези три стълба се добавят и хоризонтални мерки: нова категория „малки предприятия със средна капитализация", която цели да намали регулаторните бариери, а Европейският фонд за конкурентоспособност ще подкрепи инвестиции в иновации и намаляване на зависимостта от критични ресурси. ЕС поставя и акцент върху човешкия капитал чрез изграждането на Съюз на уменията, насочен към доживотно обучение с фокус върху STEM умения – ключови за развитието на ИИ.

„Компасът за конкурентоспособността" очертава цялостен и координиран подход, в който ИИ е интегриран в основата на икономическата трансформация на ЕС. Ако инициативите бъдат реализирани ефективно, те ще засилят иновационния капацитет, ще намалят зависимостите и ще повишат конкурентоспособността на Европа в новата дигитална и индустриална епоха.

ЕИСК

"Генеративни ИИ и базови модели в ЕС: Възприемане, възможности, предизвикателства и път напред"

Проучването на Европейския икономически и социален комитет (ЕИСК) подчертава разликата между генеративен ИИ (създаващ съдържание) и основни модели (многофункционални ИИ системи) – терминологичната неяснота крие риск от неефективна регулация. САЩ доминира с над 80% от глобалното финансиране на генеративен ИИ, докато ЕС изостава по инвестиции, инфраструктура и стратегическо координиране. Големите технологични платформи прилагат вертикална интеграция, което блокира достъпа на европейски стартъпи до пълната стойностна верига. Предимствата на ЕС са в научната база, отворените общности и образование; слабостите – пазарна фрагментация, ограничен достъп до данни и зависимост от чужда инфраструктура. Препоръките включват: създаване на „CERN за ИИ", инвестиции в суперкомпютри, улеснена регулация за МСП, етични принципи и обучение на кадри.[88] Проучването затвърждава, че ЕС има потенциал, но му липсва стратегическа оперативност. Без координирана рамка и собствена инфраструктура за големи модели, Европа рискува да остане потребител, а не създател на бъдещето на ИИ, зависещ изцяло от чуждите технологии. Приоритет трябва да бъдат интегрирани публично-частни инвестиции, технологичен суверенитет и ускорено изграждане на конкурентоспособна ИИ екосистема.

Становища в рамките на ЕИСК относно бъдещи регулации на изкуствения интелект на работното място

В условията на нарастваща автоматизация и дигитализация на труда, Европейският съюз стои пред ключова дилема: как да балансира между необходимостта от защита на работниците и стимулиране на иновациите чрез ИИ. Два противоположни подхода към тази дилема са представени в документите „**ИИ в полза на работниците: лостове за овладяване на потенциала и смекчаване на рисковете от ИИ във връзка със заетостта и политиките на пазара на труда"**,[89] на групата на работниците в ЕИСК и

[88] EESC, "Generative AI and foundation models in the EU: Uptake, opportunities, challenges, and a way forward", QE-01-25-014-EN-N, 2025

[89] EESC, "Pro-worker AI: levers for harnessing the potential and mitigating the risks of AI in connection with employment and labour market policies", SOC/803 - EESC-2024-01024-01-00-TCD-TRA - EN, 2025

„**Контрастановище**"[90] на групата на работодателите, които очертават полюсни визии за бъдещето на ИИ на работното място.

Становището „ИИ в полза на работниците" застъпва позицията, че ИИ трябва да бъде строго регулиран, особено в контекста на трудовите отношения, с цел защита на правата, достойнството и автономията на работещите. Според тази позиция, съществуващата правна рамка, включително GDPR и Акта за ИИ, е добра основа, но недостатъчна – необходимо е нейното надграждане с по-ясни изисквания за оценка на въздействието върху основните права, задължителен човешки надзор и гарантиране на принципа „human-in-the-loop" при вземане на решения от ИИ системи. В центъра на тази визия стои социалният диалог – регламентиран механизъм, чрез който работниците и техните представители участват активно във въвеждането на ИИ на работното място.[91]

От социална гледна точка, този подход предлага значителни ползи: гарантира социална стабилност, укрепва общественото доверие в технологиите и намалява риска от масова автоматизация и технологична безработица. Освен това, ЕС може да се утвърди като глобален лидер в етичната регулация на ИИ, което повишава доверието в европейските технологии и модели на управление. В дългосрочен план, обаче, твърде строга регулация крие риск от отлив на инвестиции, особено ако компаниите преценят, че административните и правни изисквания са прекомерни. Това би могло да ограничи способността на ЕС да развива конкурентоспособна ИИ екосистема, като постави в неблагоприятна позиция стартиращи фирми и МСП, които често нямат ресурси за сложна правна съвместимост.

Обратно, контрастановището на групата на работодателите защитава либерален подход, според който настоящата нормативна база – включително Акта за ИИ и GDPR – вече осигурява достатъчна защита на правата и безопасността на работещите. Вместо нови закони, се предлагат доброволни насоки за прилагане на ИИ в трудовата сфера, оставяйки на бизнеса гъвкавостта да адаптира своите системи към етични стандарти. Социалният диалог в този контекст е разглеждан като възможен, но не задължителен процес, който не бива да води до допълнителна бюрокрация. Основният фокус на контрастановището е икономическата ефективност. Според този модел, прекомерната регулация би ограничила инвестициите в ИИ, би забавила адаптацията на нови технологии и би засилила изоставането на ЕС спрямо конкуренти като САЩ и Китай.[92]

Поддържането на по-гъвкава регулаторна среда се възприема като ключово условие за стимулиране на стартъпи, навлизане на пазара и експериментиране с иновативни бизнес модели. От тази гледна точка, ЕС има възможност да се позиционира като глобален хъб за развитие на ИИ, ако избегне прекомерно рестриктивен подход. Въпреки потенциала за растеж, този модел не е лишен от рискове, които се крият в неетичното използване на алгоритми за наблюдение, дискриминация или автоматизирани уволнения. Недостатъчната прозрачност може да подкопае доверието на работниците и обществото в ИИ, което в крайна сметка също да забави неговото внедряване.

В основата си, двете становища представят два различни подхода за развитие на ИИ, търсейки баланса между социална защита и икономическа гъвкавост. Становището „ИИ в полза на работниците" акцентира върху правата на човека, прозрачността и справедливостта, дори за сметка на темпото на иновации. Докато контрастановището

[90] EESC, "Counter Opinion on "Pro-worker AI: levers for harnessing the potential and mitigating the risks of AI in connection with employment and labour market policies"", SOC/803 - EESC-2024-01024-00-02-AMP-TRA, 2025
[91] EESC, "Pro-worker AI: levers for harnessing the potential and mitigating the risks of AI in connection with employment and labour market policies" , SOC/803 - EESC-2024-01024-01-00-TCD-TRA - EN, 2025
[92] EESC, "Counter Opinion on "Pro-worker AI: levers for harnessing the potential and mitigating the risks of AI in connection with employment and labour market policies"", SOC/803 - EESC-2024-01024-00-02-AMP-TRA, 2025

залага на свободата за експериментиране и растеж, като приема, че съществуващите механизми са достатъчни за предотвратяване на злоупотреби.

Истинското предизвикателство пред ЕС е да формулира балансирана политика, която съчетава ефективна защита на основните права с подкрепа за иновации и икономическо развитие. Подход, който осигурява както социална справедливост, така и икономическа ефективност, е: въвеждането на гъвкави, секторно адаптирани рамки за ИИ на работното място, които предвиждат задължителен социален диалог единствено при прилагане на чувствителни ИИ технологии; облекчени механизми за съответствие за МСП и иновативни стартиращи компании; ясни, балансирани стандарти за обяснимост, прозрачност и право на човешко преразглеждане на автоматизирани решения, при определени условия; фокус върху устойчивост и справедливост, без да се блокира пазарната динамика. Такъв подход би позволил на ЕС да реализира както своята етична визия, така и да поддържа своята конкурентоспособност в глобалната технологична надпревара.

Регулаторната рамка на ЕС за изкуствен интелект и въздействието ѝ върху конкурентоспособността

Настоящият анализ цели да оцени въздействието на настоящите и бъдещи регулации в Европейския съюз, засягащи изкуствения интелект, върху иновациите, конкурентоспособността и икономическото развитие в ЕС. Анализът стъпва върху съществуващата правна рамка – Акт за ИИ, GDPR, Акт за цифровите услуги, Акт за цифровите пазари, Акт за киберустойчивост и Директива относно подобряването на условията на труд при работата през платформа – както и върху предстоящото вторично законодателство заложено в Акта за ИИ и стратегически инициативи като „Компас за конкурентоспособността ".

Системно въздействие на съществуващата регулаторна рамка

Съчетаното действие на всички тези регулации очертава нов европейски модел за цифровата икономика, който често се описва с думи като "ориентиран към човека", "етичен", "устойчив". Този модел безспорно отличава ЕС глобално – докато други икономики понякога възприемат подход "първо иновацията, после регулацията", Европа поставя рамки още от зародиша на технологиите, за да насочи развитието им в обществен интерес. В краткосрочен план, това създава усещането, че ЕС е *"сериен регулатор"* (последователно: поверителност – GDPR; онлайн платформи – DSA/DMA; ИИ – Акт за ИИ; киберсигурност – CRA; трудови права в дигиталния сектор – Директива относно работа през платформи). Натоварването върху бизнес средата е значително: фирмите трябва паралелно да приведат дейността си в съответствие с множество нови изисквания, често в кратки срокове. Това може временно да забави въвеждането на нови продукти или услуги в ЕС – например, големи технологични компании публично сигнализираха, че ще забавят пускането на генеративни ИИ функции в Европа, заради неяснотата около Акта за ИИ.[93] Други компании, като Apple, бавят пускането на своите ИИ технологии в ЕС с месеци, поради регулаторните изисквания, свързани с DMA.[94] Регулации като GDPR имаха същия ефект.[95] Това ограничава достъпа както на бизнеса, така и на гражданите на ЕС до най-новите ИИ технологии. С други думи, съществува риск ЕС да се превърне в пазар от второ ниво (second-tier) за глобалните технологични

[93] Browne, R. "Google, Meta execs blast Europe over strict AI regulation as Big Tech ups the ante", 2025 https://www.cnbc.com/2025/02/21/google-meta-execs-blast-europe-over-strict-ai-regulation.html 19.03.2025
[94] DigWatch Team, "Apple Intelligence expands to the EU amid regulatory changes", 2024, https://dig.watch/updates/apple-intelligence-expands-to-the-eu-amid-regulatory-changes 19.03.2025
[95] Lee D., "EU's Hard Line on Tech Keeps Users Off the Cutting Edge", 2024, https://www.bloomberg.com/opinion/articles/2024-07-03/apple-and-meta-users-lose-out-in-eu-s-hard-line-on-tech 19.03.2025

иновации, където новото идва със закъснение, и това не просто да забави икономическото развитие на ЕС, а да го откъсне от достъп до най-новите технологии.

Показателно е сравнението в инвестициите в ИИ: между 2018 г. и 2023 г. компаниите в ЕС привлякоха около 32.5 млрд евро инвестиции в ИИ, докато в САЩ – над 120 млрд. евро[96] Само за 2025 г. се очаква четирите технологични гиганта в САЩ (Meta, Microsoft, Amazon и Alphabet) да инвестират 325 млрд. щ.д.[97], а инвестициите в САЩ в изчислителни мощности до 2030 г. да достигнат 2.35 трлн. щ.д до 2030 г.[98] Пример за това е обявеният в началото на годината съвместен проект между Microsoft, Oracle, SoftBank, MGX и OpenAI, наречен **Stargate**, който ще получи инвестиции от 500 млрд. щ.д. до 2028 г. за изграждане на мощни центрове за данни, изцяло насочени към развитието на OpenAI, с пълната подкрепа на настоящата администрация в САЩ.[99] За сравнение, в началото на 2025 г. ЕК обяви инициативата InvestAI, която ще осигури 50 млрд. евро[100] инвестиции за развитието на ИИ технологии в ЕС, в които се включени и 4 AI гигафабрики в ЕС, но не е известно за какъв период от време. Частната инициативата EU Champions AI[101] чрез над 20 ключови международни инвеститори са заделили 150 млрд. евро нов капитал и предварително привлечени средства за възможности, свързани с ИИ в Европа през следващите пет години, при условие че ЕС успее да създаде прозрачна и целенасочена, основана на конкуренцията рамка за ИИ. Тези цифри недвусмислено показват мащабните разлики в инвестициите в ИИ технологии между САЩ и ЕС. Регулаторната тежест, съществуваща на територията на ЕС, може да продължи тази тенденция, тъй като капиталът търси среди с по-голяма свобода за експериментиране. Темата за инвестициите ще бъде разгледана подробно в следващият раздел на настоящия анализ.

В дългосрочен план, обаче, комбинираният ефект от тези мерки може да повиши устойчивата конкурентоспособност на ЕС. Ключовият въпрос е: *Конкурентоспособност на каква цена и в какви сектори?* Европа очевидно няма амбицията да се състезава за доминация в социалните мрежи или масовите потребителски приложения – там други региони водят. Но ЕС се позиционира като лидер във висококачествени, надеждни технологии, което е все по-търсено. Например, с Акта за ИИ Европа ще има удостоверение за "надежден ИИ" (подобно на знак за качество), което потенциално ще бъде предпочитано в чувствителни индустрии (здравеопазване, автомобилостроене) и от потребители, ценящи етиката. С Акта за киберустойчивост европейските продукти ще са с репутация на по-сигурни – един IoT[102] уред, който отговаря на изискванията за киберустойчивост, ще бъде по-малко вероятно хакнат. Това може да направи европейските продукти по-желани на глобалния пазар, особено когато инциденти с

[96] Madiega T.and Ilnicki R. "AI investment: EU and global indicators", , European Parliament Research Service, 2024, retrieved from https://www.europarl.europa.eu/RegData/etudes/ATAG/2024/760392/EPRS_ATA(2024)760392_EN.pdf

[97] Bratton, L. "Big Tech set to invest $325 billion this year as hefty AI bills come under scrutiny", 2025, https://finance.yahoo.com/news/big-tech-set-to-invest-325-billion-this-year-as-hefty-ai-bills-come-under-scrutiny-182329236.html 19.03.2025

[98] Smith, K. et al, "AI power surge: Growth Scenarios for GenAI Datacenters Through 2030", Center for Strategic and International Studies, 2025, retrieved from https://www.csis.org/analysis/ai-power-surge-growth-scenarios-genai-datacenters-through-2030

[99] "Announcing the Stargate Project", OpenAI, 2025 https://openai.com/index/announcing-the-stargate-project/ 11.03.2025

[100] Speech by President von der Leyen at the Artificial Intelligence Action Summit, France, 11-12 February 2025 - https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_25_471

[101] https://aichampions.eu/, прессъобщение https://files.elfsightcdn.com/eafe4a4d-3436-495d-b748-5bdce62d911d/9dc25764-5c0a-4c04-b55e-6be1d26f625f/EU-AI-Champions-Initiative-Media-Release.pdf

[102] IoT, Internet of Things. От английски – интернет на нещата. Интернет на нещата е мрежа от физически устройства, като сензори, уреди и автомобили, които са свързани към интернет и могат да обменят данни помежду си. Тези устройства събират информация, анализират я и автоматично предприемат действия без нужда от човешка намеса. IoT технологията намира приложение в различни области, като умни домове, здравеопазване, транспорт и индустриализация. Заедно с изкуствения интелект и големите информационни масиви, интернет на нещата е в центъра на цифровизацията на световната икономика.

несигурни устройства зачестяват. По сходен начин, DSA и DMA полагат основите на по-разнообразен дигитален пазар – ако те успеят, след няколко години ЕС може да има процъфтяваща екосистема от конкурентни платформи, интероперативни услуги и стартиращи компании, които се възползват от отворените пазари. Това би била конкурентоспособност, изразяваща се не в един двоен еднорог монополист, а в *множество средни играчи, предлагащи иновации и качество*. Традиционно, успехът на европейската икономика е бил именно в средния технологичен бизнес и индустрии с висока добавена стойност (например, индустриална автоматизация, автомобилни части, B2B софтуер). Регулациите като Акта за ИИ и Акта за киберустойчивост съответстват добре на тези силни страни – те насърчават индустриален ИИ, който е безопасен, и умни устройства, които са сигурни, което е точно полето, където европейските фирми могат да превъзхождат, конкурирайки се с качество, а не с масов обем.

Друг важен аспект е доверието на обществото. Европейските потребители, защитавайки се чрез DSA, GDPR и Акта за ИИ, ще са по-склонни да приемат и ползват нови технологии (например, автономни автомобили или диагностичен ИИ в болницата), защото знаят, че са налице защити за тях. Това "вътрешно търсене" на иновации, подплатено с доверие, е ключово за разрастване на местните шампиони. В страни, където липсва регулиране, може да се стигне до обществен отпор срещу технологии (напр., протести срещу разпознаване на лица, страх от ИИ), докато в ЕС има проактивна работа тези страхове да се адресират нормативно. По този начин, *социалният договор* относно технологиите в ЕС е по-здрав – гражданите приемат иновациите не като заплаха, а като контролирано благо, което увеличава социалната приемливост на новите продукти.

Разбира се, остава рискът ЕС да се окаже твърде бавен и скъп за разработване на определени авангардни технологии (напр., AGI, метавселени), които ще се реализират другаде и после Европа ще трябва да ги внася. Това би навредило стратегически на конкурентоспособността. За да се избегне такъв сценарий, ЕС може да използва предимството си в правилата като експорт на ценности: ако чрез дипломатически и търговски лостове Европа успее да наложи своите стандарти глобално (т.нар. "Брюкселски ефект"), тогава чуждите компании ще адаптират продуктите си към европейските изисквания и разликата в иновациите ще се стопи. Това обаче може да се случи единствено чрез развитие на конкурентоспособна икономика, която да легитимира тези стандарти. Ако това се реализира, европейският регулаторен подход може да се превърне от бреме в ключово конкурентно оръжие – задавайки правилата на играта, ЕС защитава не само гражданите си, но и своя бизнес от нелоялна конкуренция на играчи, които иначе биха пестили за сметка на безопасността или етиката.

Краткосрочно, регулациите леко охлаждат темпа на "дива" дигитална експанзия в ЕС, с риск от изоставане в някои ключови сектори, но дългосрочно те градят основи за по-сигурна, качествена и обществено приета технологична среда, в която европейските фирми имат шанс да просперират, именно защото отговарят на тези високи стандарти. Конкурентоспособността вече няма да се мери просто в брой платформи - "еднорози", а и в *способността на една икономика да внедрява иновации, без да руши социалната тъкан. За да се случи това, ЕС трябва да бъде технологично независим*.

Съществуващите ключови европейски регулации, отнасящи се до ИИ – Акт за ИИ, Директивата относно работа през платформи, DSA, DMA, CRA и GDPR – показват един цялостен регулаторен подход, поставящ етиката, правата и сигурността като равноправни цели, наред с икономическото развитие. Тези актове въвеждат строги изисквания, които ограничават известна бизнес свобода и увеличават разходите за съответствие, но адресират конкретни проблеми: Акт за ИИ намалява риска ИИ да навреди на хората, Директивата относно работа през платформи предпазва трудовите права в новата "гиг" икономика, Актовете за цифровите пазари и услуги създават по-справедлив и безопасен онлайн пазар, Акта за киберустойчивостта повишава

минималната киберсигурност на всички продукти, а GDPR утвърждава личната неприкосновеност в дигиталната ера. Социалните ползи – от по-малко дискриминация и по-прозрачни алгоритми до по-малко данни в ръцете на киберпрестъпници – са реални и значими. В същото време, регулаторната тежест не е за подценяване: Европа се движи по-тясна пътека, за разлика от по-широкия път на по-слабо регулирани пазари, и това влияе на инвестиционните решения и бизнес динамиката. Комбинирано, тези регулации формулират *"Европейска конкурентна ниша"* – доверени цифрови продукти и услуги, при които качеството, безопасността и етичността са ключови, наред с функционалността. За да успее тази стратегия, е необходимо внимателно и съобразено с икономическата реалност управление на изпълнението на регулациите и готовност за адаптация, ако се установи, че някъде ЕС се е "престарал" или не е дооценил риска.

Въздействие на предстоящото вторичното законодателство и инициативи

Разгръщането на вторичното законодателство по Акта за ИИ представлява едновременно възможност и риск за икономиката на ЕС. Въпреки че делегираните и изпълнителните актове създават необходимата адаптивност към бързо развиващия се технологичен пейзаж, те също пораждат сериозна краткосрочна правна несигурност. Ключови определения, като „високорискови" модели и „системно въздействие", все още не са уточнени в детайли. Бавната реализация на тези актове може да парализира пазарния достъп за ИИ продукти и стартиращи компании, поради неясни критерии за съответствие, одит и санкции. Необходима е координация между Комисията, индустрията и техническите органи за бързо разработване на хармонизирани стандарти.

Предстоящите инициативи в рамките на „Компаса за конкурентоспособност", включително Европейския закон за иновациите и ИИ Фабриките, носят потенциал за ускоряване на технологичната трансформация. Въпреки това, ако регулаторната тежест, произтичаща от Акта за ИИ, Кодекса за поведение и свързаните актове, не бъде облекчена за МСП чрез механизми като стандартни шаблони за съответствие, регулаторни пясъчници и дигитална сертификация, Европа рискува да загуби иновационно темпо в полза на по-гъвкави юрисдикции като САЩ и Югоизточна Азия. Делегираните и изпълнителните актове, предвидени в Акта за ИИ, имат стратегическо значение за оформянето на регулаторната реалност в следващите години. Предстои създаването на множество вторични актове, които ще определят как и до каква степен ще може да се развива технологията в рамките на ЕС. Макар това да осигурява адаптивност спрямо технологичния напредък, в краткосрочен план съществува висока степен на несигурност, която възпира инвестиции и стратегическо планиране от страна на бизнеса.

Социално-икономическите ефекти на европейския регулаторен модел се проявяват с особена сила в контекста на дигиталната трансформация и ИИ. От една страна, социалният акцент – гарантиране на права, прозрачност, етика и колективна защита – създава стабилност, доверие и предвидимост, което е от критично значение за изграждането на „етична конкурентоспособност ", която Европа се стреми да въведе като глобален стандарт. От друга страна, обаче, същата тази амбициозна социална рамка води до значителни бариери за навлизане на нови участници и ограничава динамиката на експериментиране. Това особено важи за стартиращи компании в сектори като образование, транспорт, здравеопазване и HR технологии, където алгоритмичният мениджмънт е основен двигател на ефективност. Налагането на двойна регулаторна тежест от Акта за ИИ и Директивата относно работа през платформи в тези сектори може да отблъсне рискови инвестиции и да създаде рискове от регулаторен арбитраж, при който бизнесът изнася операции извън ЕС. Възможността за стратегическо използване на делегираните актове за гъвкава адаптация съществува, но трябва да се управлява чрез публичен диалог и прозрачност, за да се избегне т.нар. „регулаторна парализа ". Ако съществуващата несигурност не бъде адресирана чрез предвидим график на вторичното законодателство и междинни указания, Европа ще изпадне в ситуация на

забавена реализация, забавено инвестиране и още по-сериозно изоставане от световните лидери не само в развитието на ИИ, но на всички технологии на бъдещето.

Социалната страна на регулацията е насочена към предотвратяване на дискриминация, засилване на човешкия надзор, и защита на личните данни. Това подкрепя общественото доверие и устойчивата интеграция на ИИ. От икономическа гледна точка, натискът върху МСП и стартиращите компании е сериозен – необходимо е съответствие със сложна и фрагментирана нормативна рамка, което повишава разходите и намалява инвестиционната привлекателност на ЕС спрямо други юрисдикции като САЩ и Азия. За да превърне своята амбициозна регулаторна архитектура в инструмент за глобално лидерство и да се избегне превръщането на Европа в „регулаторен остров", Европейският съюз трябва да предприеме стратегически стъпки за подобряване на условията за развитие на своята икономика.

В *Приложение 2* са изложени предложения за стратегически важни политики, които България да отстоява на ниво ЕС при разработването на бъдещи регулации с цел подобряване на конкурентоспособността на ЕС и съответно на България.

Какво е заложено на карта?

Европейският съюз прави **смел опит** да демонстрира, че е възможно да се постигне технологичен напредък, който не жертва основните ценности на обществото. Това е деликатен баланс – твърде строгият контрол би задушил предприемачеството, твърде слабият би подкопал доверието и правата. С горепосочените препоръки за гъвкавост, подкрепа и международна координация, ЕС може да превърне регулаторния си подход в **конкурентно преимущество**: марка за качество "Произведено/ разработено в ЕС" да означава не само иновация, а *иновация, на която можем да вярваме*. Това е стратегическата визия, която законодателите следва да продължат да следват – с отворени очи за корекции, така че Европа да процъфтява като глобален център за **отговорни и устойчиви цифрови технологии**.

Гъвкавост, прозрачност и синхронизация остават ключовите принципи, чрез които Европейският съюз може да превърне своята амбициозна регулаторна рамка за изкуствения интелект в катализатор на доверие и устойчив растеж. Актът за ИИ, заедно със съпътстващите регламенти (DSA, DMA, CRA, GDPR) и директиви (напр., за платформената заетост), предлага комплексен и правозащитен модел на иновация, основан върху европейските ценности.

Въпреки това, ако ЕС не постигне **баланс между социалната защита и икономическата реалност**, регулацията рискува да се превърне в **тежест**, а не в **рамка за растеж**. Комбинираният ефект от прекомерна регулаторна сложност, забавени делегирани актове и непропорционално въздействие върху стартиращите компании би могъл да доведе до следните негативни сценарии:

- **Продължаващ отлив на ИИ инвестиции към САЩ, Обединеното кралство и Азия**, където екосистемите са по-гъвкави и предсказуеми;
- **Изместване на иновативни проекти и изследвания извън ЕС**, включително обучение и внедряване на базови ИИ модели (Foundation Models);
- **Консолидация на пазарната сила в ръцете на големи компании**, способни да се справят със сложната административна тежест, което задълбочава дисбаланса спрямо МСП;
- **Фрагментация на вътрешния пазар**, ако държавите-членки започнат да прилагат различно вторично законодателство или допълнителни национални изисквания;
- **Изоставане в темпа на внедряване на ИИ в ключови сектори** (здравеопазване, образование, транспорт), което ще забави цифровизацията и ще подкопае икономическата трансформация.

В дългосрочен план, ако ЕС не успее да адаптира регулаторната си политика към реалностите на глобалната конкуренция и технологично развитие, съществува реален риск да бъде възприеман **не като лидер в иновациите**, а **като пазител на забавянето** – със стабилна правна рамка, но с ниска практическа ефективност и слабо въздействие върху световния ИИ пазар.

ЕС има историческия шанс да създаде „европейски модел" на изкуствен интелект – сигурен, прозрачен, социално отговорен и иновативен. Но, за да бъде този модел конкурентоспособен, той трябва да се изгражда с постоянен диалог с индустрията, със смели инвестиции и с гъвкави, адаптивни механизми на управление, които подкрепят и насочват визионерския дух на европейските иноватори, без да възпрепятстват тяхната изобретателност чрез тежки регулации.

България

"Концепция за развитие на изкуствения интелект в България до 2030 г."

Концепцията за развитие на изкуствения интелект в България до 2030 г. очертава стратегическа визия за трансформиране на страната в устойчива и високотехнологична ИИ екосистема, целяща насърчаване на технологичния трансфер и създаване на собствени ИИ продукти.[103] Крайната цел до 2030 г. е изграждане на развита научна инфраструктура, наличие на висококвалифицирани човешки ресурси, интензивно международно сътрудничество и стабилна правна рамка в областта на ИИ.

Концепцията идентифицира ключови области на въздействие, включващи развитие на изследователския капацитет чрез подкрепа за научни изследвания и създаване на връзки между академичните среди и бизнеса. Друг важен аспект е създаването на знания и умения чрез адаптиране на образователната система към STEM дисциплините и подготовка на ИИ специалисти. Предвижда се изграждане на необходимата инфраструктура, включително инвестиции в изчислителни мощности и национални пространства за данни. Концепцията акцентира върху подкрепата за иновации в индустрията чрез финансови стимули и дигитални иновационни хъбове, както и върху повишаване на осведомеността и доверието чрез разработване на етична и регулаторна рамка, съобразена с европейските принципи.[104]

Концепцията откроява съществени предизвикателства пред България, сред които са недостатъчният брой висококвалифицирани кадри и ограниченото финансиране за научни изследвания. Констатира се слаба връзка между науката и бизнеса, което възпрепятства внедряването на иновации. Налице е и ниско ниво на дигитални умения сред населението и бизнеса, което забавя приемането на ИИ технологии, както и липса на мащабни национални данни, необходими за обучението на конкурентни ИИ модели.[105]

За преодоляване на тези предизвикателства и реализиране на целите, концепцията предлага план за действие и изпълнение, който включва създаване на междуведомствена работна група за координация и разработване на национален план/ пътна карта с конкретни мерки, срокове и индикатори. Финансирането се предвижда да бъде осигурено чрез национални и европейски програми, както и чрез публично-частни партньорства.

"Концепция за развитие на изкуствения интелект в България до 2030 г." представлява амбициозна рамка, която идентифицира ключови области за развитие и

[103] МС, "Концепция за развитие на изкуствения интелект в България до 2030 г." , 2020, retrieved from https://egov.government.bg/wps/portal/ministry-meu/strategies-policies/digital.transformation/itis-national-strategic-documents/ai.development.concept.2030
[104] МС, "Концепция за развитие на изкуствения интелект в България до 2030 г." , 2020, retrieved from https://egov.government.bg/wps/portal/ministry-meu/strategies-policies/digital.transformation/itis-national-strategic-documents/ai.development.concept.2030
[105] Ibid

предизвикателства, които страната трябва да преодолее, за да се утвърди като значим участник в глобалния ИИ пейзаж. Успешното изпълнение на предложените мерки ще изисква координирани усилия от страна на държавните институции, бизнеса и научните среди, както и осигуряване на адекватно финансиране и преодоляване на съществуващите структурни слабости.

Българската концепция за ИИ е добре структурирана и съобразена с европейските приоритети, но нейният успех зависи от ефективността на изпълнението. Страната трябва да ускори инвестициите си в наука, дигитални умения и бизнес иновации, за да не остане назад в европейския технологичен пейзаж. За да догони водещите европейски държави в сферата на изкуствения интелект, България трябва да се фокусира върху няколко стратегически насоки, а именно инвестиции в ИИ, научноизследователска и развойна дейност, внедряване на ИИ в индустрията и развитие на таланти.

Анализът на AI Index Report 2024 / Анализираните международни индекси и данни подчертават неотложната нужда от целенасочена национална политика за изкуствен интелект в България, фокусирана върху инвестиции, човешки капитал, индустриална трансформация, балансирана регулация за икономически растеж и дигитална независимост. За България е критично да се интегрира в глобалната научна екосистема, подкрепяйки научни изследвания и сътрудничеството между индустрията и академичния сектор чрез стимулиране на съвместни проекти и фондове за публикации. Също така е необходимо създаването на благоприятна среда за развитие на местни ИИ компании и привличане на чуждестранни инвестиции чрез иновационни хъбове, облекчаване на административни и данъчни тежести и подкрепа за трансфер на технологии, активно участвайки в европейски инициативи. Внедряването на ИИ в различни сектори на българската икономика, особено в МСП, с цел повишаване на производителността и качеството на работа, трябва да бъде приоритет, заедно с инвестициите в образованието и обучението за придобиване на ИИ умения на всички нива. Насърчаването на автоматизацията в производствения сектор чрез подкрепа и специализирани програми за обучение е ключово за повишаване на конкурентоспособността, като България трябва активно да инвестира в ИИ и да развива иновационна екосистема.

Научната активност в областта на изкуствения интелект е пряк индикатор за капацитета на България да се утвърди като иновативна икономика. Въпреки наличието на база и активни звена, България трябва да насочи усилията си не само към увеличаване на обема на публикациите, а към изграждане на международно конкурентноспособна научна общност, която генерира значимо въздействие. Това изисква политическа воля, стратегическо финансиране и координирани действия между академичния, частния и публичния сектор. Само така страната може да се позиционира устойчиво сред водещите играчи в новата дигитална ера.

България стои пред стратегически императив да възприеме проактивен и всеобхватен подход към развитието на изкуствения интелект. За да се превърне страната от наблюдател в активен участник на глобалната ИИ сцена и да се гарантира устойчив икономически и социален прогрес, е необходимо незабавно и координирано действие на всички нива – държавна администрация, научна общност, бизнес и представителите на работниците. Само чрез целенасочени инвестиции, стимулиране на иновациите, развитие на човешкия капитал и създаване на благоприятна и ефективна регулаторна среда, България може да се възползва от огромния потенциал на ИИ и да се позиционира като конкурентоспособен и значим фактор в дигиталната ера. Времето за стратегически решения е сега, за да не пропусне страната историческия си шанс.

Франция

През март 2024 г. Франция очерта амбициозна стратегия за утвърждаване като водеща сила в областта на ИИ чрез доклад на Комисията по изкуствен интелект, съдържащ 25 препоръки.[106] Ключови елементи от стратегията включват национален план за повишаване на осведомеността и обучение по ИИ с цел обществена готовност, съществени инвестиции в цифрови предприятия за укрепване на френската ИИ екосистема и изграждане на мощна изчислителна инфраструктура в Европа. Стратегията предвижда подобряване на управлението на личните данни с акцент върху защитата и улесняване на иновациите за обществена полза, както и популяризиране на френската култура чрез достъп до културно съдържание при спазване на правата върху интелектуалната собственост. Франция, също така, планира да подкрепи научните изследвания в областта на ИИ и да инициира международно сътрудничество за създаване на глобално управление на ИИ .[107]

В допълнение, Франция създава своя Национален институт за оценка и сигурност на ИИ - INESIA,[108] чиито ключови цели са анализ на системните рискове за националната сигурност и оценката на ефективността и безопасността на ИИ моделите. Мисията му е да укрепи управлението на ИИ и да смекчи рисковете чрез научни изследвания, оценки и международно сътрудничество. Институтът ще постигне това чрез обединяване на национални експерти, провеждане на научни изследвания, разработване на технически инструменти и регулаторни рамки, активно участие в международни мрежи за безопасност на ИИ и създаване на рамка на споделено доверие за отговорно развитие на ИИ. За тази цел, без да създава нова структура, институтът обединява националните участници в оценката и сигурността, и по-специално: Националната агенция за сигурност на информационните системи (ANSSI), Националният институт за научни изследвания в областта на цифровите науки и технологии (Inria), Националната лаборатория по метрология и изпитване (LNE), и Експертният център за цифрово регулиране (PEReN).[109] Това показва ангажираността и проактивните политики по отношение на сигурното развитие на ИИ технологиите в страната.

Френският подход разкрива стратегически ангажимент към балансирано развитие на ИИ, съчетаващо технологичен прогрес с етични съображения и защита на индивидуалните права. За разлика от очертаващия се американски подход под администрацията на Тръмп, който акцентира върху дерегулацията и индустриалното лидерство с фокус върху националната сигурност, френската стратегия демонстрира по-широк поглед, включващ обществено осведомяване, културно влияние и многостранно международно сътрудничество, като пример за това беше домакинстването на тазгодишното издание на поредицата форуми на високо равнище, посветени на ИИ - AI Action Summit 2025. Френската инициатива за глобално управление на ИИ е опит за насърчаване на координиран международен подход, който би могъл да смекчи потенциални конфликти и да насърчи сътрудничеството в тази ключова технологична област. В крайна сметка, различните подходи на водещите световни сили ще оформят бъдещата траектория на развитие и регулиране на ИИ в глобален мащаб, с потенциални последици за конкуренцията, сътрудничеството и установяването на международни стандарти.

---

[106] "25 recommandations pour l'IA en France", 2024 https://www.info.gouv.fr/actualite/25-recommandations-pour-lia-en-france , 17.03.2025

[107] "25 recommandations pour l'IA en France", 2024 https://www.info.gouv.fr/actualite/25-recommandations-pour-lia-en-france , 17.03.2025

[108] "La France se dote d'un Institut national pour l'évaluation et la sécurité de l'intelligence artificielle (INESIA) ",2025https://www.economie.gouv.fr/actualites/la-france-se-dote-dun-institut-national-pour-levaluation-et-la-securite-de-lintelligence 17.03.2025

[109] Ibid

5.3.2 САЩ

САЩ утвърждава водещата си позиция в глобалния развой и внедряване на изкуствен интелект, като активно започва оформянето на регулаторна рамка за управление на възможностите и предизвикателствата, произтичащи от тази трансформираща технология. Въпреки че администрацията на президента Байдън създаде основите за всеобхватно управление на ИИ,[110] се очаква администрацията на президента Тръмп да затвърди тази траектория с по-прагматичен и напорист подход, особено по отношение на контрола върху износа на стратегически технологии и международната конкуренция в сектора. Настоящият анализ ще се фокусира основно върху политиките и регулациите, провеждани от президента Тръмп, като тези на президента Байдън ще бъдат разглеждани в контекста на тяхната релевантност към настоящите насоки за запазване на американското лидерство и стратегическото адресиране на националните интереси и сигурността в динамичната ИИ екосистема.

Президентска заповед - януари 2025 г.
Премахване на пречки пред американското лидерство в ИИ

Основната цел на първата заповед по отношение на ИИ, подписана от президента Тръмп в първите дни от встъпване в длъжност,[111] е да се запази и засили глобалното доминиране на САЩ в областта на ИИ, като тази технология се разглежда като средство за насърчаване на човешкия просперитет, икономическата конкурентоспособност и националната сигурност. Друга важна цел е разработването на ИИ системи, които са свободни от идеологически пристрастия или инженерно създадени социални програми. Заповедта също така цели премахването на съществуващи ИИ политики и директиви, които възпрепятстват американските ИИ иновации.

За постигане на тези цели се предприеха следните основни мерки:

- **Разработване на План за действие за ИИ,** който ще бъде обсъден по-долу**.** В рамките на 180 дни се цели да се създаде план, ръководен от Помощника на президента по наука и технологии (APST), Специалния съветник по ИИ и крипто, и Помощника на президента по въпросите на националната сигурност (APNSA), в координация с други ключови съветници и ръководители на агенции. Създаването на новата позиция Специален съветник по ИИ и крипто (AI and crypto Czar) и включването на APNSA е показателно за изключителното значение, което САЩ отдават на ИИ за своята национална сигурност, технологично лидерство и цялостна икономическа доминация. Тази стъпка сигнализира за стратегическо осъзнаване на факта, че ИИ е ключов фактор за геополитическото влияние и икономическия просперитет на страната. Фокусът върху ИИ на най-високо държавно ниво подчертава ангажимента на САЩ да запазят своето технологично превъзходство и да гарантират, че развитието и внедряването на ИИ се случва в рамките на техните интереси за национална сигурност и икономическа мощ.
- **Отмяна и преглед на съществуващи политики:** Незабавно ще бъде извършен преглед на всички политики, директиви, разпоредби и други действия, предприети съгласно отменената Изпълнителна заповед 14110 (относно безопасен, сигурен и надежден ИИ), подписана от президента Байдън

[110] Executive Order 14110 "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence", 2023 https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence 09.03.2025
[111] Executive Order 14179, "Removing Barriers to American Leadership in Artificial Intelligence ", 2025, https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/ 09.03.2025

през 2023 г., което затвърждава подхода на президента Тръмп към дерегулации по отношение на свободата на развитие на иновациите, докато запазването на Рамката за разпространение на изкуствен интелект (Framework for AI Diffusion), отново по заповед на президента Байдън, която ще бъде обсъдена в точка "Експортен контрол" от настоящия анализ, подчертава агресивният подход, който САЩ предприема с цел запазване на технологично лидерство и засилване на националната сигурност.

Останалите акценти на заповедта включват премахване на несъответствия с новата политика на страната за глобално доминиране в областта на ИИ, а докато не бъдат финализирани премахванията на несъответствия, ще се предприемат стъпки за предоставяне на налични изключения съгласно съществуващите правила и политики. Също така, заповедта предвижда да се преразгледат меморандуми на Службата за управление и бюджет (OMB) към Белия дом.

Заповедта демонстрира ясно изразен приоритет за постигане и поддържане на глобално лидерство на САЩ в областта на изкуствения интелект. Акцентът върху премахването на "идеологически пристрастия или инженерно създадени социални програми" може да сигнализира за промяна в подхода спрямо предишни администрации, като се поставя по-голям фокус върху иновациите и конкурентоспособността, вероятно с по-малко внимание към етични или социални аспекти, които биха могли да се възприемат като "пристрастия". Създаването на конкретен план за действие в сравнително кратък срок (180 дни) показва решимост за бързо прилагане на новата политика. Включването на ключови фигури от сферата на науката, националната сигурност и икономиката в процеса на разработване на плана подчертава стратегическото значение на ИИ за администрацията. Отмяната на съществуваща изпълнителна заповед и прегледът на всички свързани действия предполага фундаментална промяна в подхода към регулирането на ИИ, като целта е да се създаде по-благоприятна среда за американските ИИ компании, като се премахнат бариери пред иновациите. Въпреки това, липсват конкретни мерки за това как ще се балансира стремежът към доминация с потенциалните рискове, свързани с ИИ.

Инвестиционна политика „Америка на първо място" по отношение на изкуствения интелект

Президентският меморандум за национална сигурност "Инвестиционна политика „Америка на първо място"" ("America First Investment Policy") от февруари 2025 г. разглежда изкуствения интелект в контекста на националната и икономическата сигурност на САЩ.[112]

В Раздел 2 (Политика), подточка (a) се заявява, че "Политика на САЩ е да се запази отворена инвестиционна среда, за да се гарантира, че **изкуственият интелект** и други нововъзникващи технологии на бъдещето се изграждат, създават и развиват точно тук, в Съединените щати". Това подчертава стратегическото значение на ИИ за бъдещето на страната и желанието тя да остане център на иновациите в тази област.

В подточка (f) от същия раздел се посочва, че САЩ ще засилят правомощията на Комитета за чуждестранни инвестиции в САЩ (CFIUS), за да "ограничат достъпа на чуждестранни противници до американски таланти и операции в чувствителни технологии (**особено изкуствен интелект**)". Това показва, че администрацията разглежда ИИ като критична и чувствителна технология, за която достъпът на определени чуждестранни субекти трябва да бъде ограничен с цел защита на националните интереси.

[112] NSPM "America First Investment Policy", 2025, https://www.whitehouse.gov/presidential-actions/2025/02/america-first-investment-policy/ 19.03.2025

В подточка (j) се споменава, че ще бъде извършен преглед на изходящите инвестиции на САЩ в Китай в сектори като "полупроводници, **изкуствен интелект**, квантови технологии, биотехнологии, хиперзвукови технологии, аерокосмическа индустрия, напреднало производство, насочена енергия и други области, засегнати от стратегията на КНР за военно-гражданско сливане". Това подчертава опасенията относно възможността американски инвестиции да подкрепят развитието на военния потенциал на Китай чрез напредъка в ИИ и други ключови технологии.

Меморандумът "Инвестиционна политика „Америка на първо място"" разглежда изкуствения интелект като жизненоважна технология за икономическия растеж и националната сигурност на САЩ. Той предвижда мерки за насърчаване на вътрешните иновации и ограничаване на достъпа на чуждестранни противници, особено Китай, до американски ИИ технологии, таланти и инвестиции. Това показва, че **ИИ заема централно място в инвестиционната политика на САЩ**, насочена към запазване на технологичното превъзходство и защита от потенциални заплахи.

Президентски съвет на съветниците по наука и технологии

Според Раздел 2(b) на заповедта,[113] "Специалният съветник по ИИ и крипто" ще бъде член на Президентския съвет на съветниците по наука и технологии (PCAST). Освен това, Раздел 2(c) уточнява, че "Специалният съветник по ИИ и крипто" ще служи като съпредседател на PCAST, заедно с Помощника на президента по наука и технологии (APST). Това ясно показва, че ще има високопоставен служител, пряко отговорен за въпросите, свързани с изкуствения интелект, в рамките на този съвет.

Тези изказвания ясно показват, че изкуственият интелект се разглежда като **централен елемент** в политиката на президента Тръмп, особено в контекста на националната сигурност и глобалната технологична конкуренция. Създаването на специална позиция на съпредседател на PCAST, отговорна за ИИ и крипто, допълнително подчертава стратегическото значение, което се отдава на тази технология в неговата администрация. Целта за постигане и поддържане на "безспорно и безпрецедентно глобално технологично господство" ясно позиционира ИИ като ключов фактор за реализирането на тази амбиция.

Създаването на План за действие за ИИ

Във връзка с изпълнението на дейностите, заложени в президентския указ за премахване на пречките пред американското лидерство в ИИ, бе публикувано запитване за информация относно разработването на План за действие в областта на ИИ (AI Action Plan).[114] Това запитване, чийто срок изтече на 15 март 2025 г., цели да събере мнения и предложения от различни заинтересовани страни относно ключови аспекти, които трябва да бъдат включени в бъдещия план за действие на САЩ в областта на ИИ. В отговор на това запитване, редица водещи компании в сферата на ИИ публикуваха своите позиции и препоръки. Сред тях са Google, OpenAI, Anthropic и Andreessen Horowitz (a16z) , (водещ венчър фонд). Публично достъпните им коментари показват, че въпреки специфичните акценти на всяка компания, се забелязват общи точки по отношение на ключови приоритети за развитието на ИИ в САЩ, което увеличава шанса тези

[113] Executive Order 14177, "Presidential Council on Science and Technology", 2025, https://www.whitehouse.gov/presidential-actions/2025/01/presidents-council-of-advisors-on-science-and-technology/ 19.03.2025

[114] National Science Foundation, "Request for Information on the Development of an Artificial Intelligence (AI) Action Plan",2025 https://www.federalregister.gov/documents/2025/02/06/2025-02305/request-for-information-on-the-development-of-an-artificial-intelligence-ai-action-plan 15.03.2025

предложения да бъдат приети и водещи в Плана за действие за ИИ.[115] [116] [117] [118] Откроява се следните общи тенденции в позициите на компаниите:

- **Приоритет на американското лидерство:** Всички компании единодушно поставят като първостепенна цел запазването и укрепването на водещата роля на Съединените щати в глобалния ИИ пейзаж.
- **Акцент върху иновациите:** Насърчаването на иновациите е ключова тема във всички предложения, макар и с различни нюанси относно степента и формата на държавна намеса.
- **Разпознаване на значението на националната сигурност:** Google, OpenAI и Anthropic изрично подчертават критичната роля на ИИ за националната сигурност, което предполага, че бъдещият план ще включва мерки за защита на тази сфера, имайки предвид, че това е вече възприета позиция на Белия дом.
- **Необходимост от балансиран подход:** Въпреки различните гледни точки относно регулацията, съществува разбиране за нуждата от балансиране между стимулирането на иновациите и адресирането на потенциални рискове и етични съображения.
- **Важност на инфраструктурата:** Инвестициите в необходимата инфраструктура, включително изчислителни ресурси и енергетика, се разглеждат като съществен фактор за поддържане на конкурентоспособността.

Този подход може да доведе до ускоряване на развитието и внедряването на ИИ технологии в САЩ, укрепвайки тяхната глобална конкурентоспособност. Акцентът върху националната сигурност може да доведе до по-строги мерки за контрол върху трансфера на ИИ технологии и знания към определени страни. Приемането на такъв план за действие от страна на САЩ би имало значителни последици за останалите страни, вкл.:

- **Засилена конкуренция:** Други държави ще бъдат подтикнати да ускорят собствените си стратегии за развитие на ИИ, за да не изостанат в технологичната надпревара.
- **Потенциално фрагментиране на глобалната ИИ екосистема:** Различните регулаторни подходи и ограничения върху трансфера на технологии могат да доведат до фрагментиране на глобалната ИИ екосистема, затруднявайки международното сътрудничество в определени области.
- **Необходимост от адаптиране:** Другите страни ще трябва да адаптират своите политики и инвестиции в ИИ, за да отговорят на американската стратегия и да запазят своята конкурентоспособност.
- **Възможности за сътрудничество в рамките на съюзнически отношения:** Страни, които са съюзници на САЩ, могат да намерят възможности за по-тясно сътрудничество в областта на ИИ, възползвайки се от американските инвестиции и технологичен напредък.

[115] "Google's comments on the U.S. AI Action Plan", 2025 https://blog.google/outreach-initiatives/public-policy/google-us-ai-action-plan-comments/ 17.03.2025
[116] "OpenAI's Proposals for the U.S AI Action Plan", 2025, https://openai.com/global-affairs/openai-proposals-for-the-us-ai-action-plan/ 17.03.2025
[117] "Anthropic's Recommendations to OSTP for the U.S. AI Action Plan", 2025, https://www.anthropic.com/news/anthropic-s-recommendations-ostp-u-s-ai-action-plan 17.03.2025
[118] "a16z's Recommendations for the National AI Action Plan
", 2025, https://a16z.com/a16zs-recommendations-for-the-national-ai-action-plan/ 17.03.2025

Предвид силната ориентация на президента Тръмп към тясно сътрудничество с индустрията, неговата подкрепа за дерегулация и убеждението, че предоставянето на свобода за иновации е ключов фактор за националната сигурност и икономическия растеж на САЩ, може да се заключи, че бъдещият План за действие за ИИ ще бъде силно повлиян от тези принципи. В този контекст, предложенията на Google, OpenAI, Anthropic и Andreessen Horowitz залагат на създаване на благоприятна среда за иновации с минимални регулаторни ограничения, противоположно на европейския подход. Планът за действие най-вероятно ще отрази желанието на администрацията на Тръмп да даде приоритет на бързото развитие и внедряване на ИИ технологии, като се разглеждат **регулациите като потенциални бариери пред прогреса**. Акцентът ще бъде поставен върху насърчаване на предприемачеството и конкуренцията, особено сред по-малките технологични компании, в съответствие с вижданията на Andreessen Horowitz.[119]

Въпреки че въпросите за националната сигурност, повдигнати от Google,[120] OpenAI[121] и Anthropic,[122] ще бъдат взети под внимание, подходът вероятно ще се фокусира върху мерки, които не задушават иновациите, като например улесняване на сътрудничеството между индустрията и правителствените агенции за сигурност, вместо налагане на строги регулаторни рамки, като индустрията вече даде ясен сигнал за засилване на сътрудничеството както с военно-индустриалния комплекс и така и с федералните агенции за сигурност.[123] [124] [125]

Следователно, може да се очаква, че Планът за действие при администрация на Тръмп ще бъде по-скоро насочен към създаване на условия за безпрепятствено развитие на ИИ, отколкото към въвеждане на обширни регулаторни механизми, като се разчита на индустрията да води иновациите. Този подход ще бъде в пряко съответствие с основните принципи на президента Тръмп за стимулиране на икономиката и укрепване на националната сигурност чрез подкрепа на американския бизнес и технологично лидерство.

Меморандум за националната сигурност, издаден от президента Байдън през октомври 2024 г.

На 24 октомври 2024 г. Белият дом публикува Меморандум за насърчаване на лидерството на Съединените щати в областта на изкуствения интелект (ИИ), озаглавен "Използване на изкуствения интелект за постигане на целите на националната сигурност и насърчаване на безопасността, сигурността и надеждността на изкуствения интелект".[126]

---

[119] "a16z's Recommendations for the National AI Action Plan
", 2025, https://a16z.com/a16zs-recommendations-for-the-national-ai-action-plan/ 17.03.2025
[120] "Google's comments on the U.S. AI Action Plan", 2025
https://blog.google/outreach-initiatives/public-policy/google-us-ai-action-plan-comments/ 17.03.2025
[121] "OpenAI's Proposals for the U.S AI Action Plan", 2025, https://openai.com/global-affairs/openai-proposals-for-the-us-ai-action-plan/ 17.03.2025
[122] "Anthropic's Recommendations to OSTP for the U.S. AI Action Plan", 2025,
https://www.anthropic.com/news/anthropic-s-recommendations-ostp-u-s-ai-action-plan 17.03.2025
[123] "Lockheed Martin and Google Cloud Announce Collaboration to Advance Generative AI For National Security", 2025,https://news.lockheedmartin.com/2025-03-27-Lockheed-Martin-and-Google-Cloud-Announce-Collaboration-to-Advance-Generative-AI-For-National-Security 28.03.2025
[124] "Anduril Partners with OpenAI to Advance U.S. Artificial Intelligence Leadership and Protect U.S. and Allied Forces", 2024, https://www.anduril.com/article/anduril-partners-with-openai-to-advance-u-s-artificial-intelligence-leadership-and-protect-u-s/ 19/03.2025
[125] "Azure for US Department of Defense", https://azure.microsoft.com/en-us/explore/global-infrastructure/government/dod 19.03.2025
[126] National Security Memorandum, "Memorandum on Advancing the United States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence", 2024, https://bidenwhitehouse.archives.gov/briefing-room/presidential-actions/2024/10/24/memorandum-on-advancing-the-united-states-leadership-in-artificial-

Основни акценти на меморандума:

1. **Укрепване на лидерството на САЩ в ИИ, което ще бъде постигнато чрез** разработването и внедряването на ИИ технологии, особено в контекста на националната сигурност.
2. **Интеграция на ИИ в националната сигурност, чрез насочване на** федералните агенции да използват авангардни ИИ технологии за постигане на цели, свързани с националната сигурност, като същевременно се гарантира защита на човешките права и демократичните ценности.
3. **Международно сътрудничество, с цел** да се установят глобални норми и стандарти за безопасно и етично използване на ИИ, близки до американския модел.
4. **Защита на правата и свободите, чрез** акцентиране върху необходимостта от защита на гражданските права, свободите, личната неприкосновеност и безопасността при използването на ИИ в дейности, свързани с националната сигурност. Възможно е този аспект от меморандума да бъде преразгледан от настоящата администрация.

Политиките, разработени от предходната администрация, и тези, които гради настоящата, очертават сравнително последователен и напорист подход на САЩ за утвърждаване и запазване на глобално лидерство в областта на изкуствения интелект. Наблюдава се еволюция от по-обхватен подход при администрацията на президента Байдън към по-прагматична и фокусирана върху националната сигурност и икономическото превъзходство стратегия при администрацията на президента Тръмп. Ключова характеристика на очертаващия се подход е **силната връзка с индустрията и акцентът върху дерегулацията като средство за стимулиране на иновациите**. Назначаването на представител от средите на венчър капитала - Дейвид Сакс, за специален съветник по ИИ и крипто, като съпредседател на PCAST[127] и премахването на политики, възпрепятстващи иновациите, сигнализират за водещата роля на частния сектор в оформянето на бъдещите ИИ политики.

Основната цел за неоспорима глобална технологична доминация се разглежда като императив за националната сигурност и икономически просперитет. Това се проявява в мерки за засилване на експортния контрол, ограничаване на инвестициите от чуждестранни противници (особено Китай) и насърчаване на инвестиции от съюзници.

Въпреки че се признава значението на социалните аспекти на ИИ, фокусът на администрацията на Тръмп изглежда е изместен към бързото развитие и внедряване на ИИ технологии, като регулациите се разглеждат предимно като потенциални бариери. Този подход може да доведе до ускоряване на иновациите и укрепване на конкурентоспособността на САЩ в краткосрочен план. Въпреки това, в дългосрочен план, прекалено силният акцент върху дерегулацията и националния интерес може да има двойствен ефект върху американското лидерство. От една страна, бързите иновации могат да затвърдят технологичното превъзходство. От друга страна, потенциалното пренебрегване на международното сътрудничество и етичните стандарти може да отслаби глобалното влияние на САЩ в определянето на бъдещите норми и стандарти за ИИ, ако не се намери баланс с глобалните етични и социални измерения на развитието на ИИ.

intelligence-harnessing-artificial-intelligence-to-fulfill-national-security-objectives-and-fostering-the-safety-security/ 20.03.2025

[127] "What Trump's New AI and Crypto Czar David Sacks Means For the Tech Industry
" 2024 https://time.com/7200518/david-sacks-new-white-house-ai-crypto-czar-trump-administration/ 20.03.2025

5.3.3 Китай

Въпреки че Китай демонстрира ускорен технологичен напредък в областта на изкуствения интелект през последните години, амбициозно преследвайки лидерска позиция в глобалната надпревара със САЩ, съществуват важни нюанси, произтичащи от неговия авторитарен режим. Докато публикуваните етични кодекси, регулации за алгоритми и временни мерки за генеративен ИИ дават представа за намеренията на правителството, ограниченият достъп до информация и специфичният контрол върху медиите пораждат неясноти относно истинското състояние на китайската ИИ екосистема. Това се отнася както до реалната конкурентоспособност на местните ИИ компании на международната сцена, така и до потенциалните слаби страни и предизвикателства, пред които те се изправят. Анализът на публикуваните документи разкрива силен акцент върху държавния контрол, националната сигурност и социалистическите ценности, което може да има както позитивни, така и негативни последици за развитието на иновациите и конкурентоспособността в дългосрочен план.

"Етичен кодекс за новото поколение изкуствен интелект"

Издаденият на 25 септември 2021 г. "Етичен кодекс за новото поколение изкуствен интелект" представлява рамков документ, целящ интегрирането на етични съображения в целия жизнен цикъл на изкуствения интелект (ИИ). Кодексът предоставя насоки за физически и юридически лица и други институции, ангажирани с дейности, свързани с ИИ, с основен фокус върху етика, справедливост, защита на лични данни и отговорност, като за целта са представени задължителни етични норми.[128]

Основната цел на кодекса е да насърчи етичното развитие и прилагане на ИИ технологиите. Той се базира на няколко основни принципа, включващи подобряване на благосъстоянието на хората, подкрепа на справедливостта и равенството, защита на личните данни и сигурност, осигуряване на контролируемост и доверие в ИИ системите, признаване на човешката отговорност и повишаване на етичната грамотност в областта на ИИ (чл. 3).

Кодексът очертава специфични етични изисквания, приложими в различни етапи от жизнения цикъл на ИИ. В областта на управлението на ИИ се акцентира върху гъвкавостта, зачитането на правата на всички заинтересовани страни и избягването на необмислени решения (глава 2). За изследванията и разработките се препоръчва внедряване на етика в научния процес, гарантиране на качеството на данните и предотвратяване на пристрастия (глава 3). В контекста на доставките и пазара се насърчава конкурентоспособността, спазването на пазарните правила и защитата на потребителските права (глава 4). По отношение на използването на ИИ фокусът е върху добросъвестно приложение, избягване на злоупотреби и нарушения на закона (глава 5). Кодексът, също така, подчертава ролята на държавата и други организации в следването на тези етически стандарти и създаването на специфични мерки, съобразени с обществените нужди (глава 6).

Етичният кодекс поставя акцент върху защитата на личните данни и сигурността. Принципът за контролируемост и доверие в ИИ системите подчертава необходимостта от запазване на човешкия надзор и възможността за взаимодействие и отказ от автоматизирани процеси, но предвид авторитарният режим в Китай, "човешки контрол" следва да се разглежда като партиен контрол, под ръководството на президента Си. Насърчаването на етичната грамотност, в съответствие със социалистическите ценности, и активното решаване на етични въпроси показва стремеж към изграждане на култура на етично осъзнатост в областта на ИИ.

[128] "新一代人工智能伦理规范", 2021, https://www.most.gov.cn/kjbgz/202109/t20210926_177063.html 21.03.2025

"Разпоредби относно администрирането на алгоритмични препоръки в интернет информационни услуги"

"Разпоредби относно администрирането на алгоритмични препоръки в интернет информационни услуги" представлява регулаторен акт, основан на законите за киберсигурността и защита на личните данни, целящ регулирането на алгоритмичните препоръки в интернет услугите в Китай. Крайната цел е защитата на националната сигурност, обществения интерес и правата на гражданите.[129]

Регулациите се прилагат за всички интернет услуги в Китай, които използват алгоритмични технологии за препоръчване на информация, като се подчертава че алгоритмичните препоръки трябва да се основават на социалистическите основни ценности и не трябва да бъдат използват за незаконни дейности или разпространение на вредна информация (чл.2). Налага се изискване за осигуряване на безопасност на алгоритмите, защита на данни и защита срещу киберпрестъпления. Препоръчва се редовно преглеждане и тестване на алгоритмите за етика и съответствие със законодателството (глава 2). Правилата предвиждат потребителите да бъдат информирани за принципите и целите на алгоритмичните услуги и да имат възможност да откажат алгоритмичните препоръки, както и да им бъде предоставена опция за контрол върху личните им данни. Предвидени са специални разпоредби за защита на непълнолетни и възрастни хора, ползващи тези услуги (глава 3). За надзор и управление е въведен механизъм за регистриране на провайдерите на алгоритмични услуги в съответните държавни органи (глава 4). Провайдерите трябва да преминат през процедури за безопасност и оценка на рисковете, особено ако влияят на общественото мнение (глава 4). При нарушения са предвидени сериозни санкции, включително глоби и забрана за извършване на услуги (глава 6).

Основният акцент на регулацията е върху етичността, безопасността и прозрачността на алгоритмичните препоръки в интернет услугите. Фокусът е силно насочен към защитата на правата на потребителите. Правилата изискват яснота относно принципите на работа на алгоритмите и тяхното въздействие върху потребителите, като същевременно осигуряват механизми за надзор и контрол от страна на държавните органи. Съществен аспект е и предотвратяването на злоупотреби с алгоритмите, които биха могли да нарушат социалния ред, да манипулират потребителите или да създадат изкривена публична информация, а изискването за основаване на препоръките на социалистическите основни ценности подчертава идеологическия контекст на регулацията.

"Разпоредби относно администрирането на алгоритмични препоръки в интернет информационни услуги" представляват комплексен регулаторен механизъм, целящ да осигури по-етично, безопасно и прозрачно функциониране на алгоритмичните препоръчителни системи в Китай. Фокусът върху защитата на потребителските права и предотвратяването на злоупотреби подчертава ангажимента на държавата към регулирането на информационното пространство в съответствие със своите ценности и интереси.

[129] "互联网信息服务算法推荐管理规定", 2022, http://www.cac.gov.cn/2022-01/04/c_1642894606364259.htm 21.03.2025

"Разпоредби относно управлението на задълбочен синтез на интернет информационни услуги" („дъблок синтез" е термин, използван в КНР, еквивалентен на „deepfake")

Разпоредбите[130], които влязоха в сила на 10 януари 2023 г. (глава 5), въвеждат регулиране на използването на технологии за дълбок синтез за интернет информационни услуги, като се цели укрепване на управлението, утвърждаване на социалистически ценности, национална сигурност и защита на права, и се прилагат за всяко използване на дълбок синтез, като отговорността за това е на държавните органи. Забранява се незаконната информация, която следва да се интерпретира като цензура. Отговорността за това е на доставчиците, които са задължени да регистрират потребители, да преглеждат алгоритми и етични стандарти, да извършват контрол на информацията, защита на данни, предотвратяване на измами, реагиране при извънредни ситуации (глава 1). Разпоредбите целят укрепване на управлението на данните за обучение, гарантиране на сигурността, спазване на правилата за защита на личната информация, както и изискването на съгласие при редактиране на биометрични данни. В допълнение трябва да се извършва Периодичен преглед и оценка на алгоритмите и на сигурността (глава 2). Заложени са технически мерки за добавяне на символи към синтезирано съдържание, съхраняване на логове, както и ясно етикетиране на подвеждащо съдържание, забрана за премахване на етикети (глава 3). Регистрацията на влиятелни доставчици и технически поддръжници е задължителна, както и оценките за сигурност при нови продукти/функции. Предвиждат се надзорни инспекции от съответните органи и вземането на възможни мерки при рискове за сигурността, чрез спиране на услуги и налагане на наказания при неспазване на регулацията (глава 4).

Разпоредбите за дълбок синтез в Китай разкриват силно централизиран и контролиран подход към регулирането на ИИ. Основната цел е укрепване на държавното управление и утвърждаване на социалистическите ценности, а националната сигурност и защитата на права са поставени като ключови приоритети. Регулацията налага широк спектър от задължения на доставчиците, вкл. строг контрол върху потребителите, съдържанието и алгоритмите, което ефективно въвежда ново ниво на цензура. Акцентът върху управлението на данни и технологии, задължителното етикетиране и стриктният надзор демонстрират намерението на китайското правителство да установи пълен контрол върху използването на дълбок синтез в интернет пространството, като се предвиждат сериозни санкции за неспазване на тези изисквания.

"Временни мерки за управление на услуги за генеративен изкуствен интелект"

Временните мерки[131], които влязоха в сила на 15 август 2023 г., целят развитие и регулиране на ИИ, национална сигурност, обществен интерес, защита на права, като се прилагат за услуги, предоставящи съдържание на обществеността, с изключение на изследвания. Мерките изискват спазване на социалистическите ценности, предотвратяване на дискриминация, зачитане на интелектуалната собственост, прозрачност (Глава 1).

Изискванията към доставчиците за обработка на данни са в множество области като законен произход, защита на интелектуална собственост и лични данни, качество на данните, правила за етикетиране (Глава2)**.** Отговорностите на доставчиците, от друга страна, са свързани със сключване на споразумения с потребителите, насочване към законно използване, предотвратяване на зависимост при непълнолетни, защита на потребителска информация, забрана за събиране на ненужни лични данни, обработване

[130] "互联网信息服务深度合成管理规定", 2022, https://www.cac.gov.cn/2022-12/11/c_1672221949354811.htm 22.03.2025

[131] "生成式人工智能服务管理暂行办法", 2023, https://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm 21.03.2025

на искания за достъп, коригиране или изтриване на данни, маркиране на генерирано съдържание, осигуряване на стабилни услуги, мерки при незаконно съдържание и докладване (Глава 3)**.** Ролята на държавните органи е предвидена да бъде свързана с надзор и правна отговорност чрез разработване на специфични правила, извършване на оценки за сигурност и регистриране на алгоритми за влиятелни услуги (Глава 4).

Акцентът на временните мерки е върху балансиране на иновациите със сигурност и обществен интерес, като се фокусират върху насърчаване на развитието, регулиране на доставчиците, надзор, защита на ценности, защита на непълнолетни и регулиране на трансгранични услуги.

"Рамка за управление на безопасността на ИИ"

Документът очертава всеобхватна рамка за управление на безопасността на изкуствения интелект и цели да подкрепи Глобалната инициатива за управление на ИИ на Китай и да стимулира съгласувани усилия между различни заинтересовани страни, включително правителства, международни организации, компании, изследователски институти, граждански организации и физически лица, за ефективното управление на безопасността на ИИ.[132]

Рамката се основава на редица фундаментални принципи, които ръководят подхода към управлението на безопасността на ИИ, като намиране на баланс между развитието и сигурността, отдаване на приоритет на иновативното развитие на ИИ, ефективно предотвратяване и ограничаване на рисковете за безопасността на ИИ, създаване на механизми за управление, осигуряване на отговорност от всички участници, изграждане на верига за управление, ефективна защита на националния суверенитет, защита на законните права и интереси на граждани, юридически лица и други организации(раздел 1). [133]Тези принципи формират етична и прагматична основа за разработване и внедряване на ИИ технологии, като акцентират върху отговорността и сътрудничеството между различните участници**.**

Самата рамка очертава конкретни мерки за контрол, насочени към справяне с разнообразни рискове за безопасността на ИИ посредством комбинация от технологични и управленски стратегии. Мерките са свързани с идентифициране на рисковете за безопасност и сигурност, предлагане на технически контрамерки, прилагане на всеобхватни мерки за управление и предоставяне на насоки за безопасност при разработването и прилагането на ИИ (раздел 2). [134] Тази структура осигурява систематичен подход към управлението на безопасността, като обхваща както техническите аспекти на ИИ, така и по-широките управленски и регулаторни аспекти

Рамката детайлно класифицира рисковете за безопасността на ИИ в две основни категории (раздел 3)[135]:

- **Присъщи рискове за безопасността на ИИ:** Тези рискове произтичат от вътрешни технически ограничения и уязвимости на ИИ системите. Те включват рискове от модели и алгоритми, рискове от данни, рискове от ИИ системи. Подкатегорията "рискове от ИИ системи" акцентира върху вътрешните слабости на ИИ системите, които могат да бъдат експлоатирани или да доведат до непредвидени и нежелани резултати.

[132] "AI Safety Governance Framework", 2024, https://www.tc260.org.cn/upload/2024-09-09/1725849192841090989.pdf
[133] Ibid., 1
[134] Ibid., 4
[135] "AI Safety Governance Framework", 2024, https://www.tc260.org.cn/upload/2024-09-09/1725849192841090989.pdf 5-12

- **Рискове за безопасността при приложенията на ИИ:** Тези рискове възникват в резултат на потенциална злоупотреба, неправилна употреба или злонамерено използване на ИИ технологиите. Те включват рискове в киберпространството , реални рискове когнитивни рискове и етични рискове. Тази категория подчертава външните заплахи и етичните дилеми, свързани с разпространението и използването на ИИ в различни сфери.

В раздел 4 от Рамката е предложен набор от технологични мерки, насочени към предотвратяване на рискове в ключови области. Тези мерки адресират както присъщите рискове за безопасността на ИИ, така и рисковете за безопасността при приложенията на ИИ. Това показва холистичен подход, който интегрира сигурността в целия жизнен цикъл на ИИ системите.[136]

Рамката акцентира върху необходимостта от разработване и усъвършенстване на всеобхватни механизми и разпоредби за управление на рисковете за безопасността и сигурността на ИИ, които да включват участието на множество заинтересовани страни. Част от мерките включват прилагане на многостепенно и базирано на категории управление за ИИ приложения, разработване на система за управление на проследимостта за ИИ услуги, създаване на система за отговорна научноизследователска и развойна дейност и приложение на ИИ, укрепване на сигурността на веригата за доставки на ИИ, споделяне на информация и реагиране при извънредни ситуации при рискове и заплахи за безопасността на ИИ, подобряване на обучението на таланти в областта на безопасността на ИИ. Тези мерки подчертават значението на регулаторната рамка, сътрудничеството и развитието на експертиза за ефективното управление на ИИ безопасността (раздел 5).[137]

Рамката предоставя и конкретни насоки за безопасност, насочени към различни участници в процеса на разработка и внедряване на ИИ, включително разработчици на модели и алгоритми, доставчици на ИИ услуги, потребители в ключови области и общи потребители. Тези насоки обхващат широк спектър от аспекти, включващи етика, сигурност на данните, защита на личната информация, оценка на пристрастията, тестове за оценка на безопасността и сигурността, управление на външни смущения. Тези детайлни насоки са от съществено значение за практическото прилагане на принципите и мерките, заложени в рамката (раздел 6).[138]

Акцентът на регулациите е ясно насочен към проактивното управление на рисковете, свързани с развитието и прилагането на ИИ. Фокусът е върху създаването на многостранен подход, който включва както технологични, така и управленски мерки, и ангажира всички заинтересовани страни. Рамката обръща специално внимание на класификацията на рисковете, което позволява по-целенасочено разработване на контрамерки Регулациите се стремят да постигнат баланс между насърчаването на иновациите и гарантирането на безопасността, като признават потенциала на ИИ за развитие, но същевременно подчертават необходимостта от предпазливост и отговорност. Наблюдава се и силен акцент върху международното сътрудничество и обмена на информация, което е от съществено значение за ефективното управление на глобалния характер на ИИ технологиите. Тази рамка представлява зрял и всеобхватен опит за структуриране на управлението на безопасността на ИИ.

Регулаторният подход на Китай към изкуствения интелект е силно централизиран и рестриктивен. Основна цел на регулациите е запазването на националната сигурност, обществения интерес и утвърждаването на социалистическите основни ценности. Тези приоритети често доминират над други съображения, като например, пълната свобода

[136] Ibid., 13-16
[137] "AI Safety Governance Framework", 2024, https://www.tc260.org.cn/upload/2024-09-09/1725849192841090989.pdf 17 -20
[138] Ibid., 21 - 26

на иновациите или неограничената защита на индивидуалната поверителност в западния смисъл. Наблюдава се ясна тенденция за установяване на строг държавен контрол върху всички аспекти на развитието и използването на ИИ – от етичните насоки до технологичните спецификации и разпространението на информация. Въвеждат се множество изисквания и задължения за доставчиците на ИИ услуги, включително регистрация, контрол на съдържанието, защита на данни и спазване на идеологически норми. Предвидени са сериозни санкции за нарушения, което допълнително подчертава рестриктивния характер на регулаторната рамка.

В контекста на авторитарния режим в Китай, този регулаторен подход е логичен. Той позволява на правителството да насочва развитието на ИИ в съответствие със своите стратегически цели и да минимизира потенциалните рискове, които тази технология може да представлява за социалната стабилност и политическия контрол. Докато насърчаването на иновациите също е декларирана цел, тя винаги е поставена в рамките на тези по-широки държавни приоритети.

5.4. Етичен и надежден изкуствен интелект

Международният доклад за безопасност на ИИ за 2025 г. на Обединеното кралство, изготвен от 100 експерти от 33 държави, представлява ключов документ за разбиране и управление на рисковете, свързани с напредналите възможности на изкуствения интелект. [139] Докладът подчертава непредсказуемостта на развитието на ИИ и необходимостта от основан на доказателства подход при вземането на политически решения. Той очертава редица потенциални рискове, включително злонамерено използване, неизправности и системни сривове, и призовава за спешни и адаптивни действия от страна на политиците и заинтересованите страни. Докладът препоръчва разработването на системи за ранно предупреждение и международни рамки за управление на риска, както и задълбочени изследвания в областта на оценката и управлението на риска при ИИ, с цел да се гарантира, че ползите от ИИ се реализират безопасно чрез информирано управление на риска и международно сътрудничество.[140]

В епохата на експоненциален напредък в областта на изкуствения интелект, осигуряването на неговата етичност и надеждност се превръща в ключов императив за устойчивото и ползотворно интегриране на тази технология в обществото. Множество международни организации, правителствени институции, индустриални лидери и академични среди активно работят по дефинирането на принципи, създаването на регулаторни рамки и разработването на инструменти, които да гарантират отговорното развитие и използване на ИИ. Настоящият анализ ще разгледа в дълбочина информацията, предоставена в редица авторитетни източници, за да очертае сложния пейзаж на етиката и надеждността в областта на ИИ.

Европейският съюз е пионер в разработването на всеобхватни етични насоки за надежден ИИ, които се базират на фундаментални права и ценности. Тези насоки дефинират надеждния ИИ като такъв, който е едновременно етичен и технически стабилен. Етичността се измерва чрез спазването на седем ключови изисквания: човешка намеса и надзор, разнообразие, недискриминация и справедливост, обществено и екологично благосъстояние, и отчетност. Техническата стабилност пък обхваща аспекти като техническа стабилност и безопасност, поверителност и управление на данните, и прозрачност.[141] Тези детайлни насоки отразяват амбицията на ЕС да създаде

[139] "International AI Safety Report 2025", 2025, retrieved from https://www.gov.uk/government/publications/international-ai-safety-report-2025
[140] "International AI Safety Report 2025", 2025, retrieved from https://www.gov.uk/government/publications/international-ai-safety-report-2025
[141] "Ethics guidelines for Trustworthy AI", 2019, retrieved from https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

строга регулаторна среда, която да гарантира, че развитието на ИИ в Европа ще бъде в съответствие с високи етични стандарти.

Изследванията в областта на ИИ и модерацията на съдържание също подчертават етичните дилеми, свързани с използването на тези технологии.[142] Въпреки че ИИ може да бъде ефективен за откриване и премахване на вредно съдържание, съществуват рискове от пристрастия и грешки, водещи до цензура на легитимно съдържание и ограничаване на свободата на словото.[143] Това налага разработването на прецизни алгоритми и човешки надзор, за да се гарантира справедливо и безпристрастно модериране на онлайн пространството.

В глобален мащаб се полагат усилия за стандартизиране на етичните принципи в областта на ИИ. Международната организация по стандартизация (ISO) разработва стандарти, които да предоставят насоки за отговорно и етично проектиране и използване на ИИ системи.[144] Тези стандарти имат за цел да помогнат на организациите да внедрят етични съображения в своите ИИ практики.

ООН, ЮНЕСКО и МОТ имат разработки и препоръки по отношение на етичния ИИ. „Управление на ИИ за човечеството" на ООН[145] е рамка, която се фокусира върху етичното и отговорно развитие и използване на изкуствен интелект в полза на човечеството. Тя подчертава необходимостта от глобално сътрудничество и управление, за да се гарантира, че ИИ се използва по начин, който зачита човешките права, насърчава устойчивото развитие[146] и предотвратява потенциални вреди. Препоръките на ЮНЕСКО за етичен ИИ са международен инструмент, който цели да насочи развитието и използването на изкуствен интелект в съответствие с човешките права, основните свободи и етичните принципи[147]. През 2024 г. МОТ[148] стартира специализирана обсерватория, която цели да осигури надеждна и актуална информация за влиянието на ИИ върху работните места, уменията, условията на труд и социалното осигуряване. Тя се стреми да насърчи диалога между правителства, работодатели, работници и други заинтересовани страни за разработване на политики, които да гарантират справедлив и устойчив преход към дигиталната икономика. Обсерваторията се фокусира върху различни аспекти на връзката между ИИ и труда, включително: автоматизация и загуба на работни места; създаване на нови работни места и умения; промени в условията на труд и организациите; етични и социални последици от ИИ.

Индустриалните лидери също играят важна роля в насърчаването на етичен ИИ. Компании като Microsoft, Amazon и Google са формулирали свои собствени принципи за отговорен ИИ, които обхващат ключови аспекти като справедливост, надеждност, безопасност, поверителност, прозрачност и отчетност.[149] [150] [151] [152]Тези принципи служат

---

[142] Ahmed, Abdullah & Khan, Mudassar, "AI and Content Moderation: Legal and Ethical Approaches to Protecting Free Speech and Privacy." 2024, 10.13140/RG.2.2.23619.41767; DOI:10.13140/RG.2.2.23619.41767

[143] Ibid

[144] "Building a responsible AI: How to manage the AI ethics debate", https://www.iso.org/artificial-intelligence/responsible-ai-ethics 20.03.2025

[145] Governing AI for Humanity, September 2024; https://www.un-ilibrary.org/content/books/9789211067873

[146] UN Resolution on Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development, 11.03.2024 - https://docs.un.org/en/A/78/L.49

[147] Ethics of Artificial Intelligence - The Recommendationhttps://www.unesco.org/en/artificial-intelligence/recommendation-ethics Retrived 3.4.2025

[148] ILO's Observatory on AI and Work in the Digital Economy - https://www.ilo.org/artificial-intelligence-and-work-digital-economy

[149] "What is Responsible AI?", 2025, https://learn.microsoft.com/en-us/azure/machine-learning/concept-responsible-ai?view=azureml-api-2 09.03.2025

[150] "Responsible AI Transparency Report", retrieved from https://www.microsoft.com/en-us/ai/principles-and-approach, 09.03.2025

[151] "Transform responsible AI from theory into practice", https://aws.amazon.com/ai/responsible-ai 09.03.2025

[152] "Responsible AI", https://cloud.google.com/responsible-ai?hl=en 09.03.2025

като основа за вътрешни политики и практики, насочени към минимизиране на рисковете и максимизиране на ползите от ИИ.

Инициативи като Глобалното партньорство за изкуствен интелект (GPAI) обединяват усилията на множество държави и организации за насърчаване на отговорния ИИ чрез конкретни проекти и дискусии. Доклади като GIRAI Report 2024 и AI Index Report на Stanford HAI предоставят ценни анализи и данни за текущото състояние на развитието на ИИ и свързаните с него етични и социални последици.[153] [154] Тези ресурси помагат за по-доброто разбиране на предизвикателствата и възможностите, свързани с ИИ.

Докато различни региони и държави се опитват да регулират тази технология, е от съществено значение да се разбере глобалният контекст на тези заплахи и как изключенията в регулаторните рамки могат да подкопаят демократичните ценности в световен мащаб. Настоящият анализ се фокусира върху тези опасности, като използва информация от предоставените линкове, за да очертае сложната връзка между ИИ, властта и потенциалното ограничаване на основните свободи в глобален мащаб.

Докладът на CSIS "Government Deepfakes" подчертава глобалния риск от използване на deepfakes от правителства за кампании за дезинформация и влияние. Тази заплаха не е ограничена до един регион – правителства по целия свят могат да използват тези технологии за манипулиране на общественото мнение в други държави, за намеса във вътрешните работи на чужди страни или за консолидиране на властта у дома чрез създаване на фалшиви наративи и дискредитиране на опозицията.[155]

Публикацията относно ИИ и модерацията на съдържание разглежда глобалните последици от използването на ИИ за цензура.[156] Правителства по целия свят могат да внедряват ИИ-базирани системи за автоматично филтриране и премахване на онлайн съдържание, което те считат за вредно или нежелателно. Въпреки че това може да бъде оправдано с борбата срещу тероризма или езика на омразата, съществува значителен риск от злоупотреба за потискане на свободата на изразяване и заглушаване на политическата опозиция в различни страни.

Докладът на Freedom House "Репресивната сила на изкуствения интелект" документира широкото използване на ИИ за наблюдение, цензура и репресии от правителства по света. Тези практики представляват сериозна заплаха за демократичните ценности и основните свободи в глобален мащаб.[157] Обезпокоителен е фактът, че Регламентът за изкуствения интелект на ЕС, въпреки че е насочен към установяване на етични стандарти, съгласно член 2, параграф 2, изрично изключва от своя обхват ИИ моделите, разработени или използвани изключително за целите на националната сигурност.[158] Това изключение крие потенциални рискове, тъй като позволява използването на мощни ИИ инструменти за репресивни цели без съответния надзор и контрол, предвиден в регламента. Възможността за масово наблюдение, автоматизирано профилиране и целенасочени мерки, основани на ИИ, без строги етични

---

[153] Maslej, N. et al., 2024. *Artificial Intelligence Index Report 2024*, Human-Centered Artificial Intelligence. United States of America. Retrieved from https://hai.stanford.edu/ai-index/2024-ai-index-report

[154] "Global Index on Responsible AI", 2024, retrieved from https://girai-report-2024-corrected-edition.tiiny.site/

[155] Byman, D. et al, "Government Use of Deepfakes: Questions to Ask", Center for Strategic and International Studies, 2024, retrieved from https://csis-website-prod.s3.amazonaws.com/s3fs-public/2024-03/240312_Byman_Government_Deepfakes.pdf?VersionId=OAFKvrOi9ojseIjOKhzqi672claqU.if

[156] Ahmed, Abdullah & Khan, Mudassar, "AI and Content Moderation: Legal and Ethical Approaches to Protecting Free Speech and Privacy." 2024, 10.13140/RG.2.2.23619.41767; DOI:10.13140/RG.2.2.23619.41767

[157] Funk, A. et al., "The Repressive Power of Artificial Intelligence", Freedom House, 2023 https://freedomhouse.org/report/freedom-net/2023/repressive-power-artificial-intelligence

[158] "Ethics guidelines for Trustworthy AI", 2019, retrieved from https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

и правни ограничения на ниво национална сигурност, представлява сериозна заплаха за основните права и свободи на гражданите не само в Европа, но и по света.

Етичният и надежден изкуствен интелект е фундаментален императив в съвременния свят, като неговата важност надхвърля сферата на бизнеса и иновациите. За да се гарантира, че тази технология служи на обществото, а не се превръща в инструмент за потискане, е необходимо стриктно придържане към етични принципи и стандарти за надеждност не само при разработването и внедряването на ИИ в търговския сектор, но и в неговите приложения за целите на националната сигурност, така че да се предотврати използването на ИИ от държави срещу собствените им граждани, като по този начин се запазят основните демократични ценности и човешки права. Само чрез съвместни усилия и ангажираност към отговорно развитие и използване на ИИ може да се гарантира, че тази технология ще допринесе за просперитета и благосъстоянието на човечеството, без да застрашава неговите свободи.

5.5. ИИ и процесите на вземане на политически и регулаторни решения

В контекста на стратегическото развитие на изкуствения интелект, рамката на "Триадата на ИИ", предложена през 2020 г. от Бен Бюканън,[159] признат експерт по киберсигурност и съветник по ИИ политики в администрацията на президент Байдън, представлява ценен аналитичен инструмент. Тази концепция, дефинираща основните компоненти на ИИ като данни, алгоритми и изчислителна мощност, първоначално разработена с фокус върху националната сигурност, демонстрира своята приложимост и при формулирането на всеобхватни национални политики, насочени към стимулирането на цялостна ИИ екосистема.[160] Същността и комплексността на ИИ могат лаконично да бъдат обобщени, както следва: Системите за машинно обучение оперират чрез използване на изчислителна мощност за изпълнение на алгоритми, които се базират на обучение от данни. Всеки един от трите елемента е ключов, въпреки че тяхната относителна важност се променя с развитието на технологията. Всяка част от триадата предлага свои собствени инструменти за политическо влияние, като напредъкът в алгоритмите зависи от таланти, данните изискват обмислени решения за пристрастия и поверителност, а изчислителната мощност може да бъде лост за експортен контрол, който много добре се използва от САЩ, в днешната динамична геополитическа ситуация. За да използват разумно тези инструменти, политиците трябва да разберат самата технология, като концепцията за триадата на ИИ е един от начините за това и за разработване на ефективни политики за развитието на ИИ и икономиката.[161]

Политическите решения трябва да подкрепят развитието на изследователи в тази областта на алгоритмите, тъй като те са ключови за ИИ и стратегическото планиране. Данните са начинът, по който системите с машинно обучение учат, и по-големите представителни набори водят до по-добри резултати. Правните ограничения върху данните могат да попречат на ефективността на системите, затова политическите решения трябва да балансират достъпа до данни с поверителността и етиката, за да се насърчи развитието на обективни ИИ системи.[162] Изчислителната мощност е все по-важна за използването на сложни алгоритми и на големи данни, като напредъкът в ИИ е директно свързан с нуждата от големи изчислителни мощности, както беше посочно в предни точки на анализа. Политическите решения трябва да водят до осигуряване на

[159] Ben Buchanan, "The AI Triad and What It Means for National Security Strategy", Center for Security and Emerging Technology, 2020. https://doi.org/10.51593/20200021
[160] Ibid
[161] Ibid
[162] Ben Buchanan, "The AI Triad and What It Means for National Security Strategy", Center for Security and Emerging Technology, 2020. https://doi.org/10.51593/20200021

достъп до достатъчна изчислителна мощност за изследвания и разработки в областта на ИИ, включително чрез инвестиции и насърчаване на енергийна ефективност.

Триадата на ИИ ясно показва, че ефективното развитие на изкуствения интелект изисква холистичен подход при вземането на политически решения. Не може да се пренебрегне нито един от трите компонента – алгоритми, данни и изчислителна мощност. Успешните политики трябва да стимулират иновациите в алгоритмите чрез подкрепа на изследователски таланти, да осигуряват достъп до качествени и представителни данни при спазване на етичните и правни норми, и да гарантират наличието на необходимата изчислителна инфраструктура. Лицата, отговорни за вземането на такива решения, трябва да осъзнаят взаимозависимостта на тези елементи и да разработят стратегии, които да адресират всеки един от тях, за да се реализира пълният потенциал на ИИ за икономически растеж и национална сигурност.

В следващите точки от настоящия анализ разглеждаме именно ефектът, който ИИ оказва при разработването на политики относно експортен контрол, привличането и задържане на таланти, и развитието на енергийна инфраструктура, с цел икономическо развитие и сигурност.

5.5.1 Експортен контрол

Глобалната геополитическа среда се характеризира с нарастваща надпревара за влияние, в която технологичното превъзходство, особено в областта на изкуствения интелект, играе ключова роля. Съединените щати, Китай и Европейският съюз възприемат различни стратегии и прилагат политики, включително експортен контрол върху критични суровини и технологии, за да укрепят своите позиции в този нов световен ред. Тези мерки имат за цел не само да защитят националната сигурност и икономическите интереси, но и да оформят бъдещата траектория на технологичното развитие и глобалното влияние.

Съединените щати активно използват експортния контрол като инструмент за запазване на технологичното си предимство, особено в областта на ИИ. Рамката за разпространение на изкуствен интелект (Framework for Artificial Intelligence Diffusion) публикувана през януари 2025г. подчертава необходимостта от балансиране между насърчаването на иновациите и защитата на националната сигурност. Стратегията на САЩ се фокусира върху ограничаване на достъпа на Китай до високопроизводителни чипове и друго оборудване, необходимо за развитието на напреднали ИИ системи.[163] Това се вижда в различни мерки, предприети от Търговския департамент на САЩ, които ограничават износа на определени полупроводници и ИИ софтуер за Китай и държави като Русия и Иран. Включително се ограничава правото не неограничена продажба на най-мощните чипове за ИИ само до 18 държави, като голяма част от страните от ЕС (вкл. и България) попадат в групата, за която има наложена квота от максимален брой изчислителна мощност, която може да бъде закупена в еквивалент от чипове.[164] Тези мерки се очаква да бъдат засилени от президента Тръмп по време на втория му мандат, който започна на 20 януари 2025 г.

Мерките, които САЩ предприема в ограничаването на разпространението на стратегически технологии неизменно, ще окажат влияние върху развитието на ИИ в Китай. Ограниченията на САЩ върху износа на високопроизводителни графични процесори (GPU), специализирани ИИ чипове и оборудване за производство на полупроводници директно възпрепятстват способността на китайските компании и

[163] Department of Commerce, BIS, "Framework for Artificial Intelligence Diffusion". 2025, retrieved from https://www.federalregister.gov/documents/2025/01/15/2025-00636/framework-for-artificial-intelligence-diffusion
[164] Ibid

изследователски институции да разработват и внедряват най-съвременни ИИ модели. Това ясно се вижда и от новите ИИ модели, които Китай пуска на пазара. DeepSeek е типичен пример за значението на експортния контрол. Възможността Китай да разработи този модел беше благодарение на неуспешно разработените мерки за експортен контрол, обявени през октомври 2022 г. от администрацията на Байдън,[165] което позволи Китай да се сдобие с чиповете H800 на NVIDIA, но същевременно показва и забавения ефект на мерките за експортен контрол. Ефектът от мерките, обявени през 2023 г., и обявените през януари тази година, тепърва предстои да се наблюдава. На 13 февруари 2025 г. Li Guojie, учен в Китайската академия по инженерство, заявява, че успехът на DeepSeek е напредък в китайските ИИ възможности но също така пояснява, че „поради блокадата на САЩ правителство, Китай в момента не е в състояние да получи най-модерната технология за чипове."[166]

Това е ясен пример за това колко важен е достъпът до стратегически технологии за технологичното развитие на една страна. Такива мерки принуждават Китай да търси алтернативни източници на хардуер, да инвестира значително в развитието на собствени полупроводникови технологии и да оптимизира съществуващите ресурси, което може да забави темповете на напредък в определени области на ИИ. Въпреки това, тези ограничения могат, също така, да стимулират вътрешните иновации и да намалят дългосрочната зависимост на Китай от чуждестранни технологии.

Влиянието на експортния контрол на САЩ върху развитието на ИИ в ЕС е по-непряко, но също съществено. Европейските компании и изследователи, които разчитат на американски високопроизводителни чипове и софтуер за своите ИИ проекти, могат да бъдат засегнати от ограниченията, наложени на Китай, особено ако тези ограничения водят до недостиг или повишаване на цените на тези компоненти на световния пазар. Освен това, съществува риск европейските компании, които изнасят за Китай ИИ технологии или продукти, съдържащи американски компоненти, да попаднат под обхвата на американските санкции. Това може да принуди ЕС да търси по-голяма технологична независимост в областта на полупроводниците и ИИ, което е и цел на инициативи като европейския Акт за чипове.

Китай също разглежда експортния контрол като важен инструмент в своята геополитическа стратегия. Наскоро въведените от Китай ограничения върху износа на определени критични минерали, като галий и германий,[167] показват готовността на страната да използва своето доминиращо положение в добива и преработката на тези суровини като средство за натиск в международните отношения и в отговор на американските експортни ограничения. Анализът на CSIS подчертава, че тези мерки са най-строгите ограничения, налагани някога от Китай върху износа на критични минерали, и че те са пряко свързани с технологичната надпревара със САЩ.[168]

[165] Department of Commerce, BIS, "Commerce Implements New Export Controls on Advanced Computing and Semiconductor Manufacturing Items to the People's Republic of China (PRC)", 2022, https://www.bis.doc.gov/index.php/documents/about-bis/newsroom/press-releases/3158-2022-10-07-bis-press-release-advanced-computing-and-semiconductor-manufacturing-controls-final/file

[166] Allen, Gregory C., "DeepSeek, Huawei, Export Controls, and the Future of of the U.S.- China AI Race", Center for Strategic and International Studies, 2025, retrieved from https://csis-website-prod.s3.amazonaws.com/s3fs-public/2025-03/250307_Allen_AI_Race.pdf?VersionId=8Qpp5OWwGj9OKB25qzQX4tQ_DzIZ9NS_

[167] Baskaran, G. and Schwartz, M., "China Imposes Its Most Stringent Critical Minerals Export Restrictions Yet Amidst Escalating U.S.-China Tech War", Center for Strategic and International Studies, 2024, https://www.csis.org/analysis/china-imposes-its-most-stringent-critical-minerals-export-restrictions-yet-amidst

[168] Ibid

Официалните документи на Министерството на търговията на Китай[169] [170] представляват нормативната база за контрол на износа на артикули с двойна употреба. Тези документи определят списъци на стоки и технологии, чийто износ е регулиран, включително такива, които могат да бъдат използвани в ИИ и други стратегически сектори. Въпреки че конкретните ИИ технологии може да не са изрично споменати в тези общи документи, те осигуряват правната рамка за бъдещи ограничения, насочени към специфични технологии, ако е необходимо.

Експортният контрол на Китай върху критични минерали, като галий и германий, които са ключови за производството на полупроводници и други технологии, свързани с ИИ, може да има значителни последици за САЩ и ЕС. Ограниченията на Китай върху износа на тези критични минерали могат да доведат до недостиг и повишаване на цените на суровините, необходими за производството на чипове и други компоненти, използвани в ИИ системите в САЩ.[171]Анализът на CSIS, подчертава, че тези минерали са от съществено значение за производството на полупроводници, телекомуникационно оборудване и други високотехнологични продукти. [172] Това може да забави производството и да увеличи разходите за компаниите в САЩ, занимаващи се с разработка и внедряване на ИИ. Въпреки че САЩ търсят алтернативни източници и инвестират във вътрешно производство, зависимостта от Китай в тази област остава предизвикателство.

Подобно на САЩ, ЕС също е зависим от Китай за доставките на определени критични минерали, включително тези, които са обект на експортен контрол. Ограниченията от страна на Китай могат да доведат до същите проблеми с недостиг и повишаване на цените, засягайки европейските производители на полупроводници и други технологии, използвани в ИИ. Това подчертава уязвимостта на ЕС по отношение на веригите за доставки на ключови суровини и необходимостта от диверсификация на източниците и укрепване на вътрешния капацитет.

Европейският съюз възприема по-нюансиран подход към експортния контрол, балансирайки между икономическите интереси и необходимостта от защита на сигурността и ценностите си. ЕС разполага с регулация за контрол на износа на артикули с двойна употреба, [173] [174] която има за цел да предотврати разпространението на технологии, които могат да бъдат използвани за военни цели или за нарушаване на човешките права. В контекста на технологичната надпревара, ЕС се стреми да укрепи своята технологична независимост и да развие собствен капацитет в ключови области като ИИ и полупроводници. Случаят с частичното отнемане на лиценз за износ на ASML[175] показва, че ЕС е готов да предприеме мерки за ограничаване на износа на високотехнологично оборудване към страни, които могат да представляват риск за сигурността. Въпреки че за ограничението на износа на конкретни модели литографски машини на ASML не дава подробности за конкретните причини за налагането му,[176] то е показателно за нарастващото напрежение в областта на технологичния износ.

---

[169] "商务部 工业和信息化部 海关总署 国家密码局公告 2024 年第 51 号 关于发布《中华人民共和国两用物项出口管制清单》的公告", 2024, https://aqygzj.mofcom.gov.cn/qdml/art/2024/art_8bf88382af1143168e2b11beadf000ba.html 14.03.2025

[170] "中华人民共和国 两用物项出口管制清单", 2024, retrieved from https://aqygzj.mofcom.gov.cn/qdml/art/2024/art_8bf88382af1143168e2b11beadf000ba.html

[171] Ibid., CSIS report

[172] Ibid., CSIS report

[173] OJ L 206, 11.6.2021, ELI: http://data.europa.eu/eli/reg/2021/821/oj

[174] "Exporting dual-use items", https://policy.trade.ec.europa.eu/help-exporters-and-importers/exporting-dual-use-items_en 21.03.2025

[175] "Statement regarding partial revocation export license", 2024, https://www.asml.com/en/news/press-releases/2023/statement-regarding-partial-revocation-export-license 20.03.2025

[176] Ibid., ASML

Важна стъпка, която ЕС предприема, е и координирането на политиките за експортен контрол между държавите-членки, за да осигури по-ефективен и съгласуван подход. Европейският съюз, чрез компанията ASML, която е водещ производител на литографски машини, играе ключова роля в глобалната верига за доставки на полупроводници. Експортният контрол на ЕС върху тези машини има пряко и значително влияние върху развитието на ИИ в Китай. Изявлението на ASML относно частичното отнемане на лиценз за износ показва, че ЕС е готов да ограничи достъпа на Китай до най-съвременните технологии за производство на полупроводници. Литографските машини на ASML са от съществено значение за производството на усъвършенствани чипове, които са критични за развитието на напреднали ИИ системи. Ограничаването на достъпа до тези технологии значително затруднява способността на китайските производители на чипове да произвеждат най-модерните процесори, необходими за обучението и изпълнението на сложни ИИ модели. Това може да забави напредъка на Китай в ключови области на ИИ, като например автономни системи, усъвършенствано разпознаване на образи и обработка на естествен език, които изискват висока изчислителна мощност. Въпреки че Китай инвестира значително в развитието на собствени литографски технологии, се очаква да минат години, преди да успее да настигне технологичното ниво на ASML.

Експортният контрол, прилаган от САЩ, Китай и ЕС, представлява сложен и динамичен фактор, който оказва значително влияние върху глобалната надпревара в областта на изкуствения интелект. Тези мерки не само оформят технологичното развитие във всяка от тези сили, но и водят до пренареждане на веригите за доставки и стимулират търсенето на по-голяма технологична независимост. В крайна сметка, взаимодействието между тези различни режими на експортен контрол ще продължи да играе ключова роля в определянето на бъдещия баланс на силите в геополитическата и технологична сфера.

Експортният контрол се превръща във все по-важен инструмент в геополитическата и технологична надпревара. Той може да има значително влияние върху скоростта и посоката на технологичното развитие, както и върху баланса на силите между различните държави. Експортният контрол може ефективно да забави развитието на определени технологии в страните, към които са насочени ограниченията. Ограниченията на САЩ върху износа на високопроизводителни чипове за Китай могат да затруднят напредъка на китайските компании в разработването на напреднали ИИ системи. Такива ограничителни мерки пред износа на технологии свързани с ИИ, могат да доведат до пренасочване на глобалните вериги за доставки, тъй като компаниите се стремят да заобиколят ограниченията или да намерят алтернативни източници на критични суровини и технологии. Китайските ограничения върху износа на галий и германий могат да насърчат други страни да инвестират в собствени производствени мощности или да търсят алтернативни доставчици. Експортният контрол може, също така, да стимулира вътрешните иновации в страните, които са обект на ограниченията, защото странинте ще са принудени да развият собствени технологии и да намалят зависимостта си от чуждестранни доставчици. Нарастващото използване на експортния контрол може да доведе до фрагментиране на глобалната технологична екосистема и до създаването на отделни технологични блокове, водени от различни геополитически сили. Това може да има негативни последици за глобалното сътрудничество в областта на науката и технологиите и да забави общия напредък. Въпреки че целите и подходите на САЩ, Китай и ЕС се различават, общата тенденция е към нарастващо използване на експортния контрол като средство за защита на националните интереси и за оказване на влияние върху глобалната технологична и геополитическа среда. В дългосрочен план, ефективността и последиците от тези политики ще зависят от много фактори, включително способността на всяка страна да иновира, да адаптира веригите си за доставки и да поддържа международни съюзи и партньорства.

5.5.2 ИИ таланти

В съвременния глобален пейзаж изкуственият интелект се очертава като ключов двигател на икономическия растеж, иновациите и националната конкурентоспособност. В тази ера на бърз технологичен прогрес, развитието, изграждането, привличането и задържането на висококвалифицирани ИИ таланти и работна ръка се превръща във фундаментален приоритет както за развитите икономики, така и за развиващите се. От страна на световните лидери се наблюдава все по-ясно осъзнаване на стратегическото значение на човешкия капитал в областта на ИИ. Тези държави и региони активно разработват и прилагат разнообразни политики и мерки, насочени към преодоляване на съществуващите предизвикателства и създаване на благоприятна екосистема за процъфтяването на ИИ талантите, което е от съществено значение за бъдещия им икономически успех и технологично лидерство.

Според доклад за ИИ талантите, изготвен в началото на 2025 г., Съединените щати признават критичната роля на ИИ талантите за икономическия растеж и националната сигурност и целят значително увеличаване на броя на квалифицираните специалисти чрез инвестиции в образованието и обучението. Докладът подчертава необходимостта от преодоляване на съществуващите бариери пред навлизането в ИИ професиите и създаване на по-приобщаваща екосистема за развитие на таланти.[177] Според Андресен Хороуиц [178] и Центъра за сигурност и нововъзникващи технологии (CSET) към университета Джорджтаун,[179] за да се запази лидерството на САЩ в областта на ИИ, е необходима целенасочена политика, включваща значителни инвестиции в ИИ таланти и инфраструктура. За тази цел се препоръчва увеличаване на финансирането за фундаментални и приложни ИИ изследвания в университетите и държавните лаборатории, както и създаване на стипендии и програми за обучение, насочени към ИИ Също така, се отчита и ключовото значение на инвестициите в изчислителна инфраструктура, необходима за обучението и внедряването на ИИ модели, и да се улесни достъпът до данни за изследвания и разработки.[180] Това включва мерки за улесняване на достъпа на чуждестранни ИИ специалисти до работни визи и зелени карти, както и създаване на стимулираща среда за работа и изследвания, която да ги мотивира да останат в страната. Според CSET Джорджтаун, също така е важно да се инвестира в развитието на местни таланти чрез образователни програми и научни инициативи, за да се осигури дългосрочна конкурентоспособност на САЩ в областта на ИИ.[181]

Според Европейската комисия, основен проблем в развитието на ИИ екосистема е фрагментираният подход към ИИ политиките в отделните държави-членки, което затруднява създаването на единен пазар за ИИ технологии и таланти. В тази връзка ЕС се стреми да създаде хармонизирана регулаторна рамка, която да насърчава иновациите, като същевременно гарантира безопасността и етичното използване на ИИ.[182] Според програмата за цифрово десетилетие на ЕС, една от конкретните мерки е да се увеличи броят на ИКТ специалистите до 20 милиона до 2030 г., като се обърне специално внимание на развитието на ИИ умения. Според пакета за състоянието на

[177] "AI Talent Report", 2025, https://bidenwhitehouse.archives.gov/cea/written-materials/2025/01/14/ai-talent-report/#:~:text=The%20United%20States%20could%20increase,and%203 22.03.2025
[178] Perault, M. "A Policy Blueprint for US Investment in AI Talent and Infrastructure", 2025, https://a16z.com/a-policy-blueprint-for-us-investment-in-ai-talent-and-infrastructure/ 21.03.2025
[179] Zwetsloot, R., "Keeping Top AI Talent in the United States", 2019 https://cset.georgetown.edu/publication/keeping-top-ai-talent-in-the-united-states/ 21.03.2025
[180] Ibid., Perault, M.
[181] Ibid., Zwetsloot, R.
[182] "AI Excellence: Ensuring that AI works for people", https://digital-strategy.ec.europa.eu/en/policies/ai-people 21.03.2025

цифровото десетилетие за 2024 г., [183] ЕС отчита недостатъчен напредък в инвестициите в ИИ и призовава за по-активни мерки за стимулиране на публичните и частните инвестиции в тази област. Според Стратегията за Европа за ИИ умения, [184] за преодоляване на недостига на ИИ таланти е необходимо да се създадат гъвкави и адаптивни програми за обучение, които да отговарят на бързо променящите се нужди на индустрията, както и да се насърчи ученето през целия живот в областта на ИИ.

Основен проблем за Китай е превръщането на голям брой ИИ изследователи във висококвалифицирани експерти с практически опит и умения за комерсиализация.[185] Според доклад на университета Джорджтаун,[186] Китай предприема мащабни мерки за обучение на ИИ специалисти, включително чрез специални програми в университетите и подкрепа за изследователски институти, но все още се нуждае от подобряване на качеството на обучението и създаване на по-привлекателна среда за задържане на таланти. Според McKinsey,[187] недостигът на ИИ таланти в Китай се отразява негативно на способността на компаниите да внедряват иновативни ИИ решения, което може да забави икономическия растеж. Според друг анализ,[188] една от мерките, предприети от Китай, е значителното увеличаване на приема на студенти в университетите по специалности, свързани с ИИ, като се акцентира върху практическите умения и сътрудничеството с индустрията.

Министерството на науката и ИКТ на Република Корея разработи национална стратегия за ИИ, която включва конкретни мерки за развитие на ИИ таланти чрез създаване на специализирани учебни програми,[189] подкрепа за научни изследвания в ключови области на ИИ и насърчаване на сътрудничеството между университети, изследователски институти и компании. В допълнение се предвиждат и мерки за етично и безопасно използване на ИИ, както и за внедряването му в различни сектори на икономиката с цел повишаване на конкурентоспособността. Управата на столицата Сеул осъзнава необходимостта от създаване на силна ИИ екосистема и за целта предприема специални мерки за развиването на таланти в столицата[190] чрез политики за привличане на инвестиции и насърчаване на иновациите в тази област, включително чрез регулаторни облекчения и финансова подкрепа за стартиращи компании.

Политиките на водещите световни икономики показват единодушно признание за критичната роля на ИИ талантите и квалифицираната работна ръка за развитието на сектора на изкуствения интелект и за икономическия растеж като цяло. Всички разгледани региони се сблъскват с различни предизвикателства, свързани с недостига на ИИ специалисти, необходимостта от подобряване на качеството на обучението и задържането на таланти, както и с изграждането на необходимата инфраструктура. Предприетите политики включват широк спектър от мерки, насочени към инвестиции в образованието и обучението, създаване на специализирани учебни програми и

---

[183] "2024 State of the Digital Decade package", https://digital-strategy.ec.europa.eu/en/policies/2024-state-digital-decade-package 21.03.2025

[184] "AI Skills Strategy for Europe", 2023, retrieved from https://aiskills.eu/wp-content/uploads/2024/01/AI-Skills-Strategy-for-Europe.pdf

[185] Yang, Z.,"Four things you need to know about China's AI talent pool ", 2024, https://www.technologyreview.com/2024/03/27/1090182/ai-talent-global-china-us/ 21.03.2025

[186] Paterson, D. et al., "Assessing China's AI Workforce Regional, Military, and Surveillance Geographic Clusters", 2023, https://cset.georgetown.edu/publication/assessing-chinas-ai-workforce/ 21.03.2025

[187] "China's AI talent gap", 2023, https://www.mckinsey.com/featured-insights/sustainable-inclusive-growth/charts/chinas-ai-talent-gap 21.03.2025

[188] "China expands university enrolment to boost AI talent", 2025, https://dig.watch/updates/china-expands-university-enrolment-to-boost-ai-talent 21.03.2025

[189] "National AI Strategy Policy Directions", 2024, retrieved from https://www.msit.go.kr/bbs/view.do?sCode=eng&mId=4&mPid=2&bbsSeqNo=42&nttSeqNo=1040

[190] "Seoul vows to nurture 10,000 AI talents annually", 2025, https://www.koreaherald.com/article/10417202#:~:text=The%20Seoul%20city%20government%20on,living%20and%20 pleasure%20are%20 combined 22.03.2025

стипендии, насърчаване на научните изследвания и развитие, подкрепа за стартиращи компании и иновации, както и изграждане на необходимата изчислителна инфраструктура и улесняване на достъпа до данни.

За държави като България, които се стремят да развиват своя ИИ сектор и да бъдат конкурентоспособни, е изключително важно да се вземат предвид тези глобални тенденции и да се разработят адекватни политики. България може да се фокусира върху инвестиции в образованието, насочени към развитие на ИИ умения, насърчаване на сътрудничеството между академичните среди, изследователските институти и бизнеса, осигуряване на подкрепа за стартиращи компании, привличане и задържане на таланти, инвестиции в изграждане на изчислителна инфраструктура за обучение и внедряване на ИИ модели и създаване на стимулираща иновациите регулаторна рамка, която да гарантира безопасното използване на ИИ. Чрез целенасочени и координирани политики, България може да създаде условия за развитие на силен ИИ сектор, който да допринесе за икономическия растеж и повишаване на конкурентоспособността на страната.

5.5.3 ИИ и нуждата от енергия

Политическите решения, свързани с развитието на енергийния сектор в световен мащаб, са от ключово значение за стимулирането на разработването, внедряването и използването на изкуствения интелект и съпътстващото икономическо развитие. Според **Световния икономически форум (СИФ)**, [191] експоненциално нарастващото енергийно потребление, предизвикано от генеративния изкуствен интелект в глобален мащаб, налага неотложна необходимост от стратегически политически мерки, насочени към осигуряване на устойчиво енергоснабдяване. СИФ подчертава, че без адекватно планиране и инвестиции в енергийната инфраструктура в световен мащаб, може да се стигне до ограничения в развитието на ИИ секторите на отделните държави.

Анализ на **MIT Sloan** показва,[192] че високите енергийни разходи на центровете за данни за ИИ имат значителни финансови последици за компаниите по целия свят. Политическите решения, насърчаващи енергийната ефективност и достъпа до конкурентни цени на енергията на глобално ниво, могат да смекчат тези финансови тежести и да направят отделните региони по-привлекателни за инвестиции в тази област. Според McKinsey,[193] прогнозираният "Енергиен пир на ИИ" изисква проактивни политически мерки в световен мащаб за предотвратяване на бъдещи енергийни дефицити, които биха могли да забавят развитието на ИИ и свързаните с него икономически ползи за глобалната икономика.

Според организации като СИФ и MIT Sloan, политическите инициативи, подкрепящи научните изследвания и внедряването на енергийно ефективни ИИ алгоритми и хардуерни решения в световен мащаб, са от съществено значение. Насърчава се държавната подкрепа за развитието и използването на специализирани ИИ ускорители и иновативни компютърни архитектури, което може да допринесе за намаляване на общото енергийно потребление на ИИ центровете за данни в глобален мащаб. Освен това, политическите решения, стимулиращи прехода към възобновяеми енергийни източници за захранване на тези центрове в световен мащаб, са ключови за постигане

[191] "AI and energy: Will AI help reduce emissions or increase demand? Here's what to know", 2024, https://www.weforum.org/stories/2024/07/generative-ai-energy-emissions/ 22.03.2025
[192] "AI has high data center energy costs — but there are solutions", 2025, https://mitsloan.mit.edu/ideas-made-to-matter/ai-has-high-data-center-energy-costs-there-are-solutions - 22.03.2025
[193] "AI's power binge", 2024, https://www.mckinsey.com/featured-insights/sustainable-inclusive-growth/charts/ais-power-binge 22.03.2025

на устойчиво развитие и намаляване на въглеродния отпечатък на глобалния ИИ сектор, както посочват Световният икономически форум и MIT Sloan.[194] [195]

Според **Международната агенция по енергетика**,[196] политическите решения, насърчаващи интегрирането на ИИ в енергийния сектор в световен мащаб, могат да доведат до оптимизация на енергийните системи, намаляване на загубите и повишаване на ефективността в глобален мащаб. Това може да създаде допълнителни възможности за икономически растеж и да укрепи енергийната сигурност на световната икономика.

Голдман Сакс обръща внимание на важността на ролята на ядрената енергия като стабилен и нисковъглероден източник за захранване на ИИ центровете за данни в световен мащаб. Политическите решения, подкрепящи развитието на ядрената енергетика в глобален мащаб, могат да осигурят дългосрочна енергийна сигурност и да привлекат инвестиции в сектора на ИИ. Според организацията, ядрената енергия може да предостави необходимата базова мощност за задоволяване на нарастващите енергийни нужди на ИИ в световен мащаб.[197]

В началото на 2025 г. беше създаден **Европейския бизнес ядрен алианс (EBNA),** като част от него е и Българската стопанска камара (БСК). Този алианс е стратегическа инициатива, насърчаваща сътрудничеството между ЕК, държавите и бизнеса в областта на ядрената енергия, което е от съществено значение за осигуряване на устойчиво енергоснабдяване за енергоемки технологии като ИИ в глобален план.[198]

Политическите решения трябва да позиционират енергийния сектор като ключов елемент за развитието на ИИ и икономиката, защото без необходимите енергийни ресурси тази технолгия не може да бъде използвана, което би довело до липса на конкурентоспособност и икономически срив. Бъдещите политически решения трябва да залагат на мерки за насърчаване на енергийна ефективност, инвестиции във възобновяеми и ядрени енергийни източници, подкрепа за научни изследвания в областта на енергийно ефективен ИИ и международно сътрудничество в енергийния сектор.

В този контекст, за България е препоръчително да следва тези глобални тенденции и да инвестира и развива своя енергиен сектор. Участието на БСК в Европейския бизнес ядрен алианс (EBNA) е стъпка в правилната посока, демонстрираща ангажимент към осигуряване на устойчиво енергийно бъдеще на българската икономика. За да се възползва максимално от потенциала на ИИ и да стимулира икономическия си растеж, България трябва да продължи да развива своята енергийна инфраструктура, като инвестира както във възобновяеми, така и в ядрени енергийни източници, и да създаде благоприятна политическа среда за развитието на центрове за данни и ИИ технологии.

---

[194] "AI and energy: Will AI help reduce emissions or increase demand? Here's what to know", 2024, https://www.weforum.org/stories/2024/07/generative-ai-energy-emissions/ 22.03.2025

[195]"AI has high data center energy costs — but there are solutions", 2025, https://mitsloan.mit.edu/ideas-made-to-matter/ai-has-high-data-center-energy-costs-there-are-solutions - 22.03.2025

[196] Bennett, S. and Spencer, T., "How will artificial intelligence transform energy innovation?", 2024,https://www.iea.org/commentaries/how-will-artificial-intelligence-transform-energy-innovation 22.03.2025

[197] "Is nuclear energy the answer to AI data centers' power consumption?", 2025, https://www.goldmansachs.com/insights/articles/is-nuclear-energy-the-answer-to-ai-data-centers-power-consumption - 22.03.2025

[198] "European Nuclear Business Alliance", 2025, https://www.medef.com/uploads/media/default/0020/05/16382-european-business-nuclear-alliance-fully-signed.pdf 22.03.2025

6. Икономическа ефективност

6.1. Икономическа ефективност от прилагането на ИИ в различни сектори

Изкуственият интелект се превръща в двигател на икономически растеж и трансформация в глобален мащаб. Дори през настоящото десетилетие ефектът от ИИ върху световната икономика е значителен, като проучване на PwC показва, че до 2030 г. ИИ би могъл да допринесе с 15.7 трилиона щатски долара към глобалния БВП. Това представлява около 14% прираст на световната икономика спрямо базовия сценарий, или повече от текущия общ БВП на Китай и Индия взети заедно. Ефектът идва както от повишена производителност (около 6.6 трлн. долара), така и от стимулиране на потреблението с по-качествени и персонализирани продукти (9.1 трлн. долара).[199]

Поглеждайки напред, прогнозите стават още по-впечатляващи. Според анализ на McKinsey[200] има 18 области с потенциал да преоформят световната икономика и да генерират между 29 и 48 трилиона долара приходи и от 2 до 6 трилиона долара печалба до 2040 г. Това подсказва, че **през следващите две десетилетия ИИ ще се утвърди като ключов фактор за продуктивността и растежа**.

6.1.1 Глобален потенциал и регионални различия

ИИ вече демонстрира значими икономически ползи по целия свят, но мащабът на тези ползи и темпът на внедряване варират по региони. Китай и Северна Америка (предимно САЩ) са лидери по въздействие на ИИ върху БВП към 2030 г., докато Европа изостава. *Таблица 1* показва прогнозния прираст на БВП от ИИ за различни региони към 2030 г., според глобално изследване на PwC, цитирано от Глобалния икономически форум:[201]

Регион	Очакван прираст на БВП от ИИ до 2030 г.
Китай	26,0%
Северна Америка (САЩ)	14,5%
Южна Европа	11,5%
Развитите азиатски икономики	10,4%
Северна Европа	9,9%
Развиващи се страни от Латинска Америка и Африка	5,4 - 5,6%

Таблица 1: Прогнозен ръст на БВП по региони вследствие на ИИ към 2030 г.

Европейските икономики изостават спрямо САЩ и Китай по очакван икономически ефект от ИИ, което отразява по-бавното внедряване и по-ниските инвестиции в ЕС. Делът на Европа в глобалната ИИ верига на стойността по някои оценки е под 5%[202]. Най-големите европейски компании са инвестирали около 25 млрд. евро в ИИ през 2024 г., докато водещите технологични фирми в САЩ – над 150 млрд. евро. В резултат, разликата в продуктивността между Европа и САЩ се увеличава. Китай, от своя страна, следва агресивна национална стратегия в областта на ИИ и вече се превръща в световен

[199] https://www.pwc.com/gx/en/issues/artificial-intelligence/publications/artificial-intelligence-study.html
[200] The next big arenas of competition, McKinsey Global Institute, 2024, https://www.mckinsey.com/~/media/mckinsey/mckinsey%20global%20institute/our%20research/the%20next%20big%20arenas%20of%20competition/the-next-big-arenas-of-competition.pdf
[201] https://www.weforum.org/stories/2017/06/the-global-economy-will-be-14-bigger-in-2030-because-of-ai/
[202] 2025 World Economic Forum, Europe in the Intelligent Age: From Ideas to Action, https://reports.weforum.org/docs/WEF_Europe_in_the_Intelligent_Age_2025.pdf

лидер по редица показатели, като например броят на патентите в областта на ИИ в Китай нараства значително по-бързо от този в други страни.

Важно е да се отбележи, че разликите между регионите не се дължат на липса на потенциал в Европа, а на фактори като по-малки рискови инвестиции, разпокъсан пазар и строг регулаторен подход. Актът за ИИ, макар да гарантира доверие и безопасност, може временно да забави масовото навлизане на ИИ решения в сравнение с по-гъвкавите режими в САЩ и Китай. В дългосрочен план, обаче, Европа не може да си позволи да изостава: прогнозите сочат, че след 2030 г. демографските тенденции (застаряващо население) ще свиват работната сила на ЕС, което налага именно повишена продуктивност чрез технологии като ИИ, за да се запази икономическият растеж.[203]

6.1.2 Производство и индустрия

Производственият сектор е сред най-преките бенефициенти на ИИ, благодарение на автоматизацията, роботизацията и оптимизацията на процесите. В областта на машиностроенето и електрониката ИИ се внедрява за прогнозна поддръжка на оборудването, контрол на качеството чрез машинно зрение, оптимизиране на веригите на доставка и др. Според изследване на Световния икономически форум[204], ИИ има потенциал да увеличи глобалния БВП с около 2% годишно само чрез подобрения в производителността на напредналото производство. Това е огромен тласък и вече се наблюдават конкретни резултати: компании, възприели индустриален ИИ, отчитат **повишени добиви и качество**. Чрез самообучаващи се системи работниците придобиват "свръхсили", постигайки безпрецедентни нива на продуктивност, без да се прави компромис с безопасността или устойчивостта.

Роботизацията в производството намалява разходите и ускорява производствените цикли. Съвременните фабрики използват колаборативни роботи (cobots), подпомагани от ИИ алгоритми за зрение и придвижване, които могат гъвкаво да превключват между задачи. Това дава възможност за масова персонализация на продукти без загуба на ефективност.

Икономическите ползи в сектора се изразяват в **намалени производствени разходи, по-малко престои на машините и брак, както и по-бързо навлизане на иновации**. 45% от общите икономически печалби от ИИ до 2030 г. се очаква да дойдат от подобрени продукти и по-голямо търсене, благодарение на тези подобрения[205], много от които се създават именно в производствения сектор.

Потенциалните рискове и пречки в производството също трябва да се отбележат. Една основна бариера е **недостигът на квалифицирани кадри**, способни да работят с новите технологии. Близо **46% от производствените компании** посочват липсата на умения у персонала като пречка пред внедряването на умни производствени решения. Необходими са инвестиции в преквалификация и обучения. Съществува и опасение от загуба на работни места заради автоматизацията. Макар ИИ да създава нови работни роли (например, специалисти по машинно обучение, оператори на роботи), той вероятно ще елиминира някои рутинни длъжности. **Автоматизацията може да замени до 30% от трудовите роли в производството до 2030 г.**, според оценки на Световния икономически форум.[206]

Това изисква проактивни политики за смекчаване на социалния ефект, като същевременно се използва възможността хората да се насочат към задачи с по-висока

[203] The future of European competitiveness 2024, https://commission.europa.eu/document/download/97e481fd-2dc3-412d-be4c-f152a8232961_en
[204] https://www.weforum.org/stories/2024/01/industrial-ai-superpowers-advanced-manufacturing/
[205] https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf
[206] https://www.weforum.org/stories/2024/01/industrial-ai-superpowers-advanced-manufacturing/

добавена стойност. Друг риск е **киберсигурността** – свързаните фабрични машини и роботи увеличават повърхността за атака, което налага сериозни мерки срещу саботаж или промишлен шпионаж. Въпреки тези предизвикателства, тенденцията е ясна: **производствата, които бързо интегрират ИИ, ще бъдат по-конкурентоспособни, с по-ниски разходи и по-голяма гъвкавост**.

6.1.3 Селско стопанство и храни

Селското стопанство преживява **тиха революция** чрез навлизането на ИИ, превръщайки се в прецизен, управляван от данни бизнес. Чрез сензори, дронове, сателитни изображения и машинно обучение фермерите вече могат да вземат решения, базирани на данни, вместо на интуиция. Прецизното земеделие, подпомагано от ИИ, оптимизира употребата на ресурси като вода, торове, препарати и обещава по-високи добиви с по-малко вложения. Изследвания показват, че тези технологии позволяват на стопанствата да намалят потреблението на вода и торове с 20 до 40%, без това да намалява добивите. [207] Така се постига двойна полза: по-висока продуктивност и намален екологичен отпечатък на селското стопанство.

Анализ на McKinsey оценява, че цялостното приложение на ИИ в агро-хранителната верига може да създаде около 250 млрд. щ.д. годишна стойност за индустрията. Тази стойност идва от две основни направления: около 100 млрд. долара чрез подобрения "на полето" (намалени разходи за труд и ресурси, повишени добиви) и около 150 млрд. долара в бизнеса надолу по веригата (преработка, дистрибуция, по-ниски загуби и оптимизации). [208] С други думи, ИИ може да направи земеделието по-печелившо за фермерите и същевременно да намали цените за крайния потребител чрез ефективност по цялата верига от фермата до магазина.

Още конкретни примери: автономни трактори и роботи за плевене, снабдени с компютърно зрение, могат прецизно да обработват почвата и растенията, намалявайки нуждата от хербициди. Дронове с ИИ-аналитика следят здравето на посевите и откриват ранни признаци на заболявания или вредители, позволявайки навременна намеса. В животновъдството камери и алгоритми наблюдават поведението на животните, за да сигнализират за здравословни проблеми или оптимален момент за хранене/доене. ИИ не е просто модерен инструмент, а реален „партньор в растежа" на агробизнеса, водещ до осезаеми подобрения.

Един от основните проблеми е, че малките и традиционни фермери може да имат затруднения да внедрят скъпи ИИ решения и техника. Съществува риск от **дигитално разделение** в селското стопанство и големите агроконгломерати да внедряват високотехнологични решения и да увеличават производителността си, докато малките стопанства да изостават. Това може да доведе до икономически затруднения за дребните фермери и консолидация на пазара. Друг риск е **зависимостта от данни и свързаност**. Ако няма надеждна интернет връзка в селските райони (което е често срещано предизвикателство), много ИИ приложения не могат да функционират оптимално. Също така, изникват притеснения за поверителността и собствеността на данните, когато фермерските данни се събират и анализират от трети страни (например, големи платформи). Има въпроси кой притежава тези данни и как се използват.

От гледна точка на околната среда, макар ИИ да намалява употребата на химикали, една неправилна употреба (например, алгоритъм, който оптимизира добива, без да отчита дългосрочното плодородие на почвата) би могла да доведе до изтощаване на ресурси. Затова експертите призовават за отговорно използване на ИИ в земеделието, съчетано със знания по агрономия. Пример за добра практика е използването на ИИ за

[207] https://www.weforum.org/stories/2025/03/can-ai-foster-sustainability/
[208] https://www.mckinsey.com/industries/agriculture/our-insights/from-bytes-to-bushels-how-gen-ai-can-shape-the-future-of-agriculture

възстановително земеделие – модели, които препоръчват редуване на култури, покривни култури и други практики за подобряване на почвата.

6.1.4 Медицина и здравеопазване

В сферата на медицината ИИ обещава както по-добри здравни резултати, така и значителни икономии и ефективност. Машинното обучение вече превъзхожда хората в някои тясно специализирани задачи. Алгоритмите за образна диагностика могат с изключителна точност да разпознават ракови образувания или да анализират рентгенови снимки. Новите ИИ системи могат да откриват патологични изменения два пъти по-точно от специалистите в определени случаи. Така се спестява време и се повишава качеството на диагнозата, което е критично за навременната терапия.

Икономическите ползи от ИИ в здравеопазването се проявяват на няколко нива:

- **Намаляване на разходите**: чрез автоматизация на административни процеси (като водене на документация, насрочване на часове, обработка на застрахователни искове) и използване на чатботове за първична консултация, болниците и клиниките могат да спестят средства и човешки ресурс. Тези спестявания идват от по-ефективно управление на операциите, като оптимизиране на графици на лекари и операционни зали, намаляване на ненужни изследвания и др.
- **Повишена продуктивност на медицинския персонал**: ИИ асистентите помагат на лекарите, като им предоставят обработена информация, обобщават научни публикации, предлагат възможни диагнози на база симптоми или дори попълват епикризи автоматично. Това означава повече прегледани пациенти и по-малко забавяне и чакане – ефект, труден за количествено измерване, но критично важен, особено при недостиг на кадри.
- **По-добри клинични резултати**: ИИ може да предвижда пациентски усложнения (напр., риск от сепсис при болнични пациенти) или да предлага персонализирани планове за лечение. Това подобрява изхода от лечението и намалява престоя в болница. Например, предиктивните модели за повторно приемане на пациентите помагат да се вземат мерки, още преди пациентът да се влоши, спестявайки потенциални разходи от избегнати спешни интервенции.
- **Откриването на нови лекарства** се ускорява чрез ИИ. Алгоритмите пресяват огромни бази данни от съединения и биологични мишени, за да посочат обещаващи кандидати за нови медикаменти. ИИ може да повиши продуктивността в научноизследователската и развойна дейност (R&D) на фармацевтичните компании, което означава по-бързо разработване на нови терапии при по-ниска цена. Това има огромен икономически ефект: скъсяването дори с година на цикъла за разработка на едно лекарство спестява стотици милиони долари и позволява по-ранна реализация на пазара.

Въпреки потенциала, здравеопазването изостава в приемането на ИИ. Причините включват строги регулации, необходими за безопасността на пациентите, както и консервативна култура – лекарите разчитат на изпитани практики и с основание са предпазливи към решения тип "черни кутии". Един ключов риск е точността и надеждността на алгоритмите: ако един ИИ сгреши диагноза или пропусне критична находка, това може да струва човешки живот. Имало е случаи на пристрастия – напр., алгоритъм за определяне приоритета на пациентите за трансплантация, който се оказва неблагоприятен за някои етнически групи, поради исторически данни. Подобни примери

подчертават нуждата от строго валидиране на медицинските ИИ и мониторинг за справедливост.

Друг проблем е защитата на личните медицински данни. ИИ изисква огромни масиви от информация – в здравеопазването това са чувствителни данни (диагнози, образни изследвания, генетична информация). Европейските регулации (GDPR) налагат високи стандарти за поверителност и всеки пробив или неправомерно използване на данни може да доведе до глоби и, по-важното, до загуба на доверие от пациентите. Затова внедряването на ИИ трябва да е придружено от солидна киберсигурност и етични рамки.

Също така, отговорността при грешка на ИИ е правно неуредена област – ако, например, ИИ система пропусне да алармира за опасно състояние и пациентът пострада, кой носи отговорност – софтуерният разработчик, болницата или лекарят? Такива въпроси тепърва се уреждат нормативно.

В обобщение, ИИ в медицината е област с огромна възвращаемост: оценките са, че глобално ИИ може да създаде стойност за здравните системи и икономиката, равносилна на **стотици милиарди долари спестени средства и предотвратени загуби годишно**, както и да спаси човешки животи. За да се реализира този потенциал, обаче, ще е нужна не само технология, но и реформи – адаптиране на медицинските протоколи, обучение на кадрите да работят съвместно с ИИ и изграждане на доверие у пациенти и доктори, че ИИ е помощник, а не заместител на човека.

6.1.5 Научни изследвания и развойна дейност

Научноизследователската и развойна дейност се намира на прага на нова ера, задвижвана от т.нар. „научен ИИ". В традиционния модел научният прогрес често е плод на бавен процес, при който учени формулират хипотези, експериментално ги проверяват, анализират данни и постепенно натрупват знания. ИИ вече започва да **ускорява драстично този цикъл**, като поема най-трудоемките задачи по обработка на данни и дори генерира собствени предположения, които учените да изследват. ИИ може да премахне огромна част от ръчния труд (бавното лабораторно определяне на структури) и да катализира иновации.

В по-широк план, ИИ се използва като универсален изследователски помощник:

- В **химията и материалознанието** алгоритми предсказват свойства на нови съединения или материали, което насочва учените към най-обещаващите кандидати за, да речем, по-ефективна батерия или по-здрав полимер.
- Във **физиката** машинното обучение помага за откриване на сигнали в огромните данни от ускорители на частици (напр., CERN) или астрономически наблюдения. Невронните мрежи идентифицират аномалии, които може да укажат нова частица или космическо явление, пропускани от класическите анализи.
- В **биологията**, освен протеините, ИИ се прилага за анализ на геноми, прогнозиране на генна експресия, откриване на взаимодействия между биомолекули. Това помага за разбиране на болести и откриване на нови терапевтични мишени.
- В **науките за климата и Земята** ИИ моделира сложни системи като климатичните промени, прогнозира екстремни събития или оптимизира възобновяеми източници. Например, сателитни данни плюс ИИ позволяват по-прецизен мониторинг на обезлесяването и емисиите на парникови газове, подпомагайки научно обосновани политики.

McKinsey отбелязва, че **около 75% от стойността на генеративните ИИ употреби ще дойде от 4 области, една от които е R&D.**[209] Това значи, че ефектът на ИИ върху науката може да се окаже дори по-дълбок от този в маркетинг, софтуер или обслужване на клиенти. Представете си **ускорение на "метаболизма" на научния процес** – ИИ може да генерира хиляди хипотези и виртуални експерименти, да отхвърли нерентабилните и да предложи на учените само най-обещаващите. Това е качествено нов подход: вместо научните екипи да работят на сляпо или по интуиция, те разполагат с "умна навигация" през пространството на възможностите.

Макар научните открития да не се превръщат веднага в пари, в дългосрочен план те задвижват цели индустрии. Откриването на нов свръхпроводник, например, може да доведе до революция в енергетиката и трилиони икономически ефекти десетилетие по-късно. Затова **държави и корпорации инвестират значително в ИИ за наука**. Очаква се до 2040 г. значителна част от научните публикации и патенти да включват като съавтор изкуствен интелект. Европейският съюз също стратегически инвестира през програма „Хоризонт Европа" в изследвания с помощта на ИИ и изграждането на високопроизводителни изчислителни инфраструктури за учени.

Автоматизацията на научната дейност, обаче, може да повлияе на творческия процес. Някои философи на науката питат, ако ИИ генерира и тества хипотези, няма ли учените да станат твърде зависими и да загубят умението си за критично мислене? Изследване на Yale[210] предупреждава, че прекомерното ползване на ИИ може да доведе до "правене на повече, но научаване на по-малко", ако хората сляпо се доверяват на моделите. Нужно е ново равновесие – учените да използват ИИ като инструмент, но да запазят човешката си интуиция и креативност.

Друго притеснение е възпроизвеждането на научния резултат: ако научно откритие е направено от затворен ИИ модел, трудно е да се обясни как се е стигнало до него, което е нетипично за научния метод, изискващ прозрачност. Появява се и въпросът за авторството – ако ИИ открие нещо, кой е откривателят? Правната рамка за интелектуална собственост тепърва ще трябва да навакса с тези сценарии.

6.1.6 Космически технологии

Космическата индустрия също е силно повлияна от ИИ, макар често да остава извън общественото внимание. Космосът е среда, където автономността и интелигентните системи са необходими, поради огромните разстояния и невъзможността за моментална човешка намеса. Затова ИИ тук намира естествено приложение в контрола на апарати, обработка на данни и поддръжка на космическа инфраструктура.

Ежедневно спътниците генерират терабайти от изображения и измервания на Земята – за време, климат, комуникации, навигация. Човек не е в състояние да преглежда цялата тази информация, но ИИ алгоритми са се превърнали в "очите" и "мозъка" на космическите данни. Машинно обучени модели сортират спътникови снимки и откриват ключови аномалии: незаконна сеч в тропическите гори, метанови емисии от индустриални обекти, промени в ледниците и др. Това позволява бърза реакция на правителствата и бизнеса. ИИ не само оптимизира процеси, но и може да допринесе за научни открития (намирайки иглата в купата сено на огромните космически масиви от данни). Модерните ракети и кораби разчитат на ИИ за управление в реално време.

С разрастването на сателитните мрежи координацията им става непосилна без ИИ. Алгоритми управляват флота, като предотвратяват сблъсъци чрез автоматични корекции на орбитите. С увеличаването на космическия трафик тази роля само ще расте – говори се за "космически диспечерски ИИ", които в реално време ще контролират стотици хиляди обекти в околоземното пространство, гарантирайки безопасност и ефективност.

---

[209] https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/tech-forward/scientific-ai-unlocking-the-next-frontier-of-r-and-d-productivity

[210] https://futureofbeinghuman.com/p/is-ai-poised-to-suck-the-soul-out-of-science

Икономическите ползи от ИИ в космическите технологии се проявяват във **все по-ниска цена за достъп до космоса** и **нови услуги**. Това отключва "мултитрилионната космическа икономика" на бъдещето – сателитен интернет, космически туризъм, дори добив на ресурси от астероиди. Ефективното управление на спътниците чрез ИИ намалява цената на услугите (навигация, телекомуникации), което има мултиплициращ ефект за икономиката на Земята – по-евтин интернет, по-добра логистика и т.н.

Както и на Земята, в космоса ИИ носи и нови рискове. Отказ или бъг в автономна система на космически кораб може да доведе до катастрофа, тъй като няма време за ръчна човешка намеса. Ето защо надеждността на софтуера е критична – тестовете и резервните системи са част от задължителните мерки. Има и **проблем с космическия боклук** – все повече отломки обикалят Земята; ИИ може да помага за избягването им, но ситуацията се влошава и може би ще изисква и ИИ решения за активно почистване.

Киберсигурността на космическите системи също е поле за внимание: един пробив в ИИ, управляващ сателити, би могъл да доведе до смущения в комуникациите или дори насочване на сателитите към сблъсък. Затова, макар ИИ да носи ефективност, се повишава нуждата от международни норми и сътрудничество за сигурно използване на космоса.

6.1.7 Опазване на околната среда и климат

Опазването на околната среда и борбата с климатичните промени са сфери, в които ИИ може да играе ролята на ускорител на устойчивостта. Докато предишните индустриални революции често са вървели ръка за ръка с екологични щети, то дигиталната революция с ИИ може да се използва за намаляване на тези щети и по-интелигентно управление на природните ресурси.

Благодарение на сензори и сателити, днес имаме безпрецедентен поток от данни за състоянието на планетата – от качеството на въздуха във всеки град, до влажността на почвите и популациите на диви животни. ИИ е незаменим инструмент за анализ на тази информация в реално време. Това позволява на правителствата и организациите да реагират бързо срещу незаконна сеч, движението на риболовни кораби, за да открият нерегламентиран риболов в защитени зони. В градовете, умни мрежи от датчици, комбинирани с ИИ, оптимизират напояването на паркове, сметосъбирането (чрез предсказване къде контейнерите са пълни) и др., спестявайки ресурси. Екологичната трансформация чрез ИИ не е само разход, а и инвестиция, която носи възвръщаемост под формата на икономическа активност (нови индустрии, нови работни места).

Въпреки положителния потенциал, ИИ сам по себе си има екологична цена. Центровете за данни и изчислителните нужди за обучение на големи модели консумират значителна енергия, което също е разгледано в този анализ.

Друг риск е прекомерното доверие в технологични решения за екопроблеми. Някои експерти предупреждават, че **наличието на ИИ-инструменти не трябва да отлага вземането на трудни политически решения**. Например, фактът че ИИ може да оптимизира трафика, не отменя нуждата от политики за ограничаване на замърсяващите автомобили в градовете.

ИИ се очертава като двойно оръжие в екологичен план: може да е част от проблема (ако увеличи консумацията на ресурси), но много повече може да бъде част от решението – да умножи ефекта от нашите усилия за опазване на природата. Отговорното и широко приложение на ИИ в околната среда до 2040 г. би могло да помогне за овладяване на климатичната криза, докато едновременно стимулира нови сектори на икономиката, съчетаващи зелено и дигитално развитие.

6.1.8 Енергетика

ИИ вече демонстрира забележителни резултати по отношение на ефективност, надеждност и интеграция на възобновяеми източници.

- **Умни електрически мрежи (smart grids)**: ИИ идва на помощ чрез прогнозиране и оптимизация. Алгоритми предвиждат натоварването на мрежата на база час, прогноза за времето, дори анализ на навиците на потребителите, и така операторите могат проактивно да балансират производството и търсенето. AI системи за оптимизация разпределят натоварването така, че загубите по преноса да са минимални и да се избегнат пикове, които иначе биха изисквали пускане на скъпи резервни мощности.
- **Интеграция на възобновяема енергия**: ИИ модели доста точно прогнозират колко соларна или вятърна енергия ще е налична, като вземат предвид метеорологичните прогнози, данни от сензори и исторически трендове. Това позволява на диспечерите да планират предварително. ИИ, също така, управлява зарядни и разрядни цикли на електрически батерии в мрежата – когато има излишък на зелена енергия, го складира, и после освобождава електрическата енергия в пикови часове. Този интелигентен контрол увеличава дела на зелената енергия, която може да се използва ефективно, и намалява нуждата от "рязане" на възобновяемите централи в моменти на свръхпроизводство. Държави като Германия и Дания, лидери във ВЕИ, разчитат на такива системи, за да поддържат стабилност въпреки високия процент възобновяеми източници в микса.
- **Енергийна ефективност при потреблението**: Освен на производството, ИИ влияе и на търсенето на енергия. Умни термостати, климатици и промишлени контролери, захранвани с машинно обучение, научават режимите на употреба и се самооптимизират. В промишлеността ИИ може да предложи промени в графиците на работа на енергоемки машини, за да измести натоварването към часове с по-евтина тарифа или излишък на възобновяема енергия. Това намалява пиковете на търсене и води до по-ниски сметки – печелят и фирмите, и цялата система (чрез по-малко нужда от пикови централи).
- **Предиктивна поддръжка и сигурност**: ИИ модели, обучени на базата на сензорни данни (вибрации, температура, шум), могат да предскажат кога една турбина е на път да аварира. Това дава възможност за планирана подмяна или ремонт, преди да се е случила скъпа неизправност. Така се избягват непланирани спирания на подаването на електроенергия (които носят големи косвени икономически щети), удължава се животът на оборудването и се оптимизират графиците за поддръжка (ремонтните екипи се изпращат точно когато и където трябва).

Енергийният сектор често е регулиран пазар. За да се внедри ИИ, понякога са нужни промени в регулациите, като например, да се позволи динамично ценообразуване на електрическата енергия, за да могат потребителските ИИ системи да реагират на промените в цените. Или да се разработят стандарти, които да гарантират, че автоматизираните решения няма да дискриминират или ощетяват някоя група потребители.

Още множество сектори като застраховане, системи за разплащания, мобилност и други имат подобни икономически ползи и потенциални рискове от внедряването на ИИ. Няма да остане икономически сектор, незасегнат от тази технология.

Глобалната надпревара в ИИ е реална: САЩ и Китай в момента извличат лъвския дял от икономическите ползи и оформят стандартите. Европа, макар и с добри намерения и силни основи, рискува да изостане, ако не ускори темпото. **ЕС има шанс да се диференцира с "надежден, човеко-центриран ИИ",** което, ако се комбинира с достатъчно инвестиции, може да привлече и глобални потребители и партньори, търсещи високо качество и етика.

Внедряването на ИИ в един регион може да има положителен ефект и върху останалите, чрез по-евтини продукти, нови пазари и взаимна търговия. Но и обратното: ако големи части от света изостанат, това може да задълбочи неравенствата между държавите.

Икономическата ефективност на ИИ е безспорна, но не и автоматична*. Необходими са целенасочени действия, за да се отключи този потенциал до 2040г.:*

- *Инвестиране в хора и технологии днес, за да се получат ползите утре.*
- *Управление на прехода на работната сила, за да е ИИ „усилвател" на човека, а не негов конкурент.*
- *Международно сътрудничество за стандарти и справяне с глобални предизвикателства (като климат, здраве), където ИИ може да е общото оръжие за справяне с глобалните проблеми.*

ИИ **"променя правилата на играта"** за бизнеса и обществото. **Той носи възможности за тези, които се адаптират, и рискове за изоставащите**. Европейската икономика, а и световната, се намират на кръстопът: да възприемат ИИ широко и отговорно, което би означавало нова епоха на растеж, иновации и благосъстояние, или да подценят промяната, рискувайки да загубят конкурентност и благоденствие.

До 2040 г. се очаква ИИ да бъде толкова повсеместен, че да не говорим за "ИИ сектор", тъй като **всеки сектор ще е задвижван от ИИ** – от първичните индустрии до творческите професии. Това ще бъде свят, в който рутинните задачи са автоматизирани, човешкият труд е насочен към висококреативни и междуличностни дейности, продуктите и услугите са по-достъпни и персонализирани, а икономиката като цяло е по-гъвкава и устойчива. Разбира се, този оптимистичен сценарий се осъществява, само ако проактивно се справим със споменатите рискове и бариери.

6.2. ИИ във финансите: инвестиции, пазарни ефекти и бъдещи сценарии

6.2.1 Инвестиции в ИИ

Глобалният пазар на ИИ нараства значително през последните години. През 2023г. стойността на глобалния ИИ пазар надхвърля 130 млрд. евро и се очаква да достигне близо €1.9 трилиона евро до 2030 г. Частните инвестиции формират по-голямата част от вложенията в ИИ, като лидер са Съединените щати с 44 млрд. евро частни инвестиции само през 2022 г., следвани от Китай с €12.5 млрд. евро. За сравнение, целият Европейски съюз заедно с Обединеното кралство са привлекли около 10.2 млрд. евро частни инвестиции през 2022 г. Тази разлика се задълбочава, като в периода 2018 – третото тримесечие на 2023 г. в компании за ИИ в ЕС са инвестирани общо 32.5 млрд. евро, срещу над €120 млрд. евро в компании за ИИ в САЩ. Освен частния сектор, растат и публичните инвестиции – например програмата **Digital Europe** на ЕС предвижда €2.1 млрд. евро финансиране за ИИ в периода 2021–2027.[211]

[211] https://epthinktank.eu/2024/04/04/ai-investment-eu-and-global-indicators/#

Най-големите инвеститори в ИИ в глобален план са големите технологични компании и рисковите инвеститори. Бумът в генеративния ИИ през 2023–2024 г., воден от успеха на системи като ChatGPT, предизвика скок в рисковите инвестиции. Само през второто тримесечие на 2024 г. вложенията в стартиращи компании за ИИ достигат 24 млрд. долара, над два пъти повече от предходното тримесечие. ИИ секторът за пръв път изпреварва всички останали по обем на рисково финансиране като дори традиционно силните сфери като здравеопазване и биотехнологии остават след него по привлечен капитал. Пет от шестте най-големи рискови сделки в света през този период са в ИИ компании, включително 6 млрд. щатски долара набрани от стартиращата фирма на Илон Мъск (xAI), и 1.1 млрд. щатски долара за инфраструктурната ИИ компания CoreWeave. Въпреки този подем, общото финансиране на стартиращи компании глобално все още е под пиковите нива от последните три години, което е знак, че макроикономически фактори (напр., високи лихви и по-слаб пазар за IPO) оказват влияние.

Американският технологичен сектор (Microsoft, Google, Amazon, NVIDIA и др.) доминира в инвестициите, както чрез собствени вложения в научноизследователска и развойна дейност, така и чрез придобиване на ИИ стартиращи компании. Например, Alphabet (Google) обяви планове да похарчи 75 млрд. щатски долара през 2025 г. за развитието на ИИ инфраструктура и продукти, което е значително над очакванията на пазара.[212] Китай също инвестира мащабно чрез компании като Baidu, Alibaba и държавни програми, макар и с по-малък дял засега. По отрасли, освен технологичния, финансовият сектор е сред водещите инвеститори и внедрители на ИИ решения (напр., за автоматизирана търговия и анализи). Значителни инвестиции има и в здравеопазването, производството и автомобилната индустрия (напр., автономни превозни средства), както и в отбраната. В Европа се наблюдава активизиране на стартъп екосистемата за ИИ. Между 2020 и 2022 г. европейски компании набират начален капитал и започват да настигат световните лидери в специфични ниши. В областта на **генеративния ИИ** САЩ категорично водят, като са привлекли над 7.4 млрд. щатски долара рисков капитал за генеративни ИИ проекти до 2022 г., което през 2023 г. нараства експоненциално. По оценки, към 2024 г. над половината от всички инвестиции в ИИ в САЩ се насочват именно към генеративни модели и приложения. [213] Тези средства финансират разработката на големи езикови модели (LLM) и свързаната инфраструктура, която изисква значителни ресурси за обучение и внедряване.

6.2.2 Влияние на ИИ върху глобалните финансови пазари – промени в структурата, ефективността, достъпа и ролята на институциите

Интеграцията на ИИ трансформира фундаментално функционирането на финансовите пазари. **Структурата на пазара** се променя посредством навлизането на нови играчи и технологии, които разкъсват традиционните връзки между финансовите институции. Според анализ на Deloitte[214], банково-финансовата индустрия е навлязла във „фазата на ИИ" от дигиталната си еволюция – ИИ буквално отслабва връзките, които досега споявaха компонентите на традиционните финансови институции, отваряйки вратата към повече иновации и нови оперативни модели. На практика, това означава, че **нетрадиционни участници** като финтех стартъпи, големи технологични компании или дори децентрализирани автономни организации могат да предлагат финансови услуги и да конкурират банки и фондови посредници, стъпвайки на платформи, базирани на ИИ.

[212] https://www.reuters.com/technology/google-parent-alphabet-misses-quarterly-revenue-estimates-2025-02-04/#
[213] https://epthinktank.eu/2024/04/04/ai-investment-eu-and-global-indicators/#
[214] https://www.deloitte.com/ng/en/services/risk-advisory/services/how-artificial-intelligence-is-transforming-the-financial-services-industry.html

Изкуственият интелект значително повишава пазарната ефективност. Алгоритмичната търговия, задвижвана от машинно самообучение, обработва огромни масиви от данни мигновено и изпълнява сделки с безпрецедентна бързина и прецизност.[215] Регулаторите отбелязват, че технологиите през годините са допринесли за понижена волатилност при стрес, тенденция която ИИ вероятно ще засили. Например, използването на генеративен ИИ за анализ на неструктурирани данни (финансови отчети, новини, изказвания) дава възможност на инвеститорите по-бързо да осмислят нова информация и да я отразят в цените. ИИ може да автоматизира кодиране, събиране на данни и инвестиционни анализи, което снижава бариерите за навлизане на нови участници дори в по-слабо ликвидни класове активи (напр., пазари на нововъзникващи икономики, корпоративни облигации и др.). [216] Това, от една страна, подобрява ликвидността и достъпа до тези сегменти, а от друга – поставя нови предизвикателства пред стабилността, ако множество алгоритми действат сходно в тесен пазар.

По-широкото внедряване на ИИ води и до **демократизация на финансовите услуги**. С понижаването на разходите и автоматизирането на процеси, услуги като инвестиционни съвети, управление на портфейл или кредитно посредничество стават достъпни за по-малки клиенти и нови групи. Т.нар. **„***Agentic AI***"** (автономни финансови агенти) биха могли да предоставят персонализирани финансови решения на хора и бизнеси, които досега са били недостатъчно обслужвани от традиционната система. Световният икономически форум отбелязва, че подобни AI агенти могат да подобрят финансовия достъп в маргинализирани общности, стига да се преодолеят пропастите в достъпа до технологии и да има подходяща регулация.[217] Същевременно, **глобалното неравенство в ресурсите и уменията може да породи нови форми на финансово разделение**, ако напредналите моделите на финансов ИИ се концентрират само в развитите пазари.

Традиционните финансови посредници (банки, борси, фондове) са принудени да се адаптират. Банките интегрират ИИ в бекофис операции (за автоматизация на отчетността и регулаторните отчети) и в клиентските услуги (чатботове, персонализирани предложения). Големите фондови борси прилагат машинно обучение за наблюдение на търговията и откриване на аномалии или пазарни злоупотреби в реално време. Хедж фондовете и активните мениджъри все по-често използват ИИ за генериране на инвестиционни стратегии и обработка на алтернативни данни, за да надхитрят конкуренцията. В резултат, **човешкият фактор се премества към надзора върху моделите и интерпретацията на резултатите, докато рутинните решения се оставят на алгоритмите.**

Въпреки безспорните ползи, регулаторите предупреждават, че масовото навлизане на ИИ може да доведе до нов тип системни рискове. Един от тях е т.нар. риск от „**монокултура**" – ако много участници използват сходни модели или данни, пазарните им действия може да се уеднаквят и да усилят корелациите. Експертите посочват, че при широкото приложение на модели за дълбоко обучение съществува опасност множество институции да стигат до аналогични изводи и стратегии, което повишава вероятността от синхронни пазарни движения. Макар на този етап разнообразието от модели и човешкият надзор да смекчават подобни рискове, потенциалът за **„стадно" поведение** на алгоритмите остава тревожен. Допълнително, черната кутия на някои ИИ (особено моделите с дълбоко обучение) затруднява регулаторния контрол – трудно е да се обяснят решенията на модела и да се открие дали няма зародили се некоректни или манипулативни модели на търговия. Надзорните органи срещат сериозни

[215] https://www.ibm.com/think/topics/artificial-intelligence-finance#
[216] https://www.imf.org/en/News/Articles/2024/09/06/sp090624-artificial-intelligence-and-its-impact-on-financial-markets-and-financial-stability#
[217] https://www.weforum.org/stories/2024/12/agentic-ai-financial-services-autonomy-efficiency-and-inclusion/#

предизвикателства да приспособят механизмите за пазарен надзор и системи за отчетност към тази нова реалност.[218]

Международните органи като Международния валутен фонд (МВФ) и Съветът за финансова стабилност (FSB) активно изследват ефекта на ИИ. FSB отчита, че ИИ носи значителни ползи – по-висока оперативна ефективност, подобрена регулаторна съвместимост, персонализирани продукти и напреднали анализи на данни. В същото време, обаче, ИИ може да усили уязвимости във финансовия сектор и по този начин да породи риск за стабилността. Сред изтъкнатите уязвимости са: зависимостта от външни доставчици на ИИ услуги и концентрацията при малко на брой доставчици; повишените пазарни корелации при масово използване на сходни модели; нови киберрискове, свързани с ИИ; рискове при самите модели – качество на данните, прозрачност и управление. Освен това, генеративният ИИ може да улесни финансови измами и разпространение на дезинформация на пазара, ако не се управлява правилно. Зле калибриран или неправилно насочен ИИ, който оперира извън етичните и правни граници, би могъл да нанесе вреди на финансовата система. Затова регулаторите призовават за засилено наблюдение на внедряването на ИИ, запълване на информационните пропуски и, при нужда, актуализиране на регулаторните рамки, за да се обхванат новите рискове.[219]

6.2.3 ИИ като агент и финансов съветник/анализатор – приложения, регулаторна рамка, рискове, нива на автономност

Към 2025 г. изкуственият интелект вече е широко внедрен във финансовите услуги в разнообразни роли. Най-често ИИ се използва като инструмент за подпомагане на човешки решения – **финансов анализ и прогнози**, оценка на кредитоспособност, откриване на измами и аномалии, автоматизирана обработка на документи и др. В инвестиционния мениджмънт много фондове използват модели с машинно самообучение за генериране на търговски сигнали или за ребалансиране на портфейли. Т.нар. **робо-съветници** предлагат автоматизирани инвестиционни консултации и управление на активи на крайни клиенти чрез алгоритми, които съобразяват рисков профил и цели. В банковия сектор системи с ИИ анализират транзакционни данни за засичане на измами, борба с изпирането на пари (чрез разпознаване на подозрителни модели) и автоматизация на KYC[220] процедури. Клиентското обслужване все по-често включва чатботове с естествено обработване на език, които могат да отговарят на запитвания и да изпълняват прости операции 24/7, освобождавайки човешки ресурс за по-сложни казуси.[221]

Големите езикови модели (като GPT-4) вече демонстрират потенциал да действат като финансови анализатори – да обобщават финансови отчети, да отговарят на въпроси на инвеститори и дори да генерират инвестиционни идеи въз основа на огромен корпус от данни. Някои хедж фондове експериментират с генеративен ИИ за създаване на резюмета на новини и анализ на пазарни настроения от социалните медии, което да ги информира при сделки. В ролята на съветник за клиенти, ИИ системите могат да предоставят персонализирани препоръки, например, подходящи спестовни и инвестиционни продукти според профила на клиента и да обучават потребителите по финансови въпроси чрез интерактивни платформи. Важно е да се отбележи, че към днешна дата тези ИИ съветници работят под наблюдението на човешки експерти и по предварително зададени ограничения.

---

[218] https://www.sidley.com/en/insights/newsupdates/2024/12/artificial-intelligence-in-financial-markets-systemic-risk-and-market-abuse-concerns#

[219] https://www.fsb.org/2024/11/the-financial-stability-implications-of-artificial-intelligence/

[220] Know your customer – познавай клиента си. Б.а.

[221] https://www.ibm.com/think/topics/artificial-intelligence-finance#

Автономността на ИИ във финансите варира. В повечето случаи ИИ е инструмент за предоставяне на анализ или препоръка, но човек взема финалното решение (т.нар. *human-in-the-loop* модел). Например, системата може да предложи търговска идея или да сигнализира за риск, а портфолио мениджърът решава дали да действа. Наблюдават се и **полуавтономни** случаи на алгоритмични стратегии, които самостоятелно извършват определени сделки по зададени параметри, но рамката им на действие е предварително одобрена от хора и подлежи на мониторинг. Най-напредничавите експерименти са с почти автономни агенти, използващи дълбоко учене и подсилващо обучение, които могат да адаптират поведението си в движение. Въпреки научния интерес, практически напълно автономни ИИ трейдъри на финансовите пазари все още не са допуснати широко. По оценки на МВФ, пазарните участници засега не изпитват комфорт към идеята за изцяло автоматизирани системи с ИИ, които генерират и изпълняват сделки без човешки надзор. Повечето институции настояват човешкият контрол да остане ключов компонент на всяка ИИ-базирана стратегия по регулаторни, репутационни и етични причини.[222] С други думи, към 2024 г. индустрията възприема ИИ като мощен помощник, но не и като независим субект при вземането на решения.

Регулаторите третират ИИ в сферата на финансите на принципа на технологичната неутралност, т.е. съществуващите закони и правила (напр., за предоставяне на инвестиционни съвети, управление на активи, кредитна дейност) се прилагат и когато тези услуги се предоставят или подпомагат от алгоритми. Например, ако роботизиран съветник дава инвестиционни препоръки, лицензираната финансова компания, която го оперира, носи отговорност за спазването на фидуциарните й задължения към клиента. Надзорните органи започват да издават насоки относно използването на ИИ, като акцент се поставя върху управлението на моделите, необходимостта от обяснимост на решенията (напр., изискване алгоритмичните решения да бъдат обясними до степен, в която може да се обоснове пред клиент или регулатор), редовно тестване за пристрастия или дискриминация, и наличие на „спирачки" (*kill-switch*) при непредвидено поведение. В ЕС Актът за изкуствения интелект класифицира някои финансови приложения на ИИ като високорискови, които ще изискват допълнителна оценка и мониторинг.

В същото време, регулаторите признават, че **ИИ носи и ползи за самия надзор** – т.нар. *RegTech/SupTech* приложения. Надзорните агенции внедряват ИИ инструменти, които да пресяват огромни масиви от данни от финансови отчети, транзакции и пазарна информация, за да откриват аномалии или ранни сигнали за проблеми. Например, ИИ може да анализира всички сделки на даден пазар за милисекунди, което да помогне в разследването на евентуална манипулация много по-бързо от традиционните методи. Това повишава капацитета на регулаторите да реагират своевременно на пазарни злоупотреби или възникващи рискове, но поражда и зависимост от технологиите, с които самите регулатори разполагат.

Сред ключовите рискове при ИИ съветниците и агентите са **пристрастията в моделите**. Ако историческите данни са изкривени, ИИ може да взема дискриминационни решения, например, да отказва кредити на определени групи или да има оперативен риск от грешки или аномално поведение на модела. Също така, липсата на прозрачност при някои модели затруднява доверието и клиентите може да не разбират защо ИИ дава определен съвет, което да подкопае приемането му. Има и правна неяснота относно отговорността при грешка на автономен алгоритъм – в крайна сметка отговорни са институциите и техните ръководства, но изработването на вътрешни контроли за ИИ системите е ново предизвикателство. Затова надзорните органи препоръчват запазването на човешки надзор и контролни точки в процесите с ИИ, както и развитие на етични рамки за ИИ (напр., принципите за прозрачен, проследим и справедлив ИИ, лансирани от ЕС). ИИ вече действа като „умна машина" във финансите,

[222] https://www.imf.org/en/News/Articles/2024/09/06/sp090624-artificial-intelligence-and-its-impact-on-financial-markets-and-financial-stability#

но на този етап под ръководството на човека и регулатора, които поставят границите на неговата автономност.

През следващите петнадесет години развитието на изкуствения интелект има потенциала да прекрои из основи ландшафта на финансовите пазари. В *Приложение 6* разглеждаме няколко възможни сценария и аспекти на трансформацията до 2040 г.

ИИ обещава по-ефективни, достъпни и иновативни финансови пазари, но и носи рискове – от системни сривове до нови видове злоупотреби. Регулаторната рамка и човешкият надзор ще са критични за това дали ИИ ще действа в полза на стабилността, или ще я разклаща. В идеалния случай, чрез проактивно управление и международно сътрудничество, ИИ ще се интегрира като „съюзник" и мощен инструмент, който работи редом с човека за по-доброто функциониране на финансовата система. Ако обаче напредъкът изпревари контрола, ИИ може да се превърне и в източник на нови рискове, изискващи спешни мерки. **Управлението на този преход ще определи облика на глобалните финансови пазари през следващите десетилетия.**

6.3. Икономическа ефективност и разходи

Разработването на модели за изкуствен интелект (ИИ) стана все по-ресурсоемко през последните години. През последните три години разходите за разработка на авангардни ИИ модели рязко са се увеличили, тъй като размерите и сложността на моделите нараснаха, въпреки че някои единични разходи (като изчислителните операции на единица) са намалели. Гледайки напред към следващите десет години, можем да очакваме комбинация от предизвикателства и подобрения – от по-високи изисквания за изчислителна мощност и данни до нови възможности за ефективност и регулации. По-долу разглеждаме основните фактори, влияещи върху разходите, както и техните тенденции, подкрепени с данни и прогнози.

6.3.1 Разходи за изчисления и хардуер

Изчислителният хардуер, необходим за ИИ, е основен двигател на нарастващите разходи. Съвременните обучения на ИИ модели използват **масивни клъстери от GPU/TPU** (графични или тензорни процесори), които са скъпи за придобиване или наемане. Например, водещият графичен процесор на Nvidia за центрове за данни струва приблизително **10 000 долара за единица**.[223] , а обучението на големи модели изисква хиляди такива чипове, работещи в продължение на седмици. Изчислено е, че между 47% и 67% от общите разходи за най-модерните модели са отишли изцяло за хардуер.[224] В резултат на това разходите за обучение рязко са нараснали – обучението на GPT-4 на OpenAI (2023) е струвало приблизително 78 милиона долара, като в разходите влизат хардуерът за обучение, степента на използване на хардуера и продължителността на обучението, в сравнение с около 12 милиона долара за PaLM на Google през 2022 г.[225] Разходите за обучение на модела Gemini Ultra на Google (2023) нарастват на внушителните 191.4 млн. долара. Най-последният модел на Google струва между 30-190 млн. долара само в изчисления за обучение, без в разходите да са включени заплатите на екипа, които могат да достигнат между 29-49% от крайната цена.[226] Общо разходите за обучение на ИИ са скочили 44 пъти от 2020 г. насам, тласкани основно от все по-високите изисквания за изчислителна мощ.

[223] Vanian, Leswing, 2023, ChatGPT and generative AI are booming, but the costs can be extraordinary, https://www.cnbc.com/2023/03/13/chatgpt-and-generative-ai-are-booming-but-at-a-very-expensive-price.html
[224] https://www.edge-ai-vision.com/2024/09/ai-model-training-cost-have-skyrocketed-by-more-than-4300-since-2020/#
[225] https://www.visualcapitalist.com/training-costs-of-ai-models-over-time/#
[226] https://www.edge-ai-vision.com/2024/09/ai-model-training-cost-have-skyrocketed-by-more-than-4300-since-2020/#

В областта на хардуера се наблюдава бърз и устойчив напредък в производителността – тенденция, която често се обозначава като „Законът на Хуанг" (по името на главния изпълнителен директор на Nvidia). Новите поколения GPU[227], като H100 на Nvidia или TPU на Google, предоставят все повече изчислителни операции на всеки изразходван долар, което води до постепенно подобряване на ефективността на разходите. Според прогнозите на индустрията, инвестициите в хардуер за изкуствен интелект ще продължат да нарастват, като се очаква пазарът на GPU центрове за данни да достигне почти 65 милиарда щатски долара до 2032 г., спрямо приблизително 15 милиарда долара през 2024 г.[228] (средногодишен темп на растеж от приблизително 19.8%), тъй като компаниите се снабдяват с все повече GPU/TPU процесори. Това мащабиране би трябвало в дългосрочен план да доведе до понижаване на разходите за единица изчислителна мощност – доставчиците на облачни услуги и производителите на чипове ще предлагат все повече изчислителна сила за същата цена. Въпреки това, ако водещите изследователски лаборатории в сферата на изкуствения интелект продължат да се стремят към създаване на все по-големи модели, абсолютните бюджети за хардуер е възможно да продължат да нарастват. Налице са дори предупреждения, че при настоящите темпове на скалиране една-единствена тренировка на бъдещ мащабен модел може да надхвърли 1 милиард щатски долара до 2027 година.[229]

На практика е вероятно да се наблюдава определен баланс: хардуерът ще става все по-специализиран и ефективен (включително чрез разработване на собствени AI чипове и подобрения в паметта и мрежовата инфраструктура за по-добра ресурсна използваемост), което ще натиска кривата на разходите надолу. Въпреки това, свръхамбициозните проекти ще продължат да изискват значителен капитал. Очаква се по-малки организации да разчитат на облачни изчисления за гъвкавост, въпреки че наемането на облачни GPU ресурси все още остава скъпо. В обобщение – цената на хардуера за единица изчислителна мощност (на флоп) вероятно ще продължи да намалява, но общите разходи за изчисления при водещите AI модели може да продължат да нарастват, освен ако не настъпи пробивна промяна в подхода.

6.3.2 Разходи за придобиване и обработка на данни

Обучението на напреднали AI модели не се изчерпва само с изчислителната мощност – те също така „поглъщат" огромни количества данни, чието набавяне и подготовка могат да бъдат скъпи. През последните няколко години екипите се възползваха от големи публични набори от данни (напр., „уеб обхождане" - web crawling), които по същество бяха „безплатни", но стремежът към по-качествени и по-мащабни данни доведе до увеличаване на разходите за събиране и управление. Данните с високо качество и прецизно означаване са особено скъпи: например, означаването на набор от 100 000 изображения с ограничителни кутии (bounding boxes) може да струва около 200 000 щатски долара само за човешка анотация.[230]

Все повече AI разработчици заплащат на външни услуги или на платформи с краудсорсинг работници за събиране и почистване на данни, особено когато става въпрос за специализирани домейни. Пазарът за тренировъчни данни за изкуствен интелект бележи бум, отразявайки нарастващото търсене – очаква се той да нараства със средногодишен темп от около 24.3% до 2030 година.[231]

[227] Graphics Processing Unit – високопроизводителен процесор, предназначен за ускоряване на сложни изчислителни дейности, критични за ИИ, машинно обучение, анализ на данни и облачни изчисления. Б.а.
[228] https://www.verifiedmarketresearch.com/product/data-center-gpu-market/#
[229] https://hai.stanford.edu/ai-index/2024-ai-index-report
[230] По цени от интернет. Б.а.
[231] https://www.lucintel.com/artificial-intelligence-market.aspx

Компании като Scale AI (която се занимава с етикетиране/анотиране на данни) отбелязват бърз растеж, тъй като организациите инвестират все повече средства в подготовката на данни за обучение и фина настройка на AI модели.

Друга нова тенденция е, че досега безплатни източници на данни започват да въвеждат такси. Основни платформи осъзнаха стойността на своите данни за нуждите на изкуствения интелект. Например, през 2023 г. Reddit започна да таксува достъпа до своя API[232], като посочи, че неговото съдържание е било използвано за обучение на модели като GPT-4[233]. Съгласно новите условия, едно приложение, което извършва приблизително 7 милиарда заявки месечно към Reddit (нещо обичайно при мащабно извличане на данни), би генерирало такси от около 2 милиона щатски долара на месец според ценовата политика на Reddit. Twitter (сега X) също значително повиши цените за достъп до своя API. Тази тенденция означава, че разработчиците на ИИ все по-често ще се изправят пред сериозни лицензионни разходи за използване на съдържание от социални медии, новинарски източници или други частни източници, докато в миналото такива данни често са били извличани свободно. В допълнение, регулации, свързани с поверителността на данните (например, GDPR), ограничават използването на определени лични данни, което принуждава компаниите да инвестират в съответстващи на изискванията методи за събиране на данни или в генериране на синтетични данни.

През следващото десетилетие **данните ще останат ключов актив и съществен разходен център за проекти с ИИ**. С нарастването на мащаба на моделите ще се изискват все по-разнообразни и обширни набори от данни, за да могат те да се обучават ефективно. Очаква се **индустрията по подготовка и обработка на данни да продължи да расте** – както беше посочено, пазарът на тренировъчни данни се очаква да отбелязва над 20% годишен растеж[234], което означава още повече разходи за събиране, почистване и анотиране на информация. В същото време, обаче, могат да се очакват и противодействащи тенденции, които да намалят разходите за данни на модел: например, подобрени алгоритми за обучение с висока ефективност при използване на данни (изискващи по-малко примери), повторно използване на предварително обучени модели (така че не всяка организация да трябва да събира милиарди сурови токени от нулата), както и все по-широка употреба на синтетични данни за допълване на реалните набори. Технологични подходи като самообучение и активно обучение също могат да намалят зависимостта от скъпо човешко анотиране на данните, като позволяват на моделите да използват неанотирани или частично анотирани данни.

Същевременно, ако регулаторният натиск затрудни извличането на публични данни, AI разработчиците може да се наложи да плащат лицензи за повече набори от данни (от издатели, държавни институции и др.) или да изграждат собствени платформи за първични данни – всяко от които води до допълнителни разходи. **Възможно е също така да станем свидетели на възхода на специализирани пазари за данни, където подбрани набори от данни ще се купуват и продават, формализирайки още повече данните като търговски актив**.

Следователно, организациите следва да заделят съществени бюджети за данните в рамките на AI проектите. Много анализатори в индустрията очакват разходите, свързани с данни – включително за съхранение и предварителна обработка – да заемат нарастващ дял от проектните бюджети, освен ако не се разработят нови техники, които да позволят на моделите да „учат повече от по-малко".

[232] API, от английски: application programming interface - приложен програмен интерфейс
[233] https://www.techtarget.com/whatis/feature/Reddit-pricing-API-charge-explained#
[234] https://www.lucintel.com/telecommunication-market-report/artificial-intelligence.aspx#

6.3.3 Разходи за обучение и фина настройка на модели

Разходите за обучение на водещи AI модели нараснаха драстично през последните години, тъй като размерите на моделите (и необходимата изчислителна мощност) се увеличиха експоненциално. Например, според оценки, обучението на GPT-4 (2023) от OpenAI е струвало приблизително 78 милиона щатски долара, в сравнение с около 4 милиона долара за GPT-3 през 2020 г. и под 1 милион долара за ранни модели[235].

Самото обучение на AI модел – провеждането на мащабни изчисления с цел настройване на милиарди параметри – представлява една от най-скъпите фази в разработката. Както се вижда, водещите модели днес изискват десетки милиони долари за обучение от нулата. Тези разходи нараснаха рязко през последните 2–3 години, поради мащаба на модели като големите езикови модели (с по стотици милиарди параметри). Например, PaLM на Google (2022) и GPT-4 на OpenAI (2023) са изисквали в диапазона от 10^24 до 10^25 флопа (floating-point операции) по време на обучението си – еквивалентно на облачни изчисления на стойност десетки милиони щатски долари.[236]

Подобни обучения често продължават седмици и се провеждат на специализирани суперкомпютри, което допълнително увеличава разходите. Освен самия хардуер, процесът генерира и значителни разходи за електроенергия (разгледани в раздела за енергия), както и изисква сериозен инженеринг за координация, макар тук да се фокусираме основно върху преките разходи за изчисления. Основната идея е, че **обучението на авангардни модели е станало толкова скъпо, че само добре финансирани технологични гиганти и няколко стартиращи компании могат да си го позволят**. Според експертите в бранша, цената на най-напредналите обучения на модели ще струват над един милиард долара, ако тенденциите се запазят.

Една облекчаваща тенденция е нарастващото използване на предварително обучени модели и тяхната фина настройка. Вместо да се обучава нов модел от нулата (което може да струва милиони), много организации използват съществуващ голям модел и го донастройват върху свои данни за конкретна задача. Фината настройка е с порядъци по-евтина, тъй като изисква значително по-малко изчислителни ресурси – моделът вече е научил общи закономерности и се налага само леко адаптиране. Например, OpenAI предлага достъп до API за фина настройка на GPT-3.5, като разходите са относително ниски – приблизително 0.008 щатски долара на 1,000 токена входни данни. Това означава, че дори с няколко милиона токена (достатъчни за много специализирани приложения), фината настройка може да струва само няколкостотин долара за изчисления.

Фината настройка по този начин прави персонализираните AI решения достъпни и за по-малки организации, използвайки вече направените мащабни инвестиции от страна на големите технологични компании. През последните три години този подход се утвърди като стандарт – виждаме как компании фино настройват големи модели за нужди в медицината, финансите, обслужването на клиенти и др., вместо да инвестират в пълно обучение от нулата.

Разходите за пълномащабно обучение вероятно ще останат ограничаващ фактор, достъпен само за най-големите проекти през следващото десетилетие. Въпреки това, се очаква напредък в ефективността. Изследователите активно търсят начини за намаляване на изчислителните изисквания при обучение чрез алгоритмични иновации. Забележително е, че алгоритмичните подобрения исторически изпреварват хардуерните — изследване на OpenAI[237] показа, че между 2012 и 2019 г. необходимата изчислителна

[235] Stanford AI Index Report 2024 https://hai.stanford.edu/ai-index
[236] Ibid.
[237] https://openai.com/index/ai-and-efficiency/

мощност за постигане на определено ниво на точност е спаднала 44 пъти, благодарение на повишена алгоритмична ефективност, при което „времето за наполовина намаление" е било около 16 месеца.

Ако подобни тенденции в ефективността продължат (напр., по-добри алгоритми за оптимизация, по-умни стратегии за инициализация на тегла, техники за паралелизация и др.), цената за постигане на определено ниво на производителност ще продължи да намалява. Всъщност, ARK Invest изчислява, че разходите за обучение на модел със сравнима производителност на GPT-3 са спаднали от 4.6 милиона щатски долара през 2020 г. до приблизително 450 хиляди щатски долара през 2022 г. и прогнозират средногодишен спад от около 70%. Подобна траектория предполага, че до 2030 г. обучението на еквивалентен модел може да струва само няколко хиляди долара.

Все пак трябва да се направи едно важно уточнение: ако водещият модел през 2035 г. е значително по-сложен от GPT-4 (което е вероятно, ако се запази сегашната посока на развитие), неговото обучение може да остане изключително скъпо в абсолютен план. Общността обаче може да възприеме стратегии за овладяване на тези разходи, като използване на модулни или по-малки модели, които работят съвместно (вместо един монолитен модел), или техники като микс от експерти, които целят да постигнат производителност, съпоставима с гигантски модел, но с по-ниски изчислителни разходи.

Друга потенциална посока за напредък е автоматизираната настройка на хиперпараметри, която би намалила скъпия процес на проба-грешка при обучението. Трансферното обучение и продължителното обучение също могат да позволят на нови модели да стартират не само от същата архитектура, но и от вече научени параметри, спестявайки значителни изчисления.

Макар че разходите за обучение на най-големите модели рязко са се увеличили, **дългосрочната тенденция може да се измести към по-ниски разходи за единица способност,** благодарение на технологичния напредък и пазарния натиск. Фината настройка и повторната употреба на вече обучени модели ще останат ключови стратегии за широко разпространение на изкуствения интелект без необходимостта от многократни, скъпоструващи обучения от нулата.

6.3.4 Разходи за работна сила и таланти

Човешкият капитал е още един съществен компонент от разходите за разработка на изкуствен интелект. През последните няколко години наблюдаваме ожесточена конкуренция за изследователи и инженери, което доведе до значителен ръст на възнагражденията. Шестцифрените годишни заплати вече са стандарт за инженерите по машинно обучение, а най-добрите специалисти получават дори много повече. През 2024 г. 76% от компаниите съобщават, че изпитват сериозен недостиг на персонал с умения в сферата на ИИ, което тласка работодателите да наддават с по-високи компенсационни пакети.[238]

Елитните изследователи и лидери в индустрията често получават общи възнаграждения в размер на милиони долари: скорошни данни показват, че някои позиции за инженери по машинно обучение варират между 900 хил. и 2.8 млн. щатски долара, а опитни ИИ изследователи достигат между 1.7 и 4.2 млн. щатски долара под формата на заплата, акции и бонуси. Дори специалистите на средно ниво бележат значителен ръст. Заплатите на професионалисти по ИИ в САЩ са се повишили с около 10–13% само в периода 2021–2022 г., значително над средното за други технологични роли. [239] В периода 2015-2023 г. **броят обявени работни места, изискващи**

[238] https://www.informationweek.com/it-leadership/the-cost-of-ai-talent-who-s-hurting-in-the-search-for-ai-stars-#

[239] https://bidenwhitehouse.archives.gov/cea/written-materials/2025/01/14/ai-talent-report/

компетенции и умения, свързани с ИИ, само в САЩ като водеща глобална сила в областта на ИИ са нараснали с 257% при ръст на общия брой обяви за работа от 52% за същия период. [240]

Освен преките разходи за заплати, организациите инвестират в обучение и преквалификация на съществуващ персонал в посока ИИ умения, както и често наемат скъпи външни консултанти или изпълнители за специфични ИИ проекти.

Съществено е да се подчертае, че не само изследователите участват в изграждането на напреднали ИИ модели – необходими са и значителни екипи от софтуерни инженери, инженери по данни, продуктови мениджъри и експерти в конкретната област (а понякога и анотатори на данни). При водещ ИИ проект, работната сила може да представлява между 30% и 50% от общите разходи за разработка, ако се вземат предвид високите възнаграждения. Например, ако хардуерът и облачните изчисления струват 50 млн. щатски долара, персоналът, който проектира алгоритмите, пише кода, управлява процеса на обучение и оценява модела, може да добави още десетки милиони към бюджета. През последните три години много технологични компании пренасочиха ресурси към ИИ, понякога съкращавайки други екипи, за да финансират тези скъпи инициативи.

В дългосрочен план, търсенето на таланти в областта на ИИ вероятно ще остане изключително високо в следващото десетилетие. С разширяването на прилагането на изкуствения интелект в различни индустрии (финанси, здравеопазване, производство и др.), конкуренцията за опитни разработчици и изследователи ще продължи на глобално ниво. Това може да задържи високите темпове на ръст в заплатите, поне докато предлагането на квалифицирани кадри не настигне търсенето. Правителства и академични институции се опитват да обучат повече специалисти, но недостигът на кадри и умения остава значителен.

От гледна точка на разходите, компаниите трябва да планират не само за наемане, но и за задържане на талантите чрез постоянни стимули, свобода за изследователска дейност и други непреки инвестиции. Някакво облекчение може да дойде от автоматизацията в разработката на ИИ (например, чрез инструменти за автоматизирано машинно обучение), която би могла да оптимизира работния процес и да намали нуждата от големи екипи за някои задачи. **Но при разработката на водещи модели човешката експертиза ще остава незаменима.**

Една тенденция, която може да облекчи разходите, е възходът на бизнес моделът ИИ като услуга (AI-as-a-service) и облачни платформи, които поемат по-голямата част от сложността. Това позволява на компаниите да използват готови модели, без да изграждат вътрешни екипи от изследователи. По-малките фирми могат да се възползват от предварително обучени модели чрез API и да ги настроят фино спрямо своите нужди, без да поддържат пълен изследователски екип.

Въпреки това, всяка организация, която цели да бъде на предната линия или да разработва силно персонализирани решения с ИИ, ще има нужда от квалифицирани специалисти. Географските промени също могат да повлияят на разходите: AI хъбовете в Силициевата долина, Ню Йорк и други големи центрове остават скъпи, но може да станем свидетели на по-разпределени екипи или нови центрове в региони с по-ниски разходи, с разширяването на глобалния пул от таланти. В този смисъл е от изключително значение как всяка една държава ще се позиционира и дали Европа ще успее да ги привлича и задържа. Конкретно за България това ще означава дълбока трансформация на образователната система, която да бъде организирана по начин, който превръща държавата в ковачница на кадри с умения за тази нова технология.

[240] Ibid.

През следващите десет години високата „ценова надбавка" за експертиза и умения в областта на ИИ вероятно ще се запази. **Компаниите трябва да планират значителни инвестиции в човешки капитал или да търсят креативни стратегии за достъп до такъв капацитет** (например, чрез партньорства или облачни услуги), ако не могат директно да наемат най-добрите кадри. Надеждата е, че с подобряването на инструментите един инженер или изследовател ще може да постига повече с по-малко, което частично ще намали нуждата от големи екипи, но иновациите на най-високо ниво в областта на ИИ в обозримо бъдеще ще останат интензивно зависими от човешки ресурси.

6.3.5 Въздействие на регулациите за ИИ върху разходите за прилагане и съответствие

С разширяването на внедряването на изкуствен интелект правителствата въвеждат нови регулации, за да гарантират, че ИИ се използва етично, безопасно и в съответствие с обществените ценности. **Тези нови законови рамки и насоки създават допълнителен слой от разходи – такива, свързани със съответствие и нормативна уредба.** През последните три години наблюдаваме предложения за цялостни нормативни режими, като Акта за изкуствения интелект на ЕС, разгледан в този анализ, както и различни национални инициативи за надзор върху ИИ. Макар повечето от тези регулации все още да не са напълно в сила или, ако са в сила, рамката за прилагането на законодателството е в процес на подготовка, компаниите вече започват да се подготвят за спазването им.

Това често изисква **инвестиции в документация, вътрешен и външен одит и контрол на ИИ моделите**. Например, фирмите в ЕС ще са задължени да провеждат оценки на риска и пристрастията в своите модели, да поддържат подробни записи за произхода на тренировъчните данни, да внедрят функции за прозрачност към потребителя и дори да позволят външни алгоритмични одити – всички от които изискват време и средства.

Проучване на KPMG от 2024 г. показва[241], че 52% от изпълнителните директори очакват, че новите правила за защита на лични данни и ИИ ще увеличат разходите за техните организации, а 63% предвиждат допълнително затягане на регулациите в близко бъдеще. В отговор на това около 60% от компаниите вече преразглеждат или актуализират практиките си за управление на данни, а приблизително половината въвеждат технически мерки за прозрачност и справедливост на ИИ. Тези усилия водят до реални разходи, свързани с наемане на служители по съответствие, правни консултанти, както и удължаване на времето за разработка поради интеграцията на функции, свързани с нормативните изисквания.

Наред с това, регулации като Общия регламент за защита на личните данни (GDPR) вече оказват въздействие върху разработването на ИИ модели, тъй като компаниите трябва да гарантират, че **използваните за обучение данни са законно придобити и съответстват на изискванията за защита на личните данни**. Това може да ограничи използването на определени набори от данни или да наложи допълнителна анонимизация, което увеличава разходите. През 2023 г. някои AI услуги дори временно преустановиха дейност в определени региони, поради регулаторна несигурност – ChatGPT на OpenAI беше временно забранен в Италия до предоставяне на гаранции относно поверителността. Това подчертава, че съответствието с регулациите вече е ключов елемент при внедряването на ИИ.

През следващото десетилетие **се очаква регулирането на ИИ да се развие и разшири, което най-вероятно ще доведе до допълнително нарастване на**

[241] KPMG GenAI Survey 2024 https://kpmg.com/kpmg-us/content/dam/kpmg/corporate-communications/pdf/2024/kpmg-genai-survey-august-2024.pdf

разходите за съответствие. Актът за изкуствения интелект на ЕС ще влезе в пълна сила в периода 2025–2026 г., като въведе изисквания в зависимост от нивото на риск на системите с ИИ. Приложения с висок риск ще подлежат на строги процедури относно тестване, документация, а дори и задължително застраховане срещу отговорност.

Компании, които разработват ИИ за чувствителни области като здравеопазване, финанси или подбор на персонал (класифицирани като "висок риск"), ще трябва да предвидят сериозен бюджет за тези допълнителни процеси. Възможно е, също така, **външният одит на ИИ системи (по подобие на финансовите одити) да се превърне в стандартна практика, което би дало начало на нова индустрия от одитори и сертифициращи органи по ИИ** – още едно перо в разходите за разработчици. **Възможни са и такси, свързани с получаване на регулаторно одобрение или регистрация на определени системи с ИИ.**

От друга страна, ясни и предвидими регулации биха могли да намалят несигурността и да предотвратят скъпо струващи грешки, като например пускане на система с ИИ, която по-късно предизвиква юридически последици. **В дългосрочен план, инвестициите в "отговорен ИИ" още в начален етап могат да се окажат икономически изгодни, като спестяват глоби, съдебни дела или нуждата от скъпоструващи корекции впоследствие.**

Във всички случаи, обаче, привеждането в съответствие в нормативната рамка ще добави административно натоварване: очаква се удължаване на циклите по разработка с оглед на нуждата от оценка на справедливост и прозрачност, необходимост от повече експерти по право и етика в екипите, както и евентуално преработване на архитектурите на моделите, за да отговарят на новите стандарти (например, осигуряването на обяснимост на решенията на модела може да наложи използване на по-прости модели или добавяне на допълнителни модули за обяснение).

Също така, следва да се отчетат и регулациите, свързани с интелектуалната собственост и авторските права. Вече се водят правни спорове относно използването на авторско съдържание при обучението на модели с ИИ и за собствеността върху създадените от ИИ резултати. Ако законите изискват лицензиране на тренировъчни данни (например, изображения, текстове от автори и публикации), това може пряко да увеличи разходите за набавяне на данни (както вече беше отбелязано при случаите с платформи, започнали да таксуват достъпа до съдържание).

Освен това, регулациите за безопасност на ИИ могат да наложат задължителни мерки за сигурност като обширно тестване за опасни способности при напреднали модели. Това би направило процеса по разработка още по-дълъг и по-сложен, като ще изисква допълнителни ресурси за научноизследователска дейност, инженеринг и изчисления, свързани с внедряването на такива предпазни мерки.

Накратко, **привеждането в съответствие с непрекъснато променящите се регулаторни рамки за ИИ се превръща във все по-значим елемент от бюджетите за разработка на ИИ.** В индустриите вече се наблюдава консенсус около идеята, че "регулацията на ИИ ще увеличи разходите за технологии", аналогично на това как финансовият и фармацевтичният сектор отделят съществен дял от своите разходи за научно-развойна дейност за спазване на регулаторните стандарти.

През следващото десетилетие вероятно ще видим екипите по изкуствен интелект като стандарт да включват не само инженери и учени по данни, но и юристи, експерти по етика и одитори. Организациите, които планират напред и изградят устойчиви процеси за съответствие, могат да превърнат това в управляем разход за водене на бизнес в сферата на ИИ. Онези, които пренебрегнат регулациите, рискуват по-високи санкции, съдебни спорове или необходимост от скъпоструващи корекции на по-късен етап.

6.3.6 Технологичният напредък и неговото въздействие върху ценообразуването

Въпреки посочените предизвикателства, има сериозни основания за оптимизъм относно бъдещата икономика на разработването на ИИ. Технологичният напредък продължава да избутва кривата на разходите надолу. Ето обобщение на основните очаквани технологични подобрения и тяхното въздействие до 2035 г.:

- **Подобрения в хардуера**: Очаква се значителен напредък в изчислителната мощност на хардуера. Производителността на GPU спрямо цената се подобрява с всяко ново поколение, а на пазара излизат нови ИИ ускорители от компании като Google, Amazon и множество стартъпи. Това ще доведе до постепенно намаляване на разходите за единица изчисление. Например, ако днешният водещ GPU обработва X трилиона операции за 1 долар, следващото поколение може да обработва 5 или 10 пъти повече за същата цена. Специализирани чипове (ASIC), проектирани за специфични натоварвания на ИИ, могат допълнително да намалят разходите за обучение и извеждане на резултати при мащабно производство. **До 2035 г. е възможно дори да видим квантови или фотонни компютри, които поемат определени задачи за машинно обучение,** потенциално ускорявайки работните процеси с различна ценова структура (макар да е вероятно да не са все още масово разпространени в периода, за който говорим). С други думи, **индустрията очаква прогрес, подобен на закона на Мур или дори по-бърз, което ще направи всяка задача, свързана с ИИ, по-евтина за изпълнение.** Според една смела прогноза, разходите за обучение на модел от нивото на GPT-3 ще спадат със 70% годишно до 2030 г. (ark-invest.com), което би означавало, че задачи, които днес струват десетки милиони долари, може да струват само хиляди в края на десетилетието.[242]

- **Алгоритмични и софтуерни иновации**: развитието на алгоритмите за машинно обучение има потенциал драстично да намали разходите. Вече бе установено, че през 2010-те години алгоритмичната ефективност при класификация на изображения се е удвоявала приблизително на всеки 16 месеца.[243] Ако тази тенденция се запази, ефектът върху разходите ще е съществен. Новите оптимизационни методи, по-ефективните алгоритми за обучение и техниките като федеративно обучение или трансферно обучение ще намалят нуждата от огромни количества нови данни и изчислителни ресурси. Ясен пример е създаването на архитектури, които постигат същата точност с по-малко параметри – ако можем да постигнем целеви резултат с модел от 1 милиард параметри вместо 10 милиарда, това води до спестяване на около 10 пъти изчисленията при обучение. Активно се разработват и технологии за *sparsity*[244] и *pruning*[245] на модела, както и *дистилация* (процес, при който големи модели се компресират в по-малки без значителна загуба на производителност). Всички тези подходи водят до по-леки, по-бързи и по-евтини за обучение и изпълнение модели. Очакванията са, **че до 2035 г. моделите с ИИ ще бъдат значително по-ефективни откъм изчисления**

[242] https://www.ark-invest.com/big-ideas-2023/artificial-intelligence#
[243] https://openai.com/index/ai-and-efficiency/
[244] https://www.sciencedirect.com/topics/computer-science/sparsity - способност постигане на резултат с по-малко изчисления или изчислителна мощност. Оставяме тези термини изписани на английски, тъй като в българския език не съществува към момента добро решение за обозначаване на термина. Б.а.
[245] Практиката на премахване на параметри от съществуващи невронни мрежи. Б.а.

в сравнение с днешните. Това се превръща в директни икономии: по-малко време на скъп хардуер и по-нисък енергиен разход.

- **По-добри инструменти и автоматизация**: самият процес на разработване на ИИ също преминава през оптимизация. Инструменти като автоматизирано машинно обучение (AutoML) могат да търсят оптимални архитектури и хиперпараметри с по-малко човешка намеса и проби-грешки. Това води до **намаляване на нужния човешки труд и съответно до по-ниски разходи за персонал**. Също така, по-добрите инструменти за разгръщане, мониторинг и поддръжка на модели намаляват оперативните разходи. В следващото десетилетие изграждането на модел с ИИ може да се превърне в по-структуриран инженерен процес – с надеждни рамки, инструменти за отстраняване на грешки и автоматизирани системи, които съкращават работните часове, нужни за проектиране и тестване. Интересен обрат е, че **самият ИИ започва да подпомага разработването на ИИ чрез ИИ асистенти**, които пишат код, оптимизират работни потоци или предлагат подобрения. Това допълнително повишава общата ефективност на разработката и може да направи проектирането на ИИ по-достъпно и с по-голям мащаб.
- **Икономии от мащаба и конкурентен натиск**: с разрастването на пазара за изкуствен интелект се очаква **икономиите от мащаба и конкуренцията да окажат натиск надолу върху цените на много ключови компоненти.** Облачните услуги от големите доставчици вероятно ще стават по-евтини на единица изчисление, тъй като те ще продължат да инвестират в собствена инфраструктура и ще разпределят разходите между многобройни клиенти. Например, ако в момента обучението на определен модел струва на доставчик на облачни услуги 100 000 долара, при по-голям мащаб и ефективност, тази цена може да бъде намалена значително, за да остане конкурентен на пазара. Допълнително, навлизането на нови играчи – производители на AI чипове, облачни платформи и доставчици на модели увеличава конкуренцията и има потенциал да снижи цените. Вече наблюдаваме понижаване на цените за облачни GPU ресурси и въвеждане на отстъпки спрямо нивата от преди няколко години, въпреки че търсенето все още поддържа цените сравнително високи. До 2035 г. услугите с ИИ вероятно ще се превърнат в стока от широк потребителски клас, което обичайно води до значително по-ниски разходи за крайните потребители на такива технологии.

Всички горепосочени фактори сочат към ясно очертаваща се тенденция: **разходите за единица разработка на ИИ (на параметър, на изчисление, на точност и т.н.) ще намаляват значително през следващите десет години**. В същото време обаче трябва да отчетем, че „фронтът" на това, което се опитваме да постигнем с ИИ, също се измества напред. Ако през 2030 г. изследователите създават системи, които са 100 пъти по-сложни от днешните, общите (абсолютни) разходи могат да останат високи – или дори да нарастват – въпреки по-ниските единични цени. Налице е двойна тенденция: относителните разходи на единица капацитет намаляват бързо докато абсолютните бюджети могат да нарастват, ако амбициите растат още по-бързо. Някои анализатори имат оптимистична перспектива – например, ARK Invest прогнозира[246], че обучението на модели, които доскоро струваха милиони, ще се свежда до минимални суми, което би отключило мащабна вълна от внедряване на ИИ. Други

[246] https://www.ark-invest.com/big-ideas-2023/artificial-intelligence#

предупреждават, че може да стигнем предел на скалиране, при който само малък брой играчи ще могат да си позволят следващия „скок" – така наречените сценарий „твърде скъпо, за да бъде постигнато".[247]

С голяма вероятност следващото десетилетие ще бъде белязано от хибриден сценарий: много задачи с ИИ ще станат рутинни и евтини както днес е наемането на сървър в облака за няколко часа или както ползваме компютри и Интернет. В същото време, някои авангардни (frontier) проекти ще останат изключително капиталово интензивни. До 2035 г. обучението на модел със среден размер или фина настройка на голям модел вероятно ще бъде достъпна задача дори за стартъпи или академични лаборатории, благодарение на по-добрите инструменти и евтините изчисления. Паралелно, компании като Google и OpenAI може да продължат да влагат стотици милиони в модели с милиарди параметри, които правят неща извън обсега на обичайния ИИ, като разходите за това ще бъдат оправдани от трансформационния ефект на тези модели.

Успешните организации, правителства и компании ще са онези, които овладеят новите технологии, най-добрите практики и навигират регулациите стратегически. Тези, които не го направят, рискуват да останат с неустойчиви разходи и загуба на конкурентоспособност в глобален мащаб. За държавите доброто позициониране се измерва в устойчив растеж, повишено благосъстояние на населението и генериране на ползи за всички. **Следващото десетилетие ще бъде решаващо за установяване на новото равновесие между амбиция и ефективност – но едно е сигурно: изкуственият интелект ще предлага все повече стойност на все по-достъпна цена.**

6.4. Изкуствен интелект и киберсигурност: взаимодействие, приложения и бъдещи сценарии

6.4.1 Приложения на ИИ в киберсигурността

Кибератаките нанасят сериозни финансови щети на компаниите, като разходите продължават да нарастват през последните години. Според доклада на IBM за 2023 г., средната цена на пробив в сигурността е достигнала 4,45 милиона долара, което представлява увеличение от 2,3% спрямо предходната година и 15,3% увеличение спрямо 2020 г.[248] Дигиталната трансформация и киберсигурността вървят ръка за ръка и без този ключов елемент картината на икономическата ефективност на изкуствения интелект няма да бъде пълна.

Изкуственият интелект все по-активно се прилага за подсилване на киберсигурността през последното десетилетие. От 2020 г. насам има множество доказани случаи на успешно използване на машинно обучение и други техники с ИИ в защитата на мрежи и данни. Основните приложения на ИИ в киберсигурността включват:[249]

- **Откриване на проникванията и аномалии в мрежовия трафик** – системите за откриване на намесите (IDS/IPS) използват алгоритми за машинно обучение, за да анализират огромни обеми мрежови данни и да идентифицират подозрителни модели. Такива ИИ-базирани системи могат да засичат неразрешен достъп, мрежови атаки и необичайно поведение в реално време. Например, платформите за поведенчески анализ на мрежовия трафик (user

[247] https://www.edge-ai-vision.com/2024/09/ai-model-training-cost-have-skyrocketed-by-more-than-4300-since-2020/#
[248] IBM Cost of a Data Breach Report 2023,
[249] https://journalofbigdata.springeropen.com/articles/10.1186/s40537-024-00957-y#

entity and behaviour analytics (UEBA) разпознават аномалии, които биха убягнали на традиционните методи.

- **Откриване и класификация на зловреден софтуер** – дълбоките невронни мрежи се обучават върху огромни масиви от файлове, за да различават легитимен от злонамерен код. Чрез анализ на файлови характеристики, поведение в пясъчника и дори "визуализация" на бинарния код, тези модели постигат висока точност при разпознаване на вируси, рансъмуер и др. Например, системите с ИИ успешно класифицират непознат зловреден софтуер по поведение и характеристики, надминавайки скоростта на традиционните антивируси.
- **Откриване на фишинг и спам** – съвременните филтри за електронна поща и уеб трафик внедряват машинно обучение за идентифициране на фишинг атаки. Модели за обработка на естествен език (NLP) анализират съдържанието на съобщенията и URL адресите, за да откриват измамни имейли и сайтове с висока достоверност. Това помага за предотвратяване на атаки чрез социално инженерство, преди потребителите да станат жертва.
- **Анализ на поведението на потребители и вътрешни заплахи** – ИИ се използва за профилиране на нормалното поведение на потребители в една система и за алармиране при отклонения. Така могат да се хванат вътрешни заплахи (insider threats), например, служител с откраднати пароли и документи или недоволен работник, който извършва необичайни действия. Чрез непрекъснато самообучение тези решения различават злонамерени действия от легитимни, дори когато изходят от вътрешно упълномощени акаунти.
- **Прогнозиране на уязвимости и проактивна защита** – чрез анализ на исторически данни за инциденти и уязвимости моделите могат да предвидят кои системи и части от мрежата са най-вероятно да бъдат атакувани или къде има скрити уязвимости. По този начин е възможно приоритизиране и укрепване на защитата преди настъпване на атака. Предиктивните аналитични модели дори могат да сигнализират за потенциални нови уязвимости (zero-day) на база аномални индикатори, което превръща сигурността от реактивна в проактивна.
- **Автоматизиран отговор при инциденти** – някои съвременни платформи, подпомагани от ИИ, могат автоматично да реагират при откриване на заплаха. Примери за това са изолиране на компрометиран хост, блокиране на подозрителен потребител или IP адрес, или адаптиране на защитната конфигурация в реално време. Подобни автоматизирани реакции значително намаляват времето за отговор и ограничават щетите от инцидента. ИИ спомага и за намаляване на броя на фалшивите аларми, което позволява на екипите по сигурност да се фокусират върху реалните заплахи.

Тези приложения демонстрират, че интегрирането на ИИ в киберсигурността носи реални ползи – по-добро покритие на заплахите, по-бързо откриване на атаките и по-ефективна превенция. ИИ значително подобрява откриването на мрежови проникванията и позволява навременна реакция в реално време, като същевременно намалява процента на фалшивите аларми спрямо традиционните методи. Също така, системите с ИИ могат да се самообучават и адаптират към нови, непознати типове атаки, включително zero-day експлойти, което разширява обхвата на защита. Доказаните случаи от последните години сочат, че ИИ вече се е утвърдил като **необходим съюзник в борбата срещу киберзаплахите.**

6.4.2 Реалистично ли е киберсигурността да бъде изцяло поверена на ИИ?

Въпреки успехите на ИИ в киберсигурността, възниква въпросът доколко реалистично и безопасно е пълното автоматизиране на защитата, тоест поверяване управлението на киберсигурността изцяло на модели с ИИ без човешка намеса. Експертните анализи и изследвания след 2020 г. подчертават, че макар ИИ да е могъщ инструмент, **пълното елиминиране на човешкия фактор крие сериозни рискове и към момента не е препоръчително**.[250] Проучване на Takepoint от 2024 г. сочи, че основните опасения при използване на ИИ в индустриалната киберсигурност са свързани с прекалената зависимост от автоматизирани системи, уязвимост на самите системи с ИИ към манипулация и пропускане на атаки/false negatives.[251]. Данните показват също значими притеснения относно поверителността на данните и липсата на обяснимост на моделите, което подчертава нуждата от предпазливост и човешки контрол.

Специалистите по сигурност са единодушни, че **човешкият фактор остава незаменим** при тълкуването на контекста и вземането на решения въз основа на резултатите, подадени от ИИ. Автоматизираните инструменти могат да сигнализират за аномалия, но само опитен анализатор може да прецени цялостната картина, бизнес контекста и потенциалните последствия. Липсата на човешки надзор крие риск, като ИИ може да пропусне индикатори, които не пасват на обучителните му данни, или да класифицира неправилно ситуация, която изисква нюансирана преценка. Практиката показва, че оптималният подход е **баланс между автоматизация и експертен анализ**, вместо напълно автономна система. Ако една организация разчита сляпо на ИИ и смята, че той е непогрешим, това може да доведе до опасно самодоволство. Свръхразчитането на автоматизирани средства създава *"фалшиво чувство за сигурност"*, при което пропуските не се забелязват навреме. Пример за такъв риск са фалшивите аларми (false positives) – много системи с ИИ генерират огромен брой предупреждения. Ако тези предупреждения не се филтрират правилно, екипите по сигурност могат да изпитат "умора" и да започнат да игнорират или пропускат аларми, сред които може да има и истински инциденти. Същевременно, грешките с пропускане на реална заплаха остават сериозно притеснение, като сред експертите по индустриална киберсигурност това е сред топ опасенията при внедряване на ИИ.[252]

Парадоксално, но системите с ИИ, които защитават мрежите ни, също могат да станат мишена на атака. Областта на *adversarial machine learning*[253] изучава как злонамерени актьори могат да заблудят или манипулират моделите. Доказано е, че малки, специално създадени изменения в данните (напр., в мрежов трафик или файл) могат да накарат един модел да класифицира зловредното поведение като безопасно или обратно. С други думи, противник с достатъчно знания за използвания ИИ може целенасочено да го изиграе, например, да създаде *malware*, който изглежда легитимно за дадения алгоритъм, или да "отрови" данните, с които системата се обучава. През 2023 г. експертен доклад предупреждава, че ИИ системите са уязвими към широк спектър от атаки, различни от класическите уязвимости, и призовава за адаптиране на стандартните киберзащитни процеси, за да покриват и тези нови рискове.[254] Този тип

[250] https://www.isaca.org/resources/news-and-trends/industry-news/2024/overreliance-on-automated-tooling-a-big-cybersecurity-mistake#
[251] https://industrialcyber.co/ai/takepoint-research-80-of-cybersecurity-professionals-favor-ai-benefits-over-evolving-risks/#
[252] Ibid.
[253] Adversarial machine learning (AML) е процес на извличане на информация относно поведението и характеристиките на система за машинно обучение и научаване как тя да се манипулира, за да се получи предпочитан резултат.(Определение на National Cybersecurity Center of Excellence, https://www.nccoe.nist.gov/ai/adversarial-machine-learning)
[254] 2023 Center for Security and Emerging Technology, Stanford Cyber Policy Center https://cset.georgetown.edu/publication/adversarial-machine-learning-and-cybersecurity/#

уязвимости на ИИ налагат допълнителен слой проверки и защити, като например, използване на обясними модели, валидиране на решенията им и тестване на устойчивостта им. Това може да доведе до повишаване на разходите за киберзащита в компаниите и до необходимост от допълнителен човешки ресурс за надзор над системите с ИИ.

Много от съвременните ИИ модели (особено дълбоките невронни мрежи) са "черни кутии", които не дават ясна обосновка защо се стига до дадено решение. В сфера като киберсигурността, където залогът е висок, тази липса на прозрачност затруднява доверието. Ръководителите и одиторите изискват да знаят *защо* системата е блокирала определен трафик или е преценила даден потребител за злонамерен. Необяснимите решения затрудняват приемането на изцяло автоматизирани мерки. Според проучването на Takepoint, цитирано по-горе,[255] едва 28% от организациите са напълно уверени в точността и надеждността на системи за сигурност, базирани на ИИ, докато останалите се колебаят, именно заради липсата на доказана прозрачност и доказателства за ефективност. Това подсказва, че към 2024 г. пълното доверие в автономни ИИ системи още не е изградено в индустрията. На този етап изцяло поверената на ИИ киберсигурност не се счита за реалистична или безопасна стратегия сама по себе си. ИИ е блестящ "силов умножител" – той увеличава скоростта, обхвата и дълбочината на защитните способности, но не е панацея. В най-добрия случай, той трябва да действа ръка за ръка с експертите, като автоматизира рутинните задачи и предоставя препоръки, докато хората вземат финалните решения и поемат отговорността. В противен случай, рисковете – от технически (пропуски, adversarial атаки) до организационни (свръхувереност, липса на доверие) – могат да превишат ползите. **Балансираният подход, комбиниращ силните страни на ИИ (бързина, мащаб, анализ на данни) с човешкия разум (контекст, интуиция, етика), се посочва от експертите като най-надеждния път напред с извличане на най-големи икономически ползи.**[256]

Таблица 2 обобщава някои ключови предимства на ИИ в киберсигурността и съответните им предизвикателства, идентифицирани от изследователите и индустрията:

Ползи от ИИ в киберсигурността	Рискове и предизвикателства
Бързо и мащабируемо откриване на заплахи – анализ на огромни обеми данни и мрежов трафик в реално време, надминаващ човешките възможности. Позволява навременна реакция и намалява пропуските в мониторинга. Това предполага потенциални годишни спестявания от около 112 милиарда долара в световен мащаб. Очаква се доставчиците на киберсигурност да усвоят между 10-50% от тези спестявания, което представлява пазарна възможност от приблизително 34 милиарда долара.[257]	**Прекалена зависимост и самодоволство.** Сляпото доверие в автоматизирани системи може да доведе до фалшиво чувство за сигурност и игнориране на предупреждения. **Предоверяването** крие риск от пропуски и неразбиране на решенията на ИИ.
Адаптиране към нови и непознати атаки. Моделите се самообучават и могат да разпознават новопоявили се заплахи. Това повишава устойчивостта срещу еволюиращи методи на атакуващите.	Злонамерени актьори могат да манипулират входните данни или модела, за да заобиколят ИИ защитата. **Специално оформени атаки** могат да заблудят дори най-добрите модели.

[255] https://industrialcyber.co/ai/takepoint-research-80-of-cybersecurity-professionals-favor-ai-benefits-over-evolving-risks/#
[256] Ibid.
[257] https://www.watchguard.com/wgrd-news/blog/economic-impact-automation-and-artificial-intelligence

Автоматизация и ефективност – ИИ автоматизира рутинни задачи (напр. сортиране на логове, първоначален анализ), спестявайки време и ресурси на екипите. Повишава се продуктивността и се намалява човешката грешка при еднообразни дейности.	**Липса на контекст и нужда от човешка преценка.** ИИ оперира по модели и данни, без разбиране на по-широкия бизнес контекст. Някои сложни решения изискват човешка логика и етика.

Таблица 2: Обобщение на икономическите ползи и рискове от използването на ИИ в сферата на киберсигурността

6.4.3 ИИ като нападател: автономни хакерски атаки без човешка намеса

Докато досега разгледахме ИИ като средство за защита, също толкова важно е да обърнем внимание на тъмната страна – използването на ИИ за извършване на самите кибератаки. Въпросът дали ИИ може самостоятелно да хаква системи вече не е в полето на научната фантастика. През последните години има реални примери и експерименти, демонстриращи подобни възможности.

Още през 2016 г. бе показано нагледно, че компютри могат да хакват компютри без човешка намеса. В рамките на DARPA Cyber Grand Challenge за пръв път в историята няколко ИИ системи се състезаваха помежду си да откриват и експлоатират уязвимости в софтуер, защитаван от други ИИ – напълно автономно, в продължение на 10 часа.[258] Този експеримент, наблюдаван на конференцията DEF CON, демонстрира свръхчовешка скорост и мащаб на атакуване – машините претърсваха и пробиваха уязвимости със скорост и прецизност, непостижими за човека. Експертите посочват случилото се като отрезвяващ поглед към близкото бъдеще, в което ИИ "хакери" могат да намират и използват уязвимости в невероятен мащаб, без да бъдат ограничени от умора или време. Макар това състезание да бе контролирано и в лабораторна среда, то доказа на практика, че възможността за автономни хакерски агенти е напълно осъществима.

Макар все още да говорим за лабораторни условия, такива резултати показват, че в близко бъдеще софтуерни агенти биха могли самостоятелно да провеждат проникванията – от сканиране на портове, през откриване на слабости, до изпълнение на експлойт – без човек директно да решава всяка стъпка. Компании и правителствени организации вече проявяват интерес към автономни инструменти, които да проверяват автоматично сигурността на системите им, използвайки подобни подходи.

Освен чисто техническите, изкуственият интелект започва да показва и умения в областта на **социалното инженерство** – сфера, която традиционно разчита на човешка креативност. През 2023 г. е проведен показателен експеримент, при който голям езиков модел е поел изцяло ролята на фишинг нападател. Изследователи от Харвард създават инструмент, базиран на модела *Claude 3.5*, който автоматизира целия процес – от събиране на публична информация за жертвите до генериране и изпращане на персонализирани фишинг имейли.[259] ИИ сам е проучил целите онлайн (чрез сканиране на социални мрежи, уебсайтове и др.) и след това е написал убедителни имейли, съобразени с профила на всяка жертва. Резултатите са тревожни: автоматично генерираните фишинг съобщения могат да подмамят потребителя да кликне върху злонамерен линк в 54% от случаите, което е сравнимо със или дори по-добро от показателите на фишинг кампании, писани от експерти. Когато човешки оператор леко

[258] https://thedebrief.org/ai-phishing-revolution-study-reveals-bots-now-out-scheme-humans-in-cyber-deception/#
[259] https://thedebrief.org/ai-phishing-revolution-study-reveals-bots-now-out-scheme-humans-in-cyber-deception/#

напътства модела (human-in-the-loop), успеваемостта дори нараства до 56%. Този експеримент доказва, че генеративният ИИ може да извършва сложни социално инженерни атаки в голям мащаб, при това далеч по-бързо и евтино. Един злонамерен бот може да изпрати хиляди персонализирани фишинг имейли там, където на екип от хора биха били нужни дни или седмици. Подобна автоматизация на измамните схеми значително повишава нивото на киберзаплахата и представлява сериозен риск за хората и за бизнеса.

В криминалните среди вече се забелязва тенденция на разработване на **злонамерени ИИ инструменти "по поръчка" на хакери**. През 2023 г. в подземни форуми се появиха т.нар. *malicious LLMs* – езикови модели, подобни на ChatGPT, но без никакви етични ограничения, обучени изцяло за подпомагане на киберпрестъпници. Един пример е ботът WormGPT, създаден на основата на открития модел GPT-J и обучен върху данни за писане на зловреден софтуер.[260] .

Макар засега човекът да остава в центъра на кибератаките, проучванията и експериментите след 2020 г. категорично показват, че изкуственият интелект може да поеме и атакуващата роля. Той може да действа като мултипликатор на силата за злонамерените актьори, повишавайки скоростта, обхвата и съобразителността на атаките до степен, която поставя нови предизвикателства пред защитата. Това налага и защитниците да предвиждат и противодействат на атаки, задвижвани от ИИ, което до скоро не беше част от класическия модел на киберзаплаха.

В икономически измерения **това означава сериозно повишаване на разходите за киберсигурност на правителствата и компаниите от една страна, а от друга добра потенциална възможност за развитие на бизнес, ориентиран към услуги, свързани с минимизиране на рисковете**.

6.5. Потребности от ресурси

За преноса на данни и мрежите от данни се изискват, също така, множество критични и стратегически важни материали и елементи. В *Таблица 3* се виждат както критичните и стратегическите материали, необходими за преноса на данни, така и с кои от тях разполага ЕС и от кои страни-доставчици запълва недостига си:

Елемент	В производството на какво се използва?	Най-големи производители в ЕС и ЕИА[261]	Най-големи световни доставчици
Алуминий	Харддискове Захранващи блокове	Гърция Франция Норвегия	Гвинея Австралия Китай Бразилия
Антимон	Огнеупорни елементи в платките	Няма значимо производство	Китай Таджикистан Русия Турция Австралия
Арсен	Полупроводникови високочестотни и радиочестотни материали	Белгия	Китай Перу Мароко ЕС

[260] https://www.trustwave.com/en-us/resources/blogs/spiderlabs-blog/wormgpt-and-fraudgpt-the-rise-of-malicious-llms/#

[261] EU Raw Materials Information System https://rmis.jrc.ec.europa.eu/

Барий	Кондензатори, платки	**България (72% от производството в ЕС)** Германия	Индия Китай Казахстан Мароко
Ербий	Мощни оптични усилватели	Няма значимо производство	Китай Мианмар Австралия
Бисмут	Компоненти за платки	Белгия **България**	Китай Виетнам Казахстан
Галий	Полупроводници, микрочипове, оптични кабели	Германия Словакия	Китай Япония САЩ Южна Корея
Германий	Усилватели и оптични кабели	Белгия Германия	Китай Канада ЕС
Манган	Вентилатори, антени, харддискове, захранващи блокове	Румъния (от 2016 г. насам) **България (до 2016 г.)**	Южна Африка Габон Австралия
Мед	Кабели, намотки, електронни компоненти	Полша Испания **България**	Чили Перу Мексико САЩ Китай
Никел	Платки, табла, платки за разширяване	Финландия Гърция Полша	Индонезия Филипините Русия Австралия Китай
Редкоземни елементи (REEs)	Харддискове, постоянни магнити	Няма значимо производство	Китай Мианмар САЩ Австралия
Метали от групата на платината (PGMs)	Кондензатори, конектори, полупроводници, платки за харддискове, резистори	Финландия Полша	Южна Африка Русия Зимбабве Канада САЩ
Силиций	Полупроводници, платки, конвертори	Норвегия Франция Германия Испания	Китай Бразилия САЩ
Желязо и стомана	Приемници и трансмитери	Швеция Австрия Германия	Австралия Бразилия Канада Индия
Злато	Конектори, платки, микрочипове, електрически вентилатори	Франция **България** Финландия Швеция	Други държави Китай Русия Канада Австралия
Индий	Усилватели	Франция Белгия	Китай Южна Корея Япония

Калай	Спойки	Испания Португалия	Китай Индонезия Мианмар Перу Боливия
Телур	Оптични кабели	Швеция Белгия Финландия Германия **България**	Китай ЕС Русия
Сребро	Спойки за платки	Полша Швеция Португалия Финландия Испания Румъния	Мексико Китай Перу ЕС
Цинк	Кабели, антени	Швеция Португалия Ирландия Испания Финландия Гърция **България**	Русия Китай Перу Австралия

Таблица 3: Стратегически и критично важни суровини за преноса на данни.
Източник на данни: Joint Research Centre, Европейска комисия.
Графика и анализ: Българска стопанска камара

В червено са посочени критичните суровини, като под критични суровини се разбират такива елементи, които са от критична важност за икономическото развитие на ЕС, като в същото време доставката им е с много висок риск от проблеми и прекъсвания на веригите на доставки. В синьо са посочени стратегическите суровини. Критичните суровини са и стратегически по своята същност. **Ясно е, че амбициите на ЕС за развитие на ИИ няма как да бъдат посрещнати без паралелно развитие и инвестиции в добивната и преработващата индустрия**.

6.5.1 Извличане на ресурси за ИИ чрез оползотворяване на електронни отпадъци

Не без значение за икономическата ефективност е и правилното управление на електронните и електрическите отпадъци. Използването на ИИ и ускорената дигитална трансформация вървят ръка за ръка с използването на все по-нови и модерни устройства и ускорената замяна на електронни и електрически устройства с такива, които позволяват използването на ИИ с редица масови и битови приложения.

Електронните отпадъци (е-отпадъци) представляват една от най-бързо растящите категории отпадъци в света. В световен мащаб са генерирани 62 милиона тона е-отпадъци, което се равнява на средно 7,8 кг на човек годишно. От това количество само 22,3% са били официално събрани и рециклирани по екологосъобразен начин.[262] Европа е регионът с най-високо генериране на е-отпадъци на глава от населението – 17,6 кг на човек през 2022 г. Въпреки това, Европа също така има най-висок процент на официално

[262] Cornelis P. Baldé, Ruediger Kuehr, Tales Yamamoto, Rosie McDonald, Elena D'Angelo, Shahana Althaf, Garam Bel, Otmar Deubzer, Elena Fernandez-Cubillo, Vanessa Forti, Vanessa Gray, Sunil Herat, Shunichi Honda, Giulia Iattoni, Deepali S. Khetriwal, Vittoria Luda di Cortemiglia, Yuliya Lobuntsova, Innocent Nnorom, Noémie Pralat, Michelle Wagner (2024). International Telecommunication Union (ITU) and United Nations Institute for Training and Research (UNITAR). 2024. Global E-waste Monitor 2024. Geneva/Bonn. Pdf version: 978-92-61-38781-5 https://www.itu.int/en/ITU-D/Environment/Pages/Publications/The-Global-E-waste-Monitor-2024.aspx

събрани и рециклирани е-отпадъци – 42,8% или 7,5 кг на човек. Тези данни показват, че въпреки високото ниво на генериране на е-отпадъци, Европа полага значителни усилия за тяхното събиране и рециклиране. Неправилното управление на е-отпадъците води до значителни икономически загуби и екологични проблеми. През 2022 г. глобалните икономически загуби, свързани с е-отпадъците, се оценяват на 37 милиарда щатски долара. Тези загуби включват 78 милиарда долара в екстернализирани разходи за населението и околната среда, произтичащи от емисии на олово и живак, изтичане на пластмаси и принос към глобалното затопляне, особено в случаите, когато опасните вещества не се управляват правилно. Освен това, **е-отпадъците съдържат ценни ресурси като злато, сребро и редкоземни елементи, които са ключови за развитието на ИИ. Неправилното им изхвърляне води до загуба на тези ресурси и увеличава зависимостта от добив на нови материали, което има допълнителни екологични и социални последици**. Ето защо, икономическата ефективност от **все по-широкото използване на ИИ и съпровождащото го ускорено използване на все по-нови и нови електронни устройства е в пряка зависимост от иновациите и намирането на устойчиви екологосъобразни решения и технологии за рециклиране на излезлите от употреба устройства, особено по отношение извличането на ценните елементи, вложени в тяхното производство.**

6.6. Технологичен и цифров суверенитет

Поради самото естество на технологията, въздействието на изкуствения интелект върху икономическата ефективност е много тясно свързано и с пряко отражение върху икономическата сигурност. Европейската комисия посочва изрично изкуствения интелект като критична технология за икономическата сигурност на ЕС.[263] Все по-широкото **внедряване на изкуствен интелект е в пряка корелация с технологичния суверенитет**. В това отношение не само България, но и целият ЕС са силно зависими от американските технологични гиганти. При изострено геополитическо пренареждане и надпревара за ключови ресурси, голяма част от които пряко свързани с развитието на изкуствения интелект, технологичният суверенитет и свързаните с него ресурси като нови материали, енергия, критични минерали и редкоземни метали са от жизненоважно значение за извличане на максимални ползи от технологията.

Известно е, че ЕС изостава значително по отношение на почти всички ключови „технологии на бъдещето"[264] и е силно зависим от трети страни. Наличието на основни технологии е първостепенна грижа на правителствата, тъй като технологичните иновации и способности се считат за основни фактори за икономическата конкурентоспособност и, като цяло – за благосъстоянието на страната. Ето защо **основните потребности са свързани не само с обезпечаване с материални ресурси, но и с капацитет за наука, качествена развойна дейност и иновации, подплътени със съответния функционално грамотен, квалифициран високообразован човешки ресурс**.

Европейската служба за парламентарни изследвания дефинира европейския технологичен суверенитет като „способността на Европа да разработва, осигурява, защитава и запазва критичните технологии, необходими за благосъстоянието на европейските граждани и просперитета на бизнеса, както и способността да действа и взема решения независимо в условията на глобализирана среда".[265]

---

[263] Анекс към Препоръката на Комисията относно критичните технологични области за икономическата сигурност на ЕС за по-нататъшна оценка на риска съвместно с държавите членки, Страсбург 03.10.2023, C(2023) 6689 final

[264] Изкуствен интелект, квантови изчисления, геномно секвениране, космически технологии, микроелектроника, производство на чипове и полупроводници и др. Б.а.

[265] European Parliamentary Research Service, 2021, стр. 3

В Изявление относно технологичния суверенитет[266], Консултативният съвет към Европейския иновационен съвет определя три въпроса, чрез които да се измерва технологичния суверенитет:

1. Разполагаме ли с технологията в Европа?
2. Ако не, имаме ли няколко доставчици от стабилни и надеждни държави?
3. Ако все още не, разполагаме ли с гарантиран и безпрепятствен достъп до монополни или олигополни доставчици от една държава (често САЩ или Китай)?

Отговорите по отношение на изкуствения интелект предпоставят задълбочено стратегическо планиране и последователни мерки и инвестиции, тъй като изглеждат по следния начин:

1. Разполагаме ли с технологията в Европа?	Да,…	но основно чрез технологични гиганти на трети страни. ЕС не разполага с нито един глобален играч от ранга на големите американски и китайски технологични компании.
2. Ако не, имаме ли няколко доставчици от стабилни и надеждни държави?	Не,…	тъй като глобалната динамика и изострената геополитическа надпревара и свързаното с нея пренареждане не предполагат нито една трета държава да бъде считана за стабилен и надежден доставчик в дългосрочен план.
3. Ако все още не, разполагаме ли с гарантиран и безпрепятствен достъп до монополни или олигополни доставчици от една държава (често САЩ или Китай)?	Към момента да,…	но при определени ултимативни условия и все по-силен натиск върху ЕС по отношение на стандартите, регулациите и търговските дефицити (от страна на САЩ) и по отношение на демократичните ценности, свободите на гражданите и човешките права (от страна на Китай и големите мултинационални технологични гиганти).

Отбелязвайки, че „ЕС разчита на чужди държави за над 80% от цифровите продукти, услуги, инфраструктура и интелектуална собственост", докладът на Марио Драги твърди, че ЕС трябва да намали „стратегическите зависимости по отношение на критичните технологии за цифровизацията на европейската икономика". [267]

Доскоро действията на ЕС бяха насочени основно към регулация и антитръстова политика чрез бавния, но ефективен процес на изграждане на т.нар. „цифрова конституция" на ЕС, включваща Общия регламент за защита на данните (GDPR), Закона за цифровите услуги (Digital Services Act), Закона за изкуствения интелект (AI Act), Закона за данните (Data Act), Закона за киберустойчивост (Cyber Resilience Act) и Европейския закон за свободата на медиите (European Media Freedom Act), като списъкът с регулации далеч не е изчерпателен. Новата, по-амбициозна политическа цел за постигане на „технологичен суверенитет" отхвърля представата, че регулацията и нейните глобални ефекти чрез т.нар. „ефект на Брюксел" са достатъчни сами по себе си. Ето защо, може да се твърди, че **ЕС не притежава технологичен и по-конкретно цифров суверенитет, като това се отнася в още по-висока степен за България**.

[266] Анекс към изявлението на Консултативния съвет към EIC Pilot при старта на Европейския иновационен съвет

[267] Марио Драги 2024, https://commission.europa.eu/document/download/97e481fd-2dc3-412d-be4c-f152a8232961_en?filename=The%20future%20of%20European%20competitiveness%20_%20A%20competitiveness%20strategy%20for%20Europe.pdf

В Проект на доклад относно европейския технологичен суверенитет и цифровата инфраструктура (2025/2007(INI)[268] се прави предложение за резолюция на Европейския парламент, като се посочва, че „Европейският съюз е силно зависим от инфраструктура, разработена и контролирана от чужди сили, особено САЩ и Азия, което отслабва неговата конкурентоспособност, излага на риск чувствителните му данни и ограничава способността му за стратегически действия."

6.7. Използване на големи данни

Системите за изкуствен интелект възприемат околната среда чрез събиране на данни, интерпретират събраните структурирани или неструктурирани данни, разсъждават върху придобитите знания или обработват информацията, извлечена от тези данни, и вземат решение за най-доброто действие или действия, с които да постигнат зададената цел. Докато някои аспекти на ИИ като сензори, чипове, роботи, както и определени приложения като автономно управление, логистика, застраховане, финанси или медицински инструменти се отнасят до хардуерни компоненти, значителна част от ИИ е базирана на алгоритми и софтуер. Горивото на изкуствения интелект са големите данни с висок обем, скорост и разнообразие, които изискват специфични технологии и аналитични методи за превръщането им в стойност.

Ефективността на системите за изкуствен интелект пряко зависи от качеството и количеството на данните. Ето защо, **водещите икономики в света и големите глобални технологични компании увеличават инвестициите си в центрове за данни в близост до източници на енергия**, за да поддържат конкурентните си предимства.

С превръщането на „големите данни" („big data") в масова тенденция в много обществени и бизнес сфери, търсенето на капацитет за пренос на данни непрекъснато нараства и се очаква да се задържи на тези високи нива. Бизнес дейностите зависят повече от всякога от бързото предаване на данни. Основен двигател на това търсене на капацитет е продължаващата индустриална трансформация, която се подхранва от интегрирането на технологиите, основани на данни.[269]

Анализът на CBRE Group[270] за глобалните тенденции при центровете за данни за 2024 г. потвърждава нарастващото значение на инфраструктурата, необходима за поддържането на Big Data и ИИ технологии. Инвестициите в центрове за данни и облачни услуги непрекъснато нарастват, като глобалният пазар се очаква да достигне рекордни нива. Това отразява очакванията на бизнеса за значителна възвращаемост от тези технологии и иновации, базирани на ИИ. Напредъкът в изкуствения интелект се очаква значително да стимулира бъдещото търсене на центрове за данни.

Високопроизводителните изчисления ще изискват бързи иновации в дизайна и технологиите на центровете за данни, за да се справят с нарастващите нужди от висока енергийна плътност. Европейският пазар на центрове за данни е нараснал с почти 20% на годишна база през първото тримесечие на 2024 г. Най-значителен ръст е отчетен в четирите основни пазара (Франкфурт, Лондон, Амстердам и Париж), като Париж води с над 40% увеличение. Въпреки това, на континента продължават да съществуват дефицити в предлагането, особено във Франкфурт. Предварителното наемане на нови обекти е често срещано, което подчертава нуждата от допълнителни инвестиции, като ключово предизвикателство остава осигуряването на енергия. Комбинираният процент на свободен капацитет на центровете за данни на пазарите FLAP (Франкфурт, Лондон,

[268] DRAFT REPORT on European technological sovereignty and digital infrastructure (2025/2007(INI)), PE768.180v01-00, 25.02.2025
[269] Supply chain analysis and material demand forecast in strategic technologies and sectors in the EU – A foresight study, JRC, ISSN 1831-9424
[270] https://www.cbre.com/insights/reports/global-data-center-trends-2024

Амстердам, Париж) е спаднал с 2 процентни пункта на годишна база до 10,6% през първото тримесечие на 2024 г. Амстердам отчита най-голямо намаление – почти 8 пункта, достигайки 11,5%. Въпреки очакваното ново предлагане, ниският процент свободен капацитет вероятно ще се задържи поради силното търсене. Цените на европейските центрове за данни продължават да се покачват, заради високото търсене и нарастващите разходи за строителство. Цените за наем на ключови пазари като Франкфурт и Лондон отбелязват постоянен ръст, като във Франкфурт увеличението е по-рязко (15%), отразявайки позицията му на един от най-скъпите пазари в Европа.

Най-големите центрове за данни в Европа срещат сериозни предизвикателства, въпреки че европейските правителства и Европейската комисия разпознават значението на развитието на собствени центрове за големи данни. **На ниво Европейски съюз, обаче, доставчиците на центрове за данни се сблъскват с увеличени регулации, които ограничават зоните, подходящи за ново строителство, и допълнително увеличават и без това високите разходи за изграждане.** Един от четирите най-големи центрове на данни в Европа (Амстердам), например, може да срещне затруднения при привличането на инвестиции за центрове за данни, поради регулаторни ограничения, вкл. национален мораториум върху изграждането на центрове с максимална изчислителна мощност над 70 MW.[271] Доставчиците срещат нарастващи трудности при осигуряването на енергийни източници за центровете за данни, което може да доведе до изместване на инвестициите на големите технологични компании към други региони, поради значителните затруднения в реализацията на нови проекти.

Макар че първоначалните инвестиции в адекватна инфраструктура за съхранение и обработка на големи данни са високи, **дългосрочната възвръщаемост на инвестицията се очаква да надхвърли значително първоначалните разходи, особено с намаляване на цената за съхранение и обработка на данни с течение на времето**. Това означава, че въпреки високите първоначални капиталови вложения, ефективността и ползите от прилагането на ИИ са устойчиви в дългосрочен план.[272]

6.7.1 Данни и базови технологии за развитие на изкуствения интелект

Без данни и облачни технологии, развитието на изкуствения интелект е немислимо. Тъй като по-голямата част от данните се съхраняват и хостват извън територията му, ЕС остава силно зависим по отношение на облачните услуги. Европейският пазар на облачни технологии е доминиран от американски компании: Amazon Web Services, Microsoft Azure и Google Cloud притежават около 69% от пазарния дял на облачната инфраструктура в Европа. Европейските доставчици като OVHcloud и Deutsche Telekom притежават едва 13%. В допълнение, 92% от данните на западните държави се съхраняват в САЩ в инфраструктура, притежавана и управлявана от американски компании. Европа не разполага с високотехнологични заводи, способни да произвеждат модерни полупроводници (<10 nm). Европа произвежда само 10% от полупроводниците в световен мащаб, много под 54%, произвеждани в Тайван (главно от TSMC), и 16%, произвеждани в Китай. Тази зависимост излага ЕС на геополитическо напрежение и рискове от прекъсване на доставките.[273]

[271] Ibid.
[272] Ibid.
[273] Ibid.

6.8. Енергия

Използването на големи данни и тренирането на моделите на ИИ изискват огромни количества енергия, свързани с обезпечаването на центровете за данни. *Таблица 4* очертава тенденциите, свързани с енергийните потребности и потребление на центровете за данни и разработката на модели с ИИ. Тенденциите са очертани от Международната енергийна агенция, като трябва да се подчертае, че прогнозите относно енергийното потребление на центровете за данни се отличават с висока степен на несигурност, поради непълни данни към момента, ефекти от използването на ИИ (като например напредък в изследванията и развойната дейност, свързана с нови източници на енергия) и скокове в търсенето. Според данни на Международната енергийна агенция (IEA), потреблението на електроенергия от центрове за данни е останало относително стабилно през периода 2010–2022 г., въпреки значителния ръст в интернет трафика и изчислителните натоварвания. Това се дължи основно на подобрения в енергийната ефективност и преминаването към по-ефективни облачни и *hyperscale* центрове за данни.

Тенденции в енергийните потребности и потребление на центрове за данни и ИИ	
Брой центрове за данни (2024)	Над **11 000** в световен мащаб
Годишен ръст на сървърите (2010–2020)	**~4%** годишно
Потребление на електроенергия (2022)	Между **240 и 340 TWh** (около **1%–1.3%** от глобалното електропотребление)
Инвестиции в AI стартъпи (2018–2023)	Над **225 млрд. USD**
Инвестиции в стартъпи за чиста енергия (2018-2023)	**143 млрд. USD**
Капиталови разходи на Alphabet, Meta, Amazon, Microsoft (2023)	**~150 млрд. USD** (повече от двоен ръст спрямо 2019 г.)
Ново поколение чипове (NVIDIA Blackwell)	**~75%** по-ниска консумация на енергия спрямо предходни GPU

Таблица 4: Тенденции в енергийните потребности и потребление на центрове за данни и ИИ, Данни: International Energy Agency[274]

Основните предизвикателства са свързани с растежа без пълна компенсация от ефективност, като въпреки големите технологични подобрения (напр., нови GPU), търсенето расте по-бързо от спестената енергия. Реални бариери пред растежа има и по отношение веригите на доставки, свързани с чипове, както и по отношение на мрежовия капацитет и забавени разрешителни за строеж.

В началото на 2024 г. в света са регистрирани над 11 000 центъра за данни. Броят на инсталираните сървъри се е увеличавал с около 4% годишно в периода 2010–2020 г., но търсенето на електроенергия от страна на центровете за данни като цяло се е задържало на едно ниво, тъй като нарастващите нужди от информационни технологии са били компенсирани от подобрения в ефективността на охлаждането и от преместване на работни натоварвания към по-големи и по-ефективни облачни центрове за данни. През последните години, обаче, потреблението на електроенергия отново започва да расте, тъй като подобренията в енергийната ефективност вече не успяват да компенсират бурното нарастване на търсенето на услуги от центровете за данни. През 2022 г. потреблението на електроенергия от центровете за данни се оценява на между 240 и 340 тераватчаса (TWh), или около 1% до 1.3% от общото световно потребление на електроенергия (изключвайки мрежите за данни и копаенето на криптовалути).[275]

[274] World Energy Outlook 2024 Report https://www.iea.org/reports/world-energy-outlook-2024
[275] Ibid.

Въпреки че **ИИ понастоящем заема относително малък дял от глобалното електропотребление на центровете за данни, той се очертава като нов двигател на растежа**. Инвестициите в ИИ и нови центрове за данни бележат истински бум. През последните пет години глобалният рисков капитал е инвестирал над 225 милиарда щатски долара в стартиращи компании в сферата на ИИ, в сравнение със 143 милиарда щатски долара в стартиращи компании, работещи в различни области на чистата енергия.

Трябва да се има предвид и фактът, че ИИ има двоен ефект върху енергийното потребление. **Едновременно с натоварването на инфраструктурата, ИИ може да я оптимизира**, което зависи в голяма степен от алгоритмите и приложенията му. Ето защо, не могат да се направят точни прогнози относно енергийната ефективност, свързана с развитието на ИИ. Възможни са множество сценарии, всеки един от които с достатъчно висока степен на вероятност да се случи.

6.8.1 Потребление на енергия и съображения за устойчивост

Създаването и използването на модели с изкуствен интелект изисква значителни количества енергия, което оказва влияние както върху разходите, така и върху околната среда. През последните години енергията, необходима за обучението на големи модели, се превърна в предмет на сериозна загриженост. Обучението на един единствен голям модел може да изразходва огромно количество електроенергия. Оценява се, че обучението на GPT-3 през 2020г. (със 175 милиарда параметри) е изисквало около 1287 мегаватчаса (MWh) електроенергия, което се равнява на приблизително 502 метрични тона CO_2 емисии (еквивалентно на това 112 бензинови автомобила да изминат пробег за една година).[276] По-нови модели с още повече параметри вероятно използват равни или по-големи количества енергия. OpenAI не е публикувала данни за потреблението на енергия при GPT-4, но се предполага, че е по-високо, поради по-дългото обучение. Освен това, за разлика от еднократните разходи за производство, **енергийните разходи се повтарят при всяко ново обучение или експеримент**, така че при всяка итерация на даден модел потреблението на електроенергия се натрупва.

През последните три години се наблюдава засилващо се осъзнаване на този проблем. Разработчиците на ИИ все по-често си задават не само въпроса „Можем ли да обучим този модел?", но и „Колко ще струва това в електроенергия и въглеродни емисии?". Важно е да се отбележи, че дори **след като един модел бъде обучен, енергийната консумация продължава при внедряването му**. Изпълнението на модела за милиони потребителски заявки (свързано с т.нар. *inference*[277]) всъщност може да изразходва повече енергия в дългосрочен план, отколкото самото обучение. Google изчислява, че *inference* представлява около 60% от общата енергийна консумация на ИИ, спрямо 40% за обучението.[278] С ръста на услуги, базирани на ИИ (като чатботове, преводачи, препоръчващи системи) търсенето на електроенергия в центровете за данни ще нараства. Например, всяка заявка в ChatGPT, макар и малка сама по себе си, използва значително повече изчислителна мощ от обикновено търсене в интернет, а **една заявка към ИИ може да бъде до 100 пъти по-енергоемка от търсене в Google**. Това, умножено по милиарди заявки, дава ясна представа защо операторите на центрове за данни се подготвят с допълнителен капацитет. В момента центровете за данни в световен мащаб отговарят за около 2.5%–3.7% от глобалните емисии на парникови газове, а ИИ е все по-голям дял от техния товар.[279] Накратко, **енергийният отпечатък**

[276] Carbon Emissions and Large Neural Network Training, https://arxiv.org/ftp/arxiv/papers/2104/2104.10350.pdf достъпено през https://news.climate.columbia.edu/2023/06/09/ais-growing-carbon-footprint/#
[277] Сложен процес при обучението и ползването на модел с ИИ върху дадена база данни с множество заявки, докато моделът не започне да дава точни отговори и да прави изводи. Б.а.
[278] Ibid.
[279] Ibid.

на ИИ е значителен и тенденцията е за неговото нарастване, едновременно с нарастването на използването на ИИ.

От гледна точка на разходите, консумацията на енергия се превръща в електрически сметки за компаниите, които поддържат собствен хардуер, или е включена в цената на облачните услуги. В повечето големи проекти с ИИ разходите за електроенергия са относително малка част спрямо амортизацията на хардуера или разходите за персонал, но не са незначителни, като се оценяват между 2 и 6% от общата цена за обучение на съвременни големи модели.[280] По-важното е, че **устойчивостта става приоритет и много компании поемат ангажименти за въглеродна неутралност**, така че енергийната ефективност на моделите е от ключово значение както за изпълнение на екологични цели, така и за избягване на разходи по въглеродни компенсации или глоби.

През последните години се наблюдават инициативи като „Green AI"[281], чиято цел е да се докладва и намалява въглеродният отпечатък на ИИ. Някои компании вече местят обучението на ИИ в региони с по-чиста енергия или по време на деня, когато възобновяемите източници са в изобилие.

През следващото десетилетие **енергийната ефективност ще се превърне в ключов критерий при проектирането на модели с ИИ и хардуер**. Всяко ново поколение чипове се стреми към по-висока производителност на ват. GPU и ИИ ускорителите в центровете за данни все повече включват механизми за намаляване на топлинни загуби и консумация в покой. Разработват се и иновативни архитектури като аналогови или оптични изчисления за ИИ, които обещават значителна икономия на енергия, макар че все още са в експериментален етап. Изследователите разработват и оптимизационни методи за алгоритмите, които да намалят консумацията на енергия. Изследване от 2023 г. на Университета на Мичиган представя метод, който може да намали енергията за обучение с до 75% чрез интелигентно разпределение на ресурсите, с минимален ефект върху времето за обучение.[282] По отношение на моделите, подходи като *model pruning*, квантизация и ефективни архитектури позволяват постигане на същата точност с по-малко изчисления и по-ниска енергийна консумация.

Съществуват няколко ключови показатели и тенденции, свързани с енергийната ефективност и топлинните загуби в центровете за данни. Коефицинтът за ефективност на използване на енергията (Power Usage Effectiveness, PUE) е основен показател, изчисляван чрез съотношението между общата консумирана енергия и енергията, използвана от ИТ оборудването. Идеалната стойност е 1.0, което би означавало, че цялата енергия се използва директно за ИТ оборудването. Според проучване на Uptime Institute[283] средната годишна PUE на глобално ниво е 1.58. Въпреки че този показател се е подобрил през последните години, той е останал относително стабилен от 2018 г. до сега. Една от причините е, че значителна част от изследваните центрове за данни са на възраст над 11 години. Тенденцията за новоизградените центрове е за постигане на значително подобрени стойности на показателя. Големите облачни компании, като Google, Amazon и Microsoft, твърдят, че някои от техните центрове за данни имат PUE от 1.2 или по-малко. Но това не винаги отразява реалното потребление на енергия от клиентските приложения в облака. Причините са, че те могат да се хостват в центрове с по-висок PUE или да се репликират в няколко зони, което увеличава общото потребление на енергия.

[280] https://www.edge-ai-vision.com/2024/09/ai-model-training-cost-have-skyrocketed-by-more-than-4300-since-2020/#
[281] https://carboncredits.com/green-ai-explained-fueling-innovation-with-a-smaller-carbon-footprint/
[282] https://news.umich.edu/optimization-could-cut-the-carbon-footprint-of-ai-training-by-up-to-75/#
[283] https://journal.uptimeinstitute.com/large-data-centers-are-mostly-more-efficient-analysis-confirms/

Нарастват възможностите PUE да се превърне в задължителен показател, тъй като той постепенн се включва в законодателството. Например, новият Закон за енергийната ефективност в Германия изисква центровете за данни да постигнат PUE от 1.5 от 1 юли 2027 г. и 1.3 от 1 юли 2030 г. Изискванията са още по-високи за новите центрове, които ще бъдат отворени след 1 юли 2026 г. – те трябва да имат PUE от 1,2 или по-малко, което може да се окажже трудно постижимо.

Друг съществен показател, характеризиращ енергийната ефективност на центровете за данни, е консумация на енергия за охлаждане. Охлаждането е вторият по големина консуматор на енергия в центровете за данни след сървърите. В стандартните въздушно охлаждани центрове за данни енергията за охлаждане може да представлява до 45% от общата консумация на енергия. Не се оползотворяват напълно възможностите за използване на отпадната топлина от центровете за данни като източник на енергия за отоплени на сгради, включително битови, в близост до центровете, за промишлени процеси и в селскостопански оранжерии. От регулаторна гледна точка, въвеждането на въглеродни данъци или системи за търговия с емисии би стимулирало компаниите да намаляват енергийното потребление на ИИ. Това, обаче, не е най-ефективният механизъм за въздействие. Икономическата логика предполага индустрията сама да предприема стъпки в посока намаляване на потреблението и разходите за енергия и тя вече го прави. Може да се появят и сертификати или стандарти за енергийна ефективност на продукти с ИИ. **Устойчивостта вероятно ще влияе върху решенията относно проекти с ИИ -** например, избор да се тренира по-малък модел, ако това води до многократно по-малко енергийно потребление. Големите лаборатории за ИИ вече търсят компромис между размер на модел, необходима изчислителна мощ и точност, което може да доведе до иновативни решения за по-ефективно използване на всеки киловатчас енергия. С нарастващото преминаване на натоварванията на ИИ към облачни центрове за данни, има възможност тези изчисления да се концентрират в съоръжения, захранвани със 100% възобновяема енергия. Google и Microsoft вече инвестират в такива центрове. През 2024 г. пейзажът на източниците на енергия за центровете за данни претърпява значителна трансформация, като много технологични компании (Amazon [284], Meta [285], Google [286], Microsoft [287] [288]) се обръщат или търсят възможности в ядрената енергия [289] като източник без въглеродни емисии, за да посрещнат нарастващите си нужди от електроенергия. Макар енергийните разходи да остават умерен компонент в бюджетите, значението им бързо расте заради климатичното въздействие. До 2035 г. се очаква развитието на ИИ да бъде значително по-„зелено", като ще се измерва в енергия на обучение или на единична заявка, като **самата индустрия ще натиска референтните стойности все по-надолу година след година**.

6.9. Права на интелектуална собственост

Управлението на интелектуалната собственост е един от най-важните инструменти на бизнеса за придобиване на дългосрочни конкурентни предимства, възвръщаемост на инвестициите и повишаване на икономическата ефективност. Използването на ИИ налага преструктуриране и пренастройване на съществуващите международни правила

[284] https://sustainability.aboutamazon.com/climate-solutions/carbon-free-energy Retrived on 03 April 2025
[285]Accelerating the Next Wave of Nuclear to Power AI Innovation, 3 December 2024, https://sustainability.atmeta.com/blog/2024/12/03/accelerating-the-next-wave-of-nuclear-to-power-ai-innovation/
[286] https://blog.google/outreach-initiatives/sustainability/google-kairos-power-nuclear-energy-agreement/ Retrived on 3 April 2025
[287] https://www.microsoft.com/en-us/microsoft-cloud/blog/2024/09/20/accelerating-the-addition-of-carbon-free-energy-an-update-on-progress/?msockid=323e7dd3d8f462c42425688cd96d6336 Retrived on 3 April 2025
[288] "Accelerating a Carbon-Free Future: Microsoft policy brief on Advanced Nuclear and Fusion Energy" 2023,
[289] Goldman Sachs, 23 January 2025, https://www.goldmansachs.com/insights/articles/is-nuclear-energy-the-answer-to-ai-data-centers-power-consumption

за защита на интелектуалната собственост. Към 2019 г., съгласно международния консенсус, ИИ не може да притежава права на интелектуална собственост. Към момента защитата на търговски марки и патенти не е под съществен натиск. Не същото важи, обаче, за авторското право, което поставя редица въпроси пред бизнеса, особено в областта на креативните индустрии и ИКТ сектора[290]:

1. Какви са параметрите на един справедлив пазар и модели на монетизация на авторски данни (книги, статии, публикации, снимки, авторски видеа, дизайн) при тренирането на моделите на ИИ?
2. Чия собственост са изображения, софтуерен код, аудио и видео, генерирани от ИИ по задание (*prompt*)?
3. Промптовете обект ли са на защита и авторско право?

Решаването на тези въпроси е ключово за (пре)насочването на сериозни парични потоци в зависимост от приоритизирането на творците и малките компании или на глобалните тех-гиганти.

Световната организация за интелектуална собственост (WIPO) инициира глобален разговор относно интелектуалната собственост и изкуствения интелект, където поставя редица въпроси за обсъждане, включващи аспекти, свързани със собственост на изобретенията, патентоспособност, авторско право и собственост, нарушения и изключения, авторско право при случаи на дълбока фалшификация, възникване на нови права на интелектуална собственост, свързани с данни и собственост върху данни, права на интелектуална собственост при компютърно асистиран дизайн, автономно генерирани дизайни от ИИ и др.[291]

Докато световният дебат тече, вече отдавна е пределно ясно, че въздействието на законодателството в областта на защита на правата на интелектуална собственост е от първостепенно значение за икономическата ефективност от разработването и внедряването на ИИ.

ИИ радикално променя начина, по който се създават творби и изобретения, като значително ускорява процесите на иновации. От икономическа гледна точка, ясната правна регламентация на интелектуалната собственост при използването на ИИ би стимулирала инвестициите и би дала по-голяма сигурност на бизнеса. Ясно дефинираните права върху продукти, генерирани с ИИ, биха намалили риска от съдебни спорове и биха ускорили внедряването на иновациите, което от своя страна би повишило общата икономическа ефективност. Европейската служба за интелектуална собственост (EUIPO) подчертава необходимостта от гъвкав подход, който ясно разграничава човешкия от технологичния принос в процеса на създаване на интелектуални продукти.[292]

Промените в правната рамка за защита на интелектуалната собственост ще даде най-голямо отражение върху следните сектори:

- *Технологичния сектор*: високотехнологичните компании, използващи активно ИИ, ще бъдат пряко засегнати. Новите регулации могат да създадат по-благоприятна среда за иновации и ускорено внедряване на технологии. В същото време, липсата на ясни правила може да доведе до забавяне на разработката на нови продукти.
- *Фармацевтичния и биотехнологичния сектор*: ИИ се използва интензивно за ускоряване на научните открития и патентни разработки. По-ясната регулация би позволила по-бързото комерсиализиране на иновациите, което би довело до

[290] В ЕС софтуерните продукти не се защитават с патент, а са обект на авторско право. Б.а.
[291] https://www.wipo.int/about-ip/en/artificial_intelligence/policy.html
[292] https://www.wipo.int/export/sites/www/about-ip/en/artificial_intelligence/call_for_comments/pdf/org_european_union_euipo.pdf

значителни икономически ползи, по-кратък цикъл на разработка на лекарства и по-ниски разходи.

- *Творческите индустрии*: авторските права върху творби, генерирани с ИИ, са едно от основните предизвикателства. Решаването на този въпрос би стимулирало индустриите на изкуството, киното, музиката и литературата, където потенциалът на ИИ да генерира съдържание е огромен, но в момента е правно неопределен.

В момента са възможни три основни сценария за развитие, всеки от които дава пряко отражение и има огромни последствия върху икономическата ефективност от използването на ИИ:

1. Непризнаване на права на ИИ

В този случай интелектуалните права се присъждат само на хората или компаниите, създали или управляващи ИИ системата. Този сценарий е най-близък до настоящата правна рамка, но може да намали стимулите за инвестиции и иновации в ИИ.

2. Ограничено признаване на права

При този сценарий правата на интелектуална собственост се дават на разработчиците или компаниите, стоящи зад ИИ, като ясно се разграничават техните права от автономния принос на самата технология. Подобен подход би бил прагматичен и икономически ефективен, стимулирайки развитието на технологии, без да създава значителни правни противоречия. Този сценарий може да ощети творческите индустрии, ако техни авторски творби са използвани за обучение на моделите. Това би създало ситуация, в която създателите на оригиналните произведения не получават справедлив дял от икономическите ползи, генерирани от ИИ. В такъв случай е необходимо да се разработят механизми за компенсация или лицензиране, които да гарантират справедливо разпределение на икономическите ползи между авторите и технологичните компании, за да не се обезсърчават творческите индустрии.

Ефективни механизми за компенсация могат да включват:

1. *Колективно лицензиране*: авторите и правоносителите да се обединят в организации за колективно управление на права, които да сключват лицензионни споразумения с компании, разработващи ИИ. Позволява на авторите да запазят контрол върху използването на техните творби и да сключват споразумения при благоприятни условия. Колективните организации могат да договарят условия от името на множество автори, което улеснява процеса. Гарантира, че правоносителите получават компенсация според степента на използване на техните произведения. Недостатък е, че такова договаряне на лицензи с множество разработчици на ИИ може да бъде бавен процес. Възможно е известни творци да получават по-голям дял от средствата, докато по-малки автори остават недооценени. Ако компаниите откажат да лицензират съдържанието, могат да възникнат съдебни дела, които ще забавят развитието на ИИ.
2. *Роялти схеми (възнаграждение за авторско право)*: въвеждане на система, която отчита използването на защитени произведения в тренировъчните набори и автоматично изплаща пропорционално възнаграждение на авторите. Благодарение на технологии като блокчейн и изкуствен интелект, роялти плащанията могат да се изчисляват и изплащат автоматично. Гарантира се, че творците ще бъдат мотивирани да продължават да създават съдържание, което може да бъде използвано в обучителни модели. Ако се използват алгоритми за проследяване, авторите ще могат да виждат как техните творби са били използвани. Този вариант ще изисква разработване на сложни системи за

мониторинг и разпределение на плащанията, което може да повиши разходите. Освен това, не винаги е ясно каква част от успеха на даден ИИ модел се дължи на конкретна обучителна база. Не е за подценяване и съпротивата на големите технологични компании, които могат да твърдят, че използването на публично достъпно съдържание за обучение не е нарушение на правата на интелектуална собственост.

3. *Фондове за компенсация*: създаване на специални фондове, в които компаниите, развиващи ИИ, плащат определен процент от печалбите си, разпределяни сред авторите според приноса на техните творби към обучението на моделите. Събирането на средства в общ фонд гарантира, че авторите ще получат компенсация, без да зависят от индивидуални сделки с технологични компании. Компаниите могат да внасят фиксиран процент от приходите си в общ фонд, който се разпределя между правоносителите. този подход би осигурил баланс между нуждите на творците и на индустрията, разработваща ИИ, без да възпрепятства иновациите. Трудностите на подхода е, че липсва обективна методология за определяне на кои автори каква компенсация да бъде изплатена. Ако вноските от компаниите са малки, това може да не покрие реалните загуби на авторите. Управлението на такъв фонд ще изисква значителен бюрократичен апарат, което може да намали ефективността на системата.

Най-ефективен би бил хибриден модел, който съчетава колективното лицензиране за големите търговски модели на ИИ, роялти схеми за автоматично плащане на индивидуални автори и фондове за компенсация за случаите, в които директното лицензиране е трудно или невъзможно. Това ще осигури баланс между правата на творците и нуждите на технологичните компании, като същевременно ще стимулира иновациите и ще оптимизира икономическите ползи от разработването и внедряването на ИИ.

3. Пълно признаване на права на ИИ

Този сценарий, при който ИИ се признава като автор или изобретател, би отворил нови правни възможности, но също така поставя фундаментални етични и философски въпроси. Макар да изглежда иновативен, той би могъл да създаде значителни правни и административни усложнения, намалявайки икономическата ефективност в краткосрочен план, но отваряйки неограничени възможности в дългосрочен план.

Ясната регулаторна рамка би могла значително да увеличи производителността и конкурентоспособността на бизнеса, като стимулира инвестициите в ИИ. Икономическите ползи включват намаляване на несигурността и риска за компаниите, което би ускорило инвестициите и иновациите. Биха се намалили и разходите за съдебни спорове и повишаване ефективността на административните процеси, свързани с управление на интелектуалната собственост и иновациите в предприятията. Не без значение е и подобряването на международната търговия с интелектуални продукти и засилване на позициите на ЕС на глобалния пазар. От друга страна, липсата на яснота или прекалено рестриктивна рамка може да ограничи икономическите ползи от ИИ, забавяйки развитието на иновации и увеличавайки разходите за бизнеса.

6.10. Технологичен напредък и иновации с помощ от ИИ – нови индустрии, нови бизнес модели, трансформация на съществуващи бизнес модели

Изкуственият интелект бързо се превръща в движеща сила за технологичен напредък и бизнес иновации. Според доклад на McKinsey от 2018 г. ИИ може да допринесе с около 13 трилиона щатски долара към световната икономика до 2030 г.[293], което подчертава мащаба на влиянието му. Компаниите все по-често интегрират ИИ в основните си дейности, създавайки нови индустрии и бизнес модели, както и трансформирайки съществуващите модели на работа.

6.10.1 Нови индустрии, задвижвани от изкуствен интелект

ИИ вече подпомага раждането на нови индустрии и нововъзникващи индустриални сектори. Примери за такива са:

- **Индустрия на анотацията на данни:** поради нуждата от огромни обеми *обучаващи данни* за модели с ИИ, се появи нов сектор, фокусиран върху събиране, етикетиране и анотиране на данни. Множество фирми предлагат услуги по анотиране на изображения, текст и видео за машинно обучение на алгоритми. Този сектор бележи бурен ръст. През 2022 г. глобалният пазар на анотирането на данни *(data annotation)* достига около 805 млн. USD, а прогнозите сочат експлозивен растеж до над 13,7 млрд. USD през 2030 г.[294]. Бумът на тази индустрия се движи от неутолимия апетит на ИИ и машинното обучение за висококачествени несинтетични данни. Както вече споменахме, **данните се превърнаха в новия петрол**, а бизнеси като платформи за *crowdsourcing* на анотации и автоматизирани инструменти за етикетиране вече се обособяват като самостоятелна индустрия.
- **Генеративни медии и синтетично съдържание:** напредъкът в генеративния ИИ с модели като GPT-4 и много други доведе до възникването на индустрия, занимаваща се със създаване на синтетично съдържание – текстове, изображения, видео, музика, създадени от ИИ. Тази нова индустрия обхваща компании, предлагащи инструменти с ИИ за творчески сектори (маркетинг, дизайн, развлекателна индустрия), както и студия за синтетични медии. Очаква се пазарът на генеративен ИИ да нарасне от около 20–25 млрд. щатски долара през 2024–2025 г. до над 136 млрд. щатски долара към 2030 г., с годишен ръст от около 36–37%.[295] Така наречената „синтетична реалност" вече привлича предприемачи с нови възможности в развлекателния и рекламния бизнес.[296] Пример за продукт на тази индустрия са платформи с ИИ, които автоматично генерират маркетингово съдържание или персонализирани видеоклипове. Появяват се дори изцяло виртуални инфлуенсъри и нови форми на дигитално изкуство, възможни единствено благодарение на ИИ.
- **Автономни превозни средства и роботи:** ИИ е в основата на самоуправляемите автомобили, дронове и интелигентни роботи, което формира нови сектори в транспортната и производствената индустрия. Индустрията на автономните превозни средства привлича гиганти и множество стартъпи, работещи върху ИИ за самоуправление, логистика и умни градове.

293 https://www.mckinsey.com/~/media/McKinsey/Featured%20Insights/Artificial%20Intelligence/Notes%20from%20the%20frontier%20Modeling%20the%20impact%20of%20AI%20on%20the%20world%20economy/MGI-Notes-from-the-AI-frontier-Modeling-the-impact-of-AI-on-the-world-economy-September-2018

294 https://www.remotelabeler.com/data-labeling-trends-in-2024-2030/

295 https://www.marketsandmarkets.com/Market-Reports/generative-ai-market-142870584.html

296 https://www.bvp.com/atlas/roadmap-the-rise-of-synthetic-media#

Очаква се глобалният пазар на автономни автомобили да достигне около 214 млрд. щатски долара към 2030 г.[297], докато през 2023 г. е бил под 1 млрд., което е показателно за темповете на растеж. Този нов сектор преобразява традиционния автомобилен бранш, като измества фокуса от притежание на коли към „**мобилност като услуга**" (услугата за споделено пътуване с автономни автомобили). Наред с това, ИИ задвижва и роботиката в производството (индустриални роботи с компютърно зрение, колаборативни роботи и др.), както и дронове за доставки и наблюдение. Всички те образуват нови ниши, които допреди десетилетие не съществуваха или не бяха масови.

- **Индустрия на хардуера за ИИ (специализирани чипове):** изискванията на алгоритмите с ИИ стимулираха цял нов пазар за високопроизводителни чипове и хардуерни ускорители, оптимизирани за машинно обучение и дълбоко учене. Графичните процесори (GPU) на NVIDIA станаха критична част от бума на ИИ, а се появиха и специализирани ИИ процесори – например TPU на Google, чипове от стартъпи като Graphcore, Cerebras и др. Това е нов сегмент в полупроводниковата индустрия, който расте бързо. Оценките показват, че пазарът на ИИ чипове (специализиран хардуер за ИИ) може да надхвърли 200–300 млрд. щатски долара към 2030 г., при годишен ръст от около 30%.[298] Традиционните производители (Intel, Qualcomm и др.) трансформират стратегиите си, за да се позиционират в тази нова ниша. Тази индустрия е ключова за поддържане на напредъка. Компаниите масово инвестират в *edge AI* чипове *(*за *edge computing*[299], позволяващи ИИ изчисления на самото устройство), което прави възможни приложения като ИИ в смартфони, IoT устройства и автомобили.
- Екосистемата около ИИ ражда и **нишови сектори**. Например, сферата на ИИ в здравеопазването позволява на десетки биотехнологични стартъпи да използват ИИ за откриване на лекарства и диагностика, оформяйки нов клон на фарма индустрията (т.нар. *TechBio* компании). Също така, на хоризонта се появи пазар за модели и решения с ИИ – своеобразни **магазини за ИИ**, където се обменят и продават готови модели за машинно обучение и услуги. *Global AI Marketplace* моделът има за цел да предостави общ пазар за ИИ решения от различни разработчици.[300] Пример е и платформата SingularityNET[301], която се опитва да изгради **децентрализиран пазар за ИИ алгоритми**.
- Друга ниша са **консултантските и обучителни услуги по ИИ**. Това е цяла индустрия от експерти и фирми, които помагат на традиционните бизнеси да внедрят ИИ, да обучат персонала си за работа с ИИ инструменти и да управляват етичните и правните аспекти. Налице са и иновативни услуги като доставчици на синтетични данни за обучение на модели, когато реални данни липсват или са чувствителни, което също се превръща в бизнес ниша.[302] Не на последно място, дори в сферата на законодателството и етиката се формират нови специализации – компании, предлагащи **ИИ одити, оценка на алгоритми за пристрастия и съответствие с регулациите**, отговарящи на нарастващата нужда от доверие в системите с ИИ.

Тези примери на нови индустрии демонстрират как ИИ действа като **генератор на икономическа стойност извън рамките на съществуващите отрасли**. Той не само

[297] https://www.grandviewresearch.com/press-release/global-autonomous-vehicles-market#
[298] https://techstrong.ai/articles/global-ai-chip-market-worth-305-billion-by-2030/#
[299] Edge computing (граничните изчисления) е разпределяща изчислителна рамка, която доближава бизнес приложенията до източниците на данни.
[300] https://desklib.com/study-documents/ai-talent-change-management/
[301] https://singularitynet.io/
[302] https://www.datasciencesociety.net/latest-data-annotation-trends-whats-transforming-ai-models-in-2024/#

усъвършенства процеси, но и отваря изцяло нови възможности за бизнес, което води до възникване на нови продукти, услуги и пазари.

6.10.2 Нови бизнес модели, появяващи се благодарение на ИИ

Редом с новите индустрии, ИИ стимулира и развитието на нови начини, по които компаниите създават, доставят и улавят стойност. Много от тези нови бизнес модели преосмислят традиционните принципи на генериране на приходи, взаимодействие с клиенти и изграждане на продукти. Ето някои от най-важните нови бизнес модели в ерата на ИИ:

- **„ИИ като услуга" и абонаментни модели (AI-as-a-Service, AIaaS):** По аналогия със софтуера като услуга (SaaS), множество компании предлагат възможности в областта на ИИ чрез абонамент или облачна услуга. Този модел осигурява повтарящи се приходи и мащабируемост – клиентите плащат редовна такса, за да ползват функция с ИИ, вместо еднократна покупка. По същество, организациите „наемат" способностите на ИИ според нуждите си, без да инвестират в изграждане на собствена инфраструктура.
- Големите облачни платформи (AWS, Azure, Google Cloud) също популяризират **ML-as-a-Service (MLaaS)** – клиентът плаща за обработени изображения от модел на ИИ или за използване на NLP[303] услуга на брой заявки. Така AIaaS демократизира достъпа до ИИ. Дори малки фирми могат да внедрят интелигентни функции, плащайки само за потреблението си, подобно на комунална услуга. Абонаментният модел гарантира стабилен приход за доставчиците и постоянни подобрения на услугата за клиента. Според множество анализатори, услуги тип „всичко като услуга" (XaaS), включително ИИ, ще се превърнат в сърцевина на бизнес моделите с ИИ.
- **Модел с обратна връзка/данни (Feedback Loop Model):** При този модел **потребителските данни и обратна връзка се превръщат в ключов актив**. Компанията може умишлено да предложи продукт с ИИ на много ниска цена или дори безплатно, за да привлече потребители и да събира данни от тяхното поведение. Тези данни *„затварят цикъла"*, като се използват за дообучаване и подобрение на ИИ модела, което впоследствие води до по-добър продукт, който фирмата може да монетизира. Например, Spotify предлага безплатна персонализирана плейлиста Discover Weekly, генерирана от ИИ. Потребителите се наслаждават на музикални препоръки, а в замяна Spotify получава ценни данни за вкусовете им, което усъвършенства препоръчващия алгоритъм. По подобен начин, OpenAI първоначално пусна безплатен достъп до ChatGPT за широката публика. Милиони интеракции на потребители предоставиха обратна връзка и примери, с които моделът се донастрои и подобри. Така **потребителят става косвен сътрудник** в развоя на продукта. Моделът "безплатно или отстъпка в замяна на данни" е особено ефективен при ИИ, защото данните са злато за по-добра производителност. Впоследствие компанията може да монетизира подобрения ИИ чрез премиум функции, реклами или продажба на аналитична информация. Този бизнес модел трансформира класическата динамика, като **потребителите плащат не толкова с пари, колкото с данните си**, а фирмата непрекъснато подобрява предложението си в един **самоподхранващ се цикъл**.
- **Екосистемен модел (свързани продукти и услуги):** някои фирми изграждат цели екосистеми от взаимосвързани решения, създавайки стойност чрез синергията помежду им. В този модел, вместо единичен продукт, компанията предлага пакет или платформа, където няколко ИИ

[303] Natural Language Processing

продукта/услуги се интегрират и подсилват взаимно. Целта е да се „заключи" клиентът в екосистемата, предоставяйки му цялостно решение и увеличавайки жизнения цикъл на продукта/услугата. Amazon е емблематичен пример, тъй като компанията използва ИИ в различни аспекти: от гласовия асистент Alexa, през препоръчващата система на своя сайт, до логистичните си алгоритми. Всички тези компоненти работят заедно: Alexa свързва потребителя с Amazon платформата, препоръките увеличават продажбите, а логистиката гарантира бърза доставка – цялата екосистема генерира стойност, по-голяма от сбора на отделните части. Друга илюстрация е платформа с ИИ за създатели на съдържание, която включва инструмент за генериране на текст, интегриран с инструмент за графичен дизайн и с платформа за разпространение – клиентът получава всичко в едно, а доставчикът – по-широка база потребители и диференцирани приходи. Екосистемните модели позволяват диверсификация на приходите и трудно подлежат на копиране от нови конкуренти, тъй като изискват цялостна интеграция на много фронтове.

- **Персонализиран премиум модел:** ИИ позволява безпрецедентна персонализация на продукти и услуги според индивидуалния потребител. Това дава възможност на фирмите да прилагат двустепенно ценово позициониране на продуктите и услугите си чрез безплатна или основна версия за масовия пазар и премиум версия, която е силно персонализирана (и срещу заплащане). Основната идея е да се привлече широка аудитория с универсален продукт, а след това най-взискателните или професионални потребители да заплатят за хиперперсонализирано преживяване с продукта или услугата. ИИ прави това възможно, тъй като моделите могат да се обучат върху данните на конкретен потребител или да се настройват към неговите предпочитания, предоставяйки уникална стойност, за която той е готов да плати допълнително. Други примери включват стрийминг платформи, предлагащи персонални препоръки или ексклузивно съдържание срещу заплащане, или дори електронна търговия, където VIP клиенти получават ИИ-базирани съветници. Персонализираният премиум модел капитализира върху готовността на клиентите да платят за нещо, съобразено точно с тях – една бизнес стратегия, станала много по-мащабна благодарение на ИИ.
- **Глобални решения за локални проблеми:** докато много технологии са универсални, ИИ предоставя възможност за масово адаптиране към местни нужди. Нов бизнес модел е създаването на продукт, който може да се разрасне в глобален мащаб, но който обаче има вариации или модули, решаващи специфични регионални проблеми или особености. Този модел показва, че ИИ решенията могат да бъдат **гъвкави според контекста**. Много стартъпи вече следват тази линия, като взимат успешен ИИ продукт (напр. платформа за микрокредитиране) и го адаптират към пазари в развиващи се страни, добавяйки езикови модели за местни езици, отчитайки културни специфики в данните и др. По този начин се отварят нови пазари и **се създава стойност чрез локална релевантност**.
- **Отворен код и сътрудничество (Open-Source/Collaborative Models):** В духа на общността за свободен софтуер, и в сферата на ИИ възникват модели, основани на сътрудничество и споделяне. Компании и инициативи предоставят отворен достъп до своите модели, данни или инфраструктура, изграждайки около тях екосистема от ползватели и сътрудници, а след това монетизират косвено. Google например пусна своя TensorFlow, който се превърна в стандарт за индустрията. Това привлече хиляди разработчици, академични изследователи и компании да го използват и подобряват. В резултат Google укрепи позицията си на лидер в областта на ИИ и изгради услуги и хардуер,

базирани на TensorFlow (TPU чипове, облачни услуги), които комерсиализира. Отвореният модел е особено ценен в областта на ИИ, защото **насърчава иновациите**. Общността допринася с подобрения, намират се нови приложения, расте доверието (защото кодът е прозрачен). Някои компании градят бизнес и чрез **консорциуми и партньорства**, като споделят помежду си ресурси с ИИ, за да стандартизират решения и да отворят нови пазари. Например много автомобилни производители създават съвместно платформи за автономно управление, вместо всеки да разработва отделно, като така ускоряват развитието и си поделят разходите. Бизнес моделът тук често включва предоставяне на платени премиум услуги, поддръжка или облачна инфраструктура около отворения продукт. В основата му стои идеята, че сътрудничеството може да разшири пазара многократно, след което всяка от участващите страни да печели, предлагайки допълнителна стойност.

- **„Познание в реално време" и Edge AI модели:** Появяват се бизнес модели, базирани на предоставяне на услуги в реално време чрез ИИ, без забавяния и зависимост от облака, често чрез *edge computing*. Т.нар. модел Real-Time CogniSense[304] цели мигновена интелигентност на място. Тези бизнес модели имат приложение в търговията, като например концепцията за магазини без касиери. ИИ вижда какво потребителят взима от рафта и автоматично таксува сметката му, позволявайки просто да излезе, без опашки и каси. По същия начин вътрешната навигация на летище с ИИ асистент може в реално време да води пътниците до изхода за извеждане към самолета, без да изпраща всички данни в облака. *Edge AI* (локални ИИ изчисления) е в основата на тези бизнес модели: данните се обработват локално, което дава бързина и решава проблеми като поверителност. Бизнес моделът тук често включва продаване на цялостно решение (хардуер + ИИ софтуер) за конкретна среда – например, комплект от камери, сензори и ИИ софтуер за умен магазин като услуга към търговци на дребно. Нулевите пречки в потребителското преживяване се превръщат в конкурентно предимство (*zero-touch experience*). Този модел трансформира очакванията на клиентите към удобство и персонализация във физическия свят, аналогично на революцията при онлайн услугите.
- **Модел „Кошер" или Глобална колективна интелигентност („Hive Mind"):** Благодарение на интернет свързаността, фирмите могат да впрегнат интелекта на хиляди експерти по света за разработване на решения с ИИ. Моделът на глобалната *crowd-sourcing*[305] платформа за данни и модели е нов начин за създаване на стойност. Резултатът за фирмата е, че получава оптимално решение (например алгоритъм за откриване на измами или предвиждане на търсене) разработено колективно, често много по-бързо и по-ефективно, отколкото ако биха го правили вътрешно. Бизнес моделът на такива платформи печели от организирането на състезания. Някои взимат комисионна от наградния фонд или предлагат премиум услуги (достъп до таланти, облачна инфраструктура за състезателите и т.н.). Концепцията **„hive mind"** (кошерен разум) предполага, че мрежата от умове може да реши задачи, непосилни за един екип, и така проекти в областта на ИИ се реализират чрез глобално сътрудничество. Това променя традиционния модел за изследвания и иновации – иновацията идва отвън, а не вътре в организацията.
- **Модели „Дълга опашка" (Long-tail AI):** по традиция усилията за внедряване на технологии се целят в масови пазарни потребности. ИИ обаче дава възможност за адресиране на изключително специфични, нишови

[304] https://cognisense.tech/

[305] Получаване, обикновено по Интернет, на информация или принос от голяма група хора, както платено, така и безплатно. Oxford English Dictionary

проблеми, т.е. „дългата опашка" от неизпълнени потребности в различни отрасли. Компаниите могат да процъфтяват, като доставят решения с ИИ за сравнително малки по обем, но много на брой ниши. Например, модел с ИИ, който оптимизира поливането само за лозарски стопанства, или пък алгоритъм, който предсказва повреди в редки индустриални машини. Самостоятелно тези пазари са малки, но сборът на многото нишови приложения създава значителна стойност. Според анализатори, *„Long-tail AI"* бизнес моделите ще обхванат банки, застрахователи, здравеопазване, земеделие, Индустрия 4.0 и др., решавайки специфични под-проблеми във всеки от тях.[306] Това често върви ръка за ръка с *embedded AI* – вграждане на ИИ модули в най-различни устройства и продукти, така че дори нишови функционалности да имат интелигентност. От гледна точка на бизнес модела, компаниите могат да печелят от голям брой малки клиенти, предлагайки им специализирани модели с ИИ, или от лицензиране на такива нишови модели към по-големи системни интегратори. Този подход стана възможен едва сега, тъй като изграждането на ИИ вече не е привилегия само на гиганти, а и малки екипи могат да разработят нишов продукт с ИИ и да намерят своята аудитория глобално чрез онлайн платформи.

Следва да се отбележи, че много **класически бизнес модели също еволюират под влияние на ИИ**. Например, рекламният модел (основен за интернет гигантите) става изцяло задвижван от ИИ - програмиране на рекламни кампании в реално време въз основа на поведението на потребителя. Моделът на платформите, свързващи продавачи и купувачи вече използва ИИ за динамично ценообразуване и оптимално съвпадение на оферта и търсене. Дори **производственият модел** – досега базиран на икономии от мащаба – се променя, когато ИИ позволява масова персонализация (*manufacturing on demand* с ИИ за планиране). Казано накратко, **ИИ преплита технология с бизнес стратегия**, което води до появата на хибридни модели, съчетаващи най-доброто от досегашните практики с нови подходи, обогатени с ИИ.

6.10.3 Трансформация на съществуващи бизнес модели чрез ИИ

Изкуственият интелект не само създава нови модели, но и фундаментално **преобразява традиционните бизнес модели** в множество отрасли. под влияние на ИИ фирмите преосмислят начина, по който създават стойност – от това как достигат до клиенти, през това какво предлагат, до начина, по който оперират.

По-долу разглеждаме ключови насоки, в които съществуващите бизнес модели се трансформират:

- **Трансформация на подхода към пазара и ангажирането на клиента:** ИИ революционизира начина, по който фирмите идентифицират, привличат и задържат клиенти. В традиционния модел продажбите разчитаха на човешка интуиция и ръчни процеси; сега **интелигентните продажбени процеси** променят играта. Например, преработване на целия процес при имотни сделки с ИИ, който намира най-перспективните запитвания, през ИИ помощници за брокерите на имоти (подсказващи оптимални оферти и следващи стъпки), до динамични ценови алгоритми, генериращи персонализирани оферти в реално време. Това показва как един консервативен модел (продажби на имоти) може да бъде трансформиран чрез ИИ от обемен и бавен процес към бърз, целеви и базиран на данни. В по-широк план, маркетинг моделът се изменя от масов към хипертаргетиран. Банки и застрахователи пък въвеждат ИИ за динамичен профил на риска – индивидуални оферти и условия според поведението на клиента, вместо еднотипни масови продукти. Тази трансформация на бизнес

[306] https://www.hankyung.com/article/202105256916i#

модела означава, че **маркетингът и продажбите стават по-просветени от всякога** – решенията се взимат въз основа на прогнози на ИИ и откриване на повтарящи се модели (*pattern recognition*), а не само на исторически отчети или усет. Отделно от това, обслужването на клиенти като ключова част от бизнес модела за много фирми се преобразява чрез чатботове и виртуални асистенти. Компании намаляват разходите за колцентрове и подобряват удовлетвореността, като интегрират чатботове, които решават рутинни запитвания мигновено, 24/7. Това не просто автоматизира функция; то променя стойностно предложението – клиентите вече очакват незабавна и точна помощ, което се постига чрез ИИ, а служителите могат да се фокусират върху по-сложните случаи, т.е. ИИ трансформира модела на взаимодействие с пазара към по-ефективен, персонализиран и ориентиран към данни подход.

- **Иновация и разширение на продуктовото портфолио:** Много фирми преосмислят какво продават посредством ИИ, създавайки **нови предложения и навлизайки в нови пазари**. В някои случаи ИИ позволява на компания с утвърден бизнес модел да развие изцяло ново портфолио (на практика нов бизнес модел) върху съществуващата си основа. Производители на индустриална техника като Siemens и GE добавят ИИ към машините си и преминават към модел **„продукт като услуга"** – напр. продават работни часове на машина чрез ИИ мониторинг, вместо просто да продадат машината. **ИИ позволява тази „сервитизация"**, тъй като прогнозната поддръжка и оптимизация стават възможни в реално време. Така ценовото предложение вече е **резултатът, а не самият продукт** – клиентите плащат за гарантираната производителност, което е нов бизнес модел, трансформиращ отношенията доставчик-клиент.

- **Реорганизация на оперативни модели:** ИИ навлиза дълбоко в операциите на компаниите – както административни, така и производствени – и **променя икономията от мащаба и ефективността**. Много съществуващи бизнес модели предполагат определени разходни структури и ограничения, които ИИ започва да променя. Това може да е фундаментална промяна, тъй като традиционно подобряването на качеството води до увеличение на разходите, но ИИ позволява едновременно намаляване на разходите и повишаване на качеството. В производствения сектор ИИ подпомага **прогнозната поддръжка** – заводите използват ИИ, за да предвиждат кога една машина ще се повреди и да я обслужат превантивно, вместо да чакат авария. Това трансформира модела на управление на активи: престоят и запасите от резервни части се намаляват, което променя икономиката на фабриката (повече производствено време, по-ниски разходи, по-гъвкаво производство). В търговията на дребно, веригата за доставки – сърцето на бизнес модела на търговците – става управлявана от ИИ чрез **динамично снабдяване и логистика**. ИИ прогнозира търсенето по региони и автоматично коригира поръчките и дистрибуцията като дава възможност на модели като *just-in-time* да достигнат нови висоти. Така ИИ не просто спестява труд, а **премества човешкия труд към по-стратегически роли**. В макроплан, новите правила на създаване на стойност в епохата на ИИ се различават от индустриалната ера: докато преди икономиите от мащаба, силната марка и оперативното съвършенство бяха ключови за устойчиво конкурентно предимство, сега на преден план излизат **скоростта на самообучение на ИИ (*learning loops*), дълбочината на данните и умението за оркестриране на екосистеми**. С други думи, **ИИ трансформира самата основа на конкурентоспособността и изплаща дивиденти на тези, които пренастроят моделите си.**

Компаниите, лидери в областта на цифровата трансформация, **„пренаписват правилата"** на бизнеса за месеци, вместо за десетилетия напред. Тези, които успешно интегрират ИИ в своя модел, независимо дали чрез нови продукти, нов начин за достигане до клиента или оптимизация на операциите – постигат *трансформиращи резултати*. От друга страна, организациите, които не успеят да се адаптират, рискуват техните остарели модели да станат неконкурентни в новата ера.

Изкуственият интелект се утвърждава като **технология с общо предназначение** (*General Purpose Technology*) подобно на електричеството, компютрите или интернет, която навлиза във всички сектори и променя фундаментално икономиката. Важно е да се отбележи, че успешната цифрова трансформация не е само въпрос на инсталиране на технология, а на **стратегическо преосмисляне на бизнеса**. Лидерите в тази сфера (от технологични гиганти до традиционни предприятия, възприели ИИ) демонстрират, че **готовността за експериментиране, гъвкавостта и умението да се извлече стойност от данните са определящи**.

Компаниите трябва да се питат не само *„Как да използваме ИИ за ефективност?*", но и *„Как ИИ променя нашите клиенти, продукти и модели на печалба?*" Отговорите на тези въпроси вече оформят новия бизнес пейзаж.

Технологичният напредък с помощта на ИИ не е еднократно събитие, а продължаващ процес на иновация. Пред очите ни се раждат нови индустрии и бизнес модели, а утвърдени компании пренаписват своите правила за успех, водени от данни и алгоритми. Тази еволюция носи огромен потенциал за стойност – както икономическа, така и за обществото – стига да се подходи с визия, етика и разбиране на новите динамики, които ИИ въвежда в бизнеса. **Организациите, които успеят да се трансформират навреме, вероятно ще бъдат архитектите на следващата ера на иновации, докато останалите рискуват да останат в периферията на този исторически преход**.

6.11. Потребление – промяна на навици

Изкуственият интелект се превърна в ключов фактор, преобразяващ начина, по който бизнесите разбират и взаимодействат с потребителите. С навлизането на ИИ в дигиталния маркетинг и търговията, проследяването на потребителското преживяване става все по-предизвикателно на фона на динамична среда и изобилие от данни. Компаниите са изправени пред **повишени очаквания** за изключително потребителско изживяване във всички дигитални канали. Много организации разглеждат ИИ като отговор за предоставяне на персонализирано съдържание и преживявания, извличайки ценни прозрения от нарастващите масиви потребителски данни в реално време. **Икономическата ефективност на ИИ е обвързана с потреблението** и компаниите, които успеят да реорганизират бизнес моделите си по начин, който да отговори на променящите се потребителски навици в резултат от навлизането на ИИ, ще могат да извлекат най-големи ползи от неговото приложение.

6.11.1 Тенденции в прилагането на ИИ и промени в потребителското поведение

Дигиталната трансформация променя нагласите и навиците на потребителите. На пазара излизат поколения, като Z и Alpha, за които ежедневното използване на електронни устройства и интерактивност в интернет мрежата са начин на живот. Дигиталният маркетинг вече силно се опира на ИИ технологии, за да се справи с растящия обем от данни и изискванията на съвременните потребители. Пазарът непрекъснато се разширява с безброй опции за онлайн пазаруване, а потребителите изразяват нуждите и предпочитанията си чрез множество канали. В отговор на това,

бизнесите внедряват ИИ, за да подобрят дигиталното преживяване и да предоставят персонализирано съдържание на всяка стъпка от потребителския път. Тази тенденция се подхранва от постоянно нарастващия обем извличани от потребителите данни, които ИИ системите могат да анализират и от които да вникват в реално време в потребностите на потребителите. В резултат, компаниите все по-често интегрират ИИ в маркетинговите си практики, виждайки критично предимство в създаването на изключително потребителско изживяване по време на покупката.

Паралелно с това, потребителското поведение **драматично се променя под влиянието на дигиталния маркетинг** – днешните потребители очакват **по-цялостно и персонализирано преживяване** при взаимодействието си с брандовете. Проучвания показват, че интегрирането на ИИ в маркетинга прави по-лесно достигането до точния клиент в точния момент, повишавайки удовлетвореността. Например, търговците на дребно, използващи ИИ в маркетинга, постигат резултати до пет пъти по-добри от тези на традиционните търговци.[307] В същото време, навлизането на ИИ в потребителските услуги трансформира взаимодействията – автоматизирани чатботове, гласови асистенти и други ИИ-базирани инструменти стават част от всекидневието на потребителите. Според проучване на Euromonitor International, през 2023 г. 72% от потребителите глобално са използвали технология, за да подобрят ежедневието си, 42% с готовност общуват с гласови асистенти, а 17% биха се чувствали комфортно да разговарят с бот за обслужване на клиенти. [308] Тези данни подчертават, че потребителите все повече свикват да си взаимодействат с ИИ, което ускорява приемането на подобни технологии.

6.11.2 Въздействие на ИИ върху очакванията на потребителите

С повишеното присъствие на ИИ, **очакванията на потребителите** също еволюират. Съвременните потребители очакват **незабавни отговори**, висока **удобство** и **персонално отношение**. ИИ допринася за това, като например предоставя 24/7 обслужване чрез чатботове, които незабавно отговарят на запитвания и повишават удовлетвореността на клиентите. [309] Клиентите вече разглеждат персонализираните препоръки и съдържание не като лукс, а като стандартна част от изживяването. Постоянството в комуникацията на бранда на различни платформи също е ключова. Чрез ИИ компаниите могат да осигурят унифицирано послание и преживяване навсякъде, където потребителят се докосва до марката. Наред с очакванията за по-добро обслужване, потребителите стават и по-взискателни по отношение на прозрачността и поверителността. С нарастващата осведоменост за ценността на данните им, клиентите очакват фирмите да боравят отговорно и прозрачно с личната информация. Това кореспондира с възхода на т.нар. **етично потребление** – потребителите все повече държат на това компаниите да използват ИИ по начин, който зачита тяхната поверителност и права. Компаниите, които не отговорят на тези очаквания, рискуват да загубят доверие. В тази връзка, стратегиите за комуникация на компаниите все по-често включват подчертаване на мерките за защита на данните и етичната употреба на ИИ, за да успокоят клиентите и да поддържат доверието им.

[307] Rabby., F. Chimhundu., R. & Hassan, R. (2021). Artificial intelligence in digital marketing influences consumer behaviour: a review and theoretical foundation for future research. Academy of Marketing Studies Journal, 25(5), 1-7, https://www.abacademies.org/articles/artificial-intelligence-in-digital-marketing-influences-consumer-behaviour-a-review-and-theoretical-foundation-for-future-research-11892.html#

[308] https://hortoninternational.com/generative-ai-and-consumer-behaviour-a-new-era-of-personalisation-and-efficiency/#

[309] https://www.visionfactory.org/post/the-impact-of-artificial-intelligence-on-consumer-behavior-analysis-in-digital-marketing#

6.11.3 Персонализация и потребителски опит чрез ИИ

Персонализацията е може би най-осезаемото въздействие на ИИ върху потреблението. Чрез анализ на огромни масиви данни, вкл. истории на покупки, поведение при използване на интернет търсачки и активност в социалните медии, ИИ може да извлича детайлни знания за предпочитанията на всеки отделен клиент. Така компаниите успяват да създадат **"едно към едно" маркетинг** в голям мащаб. Големите стрийминг платформи и платформите за онлайн търговия са пионери в използването на ИИ, за да **предвидят какво иска потребителят, още преди самият той да го осъзнае**. Например, Netflix анализира навиците на гледане, за да препоръчва филми и сериали, а Amazon предлага продукти на база предишни покупки и разглеждани артикули. Тези прецизни, персонални препоръки водят до по-висока ангажираност и изграждане на лоялност у клиента. ИИ алгоритмите могат да **подбират съдържание по поръчка** от персонализирани препоръки за продукти до индивидуализирани маркетингови съобщения. Това значително подобрява потребителския опит и стимулира продажбите. Генеративните ИИ системи (напр., GPT) правят възможно създаването на съобразени с индивидуалния потребител рекламни послания, които **резонират със специфични сегменти**, до автоматизирано генерирани продуктови описания, персонализирани по вкус на аудиторията. Това не само увеличава ефективността на рекламата, но и изгражда по-дълбока връзка между брандовете и клиентите чрез усещане за разбиране на техните нужди. Например, **виртуалните "пробни стаи" с добавена реалност**, захранвани от генеративен ИИ, позволяват на клиентите да пробват продукти онлайн (било то дрехи или мебели), визуализирайки как биха паснали в живота им, преди да вземат решение за покупка. Това нововъведение **заличава пропастта между онлайн и физическото пазаруване**, като прави изживяването по-реалистично и повишава увереността на потребителя в избора му.[310] Крайната цел на подобна дълбока персонализация е потребителят да се чувства разбран и обслужен индивидуално, което води до по-висока удовлетвореност, повторяеми покупки и устойчива лоялност към марката.

6.11.4 Прогнозиране на поведението и подпомагане на вземането на решения

Една от най-ценните способности на ИИ е **предсказването на бъдещо потребителско поведение** въз основа на исторически данни и сложни модели. Съвременните ИИ системи **успяват да открият "шаблони в шаблоните"**. Това са фини зависимости в големи масиви информация, чрез които ИИ моделите могат да разберат в дълбочина какво наистина искат клиентите. Чрез дълбоко обучение върху поведението на потребителите маркетолозите могат много по-точно да прогнозират какви продукти или услуги ще бъдат търсени, дори месеци предварително. Например, интегрирането на данни от социалните медии и анализ на настроенията на потребителите дава възможност да се идентифицират зародишни тенденции и предпочитания, което позволява на фирмите **да предвидят пазарното търсене** и да се подготвят навреме. [311] Предиктивните алгоритми на ИИ вече се използват за разнообразни цели – от прогнозиране на отлив на клиенти до предугаждане на намерение за покупка. Генеративните модели се отличават при обработката на сложни последователни данни и могат в реално време да вникват в потребителските нагласи и да прогнозират вероятни следващи стъпки на потребителите.

[310] https://hortoninternational.com/generative-ai-and-consumer-behaviour-a-new-era-of-personalisation-and-efficiency/#

[311] https://martech.health/articles/the-future-of-marketing-predicting-consumer-behavior-with-ai#

Същевременно, генеративни състезателни мрежи (GANs) и вариационни автоенкодери (VAEs) [312] се използват за синтез на данни (напр., симулиране на потребителски профили или поведенчески сценарии), което подпомага моделирането и предвиждането на поведението при недостиг на реални данни. Тези технологии вече доказано подобряват персонализирания маркетинг, управлението на инвентара и задържането на клиенти, като помагат на фирмите да бъдат една крачка пред потребителските нужди. За потребителите това означава, че **ИИ все по-често ще участва пряко във вземането на техните решения**, **макар и невидимо**. Когато дадена платформа успее успешно да предвиди кой продукт би се харесал или какво съдържание се търси, тя на практика **насочва** потребителя към определен избор. От гледна точка на удобството, това е положително: ИИ може да съкрати времето и усилието при пазаруване, като стесни кръга от релевантни опции. По този начин се облекчава процесът на вземане на решение, като потребителите по-лесно откриват това, което търсят, и са по-склонни да направят покупка. От друга страна, обаче, отново изниква въпросът за етичните стандарти при използването на ИИ, тъй като насочването на потребителя към точно определено действие, използвайки информация и данни от негови осъзнати и не напълно осъзнати действия, има характеристики на социален инженеринг.

ИИ подпомага и самия процес на покупка чрез автоматизация и интелигентна помощ. Интелигентните чатботове и виртуални асистенти, задвижвани от ИИ, могат бързо да предоставят детайлна информация за продукти, да отговарят на често задавани въпроси и дори да дават препоръки според нуждите на клиента. Това улеснява вземането на решение от страна на потребителя, тъй като елиминира нуждата да търси информация другаде или да чака свързване с оператор. В резултат, **онлайн пазаруването става по-гладко и удобно**, което променя навиците – **наблюдава се изместване на потребителското поведение към предпочитание за онлайн покупки и доставки до дома поради улеснения процес.** [313]

6.11.5 Етични въпроси и предизвикателства

Както вече споменахме, наред с многото ползи, внедряването на ИИ в потребителското пространство повдига значими етични въпроси и предизвикателства. Един от основните е **поверителността на данните**. За да функционират ефективно, ИИ системите разчитат на огромни количества лични данни от история на покупки до демографска и поведенческа информация. Това създава риск от злоупотреба с данни или нерегламентирано споделяне, ако компаниите не спазват стриктно нормите. Потребителите стават все по-чувствителни към начина, по който данните им се събират и използват, и изискват гаранции, че това става прозрачно и отговорно. Компаниите трябва да гарантират спазването на GDPR и да внедряват сигурни практики за съхранение и обработка, за да запазят доверието на клиентите.

Тясно свързан е проблемът с **пристрастията** в ИИ алгоритмите. Ако входните данни, на база които се обучават моделите, са непълни или отразяват определени стереотипи, ИИ може да възпроизведе и дори усили тези пристрастия. Това може да доведе до несправедливо таргетиране или изключване на определени групи потребители.

Друг аспект на етиката е **прозрачността**. ИИ системите често действат като "черна кутия" – трудно е за потребителя (а понякога и за самите разработчици) да разбере защо една препоръка е била направена. Липсата на вникване в логиката на ИИ

[312] https://www.techtarget.com/searchenterpriseai/feature/GANs-vs-VAEs-What-is-the-best-generative-AI-approach
[313] https://www.theseus.fi/bitstream/handle/10024/865395/Maharjan_Sanju.pdf?sequence=2&isAllowed=y#

може да ерозира доверието. Затова експертите призовават за **обясним изкуствен интелект**, особено в потребителските приложения, т.е. компании да предоставят поне основна информация на клиентите за това как и защо ИИ взима определени решения, свързани с тях. Ако потребителите усещат, че са манипулирани от непрозрачни алгоритми, това може да предизвика негативна реакция. Като добра практика се очертава фирмите да разкриват къде използват ИИ (например, в чатботове или препоръчващи системи) и да дават възможност на клиентите да контролират степента на персонализация.

Границата между убеждение и манипулация е тънка и свръхпрецизната персонализация може да доведе до експлоатация на слабости в поведението (напр., импулсивни покупки), което поставя въпроса за отговорното използване на подобни техники. В крайна сметка, организациите трябва да балансират комерсиалните ползи от ИИ с етична отговорност, за да отговорят на възхода на "етично осъзнатия" потребител в ерата на ИИ.

6.11.6 Стратегически изводи за бизнеса и администрацията

Влиянието на ИИ върху потребителското поведение има дълбоки стратегически последици за компаниите. Първо, фирмите, които рано интегрират ИИ в маркетинга и обслужването, получават значително конкурентно предимство. Чрез автоматизиране на процеси и по-добро вникване в потребностите на клиентите те могат да вземат по-информирани решения и да реагират по-бързо на променящите се предпочитания. Изследвания показват, че повечето компании осъзнават тази необходимост. Инвестициите в ИИ нарастват и, **който изостане, рискува да бъде изпреварен на пазара.**

От стратегическа гледна точка, бизнесът трябва да се фокусира върху изграждането на подходящите умения и екипи за работа с ИИ. Необходими са специалисти, които разбират както техническите аспекти на машинното обучение, така и тънкостите на потребителското поведение. Комбинирането на експертизата на данни и маркетинг е ключово за извличане на максимума от тези технологии. От огромно значение е развитието на **култура на адаптивност и учене**, тъй като ИИ сферата се развива бързо и изисква постоянно обновяване на знанията и гъвкавост в стратегиите. **Това поставя редица въпроси и пред администрацията, планираща и изпълняваща политики, свързани с развитие на пазара на труда, образованието и насърчаването на иновациите.**

Доверието на потребителите се оформя като стратегически актив, който ИИ може както да укрепи, така и да подкопае. Затова компаниите трябва съзнателно да изграждат стратегии за етичен ИИ, да дефинират политики за прозрачно използване на алгоритмите, за недопускане на дискриминация и за защита на данните. Проактивната комуникация по тези въпроси и дори **сертифицирането на ИИ системите за съответствие с етични стандарти** могат да се превърнат в конкурентно предимство, особено пред все по-внимателните към етиката потребители. В стратегически план, интегрирането на ИИ не е еднократен проект, а **непрекъснат процес**, който изисква мониторинг на ефективността, обучение на моделите с нови данни и подобрение на потребителското изживяване.

В дългосрочен план, икономическата ефективност от интегрирането на ИИ се очаква да нараства с развитието на генеративния изкуствен интелект, усъвършенстваното машинно обучение и високопроизводителните облачни услуги. В бъдеще компаниите, които успеят ефективно да интегрират тези технологии, ще получат значителни конкурентни предимства, намалени разходи и повишена ефективност.

Все още разпокъсаният глобален регулаторен пейзаж и липсата на глобални стандарти пречат на навлизането на пазара на играчи от среден и малък калибър, като развитието е оставено почти изцяло в ръцете на глобалните технологични гиганти. Необходимо е да се отбележат усилията на **Международната организация по стандартизация, която в последните гдини издаде няколко стандарта, засягащи изкуствения интелект.** Стандартите, които повишават прозрачността, качеството на данните и надеждността на системите, са начин да се намалят рисковете и да се максимизират ползите. Включвайки желаните обществени и етични резултати въз основа на изчерпателен принос на заинтересованите страни, те могат да помогнат за преодоляване на пропуските в регулирането.

По-специално това са:

- ISO/IEC 22989 Information technology — Artificial intelligence — Artificial intelligence concepts and terminology (2022)
- ISO/IEC 42001 Information technology — Artificial intelligence — Management system (2023)
- ISO/IEC 23894 Information technology — Artificial intelligence — Guidance on risk management (2023)
- ISO/IEC 38507 Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations (2022)
- ISO/IEC 23053 Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML) (2022)
- ISO/IEC 5338 Information technology — Artificial intelligence — AI system life cycle processes (2023).

Приложение 1: Дефиниции и категоризация

Общи дефиниции за изкуствен интелект

ЕС: Акт за изкуствения интелект (член 3):

> "Система с ИИ" означава машинно базирана система, която е проектирана да работи с различни нива на автономност и която може да прояви адаптивност след внедряването си, и която, с явна или подразбираща се цел, въз основа на въведените в нея входящи данни, извежда начина на генериране на резултати като прогнози, съдържание, препоръки или решения, които могат да окажат влияние върху физическа или виртуална среда."[314]
>
> „Модел на ИИ с общо предназначение" означава модел на ИИ, включително когато такъв модел е обучен с голямо количество данни, като се използва самонадзор в голям мащаб, който има до голяма степен общ характер и е в състояние компетентно да изпълнява широк набор от отделни задачи, независимо от начина на пускане на модела на пазара, и който може да бъде интегриран в различни системи или приложения надолу по веригата, с изключение на моделите на ИИ, използвани за научноизследователски и развойни дейности и като прототипи преди пускането им на пазара. [315]

ISO: Изкуствен интелект, като инженерна система, е набор от методи или автоматизирани обекти, които заедно изграждат, оптимизират и прилагат модел, така че системата да може, за даден набор от предварително дефинирани задачи, да изчислява прогнози, препоръки или решения.[316]

ОИСР: Една ИИ система е базирана на машина система, която, за изрични или подразбиращи се цели, извлича от получените входни данни как да генерира изходни данни, като прогнози, съдържание, препоръки или решения, които могат да повлияят на физически или виртуални среди. Различните ИИ системи се различават по нивата си на автономност и адаптивност след внедряване.[317]

Университет "Станфорд" (HAI – Human-Centered AI): Изкуственият интелект е термин, въведен през 1955 г. от Джон Маккарти, първият член на факултета в Станфорд по ИИ, който го определя като „наука и инженерство на правенето интелигентни машини." Много изследвания имат човешка програма софтуерни агенти със знанията да се държат по особен начин, като играта на шах, но днес подчертаваме агенти, които могат да учат, точно както човешките същества, навигиращи в нашата променящ се свят.[318]

Google: "Изкуственият интелект е област на науката, занимаваща се със създаването на компютри и машини, които могат да разсъждават, да учат и да действат по начини, които обикновено изискват човешки интелект, или да обработват данни в мащаби, надхвърлящи възможностите на човека. ИИ е широка научна област, която обхваща много дисциплини, включително информатика, анализ на данни, статистика, хардуерно и софтуерно инженерство, лингвистика, невронаука, философия и психология. В бизнес контекста, ИИ представлява комплект от технологии, базирани главно на машинно обучение и дълбоко обучение, които се използват за анализ на данни, прогнози,

[314] OJ L, 2024/1689, 12.7.2024, ELI: http://data.europa.eu/eli/reg/2024/1689/oj
[315] Ibid
[316] ISO,, Artificial Intelligence (AI) — Assessment of the robustness of neural networks — Part 1: Overview, ISO/IEC TR 24029-1:2021, 18.03.2025
[317] OECD, "Explanatory memorandum on the updated OECD definition of an AI system", *OECD Artificial Intelligence Papers*, 2024, No. 8, OECD Publishing, Paris, https://doi.org/10.1787/623da898-en.
[318] Stanford HAI, "AI Key Terms Glossary Definition", 2022, https://hai-production.s3.amazonaws.com/files/2023-03/AI-Key-Terms-Glossary-Definition.pdf , 18. 03.2025

класификация на обекти, обработка на естествен език, препоръки, интелигентно извличане на данни и други."[319]

NVIDIA: "В своята най-фундаментална форма ИИ е способността на компютърна програма или машина да мисли, учи и предприема действия, без да е изрично кодирана с команди. ИИ може да се разглежда като разработка на компютърни системи, които могат да изпълняват задачи автономно, поглъщайки и анализирайки огромни обеми от данни, след което разпознавайки модели в тези данни. Голямото и нарастващо поле на изследване на ИИ винаги е ориентирано към разработването на системи, които изпълняват задачи, които иначе биха изисквали човешки интелект, за да бъдат извършени - само със скорости, надхвърлящи възможностите на всеки индивид или група. Поради тази причина изкуственият интелект се разглежда като разрушителна (disruptive) и силно трансформираща технология. Ключово предимство на системите с изкуствен интелект е способността действително да се учат от опита или да учат модели от данни, като се коригират сами, когато нови входове (inputs) и данни се подават в тези системи. Това самообучение позволява на ИИ системите да изпълняват зашеметяващо разнообразие от задачи, включително разпознаване на изображения, разпознаване на реч на естествен език, езиков превод, прогнози за добив на култури, медицинска диагностика, навигация, анализ на кредитния риск, податливи на грешки скучни човешки задачи и стотици други случаи на употреба."[320]

IBM: Изкуственият интелект (ИИ) е технология, която позволява на компютрите и машините да симулират човешкото учене, разбиране, решаване на проблеми, вземане на решения, креативност и автономност. [321]

Колеж "Итън": Изкуственият интелект е технология, която позволява на компютърна програма да мисли и учи като човешкия ум. Според Джон Маккарти, ИИ е науката и инженерството за създаване на интелигентни машини.[322]

Видове изкуствен интелект според способностите

Според способностите, изкуственият интелект най-общо може да бъде категоризиран в следните три категории, като тази категоризация е най-широко използваната извън академичните среди:

- Тесен изкуствен интелект (Artificial Narrow Intelligence, ANI)
- Общ изкуствен интелект (Artificial General Intelligence, AGI)
- Суперинтелигентен изкуствен интелект (Artificial Superintelligence, ASI)

Тесен изкуствен интелект

Google: "Изкуствен тесен интелект: Всички съществуващи AI системи в момента са тесни AI. Те могат да изпълняват само специфични задачи, за които са програмирани и обучени."[323]

IBM: "Изкуственият тесен интелект, известен също като слаб ИИ, е единственият тип ИИ, който съществува днес. Всяка друга форма на ИИ е теоретична. Може да бъде обучен да изпълнява единична или ограничена задача, често много по-бързо и по-добре, отколкото човешкият ум може. Въпреки това, той не може да изпълнява задачи извън

[319] Google Cloud, "What is Artificial Intelligence (AI)?"https://cloud.google.com/learn/what-is-artificial-intelligence, 18.03.2025
[320] NVIDIA, "Artificial Intelligence" https://www.nvidia.com/en-us/glossary/artificial-intelligence/ , 18.03. 2025
[321] IBM, "What is artificial intelligence (AI)?" 2024, https://www.ibm.com/think/topics/artificial-intelligence, 18.03.2025
[322] Eaton Business School, Artificial Intelligence (AI): What it is and why it matters" https://ebsedu.org/blog/artificial-intelligence/ , 18.03.2025
[323] Google Cloud, "What is Artificial Intelligence (AI)?"https://cloud.google.com/learn/what-is-artificial-intelligence, 18.03.2025

определената. Вместо това, той е насочен към една подгрупа от когнитивни способности и напредва в този спектър."[324]

Колеж "Итън": "Тесен ИИ, наричан още "Слаб ИИ", е вид изкуствен интелект, който се фокусира върху една конкретна задача или област. Той е проектиран да изпълнява специфичен набор от задачи и има способността да се учи от данни. Може да се използва за автоматизиране на различни ежедневни дейности, помощ при вземането на решения, увеличаване на ефективността и намаляване на разходите."[325]

Avast: "Тесен изкуствен интелект, известен още като слаб AI, е единствената форма на ИИ, която съществува днес. Тесен ИИ е проектиран да изпълнява специфични задачи, като например чатботове, които са широко използвани в момента."[326]

Общ изкуствен интелект

Google: "Теоретично, общия ИИ ще може да „усеща, мисли и действа" като човек. В момента общ ИИ не съществува."[327]

IBM: "Общ изкуствен интелект, известен още като "силен ИИ" ("Strong AI"), е теоретична концепция към момента. Общ ИИ би могъл да използва предишни знания и умения, за да изпълнява нови задачи в различен контекст, без нужда от човешко обучение. Това означава, че общия ИИ би могъл да научи и изпълнява всяка интелектуална задача, която човек може."[328]

Колеж "Итън": "Общият ИИ е вид ИИ, при който машините са проектирани да мислят, разсъждават и действат като хора. Този вид ИИ може да учи, да разсъждава и да взема решения в голямо разнообразие от контексти. Засега в зараждащите се етапи, изследователите все още не са постигнали концепцията за общ ИИ."[329]

Avast: "Общ изкуствен интелект е теоретична форма на ИИ, която би могла да изпълнява всяка интелектуална задача, която човек може да извърши. Общ ИИ все още не е постигнат и остава в сферата на научната фантастика."[330]

Суперинтелигентен изкуствен интелект

Google: "Суперинтелигентен ИИ би представлявал машина, която превъзхожда човешкия интелект във всички аспекти. Все още не е постигнат и е обект на спекулации и етични дебати."[331]

IBM: "Суперинтелигентен ИИ е строго теоретичен ИИ. Ако някога бъде реализиран, "супер ИИ" ще мисли, разсъждава, учи и прави преценки, и ще притежава когнитивни способности, които надхвърлят тези на човека. Приложенията със суперинтелгентен ИИ ще се се развили отвъд точката на разбиране на човешките чувства и преживявания, за да изпитват емоции, да имат нужди и да притежават собствени вярвания и желания."[332]

---

[324] IBM, "Understanding the different types of Artificial Intelligence", 2023, https://www.ibm.com/think/topics/artificial-intelligence-types, 18.03. 2025
[325] Eaton Business School, "Understanding 7 types of AI (artificial intelligence) with examples" https://ebsedu.org/blog/7-types-of-artificial-intelligence, 18.03. 2025
[326] Avast, "Different Types of Artificial Intelligence (AI) You Need To Know", 2025 https://www.avast.com/c-types-of-ai-explained 18.03.2025
[327] Google Cloud, "What is Artificial Intelligence (AI)?"https://cloud.google.com/learn/what-is-artificial-intelligence, 18.03.2025
[328] IBM, "Understanding the different types of Artificial Intelligence" 2023, https://www.ibm.com/think/topics/artificial-intelligence-types, 18.03 2025
[329] Eaton Business School, "Understanding 7 types of AI (artificial intelligence) with examples" https://ebsedu.org/blog/7-types-of-artificial-intelligence, 18.03. 2025
[330] Ibid
[331] Google Cloud, "What is Artificial Intelligence (AI)?"https://cloud.google.com/learn/what-is-artificial-intelligence, 18.03.2025
[332] IBM, "Understanding the different types of Artificial Intelligence", 2023 https://www.ibm.com/think/topics/artificial-intelligence-types, 18.03. 2025

Колеж "Итън": "Суперинтелигентен ИИ или "Супер ИИ" (Super AI) би бил в състояние да надмине хората. Супер ИИ машините са самоосъзнати и се очаква да надминат човешкия интелект. Те могат да изпълняват всяка задача по-добре от хората. Концепцията за Супер ИИ все още е хипотетична. Суперинтелигентният ИИ се счита за най-напредналия и мощен вид ИИ. Някои от характеристиките на Супер ИИ са решаване на проблеми, мислене, вземане на решения и способност за тълкуване на човешки емоции и преживявания."[333]

Avast: "Суперинтелигентен изкуствен интелект е хипотетична форма на ИИ, която би надминала човешкия интелект и способности. Суперинтелигентен ИИ е предмет на дискусии и спекулации относно потенциалните му въздействия върху човечеството."[334]

Видове изкуствен интелект според функционалностите

Според функционалностите, изкуственият интелект най-общо може да бъде категоризиран в следните три категории, като Гугъл ги разглежда като етапи в развитието на ИИ:

- Реактивни машини (Базов ИИ) (Reactive Machine AI (Basic AI))
- ИИ с ограничена памет (Limited Memory AI)
- Теория на ума (Theory of Mind AI)
- Самосъзнателен ИИ (Self-Awareness AI)

Реактивни машини (Базов ИИ)

Google:"Реактивните машини са ограничени ИИ системи, които реагират на стимули според предварително зададени правила. Нямат памет, което означава, че не могат да се учат от нови данни."[335]

IBM: "Реактивни машини са тези ИИ системи, които нямат памет и са проектирани да изпълняват много специфични задачи. Те не могат да си припомнят предишни резултати или решения и работят само с наличните в момента данни. Реактивният ИИ произтича от статистическата математика и може да анализира огромни количества данни, за да произведе привидно интелигентен резултат". Пример за такъв тип реактивна машина е суперкомпютърът на IBM, Deep Blue. [336]

Колеж "Итън": "Реактивните машини са най-старите и основни форми на ИИ системи, които са чисто реактивни. Те имат ограничени възможности и нямат функционалност, базирана на паметта. С други думи, реактивните машини не притежават способността да се учат от предишен опит и да реагират на действия. Тези видове машини емулират способността на човешкия ум да реагира на различни видове входни данни. Реактивните машини могат да се използват само за автоматично реагиране на ограничен набор или комбинация от входни данни."[337]

Avast: "Реактивни машини са проектирани да реагират на специфични входни данни в реално време, без да използват минали преживявания за вземане на решения."[338]

---

[333] Eaton Business School, "Understanding 7 types of AI (artificial intelligence) with examples" https://ebsedu.org/blog/7-types-of-artificial-intelligence, 18.03. 2025
[334] Avast, "Different Types of Artificial Intelligence (AI) You Need To Know", 2025 https://www.avast.com/c-types-of-ai-explained 18.03.2025
[335] Google Cloud, "What is Artificial Intelligence (AI)?"https://cloud.google.com/learn/what-is-artificial-intelligence, 18.03.2025
[336] IBM, "Understanding the different types of Artificial Intelligence", 2023, https://www.ibm.com/think/topics/artificial-intelligence-types, 18.03. 2025
[337] Eaton Business School, "Understanding 7 types of AI (artificial intelligence) with examples" https://ebsedu.org/blog/7-types-of-artificial-intelligence, 18.03. 2025
[338] Avast, "Different Types of Artificial Intelligence (AI) You Need To Know", 2025 https://www.avast.com/c-types-of-ai-explained 18.03.2025

ИИ с ограничена памет

Google: "Повечето съвременни ИИ системи принадлежат към категорията "ИИ с ограничена памет ". Този вид ИИ използва памет, за да се подобрява с времето, като се обучава с нови данни. Дълбокото обучение (deep learning) също попада в тази категория."[339]

IBM: "За разлика от реактивните машини, ИИ с ограничена памет може да си припомня минали събития и резултати и да наблюдава конкретни обекти или ситуации във времето. ИИ с ограничена памет може да използва данни от минал или настоящ момент, за да вземе решение за курс на действие за вземане на решения. Въпреки това, докато ИИ с ограничена памет може да използва минали данни за определен период от време, той не може да запази тези данни в библиотека с минали преживявания, за да ги използва за дългосрочен период. Тъй като с течение на времето се обучава на повече данни, ИИ с ограничена памет може да подобри представянето си." Пример за такъв тип изкуствен интелект е генеративният ИИ, виртуалните асистенти и автономните превозни средства . [340]

Колеж "Итън": Машините с ограничена памет притежават възможностите на реактивните машини и, също така, са способни да се учат от минал опит и да използват тези данни за вземане на решения. Почти всички днешни ИИ системи като чатботове, самоуправляващи се коли, виртуални асистенти и т.н. попадат в тази категория ИИ. Машините с ограничена памет съхраняват големи обеми данни, за да формират справка за решаване на бъдещи проблеми."[341]

Avast: "ИИ с ограничена памет са системи, които могат да използват исторически данни за вземане на решения, като например, автономни превозни средства, които използват данни като разстояние до други превозни средства и ограничения на скоростта, за да вземат решения в реално време."[342]

Теория на ума

Google: "Теория на ума е вид ИИ, който все още не съществува, но изследванията в тази област продължават. Теория на ума описва ИИ, който може да разбира и имитира човешкото мислене, включително разпознаване на емоции и реагиране в социални ситуации."[343]

IBM: "Теория на ума е функционален клас ИИ, който попада в категорията общ ИИ. Въпреки че е все още нереализирана форма на ИИ, ИИ с функционалност на теорията на ума би разбрал мислите и емоциите на други субекти. Това разбиране може да повлияе на това как ИИ взаимодейства с хората около тях. На теория това би позволило на ИИ да симулира човешки взаимоотношения. Тъй като Теория на ума ИИ би могъл да изведе (infer) човешки мотиви и разсъждения, този вид ИИ ще може да персонализира своите взаимодействия с индивидите въз основа на техните уникални емоционални нужди и намерения. Теория на ума ИИ също би могъл да разбира и контекстуализира произведения на изкуството и есета,нещо което днешните генеративни ИИ инструменти не могат да направят. ИИ на емоциите (Emotion AI) е вид Теория на ума ИИ, който в момента се разработва. Изследователите на ИИ се надяват, че ще има способността да

---

[339] Google Cloud, "What is Artificial Intelligence (AI)?"https://cloud.google.com/learn/what-is-artificial-intelligence, 18.03.2025
[340] IBM, "Understanding the different types of Artificial Intelligence",2023, https://www.ibm.com/think/topics/artificial-intelligence-types, 18.03. 2025
[341] Eaton Business School, "Understanding 7 types of AI (artificial intelligence) with examples" https://ebsedu.org/blog/7-types-of-artificial-intelligence, 18.03. 2025
[342] Avast, "Different Types of Artificial Intelligence (AI) You Need To Know", 2025 https://www.avast.com/c-types-of-ai-explained 18.03.2025
[343] Google Cloud, "What is Artificial Intelligence (AI)?"https://cloud.google.com/learn/what-is-artificial-intelligence, 18.03.2025

анализира гласове, изображения и други видове данни, за да разпознава, симулира, наблюдава и реагира по подходящ начин на хората на емоционално ниво. Към днешна дата ИИ на емоциите не е в състояние да разбира и да отговоря на човешките чувства."[344]

Колеж "Итън": Теорията на ума е следващото напреднало ниво в системите с ИИ, което учените се опитват да разработят. Сега съществува само като хипотетична концепция. Теория на ума ИИ машините играят решаваща роля в психологията, тъй като те ще се фокусират върху емоционалната интелигентност. Този тип ИИ се нуждае от ясно разбиране на човешките чувства и поведение в околната среда."[345]

Avast: Теория на ума е тип ИИ, който е все още хипотетичен и би могъл да разбира човешките мисли и емоции, което би му позволило да взаимодейства по-естествено с хората.[346]

Самосъзнателен ИИ

Google: "Самосъзнателен ИИ е ИИ на теоритично ниво, при което машината има самосъзнание и интелектуални способности, равни на човешките. В момента не съществува." [347]

IBM: "Самосъзнателен ИИ е вид функционален ИИ клас за приложения, които биха притежавали способности на суперинтелигентен ИИ. Подобно на Теория на ума ИИ, самосъзнателният ИИ е строго теоретичен. Ако някога бъде постигнат, този вид ИИ ще има способността да разбира собствените си вътрешни условия и черти, заедно с човешките емоции и мисли. Освен това ще има свой собствен набор от емоции, нужди и вярвания."[348]

Колеж "Итън": Самосъзнателен ИИ е последният и най-напреднал етап на ИИ, който в момента е хипотетична концепция. Това ще бъде възможно, когато машините развият самосъзнание и имат подобно на човешкото съзнание. Машините със самосъзнателен изкуствен интелект ще имат същите нужди, емоции и желания като хората. Разработването на такива машини може да отнеме няколко десетилетия или дори векове. Самосъзнателен ИИ може да се разглежда като продължение на концепцията за Теорията на ума. Разликата ще бъде, че самосъзнателните машини ще имат собствени мисли и реакции."[349]

Avast: "Самосъзнателен ИИ е най-напредналият и все още хипотетичен етап на ИИ, при който машините биха имали самосъзнание и биха могли да разбират и чувстват емоции подобно на хората."[350]

Категоризация по тип алгоритъм (бързо променяща се категория)

- Системи, базирани на правила - (Rule Based)
- Машинно обучение - (Machine Learning/ ML)
- Дълбоко обучение - (Deep Learning / Deep ML)

---

[344] IBM, "Understanding the different types of Artificial Intelligence",2023, https://www.ibm.com/think/topics/artificial-intelligence-types, 18.03. 2025
[345] Eaton Business School, "Understanding 7 types of AI (artificial intelligence) with examples" https://ebsedu.org/blog/7-types-of-artificial-intelligence, 18.03. 2025
[346] Avast, "Different Types of Artificial Intelligence (AI) You Need To Know", 2025 https://www.avast.com/c-types-of-ai-explained 18.03.2025
[347] Google Cloud, "What is Artificial Intelligence (AI)?"https://cloud.google.com/learn/what-is-artificial-intelligence, 18.03.2025
[348] IBM, "Understanding the different types of Artificial Intelligence", 2023, https://www.ibm.com/think/topics/artificial-intelligence-types, 18.03. 2025
[349] Eaton Business School, "Understanding 7 types of AI (artificial intelligence) with examples" https://ebsedu.org/blog/7-types-of-artificial-intelligence, 18.03. 2025
[350] Avast, "Different Types of Artificial Intelligence (AI) You Need To Know", 2025 https://www.avast.com/c-types-of-ai-explained 18.03.2025

Системи, базирани на правила - (Rule Based)

Lantern: "Системи, базирани на правила, функционират чрез прилагане на поредица от изрази „ако-тогава" за решаване на конкретни проблеми. Тези системи са проектирани да подражават на способността за вземане на решения на човешки експерти в тесни области. Системите могат да работят само в границите на предварително определени правила, което ги прави твърди и трудни за адаптиране към нови ситуации. С нарастването на сложността на домейна броят на необходимите правила нарасна експоненциално, създавайки предизвикателства пред скалируемостта."[351]

Машинно обучение - (Machine Learning/ ML)

Lantern: "Машинното обучение се различава фундаментално от базирания на правила ИИ по своя подход. Вместо да разчитат на предварително дефинирани правила, ML алгоритмите научават модели и вземат решения въз основа на данни. Тази способност за учене от данни позволява на системите за машинно обучение да се адаптират и подобряват с течение на времето, което ги прави по-гъвкави и мощни."[352]

Университет "Станфорд": Машинното обучение (ML) е частта от ИИ, която изучава как компютърните системи могат да подобрят тяхното възприятие, знания, решения или действия въз основа на опит или данни. За тази цел ML черпи от компютърни науки, статистика, психология, неврология, икономика и теория на контрола.[353]

Колеж "Итън": Машинно обучение е подгрупа на ИИ, която включва разработване на алгоритми за интерпретиране, обработка и анализ на данни за решаване на различни проблеми от реалния свят. Тези програми или алгоритми са проектирани така, че да се учат и подобряват с течение на времето, когато са изложени на нови данни.."[354]

Avast: Машинно обучение позволява на машините да се учат от данни и да подобряват тяхното „разбиране" и производителност. Алгоритмите са в основата на ИИ и машинното обучение. Тези набори от инструкции диктуват как машините обработват данни и правят прогнози. Благодарение на алгоритмите, AI инструментите могат да създават уникален резултат и да решават нови проблеми без необходимост от допълнително програмиране.[355]

Microsoft: Машинното обучение позволява на машините да се научат да вземат информирани решения въз основа на анализи, алгоритми и статистически модели. Крайният резултат е софтуер, който може да се адаптира към начина, по който се използва. [356]

Google: В машинното обучение, популярна подгрупа на ИИ, алгоритмите се обучават върху етикетирани или немаркирани данни, за да правят прогнози или да категоризират информация.[357]

---

[351] Lantern Studios, "History of AI: From Rule-based Algorithms to Generative Models, 2024, "https://lanternstudios.com/insights/blog/the-history-of-ai-from-rules-based-algorithms-to-generative-models/ , 19.03.2025

[352] Lantern Studios, "History of AI: From Rule-based Algorithms to Generative Models, 2024, "https://lanternstudios.com/insights/blog/the-history-of-ai-from-rules-based-algorithms-to-generative-models/ , 19.03.2025

[353] Stanford HAI, "AI Key Terms Glossary Definition", 2022, https://hai-production.s3.amazonaws.com/files/2023-03/AI-Key-Terms-Glossary-Definition.pdf , 18. 03.2025

[354] Eaton Business School, "Understanding 7 types of AI (artificial intelligence) with examples" https://ebsedu.org/blog/7-types-of-artificial-intelligence, 18.03. 2025

[355] Avast, "Different Types of Artificial Intelligence (AI) You Need To Know", 2025 https://www.avast.com/c-types-of-ai-explained 18.03.2025

[356] Microsoft, "What are the different types of AI?", 2023, https://www.microsoft.com/en-us/microsoft-365-life-hacks/writing/what-are-the-different-types-of-ai, 18.03.2025

[357] Google Cloud, "What is Artificial Intelligence (AI)?"https://cloud.google.com/learn/what-is-artificial-intelligence, 18.03.2025

IBM: Класическото или „незадълбочено" машинно обучение зависи от човешката намеса, за да позволи на компютърната система да идентифицира модели, да учи, да изпълнява специфични задачи и да предоставя точни резултати. Човешките експерти определят йерархията на характеристиките, за да разберат разликите между входните данни, обикновено изискващи по-структурирани данни за учене.[358]

Дълбоко обучение

Университет "Станфорд": Дълбокото обучение (Deep Learning) е използването на големи многослойни (изкуствени) невронни мрежи, които изчисляват с непрекъснати (реални числа) представяния, подобни на йерархично организираните неврони в човешкия мозък. Той се използва успешно за всички видове ML, с по-добро обобщаване от малки данни и по-добро мащабиране до големи данни и изчислителни бюджети. Скорошен пробив е трансформаторът, невронна мрежова архитектура, която гъвкаво включва контекст чрез механизъм за внимание, позволявайки мощен и изчислително ефективен анализ и генериране на последователности, като думи в абзац.[359]

Колеж "Итън": Дълбокото обучение е клон на машинното обучение, който използва изкуствени невронни мрежи, за да получи прозрения от данни и да реши по-сложни проблеми. Алгоритъмът за дълбоко обучение е логиката зад разпознаването на лица и реч, виртуални асистенти като Alexa & Siri, самоуправляващи се автомобили и много други.[360]

Lantern: "Дълбокото обучение се характеризира с използването на изкуствени невронни мрежи, вдъхновени от структурата и функцията на човешкия мозък. Тези мрежи се състоят от слоеве от взаимосвързани възли (неврони), които обработват входни данни и изучават йерархични представяния. Дълбочината на тези мрежи им позволява да улавят и моделират сложни модели, което прави дълбокото обучение особено мощно за задачи, включващи високоразмерни (high-dimensional) данни."

Avast: Дълбоко обучение е усъвършенствана форма на машинно обучение, която използва многопластови алгоритми, наречени невронни мрежи, които едновременно получават входни данни, обработват данни и предоставят изход. Въпреки че дълбокото обучение имитира определени процеси в човешкия мозък, то не води до съзнание или разсъждение. Дълбокото обучение позволява на ИИ да обработва големи количества данни много бързо, което му позволява да изпълнява по-сложни задачи."[361]

Microsoft: Дълбокото обучение е вид машинно обучение, което използва изкуствени невронни мрежи, за да позволи на цифровите системи да учат и да вземат решения въз основа на неструктурирани, немаркирани данни. Като цяло машинното обучение обучава ИИ системите да се учат от придобития опит с данни, да разпознават модели, да правят препоръки и да се адаптират. По-специално с дълбокото обучение, вместо просто да отговарят на набори от правила, цифровите системи изграждат знания от примери и след това използват тези знания, за да реагират, да се държат и да се представят като хора.[362]

---

[358] IBM,"AI vs. machine learning vs. deep learning vs. neural networks: What's the difference?", 2023, https://www.ibm.com/think/topics/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks 19.03.2025

[359] Stanford HAI, "AI Key Terms Glossary Definition", 2022, https://hai-production.s3.amazonaws.com/files/2023-03/AI-Key-Terms-Glossary-Definition.pdf , 18. 03.2025

[360] Eaton Business School, "Understanding 7 types of AI (artificial intelligence) with examples" https://ebsedu.org/blog/7-types-of-artificial-intelligence, 18.03. 2025

[361] Avast, "Different Types of Artificial Intelligence (AI) You Need To Know", 2025 https://www.avast.com/c-types-of-ai-explained 18.03.2025

[362] Microsoft, "What is deep learning?", https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-deep-learning#:~:text=Deep%20learning%20is%20a%20type,%2C%20make%20recommendations%2C%20and%20adapt. , 19.03.2025

Google: Дълбокото обучение използва изкуствени невронни мрежи с множество слоеве за обработка на информация, имитирайки структурата и функцията на човешкия мозък. Чрез непрекъснато учене и адаптиране, ИИ системите стават все по-умели в изпълнението на специфични задачи, от разпознаване на изображения до превод на езици и други.[363]

IBM: Дълбокото обучение е подгрупа на машинното обучение. Основната разлика между машинното обучение и дълбокото обучение е как всеки алгоритъм се учи и колко данни използва всеки тип алгоритъм. Дълбокото обучение автоматизира голяма част от частта от процеса за извличане на функции, елиминирайки част от необходимата ръчна човешка намеса. Той също така позволява използването на големи набори от данни, печелейки титлата на мащабируемо машинно обучение. Наблюдаването на модели (patterns) в данните позволява на модел за задълбочено обучение да групира входовете по подходящ начин. [364]

Категоризация по метод на обучение

Според метода на обучение ИИ се разделя на няколко вида. Този тип категоризация е динамично променяща се, поради това ще бъдат изброени само основните категории.

- (Дълбоки) невронни мрежи ((Deep) Neural Networks)
- Контролирано (наблюдавано) обучение (Supervised Learning)
- Неконтролирано (ненаблюдавано) обучение (Unsupervised Learning)
- Обучение с утвърждаване (Reinforcement Learning)
- Полуконтролирано обучение (Semi-Supervised Learning)
- Изчисление по време на теста (*Test-Time-Compute* (TTC))

(Дълбоки) невронни мрежи ((Deep) Neural Networks)

IBM: "Невронната мрежа е програма за машинно обучение или модел, който взема решения по начин, подобен на човешкия мозък, като използва процеси, които имитират начина, по който биологичните неврони работят заедно, за да идентифицират явления, да претеглят опции и да стигат до заключения. Всяка невронна мрежа се състои от слоеве от възли или изкуствени неврони - входен слой, един или повече скрити слоеве и изходен слой. Всеки възел се свързва с други и има собствено свързано тегло и праг. Ако изходът на всеки отделен възел е над определената прагова стойност, този възел се активира, изпращайки данни към следващия слой на мрежата. В противен случай не се предават данни към следващия слой на мрежата. Невронните мрежи разчитат на данни за обучение, за да научат и подобрят своята точност с течение на времето. След като бъдат фино настроени за точност, те са мощни инструменти в компютърните науки и изкуствения интелект, позволявайки класифициране и групиране на данни с висока скорост. Невронните мрежи понякога се наричат изкуствени невронни мрежи (Artificial Neural Networks) или симулирани невронни мрежи (Simulated Neural Networks). Те са подгрупа на машинното обучение и са в основата на моделите за дълбоко обучение."[365]

Oracle: Докато някои ИИ модели използват правила и входни данни за вземане на решения, дълбоките невронни мрежи предлагат способността да се справят със сложни решения, базирани на разнообразни взаимоотношения на данни. Дълбоките невронни мрежи работят с многобройни слоеве, които идентифицират модели и претеглени връзки между точките от данни, за да правят прогнозни резултати или информирани оценки.[366]

---

[363] Google Cloud, "What is deep learning?" https://cloud.google.com/discover/what-is-deep-learning?hl=en 19.03.2025

[364] IBM, "What is deep learning", 2024, https://www.ibm.com/think/topics/deep-learning ,19.03.2025

[365] IBM, "What is a neural network?", https://www.ibm.com/think/topics/neural-networks, 19.03.2025

[366] Oracle, "What is AI model training & why is it important?" , 2023 https://www.oracle.com/sa/artificial-intelligence/ai-model-training/ , 19.03.2025

Контролирано (наблюдавано) обучение (Supervised Learning)

Университет "Станфорд": При контролирано обучение компютърът се научава да предвижда дадени от човека етикети, като например, определени породи кучета върху етикетирани снимки на кучета. [367]

Google: При контролирано (наблюдавано) обучение моделът се обучава с етикетирани данни (структурирани данни).[368]

IBM: Контролираното обучение, известно още като „класическо" машинно обучение, изисква човек-експерт да етикетира данните за обучение.[369]

Oracle: Контролирано (наблюдавано) обучение (Supervised Learning) - Това е еквивалент на обучение по зададен учебен план. При ИИ моделирането това означава използване на предварително дефинирани обучителни набори от данни и параметри, като учените по данни действат като „учители", които съставят обучителните набори, изпълняват тестове и осигуряват обратна връзка. [370]

Неконтролирано (ненаблюдавано) обучение (Unsupervised Learning)

Университет "Станфорд": Неконтролираното обучение не изисква етикети, но понякога възприема самоконтролирано обучение, конструирайки свои собствени задачи за прогнозиране, като например, опит за прогнозиране на всяка следваща дума в изречение.[371]

Google: При неконтролирано (ненаблюдавано) обучение моделът анализира неетикетирани данни (неструктурирани данни). Крайният резултат не е предварително зададен – ИИ сам търси закономерности. Използва се за клъстеризация, откриване на аномалии, откриване на модели в данни.[372]

IBM Неконтролирано обучение: За разлика от техниките за контролираното обучение, неконтролираното обучение не предполага външно съществуване на „правилни" или „грешни" отговори и следователно не изисква етикетиране. Тези алгоритми откриват присъщи модели в наборите от данни, за да групират точки от данни в групи и да информират за прогнозите.[373]

Oracle: Неконтролирано (ненаблюдавано) обучение е метод, който наподобява философията на Монтесори – на ИИ се дава възможност сам да открие модели в необозначени данни, без предварително зададени параметри. [374]

Обучение с утвърждаване (Reinforcement Learning)

Университет "Станфорд": Обучението с утвърждаване позволява автономност, като позволява на агент да научи последователности от действия, които оптимизират неговите общи награди, като спечелване на игри, без изрични примери за добри техники.[375]

---

[367] Stanford HAI, "AI Key Terms Glossary Definition", 2022, https://hai-production.s3.amazonaws.com/files/2023-03/AI-Key-Terms-Glossary-Definition.pdf , 18. 03.2025
[368] Google Cloud, "What is Artificial Intelligence (AI)?"https://cloud.google.com/learn/what-is-artificial-intelligence, 18.03.2025
[369] IBM, "What is an AI model?" https://www.ibm.com/think/topics/ai-model 19.03.2025
[370] Oracle, "What is AI model training & why is it important?" , 2023 https://www.oracle.com/sa/artificial-intelligence/ai-model-training/ , 19.03.2025
[371] Stanford HAI, "AI Key Terms Glossary Definition", 2022, https://hai-production.s3.amazonaws.com/files/2023-03/AI-Key-Terms-Glossary-Definition.pdf , 18. 03.2025
[372] Google Cloud, "What is Artificial Intelligence (AI)?"https://cloud.google.com/learn/what-is-artificial-intelligence, 18.03.2025
[373] IBM, "What is an AI model?" https://www.ibm.com/think/topics/ai-model 19.03.2025
[374] Ibid
[375] Stanford HAI, "AI Key Terms Glossary Definition", 2022, https://hai-production.s3.amazonaws.com/files/2023-03/AI-Key-Terms-Glossary-Definition.pdf , 18. 03.2025

Google: При обучение с утвърждаване ИИ участва в интерактивна среда, където учи чрез проби и грешки. Получава положително възнаграждение при добро представяне и наказание при грешки.[376]

IBM: При обучението с утвърждаване моделът се учи холистично чрез проба и грешка, чрез систематично възнаграждаване на правилния резултат (или наказване на неправилен резултат).[377]

Oracle: Обучение с утвърждаване е метод, който прилича на възнаграждаването на желано поведение с награди. Обучението на ИИ започва с експериментални решения, които водят до положително или отрицателно подсилване. След време ИИ научава най-добрите решения, като най-точните или успешните, за да се справи със ситуация и да увеличи максимално положителното подкрепление. [378]

Полуконтролирано обучение (Semi-Supervised Learning)

Google: Полуконтролирано обучение е смесен подход – използва се малко количество етикетирани данни, а останалите са неетикетирани. ИИ самостоятелно организира и структурира данните, за да постигне желаните резултати.[379]

Oracle: Използвайки принципите както на контролирано, така и на неконтролирано обучение, полуконтролираното обучение започва с обучение на модела върху малка група от етикетирани набори от данни. Оттам моделът използва немаркирани и неподготвени набори от данни, за да прецизира моделите и да създаде неочаквани прозрения (insights). Като цяло, полуконтролираното обучение използва само етикетирани набори от данни за първите няколко стъпки, като тренировъчни колела. След това процесът силно зависи от немаркирани данни. [380]

Изчисление по време на теста (Test-Time-Compute (TTC))

Cloud Security Alliance: Най-новият метод на обучение е т.нар метод на "изчисление по време на теста" (*Test-Time-Compute* (TTC)), който ефективно емулира „мислене". Колкото повече време бъде дадено на ИИ модела да „мисли", до толкова по-добри резултати достига. В контекста на изчисленията по време на тестване, моделирането на възнагражденията (Reward Modeling) се използва за оценяване и подреждане на множество генерирани от модела резултати по време на inference (виж последните абзаци от раздела). Вместо да се ограничава до един-единствен резултат, моделът създава няколко кандидат-решения, които се оценяват и класират въз основа на предварително дефинирани критерии. Това позволява на системата да избере най-подходящия резултат, подобрявайки неговото качество и уместност.[381]

Ключови аспекти на изчисленията по време на тестване са:

- Процес на inference, където по време на TTC моделът приема входни данни и прилага научените си параметри, за да произведе резултат. За невронните мрежи това включва предно разпространение (forward propagation) през слоевете на мрежата, използвайки матрични умножения и активационни функции.[382]

---

[376] Google Cloud, "What is Artificial Intelligence (AI)?"https://cloud.google.com/learn/what-is-artificial-intelligence, 18.03.2025
[377] IBM, "What is an AI model?" https://www.ibm.com/think/topics/ai-model 19.03.2025
[378] Oracle, "What is AI model training & why is it important?" , 2023 https://www.oracle.com/sa/artificial-intelligence/ai-model-training/ , 19.03.2025
[379] Google Cloud, "What is Artificial Intelligence (AI)?"https://cloud.google.com/learn/what-is-artificial-intelligence, 18.03.2025
[380] Oracle, "What is AI model training & why is it important?" , 2023 https://www.oracle.com/sa/artificial-intelligence/ai-model-training/ , 19.03.2025
[381] Huang, Ken, "Test Time Compute", 2024, https://cloudsecurityalliance.org/blog/2024/12/13/test-time-compute , 19.03.2025
[382] Ibid

- Изчислителни ресурси - ТТС оказва пряко въздействие върху отзивчивостта и ефективността на ИИ системите в реални приложения. Интензивността на изчисленията може да повлияе върху скалируемостта и удовлетвореността на потребителите, особено в среди, чувствителни към времето, като автономни превозни средства или системи за анализ в реално време.[383]

За по-дълбоко разбиране в речника от дефиниции е включена и дефиницията на термина "**inference**" или процеса, при който системата генерира резултати.

Въпреки че ЕС не дава точна дефиниция на процеса, то е дадено описание в **Акта за ИИ**, което е както следва: "Тази способност да се правят изводи се отнася до процеса на получаване на резултати, например прогнози, съдържание, препоръки или решения, които могат да повлияят на физическата и виртуалната среда, както и до способността на системите с ИИ да извличат модели или алгоритми, или и двете, от входяща информация или данни."[384]

ОИСР: Концепцията за „inference" обикновено се отнася до стъпката, в която системата генерира резултат от изчисления (output) от своите входни данни (inputs), обикновено след внедряване.[385]

ISO: Разсъждения, чрез които се извличат заключения от известни предпоставки.[386]

[383] Ibid
[384] OJ L, 2024/1689, 12.7.2024, ELI: http://data.europa.eu/eli/reg/2024/1689/oj 4
[385] OECD, "Explanatory memorandum on the updated OECD definition of an AI system", *OECD Artificial Intelligence Papers*, 2024, No. 8, OECD Publishing, Paris, https://doi.org/10.1787/623da898-en. 9
[386] ISO, "Information technology — Vocabulary — Part 28: Artificial intelligence — Basic concepts and expert systems", **ISO/IEC 2382-28:1995** , 20.03.2025

Приложение 2: Предложения за стратегически важни стъпки относно прилагането на регулациите

1. Гъвкаво прилагане на вторичното законодателство: Делегираните и изпълнителни актове трябва да бъдат приети навреме, в консултации със социалните партньори и с ясна комуникация за обхвата и графика на прилагане. Следва да се разработи пътна карта за делегираните актове, с приоритизиране на технически критичните сфери – например системен риск, методи за обяснимост и шаблони за съответствие.
    - Редовен мониторинг и гъвкава корекция на правилата: Препоръчително е да се установи механизъм за периодична оценка на въздействието на тези регулации върху иновациите и МСП с цел да се открият непредвидени отрицателни ефекти (напр., значимо изоставане в даден технологичен отрасъл заради регулацията), законодателите следва да са готови да ревизират нормативните актове или да издадат насочващи тълкувания, които да облекчат тежестта, без да компрометират целите.
    - Да се въведе „принцип на регулаторна предпазливост", при който нови тежки изисквания не се въвеждат без предварителна оценка на въздействието върху конкурентоспособността, особено на МСП и стартиращи компании.
2. Създаване на постоянни експертни групи по области (ИИ, платформи, киберсигурност, данни), които да включват представители на социалните партньори, академичните среди и гражданското общество. Например, допълнителен формат, фокусиран върху новите технологии, който да съветва как Актът за ИИ влияе върху внедряването на ИИ и дали има нужда от разяснения или дерогации в определени сфери.
3. Създаване на модулни и секторно адаптирани механизми за съответствие, вкл. опростени сертификати, дигитални шаблони и отворени библиотеки с тестови протоколи. Институциите следва да съпровождат приемането на тези регулации с конкретни мерки в подкрепа на иновациите. Това включва: практически насоки и примерни шаблони (например, "Ръководство за малкия бизнес по Акта за ИИ"), финансова помощ или данъчни стимули за покриване на разходите за съответствие (например, ваучери за киберсигурност обновления по CRA). Така ще се избегне по-дребните играчи да отпаднат, което би концентрирало още повече пазара. Примерни инструменти на ЕС, които могат да бъдат предложени: Създаване на EU AI Compliance Hubs за безплатно или субсидирано консултиране, както и развитие на AI Regulatory Tech платформи за автоматизирана подготовка на документация (напр., по Приложения IV–IX от Акта за ИИ).
4. Координация между регулациите: Необходимо е да се мисли интегрирано – много компании попадат под няколко от тези регулации едновременно (напр., една платформа за ИИ услуги има изисквания по Акт за ИИ, DSA, GDPR и CRA).
    - Създаването на едно гише или координационен център на национално ниво, където фирмите да получават комплексна помощ по всички цифрови регулации, вместо да обикалят различни ведомства, би оказало положителен ефект. В идеалния случай, контролиращите органи също трябва да координират проверките си, за да не дублират или противоречат изискванията. Това ще улесни спазването и ще намали бюрокрацията.
    - Съгласуване на секторни политики: Необходимо е стратегическо съгласуване между политиките за конкурентоспособност, индустриална трансформация и

цифровизация, за да се избегне фрагментирано прилагане на различните регулации.

- Да се прилагат секторно адаптирани регулаторни рамки, например, в здравеопазване, мобилност, образование, правосъдие, с различен обхват и тежест на изискванията.
- Подкрепа за етичен бизнес модел: стимули за компании, които разработват обясним и справедлив ИИ – чрез достъп до финансиране, данъчни облекчения или приоритетен достъп до обществени поръчки.

5. Инвестиции в капацитет и административна готовност: обучение на администрацията, изграждане на технически и правен капацитет и създаване на хъбове за подкрепа при прилагане на новите изисквания.
6. Акцент върху човешкия капитал и обучение: инвестиции в дигитални умения, научноизследователска дейност и предприемачество.
    - Например, повече програми за обучение по "privacy by design" за инженери, курсове по ИИ законодателство и етика за разработчици, стимули за мултидисциплинарни екипи (инженер + юрист + специалист по етика) при разработка на нови ИИ продукти. По този начин, спазването на регулациите ще се превърне в конкурентно предимство, а не в бреме – ще имаме кадри, които умеят да правят иновативни продукти, съвместими с изискванията още от първия ред код.
    - Да се интегрират специализирани модули за етичен ИИ, правно съответствие и регулаторна експертиза в програмите по STEM, право и публична администрация.
    - Изграждане на „Съюз на уменията за ИИ съответствие", включващ университети, бизнес и публичния сектор, който да подготвя кадри за сертифициране, одит и регулаторно управление на ИИ.
7. Международен диалог и сътрудничество: За да не се окаже ЕС изолиран остров на строги правила, институциите следва активно да популяризират своите регулаторни решения на глобалната сцена – чрез Г-7, Г-20, двустранни търговски споразумения, дипломатически форуми. Целта е да се постигне взаимно признаване или сближаване на стандартите. По този начин, европейските компании ще оперират в по-хомогенна среда, а чуждите няма да избягват европейския пазар. Особено важно е партньорство в областта на ИИ – да се споделят добри практики за етичен дизайн между ЕС, САЩ, Япония и др., за да се избегне регулаторен арбитраж.

Приложение 3: Регулаторни рамки на страните-членки на Международната мрежа на институтите по безопасен ИИ

Индия

Национална стратегия

Индийският подход към лидерство в областта на изкуствения интелект е амбициозен и всеобхватен, с акцент върху приобщаващото технологично лидерство, дефинирано като **#AIforAll**. Визията, която беше представена през 2018 г., предполага реализиране на пълния потенциал на ИИ в съответствие с уникалните нужди и стремежи на страната да използва ИИ за стимулиране на икономическия растеж, социалното развитие и приобщаващия растеж, превръщайки Индия в "Гараж" за развиващите се икономики.[387] Този стратегически императив подчертава ангажимента на Индия да стане не само потребител, но и значим двигател на ИИ иновациите в глобален мащаб.

Стратегията се фокусира върху пет ключови сектора, където ИИ може да има най-голямо въздействие за решаване на обществени нужди: здравеопазване, селско стопанство, образование, интелигентни градове и инфраструктура, интелигентна мобилност и транспорт. [388] Този селективен подход позволява концентриране на усилията за постигане на осезаеми резултати в сектори, критични за социално-икономическото развитие на Индия.

Въпреки значителния потенциал на ИИ, страната е изправена пред различни предизвикателства, свързани с широкомащабното му внедряване, включително липса на експертиза, отсъствие на благоприятстващи екосистеми за данни, високи разходи и ниска осведоменост, опасения за поверителността и сигурността, и липса на съвместен подход.[389]

За да се справи с изследователските стремежи на Индия в областта на ИИ, стратегията предлага двустепенна структура, включваща **Центрове за върхови постижения в областта на изследванията (CORE)**, фокусирани върху фундаментални изследвания, и **Международни центрове за трансформиращ ИИ (ICTAI)**, ангажирани с разработването и внедряването на приложни изследвания в сътрудничество с частния сектор. Тази структура има за цел да стимулира както създаването на нови знания, така и тяхното практическо приложение. Стратегията предлага също създаването на организация-чадър за насочване на изследователските усилия, провеждане на "изследователски проекти за пробив" и разработване на обща изчислителна инфраструктура. Тези инициативи подчертават амбицията на Индия да бъде в челните редици на ИИ иновациите.

Стратегията признава важността на квалификацията и преквалификацията на работната сила и предлага децентрализирани механизми за обучение в сътрудничество с частния сектор и образователните институции, както и насърчаване на създаването на работни места в нови области като анотиране на данни.[390] Тези мерки са насочени към ефективно адаптиране на работната сила към променящите се нужди на пазара на труда в ерата на ИИ.

---

[387] "National Strategy for Artificial Intelligence", 2018 https://www.niti.gov.in/sites/default/files/2023-03/National-Strategy-for-Artificial-Intelligence.pdf

[388] "National Strategy for Artificial Intelligence", 2018 https://www.niti.gov.in/sites/default/files/2023-03/National-Strategy-for-Artificial-Intelligence.pdf ,24

[389] Ibid., 46

[390] Ibid., 64

За ускоряване на внедряването на ИИ, стратегията предлага създаването на **Национален ИИ пазар (NAIM)**, който да улесни откриването на данни, инструменти за анотиране и разработването на модели от обществена полза. Правителството има ключова роля като катализатор за подкрепа на партньорства, осигуряване на достъп до инфраструктура и насърчаване на иновациите.[391]

На фона на нарастващото влияние на ИИ, стратегията обръща внимание на етичните съображения, базирани на рамката FAT (Fairness, Accountability and Transparency), и предлага създаването на консорциум от съвети по етика във всеки CORE. Предлагат се рамки за защита на данните и секторни регулаторни рамки, както и приемането на международни стандарти. За стимулиране на иновациите стратегията признава необходимостта от стабилна рамка за интелектуална собственост и предлага създаването на центрове за улесняване на IP правата и обучение на съответните органи.[392]

Националната стратегия на Индия за изкуствен интелект има за цел да насочи страната към лидерска позиция в тази трансформираща епоха, като балансира между насърчаването на иновациите и осигуряването на етично и отговорно развитие на ИИ технологиите. Стратегията признава необходимостта от справяне със съществуващите бариери и предлага конкретни мерки за изграждане на силна и приобщаваща ИИ екосистема, която да допринесе за икономическия и социален прогрес на Индия.

Принципи за отговорен ИИ - част 1

Документът, разработен като част от националната стратегия и публикуван 2021 г., "Към развитието на отговорен "ИИ за всички" предлага принципи за отговорното управление на ИИ системите, които могат да бъдат използвани от съответните заинтересовани страни в Индия. Принципите за отговорен ИИ са извлечени от Конституцията на Индия и различни закони, приети съгласно нея.[393] Този подход подчертава значението на етичните и правни основи за развитието на ИИ.

Казусите и съображенията в този документ са ограничени в контекста на решенията "Тесен ИИ", като са групирани в две широки категории: "*Съображения за системите*", възникващи в резултат на избора на проектиране на системата и процесите на внедряване, и имат потенциал да повлияят на заинтересованите страни, взаимодействащи с конкретна ИИ система; "*Обществени съображения*`, които са по-широки етични аспекти, отнасящи се до рисковете, произтичащи от самото използване на ИИ решения за специфични функции, и имат потенциални последици за обществото извън заинтересованата страна, взаимодействаща директно с конкретни системи.[394] Това разделение е полезно за разграничаване на непосредствените и по-широките етични предизвикателства, свързани с ИИ.

Съображения за системите, които се разглеждат в документа, включват: затруднение при надеждно и безопасно внедряване на ИИ системи, поради липса на разбиране на функционирането на ИИ системата; предизвикателства при обяснението на конкретни решения на ИИ системите, което поражда недоверие; риск пристрастеността да направи решенията предубедени срещу определени групи от населението; потенциал за изключване на граждани в ИИ системи, използвани за предоставяне на важни услуги и ползи; трудност при определяне на отговорност; рискове за поверителността; рискове за сигурността. В допълнение се разглеждат и две

[391] "National Strategy for Artificial Intelligence", 2018 https://www.niti.gov.in/sites/default/files/2023-03/National-Strategy-for-Artificial-Intelligence.pdf ,71
[392] Ibid., 85
[393] "Approach Document for India Part 1 – Principles for Responsible AI ",2021, https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf
[394] "Approach Document for India Part 1 – Principles for Responsible AI ", 2021, https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf

важни обществени съображения, а именно въздействието върху работните места и злонамерено психологическо профилиране. Документът разглежда различните съображения през призмата на Конституцията и идентифицира "Принципи за отговорно управление на изкуствения интелект в Индия".[395]

Въз основа на съображенията за системите и обществените съображения се идентифицират следните широки принципи за отговорно управление на ИИ: на безопасност и надеждност, на равенство,на приобщаване и недискриминация, на поверителност и сигурност, на прозрачност, на отчетност и на защита и подсилване на положителните човешки ценности.[396]Документът предвижда мерките да могат да бъдат подходящо калибрирани според специфичния риск, свързан с различните ИИ приложения, по начин, който е в крак с технологичния напредък.[397] Тези принципи и мерки са от съществено значение за изграждане на доверие в ИИ технологиите и осигуряване на тяхното отговорно развитие и внедряване.

Принципи за отговорен ИИ - част 2

Втората част на стратегията[398] предоставя задълбочен анализ на механизмите, чрез които тези абстрактни принципи могат да бъдат операционализирани и ефективно приложени в рамките на ИИ екосистемата. Предложените интервенции, изискващи съгласувани усилия от страна на правителството, частния сектор и изследователските институции, са насочени към предизвикване на фундаментална промяна в процеса на създаване на политики, свързани с ИИ. Тази промяна се характеризира с преминаване от регулаторна рамка, която е агностична към риска, към подход, базиран на оценка и управление на специфичните рискове, присъщи на различните ИИ приложения.

Механизми целят да постигнат деликатен баланс между стимулирането на иновациите и осигуряването на отговорност при разработването и внедряването на ИИ, като се цели разработване на подход базиран на риска - колко по-голям риск крие една система, толкова по-сериозни да са регулаторните изисквания при разработването, внедряването и използването ѝ, подобно на подхода на ЕС.[399] С тази цел се предлага създаването и Съвет по етика и технологии (СЕТ). Предлага се контролирано внедряване като инструменти за ранно откриване и ограничаване на злонамерен ИИ.[400] Освен това, се акцентира върху разработването на стандарти и бенчмаркове, които отчитат специфичните индийски социално-икономически и културни фактори, което ще позволи по-адекватно реагиране на уникалните предизвикателства на страната, като например спазването на конституционните права и преодоляването на съществуващото дигитално разделение. Подчертава се, също така, важността на продължаващите изследвания в тези области за постигане на динамично вземане на решения при възникване на нови предизвикателства в ИИ, като се гарантира, че бъдещите рискове и съображения няма да бъдат посрещнати със закъснели политически мерки.

Чрез препоръчване на задължително спазване на принципите за високорисков ИИ, придобит от правителството, се цели да се намали вероятността от грешки или злоумишлени действия при използването на ИИ за изпълнение на чувствителни функции, като същевременно насърчава иновациите и полезността. По същия начин, препоръката правителствено финансираните изследвания да включват инструменти за

[395] Ibid
[396] Ibid
[397]"Approach Document for India Part 1 – Principles for Responsible AI ",2021, https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf
[398] "Approach Document for India Part 2 – Operationalizing Principles for Responsible AI ",2021, https://www.niti.gov.in/sites/default/files/2021-08/Part2-Responsible-AI-12082021.pdf
[399] Ibid., Част 2; Глава 3
[400] Ibid., Част 2; Глава 4

оценка на въздействието демонстрира ангажимент за повишаване на капацитета на ИИ решенията за подобряване на общественото благосъстояние.[401]

Стратегията адресира настоящите предизвикателства, свързани с широкомащабното внедряване на ИИ системи в Индия, и очертава първите стъпки за адекватно справяне с тях, особено в контекста на бързото утвърждаване на Индия като център на ИИ иновациите.

IndiaAI

IndiaAI е мащабна инициатива на правителството на Индия, насочена към развитието и внедряването на изкуствен интелект в страната.[402] Тя има за цел да позиционира Индия като глобален лидер в ИИ, като същевременно гарантира, че ИИ се използва по начин, който е етичен, отговорен и от полза за всички граждани.

IndiaAI е официалният национален портал за изкуствен интелект на Индия. Той представлява съвместна инициатива на Министерството на електрониката и информационните технологии (MeitY), Националната платформа за управление (NPG) и Digital India Corporation (DIC). Мисията на IndiaAI е да бъде централен център за всички свързани с ИИ дейности в Индия, насърчавайки иновациите, знанията и растежа в тази област. Целите на IndiaAI включват изграждане на силна ИИ екосистема, насърчаване на сътрудничеството между заинтересованите страни, разпространение на знания и най-добри практики, както и подкрепа за развитието и внедряването на ИИ решения в различни сектори. Порталът служи като платформа за новини, ресурси, събития и възможности, свързани с ИИ в Индия.[403]

IndiaAI Compute Capacity е стълб от стратегията на IndiaAI, фокусиран върху осигуряването на достъп до необходимата изчислителна мощност за ИИ изследвания и разработки. [404] Чрез публично-частни партньорства тази инициатива цели да демократизира достъпа до високопроизводителни изчислителни ресурси (HPC), които са от съществено значение за обучението на сложни ИИ модели. Предоставянето на лесен и достъпен достъп до значителен изчислителен капацитет е ключово за стимулиране на ИИ иновациите в страната.

IndiaAI Innovation Centre е инициатива, насочена към насърчаване на иновациите в областта на изкуствения интелект в Индия.[405] Центърът за иновации е ангажиран с разработването и внедряването на фундаментални модели, като акцентът ще бъде поставен върху големи мултимодални модели, специфични за местния контекст, както и върху основополагащи модели, насочени към конкретни индустриални домейни. Това включва високопроизводителни модели, изискващи значителни ресурси, които да обслужват приложения в ключови сектори като управление, здравеопазване, селско стопанство, устойчиво развитие, производство и други приоритетни области. Създаването на такъв център е важна стъпка за превръщането на Индия в център на ИИ иновациите.

IndiaAI Datasets Platform има за цел да създаде централизирана платформа за достъп до висококачествени набори от данни, които са необходими за обучението и валидирането на ИИ модели.[406] Платформата IndiaAI Datasets има за цел да преобрази достъпа до неперсонализирани данни, предоставяйки възможност на индийските стартъпи и изследователска общност да катализират иновативни пробиви в областта на изкуствения интелект, като оптимизира процеса на достъп до висококачествени, готови

[401] "Approach Document for India Part 2 – Operationalizing Principles for Responsible AI ",2021, https://www.niti.gov.in/sites/default/files/2021-08/Part2-Responsible-AI-12082021.pdf
[402] "IndiaAI Portal", https://indiaai.gov.in/ 18.03.2025
[403] "IndiaAI Portal", https://indiaai.gov.in/ 18.03.2025
[404] "IndiaAI - Compute Capacity", https://indiaai.gov.in/hub/indiaai-compute-capacity 18.03.2025
[405] "IndiaAI Innovation Center", https://indiaai.gov.in/hub/indiaai-innovation-centre 18.03.2025
[406] "IndiaAI Datasets Platform", https://indiaai.gov.in/hub/indiaai-datasets-platform 18.03.2025

за ИИ набори от данни, чрез изграждане на интегрирана платформа, осигуряваща безпроблемно откриване, достъп и използване на данни. Тази инициатива ще улесни изследователите и разработчиците в Индия да създават иновативни ИИ решения, базирани на надеждни данни.

IndiaAI Application Development Initiative е насочена към стимулиране на разработването и внедряването на ИИ решения с висока добавена стойност, насочени към разрешаване на реални обществени и икономически предизвикателства. [407] Успешното разработване и внедряване на ИИ решения е ключово за реализиране на ползите от ИИ за страната .

IndiaAI Future Skills е инициатива, фокусирана върху подготовката на работната сила на Индия за ерата на изкуствения интелект.[408] Инициативата цели да увеличи курсовете за ИИ в бакалавърска, следдипломна и докторска степен. Усилията, също така, ще бъдат насочени към приобщаващия достъп до ИИ образование чрез създаване на лаборатории за данни и ИИ в градове от ниво 2 и ниво 3 в цяла Индия, за да се предоставят курсове на основно ниво.

IndiaAI Startup Financing е стълб, който се фокусира върху улесняването на достъпа до финансиране за ИИ стартъпи в Индия.[409] Тази инициатива включва държавно финансиране, насочено към подкрепа на иновативни ИИ компании, с цел развитие на футуристични ИИ проекти. Осигуряването на адекватно финансиране е критично за растежа и успеха на ИИ стартъпите, които са двигател на иновациите в тази област.

Safe & Trusted AI е важен стълб от стратегията на IndiaAI, който се фокусира върху осигуряването на безопасно и надеждно развитие и използване на ИИ технологиите.[410] Тази инициатива работи за улесняване на изпълнението на проекти, насочени към отговорно развитие и внедряване на ИИ, чрез стимулиране на създаването на локализирани инструменти и рамки, контролни списъци за самооценка, предназначени за иноватори, както и други насоки и управленски структури. Гарантирането на безопасността и надеждността на ИИ е от съществено значение за изграждане на обществено доверие и за широкото приемане на ИИ технологиите.

Индийският подход към регулирането на изкуствения интелект се характеризира като **балансиран**, с подчертан акцент върху стимулирането на иновациите при едновременно адресиране на етични и социални съображения. Стратегията "ИИ за всички" илюстрира амбицията за широкообхватно внедряване на ИИ за икономически и социален прогрес, като същевременно се признават и активно се търсят решения за съществуващите бариери като липса на експертиза, данни и опасения за сигурността. Предложената изследователска структура и инициативите за изграждане на ИИ пазар и насърчаване на квалификацията на работната сила показват проактивен подход към създаване на благоприятна екосистема. Наблягането на етични принципи и разработването на насоки за отговорен ИИ допълнително подкрепят заключението за балансиран подход, който се стреми да максимизира ползите от ИИ, като същевременно минимизира потенциалните рискове.

[407] "IndiaAI Application Development Initiative", https://indiaai.gov.in/hub/indiaai-application-development-initiative ,18.03.2025
[408] "IndiaAI FutureSkills", https://indiaai.gov.in/hub/indiaai-futureskills, 18.03.2025
[409] "IndiaAI Startup Financing", https://indiaai.gov.in/hub/indiaai-startup-financing , 18.03.2025
[410] "IndiaAI Safe & Trusted AI", https://indiaai.gov.in/hub/safe-trusted-ai , 18.03.2025

Република Корея

Национална стратегия от 2019 г.

Националната стратегия на Република Корея за изкуствен интелект от 2019 г. представлява детайлен план за трансформиране на страната в глобален лидер в областта на ИИ, надграждайки вече постигнатите успехи на страната в сектора на информационните технологии. Стратегията представя всеобхватен план за превръщането на страната във водеща сила в ИИ ерата, очертавайки амбициозна визия и конкретни цели, които да бъдат постигнати до 2030 г. и насочени към постигане на значими трансформации в икономически и социален план с положителен ефект върху качеството на живот. В основата ѝ е създаването на глобално конкурентоспособна ИИ екосистема, което предполага развитие на необходимата инфраструктура, укрепване на технологичния капацитет, въвеждане на гъвкави регулаторни рамки и правни реформи, както и активна подкрепа за стартиращи предприятия в ИИ сектора с глобален потенциал.[411]

Стратегията, също така, акцентира върху максималното оползотворяване на ИИ във всички аспекти на живота, включително развитие на висококвалифицирани кадри, насърчаване на широкото приложение на ИИ в индустрията и публичната администрация, както и създаване на етичен и безопасен ИИ ориентиран към подобряване на благосъстоянието на гражданите. За реализирането на тези цели се предвиждат ключови инициативи, като създаване на специализиран инвестиционен фонд за ИИ и провеждане на мащабен конкурс за научноизследователски и развойни проекти в областта на ИИ. Очакваните резултати включват значителен икономически растеж, създаване на нови работни места и повишаване на международната конкурентоспособност на страната в ИИ сферата.[412]

Стратегията представлява проактивен и добре структуриран подход, насочен към пълноценното използване на потенциала на ИИ за стимулиране на икономическия и социален прогрес, като същевременно се предвиждат мерки за минимизиране на възможните рискове и предизвикателства.

Национални насоки за етичен ИИ

Етичните насоки за изкуствен интелект, представени от Корейския институт за развитие на информационното общество (KISDI), имат за цел да осигурят отговорно развитие и използване на ИИ технологиите, като представят задължителни изисквания и важни принципи.[413]

Етичните насоки залагат ключови задължителни изисквания, които трябва да се спазват при разработването и използването на ИИ системи: защита на човешките права, защита на поверителността, уважение към разнообразието (недискриминация), обществено благо, предотвратяване на вредите, солидарност, управление на данни (защита на личните данни), отчетност, безопасност и прозрачност.[414]

Освен задължителните изисквания, са разработени и основни стандарти, които трябва да се спазват от всички членове на обществото, като се залага на ИИ разработен с човека в центъра и спазва принципите на уважение към човешкото достойнство, обществено благосъстояние, правомерно използване на технологията. Стандартите се фокусират върху: прозрачност (обяснимост, предпазни мерки и предварително предоставена информация за работата на ИИ), защита на човешките права

[411] "National Strategy for Artificial Intelligence", 2019, https://www.msit.go.kr/bbs/documentView.do?atchFileNo=30011&fileOrdr=1 , 19.03.2025
[412] "National Strategy for Artificial Intelligence", 2019, https://www.msit.go.kr/bbs/documentView.do?atchFileNo=30011&fileOrdr=1 , 19.03.2025
[413] "인공지능 윤리기준", https://ai.kisdi.re.kr/aieth/main/contents.do?menuNo=400029 19.03.2025
[414] Ibid,, "인공지능 윤리기준"

(човекоцентричност, гарантиране на човешките права и свободи, услуги, ориентирани към човека), защита на личните данни (защита на личните данни, минимизиране на злоупотребата с лична информация), уважение към многообразието (многообразие, гаранции за достъп, недопускане на дискриминация, минимизиране на пристрастия и дискриминация), предотвратяване на вреди (ненарушаване (non-infringement), употреба за цели, които не причиняват вреда на хората), безопасност (превенция на потенциалните опасности, гаранции за безопасност), отговорност (ясно дефинирани отговорности и отговорни участници), предотвратяване на вреди (ненарушаване, използване за цели, които не вредят на хората), управление на данните (минимизиране на пристрастия в данните, управление на качеството и риска), солидарност (солидарност между групи, гарантиране на участието на заинтересованите страни, глобално сътрудничество) и обществено благо (насърчаване на общественото благо, полза за общото благо на човечеството, максимизиране на положителното въздействие на ИИ, образование).[415]

Тези стандарти и принципи представят един всеобхватен подход към управлението на потенциалните рискове и предизвикателства, свързани с ИИ. Насоките се стремят да балансират иновациите с отговорността, като същевременно насърчават сътрудничеството и диалога между всички заинтересовани страни, а динамичният характер на тези насоки е важен за тяхната дългосрочна ефективност в бързо развиващата се област на изкуствения интелект и самото общество.

"Основен закон за насърчаване на развитието на изкуствен интелект и създаване на основа за доверие"

Основният закон на Република Корея, озаглавен "Основен закон за насърчаване на развитието на изкуствен интелект и създаване на основа за доверие", който е приет в началото на 2025 г. и влиза в сила от 24 януари 2026 г., има за цел да стимулира растежа на индустрията за изкуствен интелект и да улесни внедряването на ИИ технологии в редица ключови сектори, включително здравеопазване, енергетика, транспорт, финанси и правоприлагане[416] . За тази цел законът предвижда правителствена подкрепа за НИРД (чл.13), както и за установяване на стандартизация (чл.14) и насърчаване на международното сътрудничество в областта на ИИ (чл. 22).

Законът въвежда важна категоризация на ИИ системите въз основа на тяхното потенциално въздействие, като определя категорията "Изкуствен интелект с висока степен на въздействие" (чл.2) за системи, които се използват в критични сектори като здравеопазване, ядрена енергетика, финанси и държавно управление. Тази категоризация е ключова, тъй като налага допълнителни и по-строги изисквания за прозрачност, безопасност и отчетност върху тези ИИ системи с високо въздействие. Този подход демонстрира осъзнаване на различните нива на риск, свързани с различните приложения на ИИ.

В подкрепа на индустрията и иновациите (глава 3), законът предвижда създаването на Център за политики в областта на ИИ (чл.11), предоставянето на финансова, административна и инфраструктурна подкрепа за ИИ компании, включително малки и средни предприятия (МСП) и стартъпи (чл.17). Освен това, законът насърчава създаването на специализирани изследователски институти и ИИ индустриални клъстери (чл.11, чл. 21), които да привлекат инвестиции и

[415] "인공지능 윤리기준", https://ai.kisdi.re.kr/aieth/main/contents.do?menuNo=400029 19.03.2025
[416] "인공지능 발전과 신뢰 기반 조성 등에 관한 기본법", 2025, https://www.law.go.kr/%EB%B2%95%EB%A0%B9/%EC%9D%B8%EA%B3%B5%EC%A7%80%EB%8A%A5%20%EB%B0%9C%EC%A0%84%EA%B3%BC%20%EC%8B%A0%EB%A2%B0%20%EA%B8%B0%EB%B0%98%20%EC%A1%B0%EC%84%B1%20%EB%93%B1%EC%97%90%20%EA%B4%80%ED%95%9C%20%EA%B8%B0%EB%B3%B8%EB%B2%95/(20676,20250121)

висококвалифицирани специалисти в сектора. Тази мярка е насочена към изграждане на силна и конкурентоспособна ИИ екосистема в страната.

Законът, също така, определя етични принципи, които трябва да ръководят безопасното разработване и използване на ИИ технологиите (глава 4). Той предвижда създаването на национална комисия по ИИ (чл.7), доброволен комитет по етика на ИИ (чл.28), механизми за сертифициране и оценка на въздействието на ИИ системите (чл.30), както и въвеждането на модел за човешки надзор при приложения на ИИ, които представляват висок риск (чл.34) и оценка на въздействието на такъв тип модели (чл.35). Тези разпоредби подчертават ангажимента на Република Корея към отговорното развитие на ИИ и са насочени към повишаване на доверието в ИИ технологиите и осигуряване на възможност за човешка намеса при необходимост като се предвиждат и наказателни разпоредби (глава 6).

Законът съдържа разпоредба, която го прави приложим и за чуждестранни компании, в случай че техните ИИ продукти оказват въздействие върху вътрешния пазар на Южна Корея (чл. 36). Това показва намерението на Република Корея да регулира ИИ дейностите, които имат влияние на нейна територия, независимо от произхода на компаниите.

Институт за безопасен ИИ

Институтът за изследване на изкуствен интелект (AISI) е академичен изследователски институт, фокусиран върху фундаментални и приложни изследвания в областта на изкуствения интелект.[417] Мисията на AISI е да провежда задълбочени изследвания върху основните принципи на изкуствения интелект и да разработва технологии от следващо поколение, които са безопасни, надеждни и етични.[418]

Основната цел на AISI е да напредне в разбирането и разработването на безопасен и надежден ИИ, както чрез анализирането на ИИ политики, така и чрез валидиране на модели и сътрудничество както на международно ниво, така и на национално, заедно с индустрията и академичните институции. Институтът иска да постигне провеждане на фундаментални изследвания върху теорията и основите на безопасния ИИ; разработване на нови методологии и техники за оценка и осигуряване на безопасността на ИИ системите; изследване на етичните и социални аспекти на ИИ и разработване на насоки за тяхното отговорно управление; създаване на международна платформа за сътрудничество и обмен на знания в областта на безопасността на ИИ, както и обучение на следващото поколение изследователи и експерти в областта на безопасния ИИ.[419]

Посоката на развитие на регулациите в Република Корея в областта на изкуствения интелект може да бъде определена като целенасочен стремеж към баланс между насърчаване на иновациите и установяване на етични и безопасни рамки за развитие и използване на ИИ технологиите. Регулациите в Република Корея се характеризират със стратегически подход за насърчаване на иновациите, въвеждане на етични насоки и правни рамки за отговорно развитие, диференциран подход към регулирането в зависимост от риска, подкрепа за индустрията и иновациите успоредно с регулаторните мерки и създаване на специализирани институции.

[417] "AI Safety Institute", https://www.aisi.re.kr/kor 20.03.2025
[418] "AI 안전정책", https://www.aisi.re.kr/kor/contents/10 20.03.2025
[419] Ibid., "AI 안전정책"

Великобритания

ИИ възможности и План за действие

Документът "ИИ възможности и План за действие" на Обединеното кралство очертава визията на правителството за оползотворяване на възможностите, които предлага изкуственият интелект, и за ефективно справяне с потенциалните предизвикателства. Той представя план за действие, включващ конкретни мерки и инициативи, насочени към укрепване на позицията на Обединеното кралство като водеща световна сила в областта на ИИ.[420]

Планът за действие подчертава значителния потенциал на ИИ за стимулиране на икономическия растеж, повишаване на общата производителност и подобряване на качеството на обществените услуги. Разглежда се приложението на ИИ в широк спектър от сектори, включително здравеопазване, транспорт, финанси, образование и творчески индустрии. Документът акцентира върху ключовата роля на инвестициите в научни изследвания и разработки в областта на ИИ, както и в изграждането на необходимата инфраструктура и развитието на съответните умения. Той, също така, обръща внимание на важните етични и социални последици от развитието на ИИ и подчертава необходимостта от отговорно и приобщаващо развитие и внедряване на тази технология. Документът акцентира върху развитието на таланти и умения в областта на ИИ, като признава критичната роля на квалифицираната работна сила за реализиране на потенциала на технологията. [421]

Планът предлага конкретни мерки за насърчаване на общественото доверие в ИИ, включително осигуряване на прозрачност на системите, ясна отчетност за техните действия и гарантиране на справедливост при тяхното използване. Той разглежда и въпроси, свързани с регулирането на ИИ, като се стреми към създаване на подход, който едновременно насърчава иновациите и ефективно управлява потенциалните рискове. В заключение, документът подчертава същественото значение на активното международно сътрудничество в областта на ИИ за постигане на общ напредък., като пример за това беше домакинството на Великобритания на първия форум на високо равнище за безопасен ИИ през 2023 г. [422]

Документът очертава цялостна и амбициозна стратегия на Обединеното кралство в областта на изкуствения интелект, която обхваща както икономическите възможности, така и социалните и етични аспекти на технологията, с акцент върху балансиран регулаторен подход и международно сътрудничество.

Институт за сигурен ИИ

Институтът за сигурен ИИ (AI Security Institute - AISI) е изследователска организация в рамките на Министерството на науката, иновациите и технологиите на правителството на Обединеното кралство. [423] Следва да се отбележи, че до началото на 2025 г. Институтът носеше името "Институт за безопасен ИИ", но бе преименуван на "Институт за сигурен ИИ". Тази промяна отразява стратегическо пренасочване на фокуса към сериозните рискове, които ИИ може да представлява за сигурността, включително злонамерени кибератаки, кибер измами и други киберпрестъпления. Преименуването подчертава ангажимента на правителството да гарантира националната сигурност и да защитава гражданите от престъпления, които могат да бъдат извършени с помощта на

[420] "AI Opportunities Action Plan", 2025, https://assets.publishing.service.gov.uk/media/67851771f0528401055d2329/ai_opportunities_action_plan.pdf 14.03.2025
[421] Ibid.
[422] "AI Opportunities Action Plan", 2025, https://assets.publishing.service.gov.uk/media/67851771f0528401055d2329/ai_opportunities_action_plan.pdf
[423] "UK AI Security Institute", https://www.aisi.gov.uk/ 19.03.2025

ИИ. Преминаването от "безопасност" към "сигурност" отразява нарастващото глобално осъзнаване на потенциалните заплахи за националната сигурност, произтичащи от напреднали ИИ технологии.

Институтът има за цел да постигне мисията си чрез разработване на социотехническа инфраструктура, необходима за разбиране на рисковете от напреднал ИИ и за създаване на условия за неговото управление. Той се фокусира върху смекчаване на сериозни рискове, свързани с ИИ, като потенциалното му използване за разработване на химически и биологични оръжия, извършване на кибератаки и улесняване на престъпни дейности като измами и насилие над деца.[424]

AISI работи за изграждане на научна база от доказателства, която да информира политиците и да помогне за защитата на Обединеното кралство от нововъзникващи рискове, свързани с ИИ. Институтът ще си сътрудничи с Лабораторията за изследване на сигурността на ИИ (LASR) и Националния център за киберсигурност (NCSC) за укрепване на усилията на страната в областта на киберсигурността. За да постигне амбициозните си цели, институтът е създаден като стартъп в правителството, съчетавайки авторитета на държавата с експертизата и гъвкавостта на частния сектор.[425]

ИИ: етика, управлението и регулации

Парламентарният обзор на Обединеното кралство разглежда етичните, управленските и регулаторните аспекти на изкуствения интелект, като очертава, че ИИ предлага значителни ползи, но поражда и съществени етични предизвикателства, свързани с пристрастията, прозрачността, отчетността и въздействието върху заетостта.[426]

В контекста на етиката, обзорът подчертава риска от възпроизвеждане и усилване на съществуващи социални неравенства чрез ИИ системите, както и опасенията относно липсата на прозрачност в начина, по който някои ИИ вземат решения, като също така засяга въпросите за отговорността при грешки, причинени от ИИ, и потенциалното въздействие върху пазара на труда поради автоматизацията. По отношение на управлението се посочва, че съществуващите закони може да не са достатъчни за справяне с всички предизвикателства, породени от ИИ, и се разглеждат различни подходи за регулиране, вариращи от саморегулация до създаване на нови законодателни рамки. Обзорът отбелязва, че регулаторните усилия се стремят да балансират между насърчаването на иновациите и защитата на обществените интереси и права.[427]

Парламентарен обзор представя ИИ като технология с голям потенциал, но и със значителни етични и управленски предизвикателства, които изискват внимателно обмисляне на регулаторни мерки. Акцентът върху баланса между иновациите и защитата на обществените интереси е ключов момент в дискусията за бъдещото регулиране на ИИ в Обединеното кралство.

Хъб за записване на стандарти за алгоритмична прозрачност

Хъбът за записване на стандарти за алгоритмична прозрачност (Algorithmic Transparency Recording Standard (ATRS))[428] представлява инициатива на правителството на Обединеното кралство, целяща да повиши прозрачността при използването на алгоритми в публичния сектор. Основната цел е да се предостави структуриран начин

[424] "UK AI Security Institute", https://www.aisi.gov.uk/ 19.03.2025
[425] Ibid
[426] "Artificial intelligence: ethics, governance and regulation", 2024, https://post.parliament.uk/artificial-intelligence-ethics-governance-and-regulation/ 19.03.2025
[427] "Artificial intelligence: ethics, governance and regulation", 2024, https://post.parliament.uk/artificial-intelligence-ethics-governance-and-regulation/ 19.03.2025
[428] "Algorithmic Transparency Recording Standard Hub", 2024, https://www.gov.uk/government/collections/algorithmic-transparency-recording-standard-hub 21.03.2025

на държавните организации да документират ключова информация относно алгоритмите, които внедряват в своята дейност. Това включва данни за предназначението на алгоритъма, вида на използваните данни, логиката на вземане на решения и потенциалното въздействие върху гражданите.

Тази инициатива е важна стъпка към подобряване на отчетността и изграждане на обществено доверие в автоматизираните процеси на вземане на решения в публичната администрация. Чрез стандартизиране на информацията, която се записва, ATRS улеснява външния контрол и възможността за оценка на етичните и социални последици от използването на алгоритми. В крайна сметка, това може да допринесе за по-отговорно и прозрачно използване на ИИ и други алгоритмични технологии в държавния сектор.

Регулаторната посока на Обединеното кралство в областта на изкуствения интелект се стреми към **балансиран подход**, който едновременно насърчава иновациите и адресира етичните и социални предизвикателства, както и нарастващите опасения за сигурността. "Планът за действие и възможности за ИИ" ясно заявява намерението на правителството да стимулира икономическия растеж и производителността чрез ИИ, като същевременно подчертава необходимостта от отговорно и приобщаващо развитие и внедряване на технологията. Акцентът върху инвестициите в научни изследвания, инфраструктура и умения показва силен фокус върху подкрепата на иновациите. Същевременно, вниманието към етичните и социални последици, както и мерките за насърчаване на общественото доверие чрез прозрачност, отчетност и справедливост, демонстрират съществен ангажимент към етичните аспекти на ИИ. По този начин, може да се заключи, че подходът не набляга прекомерно на едната страна за сметка на другата, а се стреми към холистично регулиране, което да максимизира ползите от ИИ, като същевременно минимизира потенциалните вреди и рискове за сигурността.

Австралия

Австралийският регулаторен пейзаж, в контекста на който се имплементира стратегията за изкуствен интелект, се характеризира с фокус върху създаването на условия за дигитална трансформация и иновации, като същевременно се обръща внимание на етичните и социални аспекти. Съществува осъзната необходимост от балансиране между насърчаването на развитието и приемането на нови технологии, включително ИИ, и осигуряването на защита на гражданите и бизнеса. Стратегията за ИИ на Австралия се вписва в тази рамка, като цели да насочи развитието и използването на ИИ по начин, който е отговорен, етичен и носи ползи за цялото общество.[429]

Национална стратегия за ИИ - за държавния сектор

Стратегията за изкуствен интелект на Австралия има за цел да позиционира страната като лидер в разработването и приемането на тази технология. Тя се фокусира върху отключването на потенциала на ИИ за подобряване на живота на австралийците, стимулиране на икономическия растеж и укрепване на позицията на Австралия на световната сцена. Стратегията е структурирана около няколко ключови мисии, които очертават конкретни области на фокус и цели.[430]

Една от основните мисии е "Предоставяне за всички хора и бизнес", която се стреми да гарантира, че ползите от дигиталната трансформация и ИИ са достъпни за всички граждани и предприятия. Това включва подобряване на достъпа до цифрови услуги и

[429] "Data and digital policy landscape", https://www.dataanddigital.gov.au/about/landscape 14.03.2025
[430] "The Data and Digital Government Strategy", https://www.dataanddigital.gov.au/strategy 14.03.2025

технологии, както и подкрепа за развитието на умения, необходими за участие в дигиталната икономика.[431]

Друга ключова мисия е "Прости и безпроблемни услуги", която има за цел да преобрази начина, по който правителството предоставя услуги на гражданите и бизнеса, като ги направи по-лесни за използване, по-ефективни и по-персонализирани чрез използването на дигитални технологии и ИИ. Това включва рационализиране на процесите, автоматизиране на рутинни задачи и предоставяне на проактивна и навременна подкрепа.[432]

Мисията "Правителство за бъдещето" се фокусира върху изграждането на по-гъвкав, адаптивен и ориентиран към бъдещето държавен сектор, който може ефективно да използва данните и цифровите технологии, включително ИИ, за да взема по-добри решения и да предоставя по-добри резултати за обществото. Това включва инвестиране в цифрови умения в държавния сектор и насърчаване на култура на иновации и експериментиране.[433]

"Доверено и сигурно" е мисия, която подчертава важността на защитата на данните и дигиталните системи, както и изграждането на доверие в използването на технологии, включително ИИ. Това включва прилагане на силни мерки за киберсигурност, осигуряване на прозрачност и отчетност при използването на ИИ и защита на личната информация.[434]

Последната разгледана мисия е "Основи на данните и цифровите технологии", която се фокусира върху създаването на силни основи за дигитална трансформация, включително ефективно управление на данните, развита цифрова инфраструктура и подкрепяща регулаторна среда. Това е предпоставка за успешното внедряване и използване на ИИ в различни сектори.[435]

Стратегията за ИИ на Австралия е тясно свързана с по-широкия регулаторен пейзаж, като се стреми да максимизира ползите от ИИ, като същевременно управлява потенциалните рискове. Чрез своите мисии, стратегията очертава ясен път за интегриране на ИИ в различни аспекти на държавното управление и икономиката, с фокус върху приобщаването, ефективността, сигурността и изграждането на доверие.

Етичен ИИ

Австралийският подход към ангажирането с изкуствен интелект подчертава необходимостта организациите да бъдат наясно с потенциалните рискове и ползи от тази технология. Той насърчава, но не задължава бизнеса да обмисли етичните последици от внедряването на ИИ и да предприеме стъпки за минимизиране на вредите.[436]

Правителството формулирала набор от етични принципи за изкуствен интелект, които служат като насоки за отговорното проектиране, разработване и внедряване на ИИ. Тези принципи включват: полза (генериране на нетна полза за хората и планетата), непричиняване на вреда (избягване на неприемливи вреди), автономна намеса (способност на хората да контролират решенията), справедливост (справедливи и

[431] "Mission: Delivering for All people and business", https://www.dataanddigital.gov.au/strategy/missions/delivering-for-all-people-and-business 14.03.2025
[432] "Mission: Simple and seamless services", https://www.dataanddigital.gov.au/strategy/missions/simple-and-seamless-services 14.03.2025
[433] "Mission: Government for the future", https://www.dataanddigital.gov.au/strategy/missions/government-for-the-future 14.03.2025
[434] "Mission: Trusted and Secure", https://www.dataanddigital.gov.au/strategy/missions/trusted-and-secure 14.03.2025
[435] "Mission: Data and digital foundations", https://www.dataanddigital.gov.au/strategy/missions/data-and-digital-foundations 14.03.2025
[436] "Engaging with Artificial Intelligence", 2024, https://www.cyber.gov.au/resources-business-and-government/governance-and-user-education/artificial-intelligence/engaging-with-artificial-intelligence 13.03.2025

приобщаващи резултати), прозрачност и обяснимост (разбираемост и възможност за оспорване), поверителност и сигурност (защита на данните и системите) и отчетност (ясна отговорност за резултатите от ИИ). Тези принципи са предназначени да бъдат приложими за правителството, и за бизнеса.[437]

Правителствената политика за отговорно използване на ИИ в държавния сектор предоставя допълнителни насоки, фокусирани върху специфични аспекти като прозрачност, отчетност и човешки контрол при вземането на решения, базирани на ИИ. Тя подчертава необходимостта от оценка на риска и предприемане на мерки за смекчаване на потенциалните негативни последици.[438]

Агенцията за дигитална трансформация на Австралия се фокусира върху ангажимента за насърчаване на иновациите, като същевременно се гарантира отговорното и етично използване на ИИ технологиите.[439]

Етичните принципи за ИИ на Австралия, както са изложени в предоставените линкове, представляват по-скоро **насочваща рамка**, отколкото строги регулаторни изисквания към бизнеса. Те са формулирани като общи принципи, които целят да информират и ръководят разработването и внедряването на ИИ, а не да налагат конкретни забрани или задължителни технически стандарти. Акцентът е върху насърчаването на отговорно поведение и етични съображения при използването на ИИ.[440]

Компаниите, които не се съобразяват с етичните насоки, могат да се сблъскат с репутационни рискове, загуба на потребителско доверие и потенциално бъдещи регулаторни мерки, ако се появят значителни обществени или етични проблеми, свързани с ИИ. Правителствената политика за отговорно използване на ИИ, макар и насочена към държавния сектор, също може да послужи като индикатор за бъдещи очаквания към частния сектор. Етичните принципи на Австралия за ИИ в момента са **набор от препоръки и насоки**, целящи да култивират отговорна и етична практика при разработването и използването на ИИ от бизнеса, отколкото като строги регулаторни ограничения. Те предоставят основа за саморегулация и насърчават проактивното управление на етичните рискове, свързани с ИИ.

Предстоящ план за ИИ способности на Австралия

Министерството на промишлеността, науката и технологиите на Австралия обяви разработването на Национален план за капацитет в областта на изкуствения интелект. Целта на този план е да превърне Австралия във водеща нация в разработването и използването на ИИ технологии. Процесът на разработване ще включва консултации с индустрията, изследователските институции и общността, за да се гарантира, че планът отразява широк кръг от гледни точки и нужди.[441]

Планът ще се фокусира върху няколко ключови области, включително стимулиране на иновациите и научните изследвания в областта на ИИ, развитие на квалифицирана работна сила, насърчаване на приемането на ИИ в различни сектори на икономиката и осигуряване на етично и отговорно използване на технологията.[442] Акцентът върху

[437] "Australia's AI Ethics Principles", https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-principles/australias-ai-ethics-principles 13.03.2025
[438] "Policy for the Responsible Use of AI in Government", ver. 1.1, 2024, https://www.digital.gov.au/sites/default/files/documents/2024-10/Policy%20for%20the%20responsible%20use%20of%20AI%20in%20government%201.1.pdf 13.03.2025
[439] "Artificial Intelligence in Government", https://www.digital.gov.au/policy/ai 13.03.2025
[440] "Engaging with Artificial Intelligence", 2024, https://www.cyber.gov.au/resources-business-and-government/governance-and-user-education/artificial-intelligence/engaging-with-artificial-intelligence 13.03.202
[441] "Developing a National AI Capability Plan", 2024, https://www.industry.gov.au/news/developing-national-ai-capability-plan 13.03.2025
[442] Ibid

етичното и отговорно използване на ИИ показва, че правителството признава важността на тези аспекти при развитието на националния ИИ капацитет. Етичните принципи или потенциалните ограничения за бизнеса, ще бъдат взети предвид при разработването на плана.[443] Включването на консултации с индустрията предполага, че ще бъде търсен баланс между насърчаването на иновациите и осигуряването на етични и обществено приемливи практики.

Обявеният национален план за капацитет в областта на ИИ на Австралия има за цел да укрепи позицията на страната като лидер в тази област, като същевременно обръща внимание на етичните съображения и необходимостта от отговорно използване на технологията. Предстоящите консултации с индустрията ще бъдат ключови за определяне на конкретните мерки и за намиране на баланс между стимулирането на икономическия растеж и осигуряването на етични практики, което ще повлияе на степента на потенциални ограничения за бизнеса.

Регулаторният подход на Австралия към ИИ може да бъде определен като **умерен**. Той се стреми да намери среден път между предоставянето на значителна свобода на бизнеса за иновации и въвеждането на прекомерни ограничения. Акцентът е върху създаването на благоприятна екосистема за развитие на ИИ, която същевременно е съобразена с етичните норми и общественото благо. Бъдещото развитие на националния план за капацитет в областта на ИИ, включващ консултации с индустрията, ще бъде ключово за финализирането на конкретните регулаторни мерки и за поддържането на този деликатен баланс.

Канада

Канадската екосистема за изкуствен интелект (ИИ) представлява динамична и разнообразна структура, която е активно подкрепяна от правителството с цел насърчаване на отговорното развитие и използване на ИИ технологии. Един от основните аспекти е Пан-канадската стратегия за ИИ, която е инициирана за укрепване на позицията на Канада в ИИ чрез инвестиции в изследвания, развитие на таланти и комерсиализация на ИИ технологии. За осигуряване на стратегическо ръководство е създаден Консултативен съвет по ИИ, който е съставен от водещи експерти от индустрията, академичните среди и гражданското общество и предоставя насоки за развитието и внедряването на ИИ в съответствие с канадските ценности.[444]

В областта на регулацията и стандартите се наблюдава активност на законодателно ниво с предложения Закон за изкуствения интелект и данните (AIDA), който е предложен като част от Законопроекта C-27 и цели да установи рамка за отговорно проектиране, разработване и внедряване на ИИ системи в Канада. Освен това е въведен Доброволен кодекс на поведение, който има за цел насърчаване на отговорното развитие и управление на напреднали генеративни ИИ системи от канадските компании. За повишаване на безопасността е създаден Канадски институт за безопасност на ИИ (CAISI), чиято цел е насърчаване на безопасното и отговорно развитие и внедряване на ИИ чрез изследвания и международно сътрудничество.[445]

Канада, също така, обръща внимание на инфраструктурните аспекти чрез Стратегия за суверенни изчислителни ресурси за ИИ, която е насочена към увеличаване на националния капацитет за изчисления, подкрепа на ИИ екосистемата и стимулиране на икономическия растеж чрез инвестиции в инфраструктура и ресурси. Не на последно

[443] Ibid
[444] "Artificial Intelligence Ecosystem", https://ised-isde.canada.ca/site/ised/en/artificial-intelligence-ecosystem 11.03.2025
[445] "Artificial Intelligence Ecosystem", https://ised-isde.canada.ca/site/ised/en/artificial-intelligence-ecosystem 11.03.2025

място, Канада активно участва в Международно сътрудничество, като играе водеща роля на световната сцена и си сътрудничи с международни партньори чрез участието си в Глобалното партньорство за изкуствен интелект (GPAI), за да гарантира, че използването на ИИ отразява канадските ценности, като защита на човешките права и демокрацията. Чрез тези координирани инициативи, Канада демонстрира ангажимент към осигуряване на отговорно и етично развитие и използване на ИИ технологии, което е в съответствие с нейните ценности и цели да бъде в полза на обществото и икономиката .

Панканадска стратегия

Панканадската стратегия за изкуствен интелект (ИИ) представлява ключова инициатива на правителството на Канада, която е насочена към насърчаване на приемането и развитието на ИИ в икономиката и обществото на страната. Стратегията има за основен фокус свързването на световно признатите канадски таланти и изследователски капацитети с програми за комерсиализация и приложение, с крайна цел да се гарантира, че канадските идеи и знания се мобилизират и комерсиализират в страната.[446]

Един от основните акценти на стратегията е комерсиализацията. В тази връзка се оказва подкрепа за Националните институти за изкуствен интелект с инвестиция от 60 милиона долара, разпределени по 20 милиона за всеки институт за период от пет години (2021-2026). [447] Тези институти имат за задача да подпомагат превръщането на изследванията в ИИ в комерсиални приложения и да увеличат капацитета на бизнеса да приема тези нови технологии. Допълнително се предвижда инвестиране на 125 милиона долара в Канадските глобални иновационни клъстери – Digital Technology, Protein Industries Canada, Next Generation Manufacturing Canada, Scale AI и Canada's Ocean Supercluster за насърчаване на приемането на канадски ИИ технологии в ключови индустрии и обществени организации за периода 2021-2026. Друг важен аспект на стратегията е разработването и приемането на стандарти. Чрез Съвета за стандарти на Канада се подкрепят усилията за разработване и приемане на стандарти, свързани с изкуствения интелект, с инвестиция от 8,6 милиона долара за пет години (2021-2026).[448]

Стратегията обръща сериозно внимание и на талантите и изследванията. CIFAR (Канадски институт за напреднали изследвания) разширява програмите за привличане, задържане и развитие на академични изследователски таланти и поддържа центрове за изследвания и обучение, като за тази цел е предвидена инвестиция от 208 милиона долара за десет години (2021-2031).[449]

Като цяло, стратегията има за основен фокус укрепването на позицията на Канада като лидер в областта на изкуствения интелект. Това се осъществява чрез комерсиализация на изследванията, разработването и приемането на стандарти и развитието на таланти и изследователски капацитет. Чрез тези целенасочени инициативи, Канада се стреми да гарантира, че иновациите и знанията в областта на ИИ се развиват и прилагат в полза на канадското общество и икономика.

Стратегия за изчислителните ресурси

Канадската стратегия за суверенни изчислителни ресурси за изкуствен интелект (ИИ) е инициатива на правителството на Канада, която е насочена към увеличаване на националния капацитет за изчисления, подкрепа на ИИ екосистемата и стимулиране на икономическия растеж. Основната цел на стратегията е да осигури на канадските

[446] "Pan-Canadian Artificial Intelligence Strategy", https://ised-isde.canada.ca/site/ai-strategy/en 11.03.2025
[447] Ibid
[448] "Pan-Canadian Artificial Intelligence Strategy", https://ised-isde.canada.ca/site/ai-strategy/en 11.03.2025
[449] Ibid

изследователи, иноватори и бизнеси достъп до необходимите изчислителни ресурси, като същевременно защитава данните и интелектуалната собственост на страната.[450]

Един от основните елементи на стратегията е мобилизирането на частни инвестиции. Правителството предвижда инвестиция до 700 милиона долара чрез инициативата AI Compute Challenge, която има за цел да подкрепи канадската ИИ екосистема чрез увеличаване на националния капацитет за ИИ изчисления. Тази мярка е насочена към стимулиране на частния сектор да участва активно в изграждането на необходимата изчислителна инфраструктура.[451]

Друг ключов елемент е изграждането на публична а инфраструктура от супер компютри. Предвидена е инвестиция до 1 милиард долара за изграждане на публична такава инфраструктура, която да бъде достъпна за изследователи, правителствени и индустриални организации. [452] Тази инвестиция ще осигури критично важен изчислителен капацитет за нуждите на научните изследвания и публичния сектор.

Стратегията включва и Фонд за достъп до ИИ изчисления. С инвестиция от 300 милиона долара, този фонд има за цел да преодолее високите разходи за ИИ изчислителни ресурси и ограничената наличност на национален капацитет, като предоставя директна подкрепа на канадски изследователи и разработчици в областта на ИИ. Тази мярка е насочена към улесняване на достъпа до изчислителни ресурси за по-малки изследователски групи и стартиращи компании.[453]

Канадската стратегия за суверенни изчислителни ресурси за ИИ е комплексен план, който съчетава публични инвестиции и стимули за частния сектор, с цел да се създаде мощна и сигурна изчислителна инфраструктура, необходима за развитието на ИИ в Канада. Фокусът върху суверенитета на данните и интелектуалната собственост е важен аспект, който цели да гарантира националната сигурност и конкурентоспособност в тази стратегическа област. Очаква се тази стратегия да има значителен положителен ефект върху иновациите и икономическия растеж, като предостави на канадските изследователи и бизнеси необходимите инструменти за разработване и внедряване на водещи ИИ решения.

Институт за безопасен ИИ

Канадският институт за безопасност на изкуствения интелект (CAISI) представлява инициатива на правителството на Канада, която е насочена към подкрепа на безопасното и отговорно развитие и внедряване на ИИ. Основната цел на CAISI е да насърчава науката за безопасността на ИИ в сътрудничество с международни партньори, за да се гарантира, че правителствата са добре подготвени да разбират и реагират на рисковете, свързани с напредналите ИИ системи.[454]

Един от основните акценти на CAISI е неговата изследователска дейност. Институтът провежда водещи изследвания в областта на безопасността на ИИ, като си сътрудничи с индустрията и Международната мрежа на институтите за безопасност на ИИ. CAISI изучава как работят напредналите ИИ системи, какви са рисковете и предоставя инструменти, за да гарантира, че правителството на Канада е добре подготвено да адресира възникващите рискове. [455] Тази изследователска дейност е

[450] "Canadian Sovereign AI Compute Strategy", https://ised-isde.canada.ca/site/ised/en/canadian-sovereign-ai-compute-strategy 11.03.2025
[451] Ibid
[452] Ibid
[453] Ibid
[454] "Canadian Artificial Intelligence Safety Institute", https://ised-isde.canada.ca/site/ised/en/canadian-artificial-intelligence-safety-institute 11.03.2025
[455] Ibid

ключова за изграждане на задълбочено разбиране на потенциалните опасности, свързани с ИИ.

По отношение на методите на изследване, CAISI използва приложни и изследователски проекти чрез CIFAR (Канадски институт за напреднали изследвания). Програмата за изследвания на CAISI, ръководена от CIFAR, гарантира, че най-новите международни и местни изследвания информират разбирането ни за рисковете от ИИ. Освен това, CAISI изпълнява проекти в подкрепа на правителствените приоритети относно безопасността на ИИ, включително съвместни тестове с международни партньори, и предоставя насоки на разработчиците и потребителите на ИИ, както и на обществеността, относно рисковете от напредналите ИИ системи и начините за тяхното адресиране. Като основател, на Международната мрежа на институтите за безопасност на ИИ CAISI активно координира усилията с международни партньори и сътрудничи по съвместни проекти, включително разработване на насоки за безопасността на ИИ. Това подчертава ангажимента на Канада към международното сътрудничество в областта на безопасността на ИИ.[456]

Канадският институт за безопасност на изкуствения интелект (CAISI) е стратегическа инициатива, която активно работи за насърчаване на науката за безопасността на ИИ чрез изследвания, сътрудничество и подкрепа на правителствени приоритети, като същевременно играе важна роля в глобалните усилия за осигуряване на безопасното развитие на напреднали ИИ системи.

Акта за ИИ и данните[457]

Законът за изкуствения интелект и данните (AIDA), внесен е през 2022г., представлява законодателна инициатива в Канада, чиято основна цел е да установи отговорно проектиране, разработка и внедряване на ИИ системи в страната, като се акцентира върху гарантирането на тяхната безопасност, справедливост и отчетност, особено за системи с висок риск. Един от основните акценти на AIDA е отговорността на бизнеса. Законът ще задължи компаниите да внедрят механизми за управление и политики, които да оценяват и адресират рисковете, свързани с техните ИИ системи, и да предоставят на потребителите достатъчно информация за информирани решения.[458] Това означава, че компаниите ще носят отговорност за идентифицирането и смекчаването на потенциалните вреди, произтичащи от техните ИИ продукти и услуги.

AIDA предвижда поетапен подход към прилагането на изискванията и въвежда гъвкавост на политиката, като предвижда, че безопасността на ИИ системите ще бъде пропорционална на свързаните с тях рискове; системите с по-висок риск ще подлежат на по-строги задължения.[459] Този подход, базиран на риска, позволява фокусиране на регулаторните усилия върху най-потенциално опасните приложения на ИИ. AIDA също така цели подобряване на отчетността, като ще осигури рамка, допълнена с подробни регулации, разработени чрез отворени и прозрачни консултации с обществеността и заинтересованите страни.[460] Този процес на консултации е важен за осигуряване на легитимност и ефективност на регулаторната рамка.

Законът за изкуствения интелект и данните (AIDA) в Канада представлява значителна стъпка към регулирането на ИИ, като поставя акцент върху отговорността на бизнеса, безопасността, справедливостта и отчетността, особено за ИИ системите с

[456] "Canadian Artificial Intelligence Safety Institute", https://ised-isde.canada.ca/site/ised/en/canadian-artificial-intelligence-safety-institute 11.03.2025
[457] Да се има предвид, че законът мина трето четене и все още няма индикации, че ще бъде одобрен
[458] "Artificial Intelligence and Data Act", 2022, https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act - 11.03.2025
[459] Ibid
[460] "Artificial Intelligence and Data Act", 2022, https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act - 11.03.2025

+исок риск. Законът възприема гъвкав и поетапен подход, който цели да балансира между насърчаването на иновациите и минимизирането на потенциалните вреди, като същевременно осигурява прозрачен процес на разработване на регулаци.

Насоки за използване на генеративен ИИ

Ръководството за използване на генеративен ИИ е създадено от правителството на Канада с предназначението да подпомогне служителите на канадското правителство при отговорното използване на генеративен ИИ инструменти. Неговата приложност обхваща всички федерални институции и служители, които обмислят или използват генеративен ИИ в своята работа. [461]

В основата на ръководството са основни принципи за отговорно използване, които се базират на съществуващите принципи на правителството за ИИ. Тези принципи включват зачитане на човешките права и ценности, прозрачност и обяснимост, отчетност сигурност и надеждност, и конфиденциалност и защита на данните.[462]

Ръководството предоставя специфични насоки за използване на генеративен ИИ, включващи оценка на риска, която трябва да се извърши преди използване на генеративен ИИ; избор на подходящ инструмент, според конкретната цел и която отговаря на изискванията за сигурност и защита на данните; проверката на резултатите, като те трябва, винаги трябва да бъдат проверявани и валидирани от човек, преди да бъдат използвани за вземане на решения или публикуване; осъзнаване на ограниченията на генеративния ИИ; защита на чувствителна информация; спазване на авторските права; прозрачност към потребителите, и не на последно място обучението и повишаването на осведомеността на служителите за отговорното използване на генеративен ИИ.[463]

Политиката на Канада за изкуствен интелект може да бъде определена като балансирана и гъвкава, въпреки че се наблюдава движение към законодателно регулиране с предложения Закон за изкуствения интелект и данните (AIDA), който цели установяване на рамка за отговорно проектиране, разработка и внедряване на ИИ системи, особено за тези с висок риск. Наличието на Доброволен кодекс на поведение за генеративен ИИ и създаването на Канадски институт за безопасност на ИИ (CAISI), който се фокусира върху изследвания и сътрудничество, също подкрепят идеята за по-гъвкав подход, който дава възможност за иновации, но същевременно обръща внимание на етичните и безопасни аспекти.

Въпреки че AIDA представлява потенциална регулаторна рамка, фактът, че законът все още не е одобрен и че съществуват множество други, по-скоро насочващи и подкрепящи инициативи, предполага, че Канада се стреми към баланс между стимулиране на иновациите и управление на рисковете, като избягва прекомерно строга регулация на този етап. Политиката е насочена към създаване на екосистема, която подкрепя отговорното и етично развитие на ИИ, като същевременно защитава обществените интереси и ценности.

Сингапур

Национална стратегия 2.0 от 2023г.

Националната стратегия за изкуствени интелект (NAIS 2.0) на Сингапур представлява амбициозен план, чиято основна цел е да трансформира страната в глобален лидер в областта на изкуствения интелект. Стратегията се фокусира не само

[461] "Guide on the use of generative AI", https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/guide-use-generative-ai.html 17.03.2025
[462] Ibid
[463] "Guide on the use of generative AI", https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/guide-use-generative-ai.html 17.03.2025

върху технологиите, но и върху създаването на значими ползи за обществото и икономиката. Основният акцент е да се използва ИИ като двигател за иновации и обществено благо, като същевременно се осигуряват условия за безопасно и етично прилагане на технологиите. Възможностите за растеж в ИИ индустрията се свързват с генерирането на нови работни места, което ще допринесе за социално и икономическо развитие на страната.[464]

Стратегията поставя акцент върху няколко ключови компонента и системи, които обединяват усилията на индустриалния сектор, правителството и академичните среди. Тези усилия ще позволят създаването на интегрирани системи и инфраструктури, които да подкрепят ефективното внедряване на ИИ в различни области. Важен аспект на стратегията е развитието на облачни изчисления и ИИ технологии, които ще предоставят стабилна основа за бъдещите иновации. Осигуряването на безопасна и доверена среда за разработка и внедряване на ИИ ще минимизира рисковете, свързани със сигурността и етиката на новите технологии.[465]

Друг ключов елемент на стратегията е образованието и развитието на таланти. Въпреки значителния напредък в ИТ сектора, Сингапур инвестира активно в обучение и квалификация на специалисти, които да отговорят на нарастващото търсене на експерти в областта на ИИ. Стратегията включва различни инициативи за развитие на умения както за местни кадри, така и за привличане на международни таланти. Това ще способства за създаването на устойчиви работни места и ще подпомогне икономическата диверсификация чрез новите технологии.[466]

Развитието на национални бази данни и ресурси е друга основна част от стратегията. Събирането и интегрирането на данни от различни сектори ще осигури на ИИ алгоритмите възможността да се адаптират и оптимизират според нуждите на различни индустрии като здравеопазване, транспорт, енергийни решения и други. По този начин ще се създадат нови ИИ решения и услуги, които ще ускорят иновациите в ключови области на обществото. Сингапур, също така, активно участва в глобални инициативи и платформи за регулиране на ИИ, което укрепва репутацията му като лидер в тази сфера. Чрез тези партньорства страната не само ще обменя знания и опит с други нации, но и ще подпомага изграждането на глобални етични и регулаторни стандарти за ИИ. Това ще гарантира безопасното и ефективно използване на технологиите и ще съдейства за глобалното сътрудничество в областта на ИИ. [467]

Един от основните фактори за успех на стратегията е инвестицията в научни изследвания и разработки, които ще стимулират иновациите в ИИ. Подкрепата за академични изследвания и стимули за частния сектор да инвестира в новаторски технологии ще осигури основа за конкурентоспособност на Сингапур в глобален мащаб. Така страната ще може да се утвърди като водещ източник на иновации и разработки в сферата на ИИ.[468] Сингапур, също така, поставя силен акцент върху регулациите и етичните стандарти. Стратегията включва мерки за контрол върху използването на ИИ, като се осигурява, че технологиите ще се използват в съответствие с високи етични норми, като същевременно се минимизират рисковете за личната сигурност и правата на гражданите.[469]

В заключение, Националната стратегия за изкуствени интелект на Сингапур (NAIS 2.0) съчетава амбициозни цели за развитие и внедряване на ИИ с дългосрочни планове за осигуряване на етични и безопасни условия за тяхното прилагане. Стратегията

[464] "NAIS 2.0 Singapore National AI Strategy", 2023, https://file.go.gov.sg/nais2023.pdf / 19.03.2025
[465] Ibid., p.17-31
[466] Ibid., p.33-45
[467] "NAIS 2.0 Singapore National AI Strategy", 2023, https://file.go.gov.sg/nais2023.pdf / p.47-63
[468] Ibid., p.17-31
[469] Ibid., p.47-63

включва не само технологични и инфраструктурни усилия, но и акцентира върху създаването на необходимите човешки и организационни ресурси. Чрез активно участие в глобалните дискусии и сътрудничество със световните лидери в ИИ, Сингапур се позиционира като лидер в глобалната ИИ екосистема. Това ще доведе до икономически растеж, подобряване на качеството на живот и същевременно ще укрепи общественото доверие в новите технологии.

Второ издание на Рамката за управление на изкуствения интелект

Второто издание на Рамката за управление на изкуствения интелект на Сингапур има за цел да насърчи отговорното внедряване на ИИ.[470] Рамката се основава на два основни принципа, които служат като насоки за организациите, разработващи и използващи ИИ. Тези принципи са прозрачност и обяснимост на ИИ при вземането на решения и всички решения да поставят човека в центъра. По отношение на регулаторния подход, Рамката на Сингапур изглежда възприема по-лек подход, който цели подпомагане на иновациите чрез насърчаване на отговорното поведение, вместо налагането на строги, предписващи регулации. Рамката е гъвкав и адаптивен набор от принципи, който позволява на организациите да прилагат ИИ в различни контексти, като същевременно се придържат към основните ценности. Рамката на модела се фокусира основно върху четири широки области, а именно: вътрешни управленски структури и мерки; човешко участие, в процеса на вземане на решения с ИИ; управление на операциите и взаимодействие; комуникация със заинтересованите страни.[471]

Рамката за управление на ИИ на Сингапур изглежда цели да намери баланс между насърчаването на иновациите и осигуряването на отговорно и етично използване на ИИ. Въпреки че акцентът е върху гъвкавостта и подпомагането на иновациите, принципите, свързани с отговорност, справедливост, защита на данните и етика, могат да наложат определени ограничения и допълнителни изисквания към бизнеса, но с това да гарантират социална справедливост. Рамката насърчава организациите да бъдат проактивни в управлението на рисковете и етичните аспекти на ИИ, което в дългосрочен план може да доведе до по-устойчиво и обществено приемливо развитие на тази технология.

Рамка за споделяне на данни

"Рамката за споделяне на данни" на Infocomm Media Development Authority (IMDA) на Сингапур има за цел да улесни споделянето на данни по начин, който е сигурен, етичен и взаимноизгоден за всички участници. Рамката за споделяне на данни е структурирана в четири основни части, като организациите имат възможност да избират кои части да използват в зависимост от техните нужди, като е важно да се отбележи, че нито една част не трябва да бъде пропусната. Тази структура осигурява гъвкавост при прилагането на принципите за споделяне на данни, като същевременно гарантира, че всички основни аспекти на процеса са разгледани.[472]

Част 1 "Стратегия за споделяне на данни" е насочена към разбирането на какви данни е необходимо да се споделят, как да се оценят тези данни и какви модели на споделяне могат да бъдат приложени. Чрез разработването на ясна стратегия организациите ще могат да идентифицират ключовите данни, които ще им осигурят конкурентни предимства, и да създадат ефективни механизми за тяхното управление и обмен. Важно е организацията да има ясна представа не само за техническата част на споделянето на данни, но и за стойността, която тези данни могат да предоставят, като

[470] "Model Artificial Intelligence Governance Framework" ed.2, 2020, https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf 19.03.2025
[471] "Model Artificial Intelligence Governance Framework" ed.2, 2020, https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf
[472] "Trusted Data Sharing Framework", 2019, https://www.imda.gov.sg/-/media/imda/files/programme/ai-data-innovation/trusted-data-sharing-framework.pdf

се вземат предвид нуждите на различните заинтересовани страни. Тази част от рамката е препоръчителна за всички ключови вземащи решения, заинтересовани лица и вътрешни екипи, които са пряко ангажирани в процеса на споделяне на данни.[473]

Част 2 "Правни и регулаторни съображения" от рамката разглежда изискванията за съответствие и регулиране, които трябва да бъдат спазвани при споделянето на данни. Организациите трябва да имат ясни насоки относно правните отношения, които ще се изградят между страните, участвали в споделянето на данни. Това включва не само въпроси, свързани със защитата на личните данни и съответствието с регулациите за конфиденциалност, но и определянето на правата и отговорностите на всички участници в процеса. Тази част е важна за всички, които ръководят проекти, както и за екипите, които събират, управляват и използват данни, като същевременно се консултират с правни и регулаторни експерти.[474]

Част 3 "Технически и организационни съображения" разглежда техническите аспекти и механизми за прехвърляне на данни между организации. Това включва избор на подходящи технологии и инфраструктури за сигурно и ефективно прехвърляне и обработка на данни. В допълнение, организациите трябва да вземат предвид въпросите, свързани с интеграцията на различни технологични платформи и осигуряването на съвместимост между системите за обмен на данни. Осигуряването на правилната технологична подкрепа и подходящата организация на процесите е от съществено значение за успешното и безопасно споделяне на данни. Тази част е препоръчителна за потребители, които ръководят проекти и бизнес единици, които активно събират, управляват и използват данни, както и за екипи, занимаващи се с технически и рискови въпроси.[475]

Част 4 "Операционализиране на споделянето на данни" от рамката се фокусира върху действията и съображенията, които трябва да се предприемат, след като данните вече са били споделени. Това включва въпроси, свързани с мониторинга на използването на данните, осигуряване на безопасността им и оценка на ефективността на процесите за споделяне. Организациите трябва да са готови да извършват редовни прегледи и актуализации на механизмите за споделяне на данни, като същевременно запазват съответствието с всички приложими регулации и етични стандарти. Тази част е препоръчителна за потребители, които управляват проекти, както и за консултантски екипи, които се занимават с правни, технически и рискови въпроси.[476]

Общата структура на рамката предоставя на организациите задълбочено разбиране на целия процес на споделяне на данни, като се гарантира, че всички важни аспекти – от стратегическо планиране и правни съображения до техническа реализация и оперативни процеси – са покрити и анализирани. Рамката за споделяне на данни с доверие на Сингапур представлява балансиран подход, който има за цел да насърчи иновациите чрез улесняване на споделянето на данни, като същевременно се обръща внимание на важни аспекти като сигурност, етика и защита на личните данни.

Институт за безопасен ИИ

Институтът за безопасен ИИ (SGAISI) е национален институт, който има за цел да стимулира развитието и внедряването на изкуствен интелект в Сингапур. Една от основните дейности на SGAISI е провеждането на задълбочени изследвания в областта на ИИ, обхващащи широк спектър от теми.[477] Други важна дейност е разработването на

[473] "Model Artificial Intelligence Governance Framework" ed.2, 2020, https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf p.17-26
[474] Ibid., p.27-37
[475] "Model Artificial Intelligence Governance Framework" ed.2, 2020, https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf p.39-41
[476] Ibid., p.49-51
[477] "Research", https://sgaisi.sg/our-works/ 19.03.2025

политики, свързани с ИИ,чрез предоставянето на експертни съвети на правителството и други заинтересовани страни с цел изграждането на ИИ екосистема; тестване на системи и модели, както и разработване на различни инструменти за тестване с цел подобряване на безопасността. [478]

Институтът за безопасен ИИ (SGAISI) е активна организация, която изпълнява своята мисия чрез провеждане на изследвания, разработване на стратегии и политики, изграждане на екосистема и предприемане на мерки за подкрепа на развитието и внедряването на ИИ в Сингапур.

Нови инициативи за безопассност, обявени по време на AI Action Summit

В началото на 2025 г. Сингапур обяви нови инициативи за управление на ИИ с цел повишаване на безопасността на ИИ. Тези инициативи бяха обявени от министъра на дигиталното развитие и информацията Жозефин Тео на AI Action Summit, в Париж. Тези мерки включват Global AI Assurance Pilot за най-добри практики около техническото тестване на приложения с генеративен ИИ, съвместен доклад за тестване, разработен заедно с японския Институт за безопасен ИИ, и публикуване на доклада за оценка на AI Safety Red Teaming Challenge в Сингапур за 2025 г., която цели да разбере как LLMs се представят по отношение на различни езици и култури в Азиатско-тихоокеанския регион и дали защитните мерки са ефективни в тези контексти.[479]

Посоката на развитие на регулациите в Сингапур е към създаване на балансирана и адаптивна рамка, която подкрепя иновациите чрез насърчаване на етични практики и отговорно използване на ИИ, а не чрез налагане на строги, задушаващи иновациите регулации. Този подход има потенциал значително да допринесе за икономическото развитие на страната, като я позиционира като привлекателен център за ИИ изследвания, разработки и внедряване, привличайки инвестиции и таланти. Акцентът върху безопасността и доверието ще бъде от решаващо значение за устойчивото развитие на ИИ и за гарантиране, че тази технология носи реални ползи за икономиката и обществото на Сингапур в дългосрочен план.

Япония

Социални принципи за ИИ ориентиран към човека

Концепцията за изкуствен интелект на Япония, [480] ориентиран към човека, подчертава важността на социалните принципи и ключовите съображения при неговото развитие и внедряване. Визията на Япония за ИИ е такава, в която ИИ служи за благото на хората и обществото като цяло. Стратегията акцентира върху необходимостта от създаване на ИИ системи, които са безопасни, сигурни, етични и прозрачни, и подчертава, че развитието на ИИ трябва да се ръководи от принципи като уважение към човешкото достойнство и права, както и отговорност и устойчивост.[481]

Един от основните аспекти, разгледани в документа, е важността на етиката в ИИ, като се посочва необходимостта от разработване на етични насоки и рамки, които да регулират проектирането, внедряването и използването на ИИ технологии. Тези насоки трябва да гарантират, че ИИ системите не дискриминират, че зачитат неприкосновеността на личния живот и че решенията им могат да бъдат обяснени. Друг

[478] Ibid., SGAISI
[479] "Singapore Announces New AI Safety Initiatives at the Global AI Action Summit in France ", 2025, https://www.mddi.gov.sg/media-centre/press-releases/singapore-announces-new-ai-safety-initiatives/ 13.03.2025
[480] "Social Principles of Human-Centric AI ", https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/humancentricai.pdf
[481] Ibid

ключов елемент е образованието и осведомеността, като се подчертава нуждата от повишаване на разбирането на обществото за ИИ и неговите потенциални въздействия В документа се обръща внимание и на регулаторната рамка за ИИ. Отбелязва се, че съществуващите правни и регулаторни механизми може да не са адекватни за справяне с всички предизвикателства, породени от ИИ. Поради това, се предлага обмислянето на нови подходи към регулирането на тази технология, които да балансират между насърчаването на иновациите и защитата на обществените интереси. Насоките подчертават значението на международното сътрудничество в областта на ИИ, като се призовава за съвместни усилия между различни страни и организации за обмен на знания, най-добри практики и за разработване на общи стандарти и принципи за ИИ.[482]

Документът "ИИ ориентиран към човека: социални принципи и ключови съображения" (Human-centric AI: Social Principles and Key Considerations) представя визия за развитие на ИИ, която поставя човека в центъра, като набляга на етични съображения, образование, регулация и международно сътрудничество. Той очертава необходимостта от проактивни мерки за осигуряване на това, че ИИ ще служи за благото на обществото.

Институт за безопасен ИИ (J-AISI)

J-AISI е организация, създадена в Япония с цел да бъде център за стратегически изследвания и формулиране на политики в областта на изкуствения интелект.[483] AISI играе ключова роля в подпомагането на правителството, индустрията и академичните среди в Япония при разработването и внедряването на ИИ технологии.

J-AISI играе критична роля в подкрепа на правителството чрез провеждане на обширни проучвания за безопасността на ИИ, анализиране на методи за оценка и установяване на съответни стандарти. Като централно звено за безопасността на ИИ в Япония, то консолидира най-новата информация от индустрията и академичните среди, като същевременно улеснява сътрудничеството между различни свързани компании и организации. Важно е да се отбележи, че J-AISI не се занимава с научноизследователска и развойна дейност (НИРД), а вместо това се фокусира върху насърчаване на координацията и споделянето на информация в ИИ сектора.[484]

Допълнително, J-AISI цели да установи международен консенсус относно безопасността на ИИ чрез сътрудничество с организации за безопасност на ИИ в други държави, като гарантира, че подходът на Япония е в съответствие с глобалните стандарти. Обхватът на организацията е проектиран да бъде гъвкав, като обхваща разнообразие от въпроси, свързани с ИИ, и като същевременно отчита по-широките глобални тенденции. Тези тенденции включват социалното въздействие на ИИ, моделите на управление, развитието на ИИ системи, както и управлението и регулирането на съдържанието и данните, свързани с ИИ. Чрез този стратегически подход, J-AISI се позиционира като ключов играч в оформянето на бъдещето на безопасността на ИИ на глобално ниво.[485]

J-AISI е ключова институция за развитието на безопасността на ИИ в Япония, която успешно координира усилията на правителството, индустрията и академичните среди. Чрез изследвания, анализи и сътрудничество с международни организации J-AISI формира глобални стандарти и политики за безопасност на ИИ. Неговият фокус върху

[482] "Social Principles of Human-Centric AI ", https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/humancentricai.pdf
[483] "Japan AI Safety Institute", https://aisi.go.jp/ 20.03.2025
[484] "Overview of the Japan AI Safety Institute (J-AISI)", 2025, https://aisi.go.jp/assets/pdf/20250106_AISI_en.pdf
[485] "Overview of the Japan AI Safety Institute (J-AISI)", 2025, https://aisi.go.jp/assets/pdf/20250106_AISI_en.pdf

координация и обмен на знания, а не върху НИРД, го позиционира като стратегически играч в глобалното регулиране и развитие на ИИ технологиите.

Япония демонстрира стратегически ангажимент към развитието и внедряването на ИИ като ключов фактор за бъдещия си просперитет. Фокусът върху "човекоцентричен" ИИ (от текста за "Human-centric AI") показва, че Япония има за цел да развива тази технология по начин, който е в съответствие с обществените ценности и нужди, което може да доведе до по-широко приемане и доверие в ИИ системите.

Политиките и инициативите създават благоприятна среда за развитието на ИИ в Япония. Комбинацията от етични насоки, улесняващо авторско право, стратегически правителствени инициативи и специализирани институции като IPA и J-AISI, има потенциала значително да повиши иновационния капацитет и конкурентоспособността на японската икономика в глобалната ИИ надпревара. Тези усилия могат да доведат до повишаване на производителността, създаване на нови индустрии и подобряване на качеството на живот в Япония, но изискват продължителни инвестиции и стратегическо изпълнение.[486]

Кения

Кения е една от страните, които имат представители в Международната мрежа на институти за безопасен ИИ. Страната засилва своите икономически позиции чрез множество международни проекти и инициативи в областта на ИИ, което е в резултат на проактивни политики в областта. Най-яркият пример за такъв проект е проектът G42 между Microsoft и MGX – инвестиционен фонд, базиран в ОАЕ, на стойност 1 млрд щ.д., който цели създаване на цялостна дигитална екосистема в Кения.[487] Тези два международни партньора участват активно в инвестирането на високотехнологични проекти в целия свят, което е показателно за привлекателността на държави като Кения, които работят активно в посока на своята дигитална трансформация чрез про-иновационни политики.

Проектът за Национална стратегия на Кения за изкуствен интелект има амбициозната визия да превърне страната в "център за иновации, водещ в Африка, използващ ИИ за устойчиво развитие и просперитет".[488] Мисията на стратегията е „да създаде благоприятна екосистема за отговорно развитие и внедряване на ИИ в Кения, стимулирайки икономически растеж, подобрявайки обществените услуги и повишавайки качеството на живот", като в основата на стратегията стоят ключови принципи като "ориентираност към човека", "приобщаване и справедливост", "етика и отговорност", "прозрачност и обяснимост", "сигурност и безопасност", "устойчивост" и "сътрудничество".[489] Тези принципи подчертават ангажимента на Кения към отговорно и етично развитие на ИИ, което е от съществено значение за изграждане на обществено доверие и осигуряване на ползи за всички граждани.

Стратегията е структурирана около шест стратегически приоритета, които включват "развитие на ИИ таланти и умения", "създаване на благоприятна екосистема

[486] Habuka, H., "New Government Policy Shows Japan Favors a Light Touch for AI Regulaton", Center for Strategic and International Studies, 2025, retrieved from https://csis-website-prod.s3.amazonaws.com/s3fs-public/2025-02/250225_Habuka_Japan_AI_0.pdf?VersionId=aD5DJCIrllfgGup3x_ojtmE_vn4Y8NTP

[487] "Microsoft and G42 announce $1 billion comprehensive digital ecosystem initiative for Kenya ", 2024, https://news.microsoft.com/2024/05/22/microsoft-and-g42-announce-1-billion-comprehensive-digital-ecosystem-initiative-for-kenya/ 20.03.2025

[488] "Kenya National Artificial Inteliligence (AI) Strategy 2025 – 2030 (draft)", 2025 https://ict.go.ke/sites/default/files/2025-01/Kenya%20National%20AI%20Strategy%20%28Draft%29%20for%20Public%20Validation%20%20%5B14-01-2025%5D.pdf

[489] Ibid "Kenya National AI Strategy"

за иновации в ИИ", "развитие на ИИ инфраструктура и данни", "насърчаване на внедряването на ИИ в ключови сектори", "установяване на етични и правни рамки за ИИ" и "насърчаване на международно сътрудничество".[490] Тези стълбове обхващат ключови аспекти, необходими за успешното развитие и внедряване на ИИ, от образователната подготовка до регулаторната рамка и международното сътрудничество.

Стратегията идентифицира приоритетни сектори за внедряване на ИИ, които имат пряко въздействие върху социално-икономическото развитие на страната, като "селско стопанство", "здравеопазване", "образование", "финансов сектор" и "транспорт и логистика".[491] Фокусът върху тези сектори показва прагматичен подход към използването на ИИ за решаване на конкретни проблеми и подобряване на живота на хората.

Планът за действие очертава конкретни области като разработване на учебни програми и курсове за ИИ", "създаване на инкубатори и акселератори", "инвестиране в изграждане на центрове за данни и облачна инфраструктура" и "създаване на регулаторни органи" .

Предложеният "поетапен подход" и предвидените "механизми за мониторинг и оценка" са важни за осигуряване на ефективното изпълнение на стратегията и измерване на нейния успех.[492]

Като цяло, стратегията на Кения за ИИ демонстрира ясен ангажимент към използването на тази технология за постигане на устойчиво развитие и просперитет, като същевременно подчертава важността на етиката, приобщаването и сътрудничеството, с цел да положи основите за отговорно и устойчиво развитие и внедряване на изкуствен интелект, което да доведе до значителни социално-икономически ползи за страната и обществото.

[490] "Kenya National Artificial Inteliligence (AI) Strategy 2025 – 2030 (draft)", 2025 https://ict.go.ke/sites/default/files/2025-01/Kenya%20National%20AI%20Strategy%20%28Draft%29%20for%20Public%20Validation%20%20%5B14-01-2025%5D.pdf
[491] Ibid
[492] Ibid

Приложение 4: Прогнозни сценарии относно регулациите

Прогнозните сценарии са дадени единствено на база състояние към момента относно регулаторните рамки на съответните държави. Те не отразяват обезпеченост с ресурси, търговска политика, технологичен суверенитет.

Прогнозно въздействие на регулаторните рамки до 2040 г., ако тенденциите в политиките бъдат запазени:

Държава	*Пазар На Труда*	*Конкурентоспособност*	*Технологичен Капацитет*	*Внедряване*	*Дипломатическа Роля*
САЩ	Съществени трансформации, висока автомати-зация, засилена нужда от преквали-фикация	Лидер в авангарден ИИ, конкурентно предимство в генеративен ИИ	Доминиция в ИИ чипове, изчислителна мощ и частни инвестиции	Всички сектори – от здравеопазване до отбрана, водещи компании	Лидер в глобалните ИИ стандарти и алианси
ЕС	Подкрепена преквалификация, социални буфери, запазване на трудовите права	Силен в етичен ИИ и секторни стандарти, но по-бавна комерсиализация	Публична инфраструктура и открити модели, но зависимост от чужд хардуер	Публични услуги, индустрии, малки и средни предприятия	Активен в многостранни форуми, етичен хегемон
Китай	Масова автоматизация, бързо пренасочване към технологични и производствени роли	Висока скорост на внедряване, силен държавен стимул, напредък в хардуера	Мащабна инфраструктура и контрол върху ИИ екосистема	Масов секторен rollout в производство, транспорт, сигурност	Изгражда алтернативни глобални мрежи и норми
Република Корея	Фокус върху високотехнологични работни места, адаптиране на образованието	Силни позиции в роботика и индустриален ИИ, износ на решения	Целеви инвестиции в AI центрове и суперкомпютри	Здравеопазване, финанси, интелигентни градове	Регионален лидер в технологична дипломация
Великобри-Тания	Реорганизация на труда в публичния и частния сектор, растеж на дигитални умения	Висок клас изследвания, възход на ИИ стартъпи	Интегриран публично-частен подход към тестване и етика	Обществени услуги, сигурност, цифрова администрация	Инициатор на споразумения
Австралия	Умерени промени, насочване към адаптивност и етична интеграция	Конкурентност в етично управление и публични услуги	Стандартизи-рана инфраструктура и cloud-first подход	Граждански интерфейси, цифрово управление	Пример за умерен регулиращ подход, партньор в Азиатско-тихоокеанския регион

Канада	Етапно внедряване с акцент върху защита на работниците	Фокус върху отговорна комерсиализация и безопасност	Инвестиции в суверенни ИИ изчисления, държавна подкрепа	Правителствени услуги, научни институти, индустриални клъстери	Съучредител на глобалната AI мрежа за безопасност
Сингапур	Насърчаване на дигитална грамотност, селективно внедряване в ключови услуги	Гъвкаво регулиране, лидер в ИИ за управление и сигурност	Национална ИИ платформа, инфраструктура за тестване и сертифициране	Управление, киберсигурност, трансгранична търговия	Мост между изтока и запада, интегративен подход
Япония	Бавен преход с фокус върху запазване на социалния баланс	Приложна конкурентоспособност в ниши (здраве, индустрии)	Подкрепа за университетски изследвания и индустриално внедряване	Здраве, автомобилен сектор, интелигентни системи	Активна в G7, съюзник на демократичните ИИ политики
Кения	Фокус върху създаване на нови възможности и микропредприемачество	Регионално лидерство чрез партньорства и достъпност	Международни инвестиции, изграждане на дигитална инфраструктура	Селско стопанство, образование, логистика	Говорител за Глобалния Юг, застъпник за достъп и справедливост

Прогнозно развитие за геоикономическото и технологичното влияние на регионите до 2040 г.

Държава	*Геоикономическо влияние до 2030 г.*	*Геоикономическо влияние до 2035 г.*	*Геоикономическо влияние до 2040 г.*	*Технологично влияние до 2030 г.*	*Технологично влияние до 2035 г.*	*Технологично влияние до 2040 г.*
САЩ	Глобален хегемон	Запазва водеща роля, но с конкуренция	Доминиращ в повечето ИИ вериги	Глобален лидер	Фронтова линия при AGI	Доминиращ в авангардни ИИ и автономни системи
ЕС	Регионален лидер	Засилено чрез регулации и инвестиции	Консолидиран лидер в регулации и иновации	Регулаторен модел	Конкурент в публичен ИИ	Глобален стандартен архитект
Китай	Глобален контра-хегемон	Разширено чрез инициативи като BRI и износ на ИИ	Вторият глобален полюс в ИИ икономиката	Глобален технолидер с рискове	Засилване чрез вътрешни платформи	Интегрира глобални мрежи чрез ИИ
Южна Корея	Високо регионално значение	Позиционира не като сигурна алтернатива	Глобален експертен хъб по ИИ безопасност	Стабилен и адаптивен	Силна автономна индустрия	Лидер в безопасен ИИ и чипове
Великобри-Тания	Високо в англоезичната мрежа	Етичен лидер с нови коалиции	Транснационален лидер в ИИ етика и стандарти	Силен в етични и приложни системи	Първичен в ИИ тестване и етика	Регулаторна технологична сила
Австралия	Средно, но стабилно	Активен партньор в тихоокеанската зона	Междуконтинентален ИИ партньор	Фокус върху приложен ИИ	Партньор по безопасни решения	Иновативен публичен ИИ оператор

Канада	Канадска меко-влияеща сила	Нарастваща роля чрез GPAI и съюзи	ИИ хъб за североамерикан ска демократична зона	Генеративен ИИ и държавна безопасност	Лидер по социално отговорен ИИ	Научно-технологичен хъб
Сингапур	Азиатски възел	Стабилен дипломатически център	Глобален microstate за AI стандарти и безопасност	Малки, но прецизни приложения	Доверен център за тестове за генеративен ИИ	Символ на доверен ИИ
Япония	Стабилен, но ограничен	Институционално разпознаване в Азия	Институционализиран фактор в Азиатско-тихоокеанския регион	Силен в индустриални решения	Ясно диференцирани ниши	Технологичен ментор в Азия
Кения	Растящ регионален хъб	Африкански дипломатически активист	Ръководещ Африкански консорциум по ИИ	Начален, развиващ се	Възникващ генеративен капацитет	Глобален глас от глобалния юг

Прогнозни сценарии за конкурентноспобсността и технологичното развитие (2035 г.)

Държава	*Най-вероятен сценарий*	*Най-добър сценарий за ес и отражение върху останалите региони*	*Най-лош сценарий за ес и отражение върху останалите региони*	*При запазване на свръхрегулаторен подход в ес*	*При преминаване към балансиран регулаторен подход в ес*
ЕС	Бавен напредък поради регулаторна тежест, частични успехи в критични сектори	Балансирана регулация + инвестиции → глобален еталон за доверен ИИ	Свръхрегулация → технологичен отлив, спад в конкурентоспособността	Забавяне на ИИ внедряване, отслабване на индустриални капацитети	Съчетание на доверие и иновации, икономически възход
САЩ	Гъвкави политики, доминация в авангарден ИИ, укрепени позиции на пазара	ЕС партньор за иновации → задълбочено трансатлантическо сътрудничество	Ограничен достъп до ЕС → ръст на вътрешен капацитет	Укрепване на позиции чрез разлика в гъвкавост	ЕС-САЩ съвместни ИИ рамки → ко-лидерство
Китай	Глобална експанзия, интензивни инвестиции, регионално лидерство	ЕС пазарен контрапункт → конкурентност без конфронтация	Използване на вакуума → геоикономическа доминация	Възползване от регулаторната празнина, увеличен дял на пазара	Взаимна координация в стандартите и внедряването
Южна Корея	Фокус върху индустриално приложение, силна експортна ИИ база	ЕС стандарт → засилен обмен на технологии и регулации	Загуба на доверие в ЕС → ориентиране към САЩ/Китай	Засилен експорт към нерестриктивни пазари	Технологично сътрудничество ЕС-Корея

Великобри-Тания	Адаптивен модел, нишова лидерска роля, високо доверие	Съгласувани рамки → по-висока съвместимост	Изолираност на ЕС → Великобритания укрепва глобалната си роля	Позициониране като балансирана алтернатива	Съвместими етични и правни рамки → по-широк пазар
Австралия	Умерен напредък с фокус върху етично внедряване	ЕС модел → засилване на етичните стандарти в региона	Загуба на съвместимост → регионален подход	Минимизиране на съвместимостта с ЕС технологии	Интеграция на ESG и ИИ с ЕС практики
Канада	Стабилен напредък, силна интеграция в обществени системи	Координирани регулации → общи R&D усилия	Ограничено сътрудничество → ориентация към САЩ	Ограничена колаборация с ЕС	Съвместни програми и достъп до ЕС изследователск и хъбове
Сингапур	Иновативен хъб за тестване, лидер в ИИ етика в Азия	Улеснен достъп до пазара на ЕС → ръст на експорта	ЕС неефективен → Сингапур като модел за глобален Юг	Изместване към азиатски и американски пазари	ЕС като партньор за тестване на отговорен ИИ
Япония	Приложение в индустрия и услуги, стабилна научна основа	Интероперативн ост → научна синергия	Изоставане на ЕС → научна миграция към Азия/САЩ	Разминаване с ЕС → засилване на връзки с глобалния Юг	Възможности за участие в европейски проекти
Кения	Силен напредък в приоритетни сектори чрез международни партньорства	Партньорски програми → технологичен трансфер и обучение	ЕС отпада като партньор → ориентация към Китай/САЩ	Отслабване на достъпа до ЕС проекти	Разширен достъп до финансиране и инфраструктура от ЕС

Приложение 5: Сценарии относно икономиката на бъдещето, задвижвана с ИИ

1. **Напълно автономни тъмни фабрики**: В този сценарий фабриките по света оперират почти изцяло автономно, с минимална или нулева човешка намеса на място. Такива фабрики на тъмно са напълно автоматизирани производства, които могат да работят денонощно без осветление и без персонал. Роботи и ИИ-системи изпълняват всички рутинни операции от обработка и сглобяване до вътрешна логистика и контрол на качеството. Човешката роля се измества главно към дистанционен надзор, поддръжка и стратегическо управление на производството. Машините са оборудвани със сензори, взаимосвързани са в интелигентни мрежи и се саморегулират чрез алгоритми за машинно обучение. Продукцията тече непрекъснато, като „светлините са изгасени" – символ, че не са нужни работници на място. Това води до безпрецедентна ефективност: минимизиране на човешките грешки, елиминиране на трудови паузи и оптимизиране на скоростта на производствения цикъл. Автономното производство обхваща различни индустрии от машиностроене и автомобилостроене през електроника до складове и дистрибуционни центрове. **Когнитивни ИИ модели** ще наблюдават процесите и самостоятелно ще оптимизират параметрите на производство в реално време. Правителствата и индустриалните стандарти до 2040 г. вероятно ще се адаптират към тази нова реалност. Ще бъдат разработени **сертификационни режими за автономни фабрики** – например, стандарти за безопасност на роботите, които работят без наблюдение. Могат да се появят и регулации за минимално присъствие на човешки контрол при критични производства. **Икономическите ефекти** от изцяло автономното производство са значителни. От една страна, продуктивността и капацитетът на глобалната индустрия се увеличават многократно – заводите работят 24/7 без прекъсване, което повишава БВП и икономическия растеж в страните, възприели масова автоматизация. Разходите за труд на единица продукция спадат драстично, което поевтинява много стоки. Компаниите реализират по-високи печалби, благодарение на по-ниските постоянни разходи и по-малкото производствени спирания. Веригите на доставки също стават по-устойчиви, тъй като автономните заводи могат бързо да се „преконфигурират" за производство на различни компоненти, смекчавайки зависимостта от единични доставчици. В същото време, обаче, социалните последици са сложни. Масовата автоматизация води до мащабно елиминиране на работни места за оператори, техници и други нискоквалифицирани позиции в индустрията. Това поставя пред държавите въпроса за реформи на трудовата и образователната политика.

2. **ИИ като равноправен инженер**: При този сценарий ИИ навлиза повсеместно в ролята на колаборатор и коинженер – човекът и изкуственият интелект работят рамо до рамо в симбиоза, вместо ИИ изцяло да замени човека. Производството през 2040 г. се характеризира с „умни" фабрики, в които всяко производствено решение е резултат от диалог между хора и ИИ. Вместо фабрики, лишени от хора, тук имаме високотехнологични работни екипи, в които инженери, техници и ИИ системи (вградени в машини или като софтуерни асистенти) проектират и произвеждат заедно. **ИИ в този сценарий действа като равноправен член на екипа** – наричат го понякога „цифров колега". На практика това означава, че при разработката на нов продукт, например, хората инженери задават целите и ограниченията, а ИИ предлага множество варианти на дизайна за секунди. Хората оценяват и донагласят, след което автоматизирани симулации проверяват вариантите. Решението е взето съвместно: ИИ дава идеи извън рамките на човешкия опит, докато човекът прилага творческа преценка и контекст. Този сценарий предоставя балансирани икономически ползи – постига се увеличение на

производителността, но без драстична загуба на работни места, както в изцяло автоматизирания модел. Производителността расте благодарение на синергията: екип човек+ИИ може да свърши повече работа и да реши по-сложни проблеми. Разработката на нови продукти се ускорява значително, защото ИИ генерира варианти и открива оптимални решения много по-бързо.

3. **Производство като услуга, управлявано от ИИ конгломерати (глобални платформи):** този сценарий предвижда, че до 2040 г. производството се е трансформирало в глобална услуга, предоставяна от няколко могъщи управлявани от ИИ платформи или конгломерати. Представете си, вместо всяка компания да поддържа собствени фабрики, тя наема производствени капацитети и ресурси от гигантски платформи при поискване, които организират хиляди фабрики, роботи и доставчици по целия свят чрез изкуствен интелект. Производството се предоставя „като услуга" (Manufacturing-as-a-Service, MaaS) – фирмата просто подава дизайн или поръчка, а ИИ платформата намира оптималния начин да произведе и достави продукта – избира подходящи заводи в мрежата, координира доставки на материали, планира производство и логистика, контролира качеството чрез данни в реално време. В основата стоят облачни производствени мрежи, обединяващи множество разпръснати производители (от големи заводи до малки цехове и 3D принтери) в една **виртуална мегафабрична екосистема**. Тези мрежи се управляват от ИИ конгломерати – може да са корпорации, родени от днешните тех-гиганти, или нови лидери, които контролират платформите. Те притежават или най-малкото посредничат за капацитет, притежаван от други. ИИ алгоритми оптимизират натоварването: поръчките се насочват автоматично към свободни линии по света според критерии – цена, близост до клиента, специализация и др. **В този свят границите между отделните компании и заводи са размити** – голяма част от производството е насочено през тези платформи. Може да има няколко конкуриращи се глобални платформи: например, една доминирана от САЩ, друга - от Китай, или няколко специализирани по индустрии. Производството се превръща в услуга, подобно на облачните изчисления – мащабно, гъвкаво и получавано при поискване. Това позволява на стартиращи фирми бързо да влязат на пазара без голям начален производствен капитал – имат идея за продукт, използват MaaS платформа за производство и доставят глобално. ИИ конгломератите имат огромна роля – те оперират мрежите с помощта на собствени мощни ИИ. Тези ИИ агрегирани системи са способни да предвиждат глобалното търсене, да местят ресурси динамично (например, да пренасочат производство от един регион в друг при природно бедствие или търговска война). Конгломератите може да владеят и собствена инфраструктура: огромни напълно автоматизирани фабрики-хъбове, които отдават капацитет под наем. Но, също така, в мрежата влизат и независими доставчици, които се регистрират, за да получават поръчки. ИИ поддържа ранкинг и сертифициране на тези доставчици въз основа на показатели (качество, бързина, устойчивост). Регионалните различия са изгладени от платформите – поръчките се изпълняват там, където е оптимално. За клиентите по света това означава, че получават изделия бързо и евтино, без да се интересуват от коя държава идват. **„Произведено от CloudManufacturing Inc." може да замени етикетите по стоките.** Производството като услуга, също така, улеснява персонализацията: ИИ конгломератите могат да комбинират поръчки на много дребни серии за различни клиенти и да ги произвеждат ефективно в една мрежа, което единична фабрика трудно би направила. С други думи, светът се движи към „раздробено, но управлявано централизирано" производство. Това е радикална промяна в бизнес модела на индустрията – фирмите вече не инвестират в машини, а в дизайн, бранд и софтуер, оставяйки „физическата" част на мета-доставчиците, водени от ИИ. Това е

свързано с въвеждане на нови стандарти, нови митнически правила, внедряване на блокчейн, културна промяна и т.н.

4. **Когнитивно и етично регулирано производство**: В този сценарий развитието на производството до 2040 г. е ръководено от принципа „на първо място е безопасността и етиката". Изкуственият интелект се внедрява широко, но под строг регулаторен контрол и с въведени етични рамки на всяко ниво. Правителствата и международните органи са наложили ясни правила как, къде и в каква степен може да се използва ИИ в производството. „Когнитивно регулирано" означава, че самите регулаторни системи използват интелигентни технологии като държавни надзорни ИИ системи, които следят алгоритмите на заводите за съответствие с правилата. „Етично регулирано" означава, че приоритизирано е отговорното и прозрачно използване на ИИ – всяко решение, което ИИ взема във фабриката, трябва да бъде обяснимо и справедливо, а за определени чувствителни решения е задължително наличието на човешко одобрение. Производствените компании в този свят не могат просто да внедрят най-новия алгоритъм. Те трябва първо да го сертифицират пред регулаторен орган, доказвайки, че е безопасен, няма склонност към дискриминация или непредвидими действия. Това е свят, в който законите настигат технологията. Вече съществуват всеобхватни закони за ИИ (подобни на Акта за ИИ на ЕС, но по-развити), стандарти ISO за етичен ИИ в индустрията и глобални споразумения за AI безопасност. В производствените халета виждаме ИИ и роботи, но винаги с някакъв „етичен сензор": например, колаборативен робот няма да заработи, ако около него няма човешки надзорник с оторизация. В този сценарий, зоните без регулация практически не съществуват.

5. **Зони без достъп на ИИ**: В този сценарий до 2040 г. се оформят „острови" – географски или секторни зони, където изкуственият интелект е силно ограничен или изцяло забранен в производството. Тези свободни от ИИ зони могат да бъдат цели държави, отделни региони или определени индустриални сектори, които съзнателно избират да поддържат традиционни методи на производство без висока автоматизация. Мотивите зад това могат да варират: в някои случаи са политически (правителства, които защитават заетостта и се боят от социални напрежения, решават да забранят роботизацията над определено ниво), в други са културни или етични (общности, които ценят ръчния занаят и обявяват своите продукти за „създадени от човека, не от машина"), а трети – икономически прагматични (регион с изобилие от евтина работна ръка предпочита да остане трудоемък, за да се отличава на пазара). В тези зони технологичното ниво умишлено се държи по-ниско: фабриките приличат на тези от началото на XXI век, с по-малко роботи, повече работници на поточните линии или при машините. Ако има компютъризация, тя е по-скоро за подпомагане на човека (напр., базови CAD системи, конвенционална автоматизация), но не и самостоятелно вземащи решения ИИ. В крайни случаи, могат да се приемат закони, забраняващи конкретно използването на ИИ алгоритми в управлението на производството или ограничаване броя на роботите на определен брой служители. В международен план подобни зони се превръщат в нишови производствени центрове – например, една държава рекламира своя текстилен сектор като „100% произведено от хора", апелирайки към брандове, които искат да се асоциират с етична заетост. Или регион обявява, че е „свободен от ИИ" и това привлича туризъм и интерес (образ на общност, съхраняваща традиционния начин на живот, подобно на днешните екологични резервати, но за труд без машини). Вътре в тези острови, социалният договор е: работниците запазват работните си места и човешкото достойнство (не са заменени от роботи), а в замяна приемат по-ниска производителност и заплати спрямо високотехнологичните отрасли. Някои такива зони

могат да са и „лаборатории" за човешки умения, като запазват занаяти и техники, които другаде са автоматизирани. Този свят обаче е фрагментиран. Не навсякъде отказват ИИ, а само отделни „анклави" в глобалната карта. Общата икономика извън тях продължава да се движи от ИИ, но тези зони си създават собствени алтернативни вериги на доставка. Те може да търгуват повече помежду си или да служат на специфичен пазарен сегмент. Този сценарий описва свят с двоен технологичен стандарт. Глобалната икономика върви с ИИ и автоматизация, но отделни територии умишлено спират на ниво 2020-те, създавайки „резервати на човешкия труд", защитени по един или друг начин от навлизането на изкуствен интелект. В зоните, които поемат по този път, технологичното развитие до голяма степен се "замразява" или канализира в различна посока. Вместо инвестиции в ИИ и роботика, до 2040 г. те инвестират в подобряване на човеко-машинните системи от стар тип – по-ергономични инструменти, безопасност на работното място, механизация (но не автоматизация). Може да видим ренесанс на междинни технологии: например, екзоскелети за работници (които дават сила, но работникът все пак контролира), усъвършенствани ръчни устройства с малко умни функции (но не пълна автономност). Тези зони могат дори да разработят свои технологии за ефективност, които не разчитат на ИИ, като специализирани механични приспособления, "глупави" автоматични машини, които обаче все пак изискват човек оператор. В известен смисъл, това са **алтернативни технологични траектории: ако не беше дошъл модерният ИИ, индустрията щеше да върви по тях**.

В *Таблица 5* са обобщени ключови характеристики на всеки сценарий и как различните региони биха се вписали в него:

Сценарий/ Регион	*ЕС*	*САЩ*	*Китай*	*Други водещи региони*
Изцяло автономно производ-ство	Частично реализиран. Автоматизацията напредва, но ЕС запазва известен човешки контрол заради регулации и синдикати.	Силен стремеж. Корпорациите масово внедряват тъмни заводи. Регулациите са по-слаби, така че САЩ постига висока степен на роботизация в много отрасли.	Доминантен лидер. Държавни програми ускоряват пълната дигитализация. Много "тъмни фабрики" са реалност към 2040.	Япония, Корея: много високо автоматизирани. Развиващите се страни: ограничено внедряват там, където чужди инвеститори финансират.
ИИ като равноправе н инженер	Водещ подход – ЕС възприема визията за човека в центъра. Фабриките интегрират ИИ, но хората запазват ключова роля. Регулации насърчават сътрудничеството човек-ИИ.	Широко разпространен във високотехнологич ните отрасли – Big Tech осигуряват ИИ инструменти на производителите. Културата цени иновацията, но пазарът сам определя баланса човек-ИИ.	Прилага се, но под контрол. Китайските предприятия използват ИИ за дизайн и оптимизация, обаче партийна политика гарантира човешки надзор. ИИ-коефициентът на инженерите е висок, но "човекът командва".	Япония: силен фокус върху колаборативни роботи и синергия човек-машина. Южна Корея, Сингапур: бързо приемат ИИ *co-pilots* в индустрията. Индия: стреми се към това, но варира (водещи фирми – да).
Производ-ство като	ЕС фирми участват в глобални	Тех гигантите създават основните MaaS	Създава своя затворена платформа с	Япония/Корея: присъединяват се към западни

услуга (MaaS)	платформи, но ЕС се тревожи за суверенитет – опитва собствена платформа (с по-етичен/зелен акцент). Някои по-слаби страни стават звена в мрежи на чужди платформи.	платформи. САЩ печелят от това, но местните МСП стават зависими от тези облачни услуги за производство.	подкрепа на държавата. Китайската мрежа обслужва вътрешен пазар и страни под китайско влияние.	платформи или развиват нишови (за специални компоненти). Развиващи се страни: включени като възли (предоставят капацитет) – например виетнамски фабрики работят по задачи от глобалната платформа.
Етично регулирано производ-ство	*Пионер*: ЕС въвежда строги правила всяка ИИ система да е сертифицирана, човешкият надзор е задължителен. Производството е по-безопасно, но по-бавно и по-неиновативно.	*Смесен подход*: появяват се федерални стандарти за безопасност, но по-меки от ЕС. Големите фирми имат собствени етични кодекси. Някои щати имат по-строги стандарти, други – *laissez-faire*.	*Държавно контролирано*: силни регулации, но с цел партиен контрол. ИИ се използва, но непрозрачно за властта. Китайските стандарти стават задължителни (и се опитват да ги налагат глобално).	Япония приема подобни на ЕС принципи. Южна Корея балансира – регулира за безопасност, но пази иновациите. Развиващите се страни – някои копират ЕС стандарти, други не могат да ги прилагат стриктно заради слаби институции.
Зони без ИИ	Някои регионални области или малки държави на Балканите/Източна Европа защитават местния труд с протекционизъм. ЕС като блок не подкрепя това официално (противоречи на идеята за иновации).	Национално малко вероятно, но отделни щати или общности може да ограничат автоматизацията чрез местни закони. Повечето от САЩ обаче не биха изостанали технологично.	Китай няма да позволи цели провинции да се отклонят, но може да задържи автоматизацията в определени сектори/региони за заетост. Например, по-бедни западни провинции остават по-трудоемки и абсорбират работна ръка.	Развиващи се страни: някои в Африка и Южна Азия избират този път – предлагат евтин човешки труд вместо роботи, за да привлекат нишови индустрии и да избегнат безработица.

Приложение 6: Сценарии за развитието на ИИ в областта на инвестициите и финансите

1. **ИИ като самостоятелен финансов играч:** При този сценарий ИИ системите еволюират от инструменти към **самостоятелни икономически субекти** на пазара. Агентният ИИ (*Agentic AI*) би могъл да достига високо ниво на автономност, т.е. алгоритми, които **самостоятелно възприемат обстановка, разсъждават, действат и се самообучават** без постоянни човешки напътствия. До 2040 г. е възможно да видим **напълно автономни ИИ трейдъри или ИИ инвестиционни фондове**, в които решенията за покупко-продажба на активи се генерират и изпълняват от ИИ в реално време. Такива ИИ играчи биха сканирали всички налични данни – финансови, новинарски, дори сензорни IoT данни – и биха сключвали сделки за милисекунди, конкурирайки се както помежду си, така и с хората. Напълно е възможно тези ИИ през 2040 г. да имат **правен статут** (напр., да се третират като електронни лица). Регулациите може да се наложи да дефинират границите: например, да изискват всички автономни алгоритми да са регистрирани и сертифицирани за пазарна употреба, както и да имат определен "червен бутон" за спиране при необичайно поведение. Ако технологичните и правни бариери бъдат преодолени, **към 2040 г. ИИ може да се превърне в самостоятелен пазарен участник** – например, автономен хедж фонд, управляващ активи на клиенти със минимална човешка намеса, или дори ИИ, оперираш като маркет-мейкър на децентрализирани борси.

2. **Нови модели на инвестиции:** При този сценарий са възможни **ИИ-дирижирани портфейли**, където множество индивидуални инвеститори предоставят средства на общ ИИ мениджър. Този ИИ агент регулярно преразпределя активите на базата на непрекъснато обучение и оптимизация спрямо пазарните условия и цели, зададени от инвеститорите (напр., максимална доходност при определен риск). По същество традиционните взаимни фондове може да бъдат допълнени или заменени от **общи ИИ фондове**, където алгоритъмът колективно управлява капитала. Друга новост може да са **динамични смарт договори в блокчейн, управлявани от ИИ**. Те автоматично ще инвестират в различни активи, гласуват в корпоративни управленски решения или пренасочват средства към различни проекти според пазарните сигнали. ИИ може да улесни и масовото инвестиране *(crowd-investing)*: представете си платформи, където всеки човек разполага със свой „личен ИИ финансов асистент", който инвестира от негово име, координирайки се с асистентите на други хора за колективно участие в проекти (нещо като **децентрализиран инвестиционен клуб от ИИ агенти**). Генеративният ИИ вероятно ще създава и нови деривати и структурирани продукти, моделирани според специфични потребности, например, персонализирани застраховки или хеджиращи инструменти, изготвени изцяло от алгоритъм на база на профила на клиента. До 2040 г. креативността на ИИ може да доведе до финансови иновации, които днес трудно си представяме, а цикълът на разработване на нови продукти ще се ускори значително благодарение на автоматизацията.

3. **Трансформации на финансовите пазари:** При този сценарий финансовите пазари като инфраструктура и организация биха се преобразили. Високата автоматизация и свързаност може да доведе до сливане на пазари, например, 24/7 търговия в глобален мащаб, където границите между борсите в различни държави са размити от ИИ системи, които арбитрират между тях. Може да възникнат **нови типове борси**, управлявани от ИИ, които динамично настройват правилата си (напр., механизми за откриване на цената, паузи при волатилност) в отговор на пазарните условия в реално

време. Ликвидността вероятно ще стане още по-концентрирана в електронни платформи. Към 2040 г. хората трейдъри на борсата може окончателно да изчезнат, заменени изцяло от алгоритми. Това би увеличило ефективността, но и рисковете от светкавични сътресения (*flash crash*) събития, предизвикани от автоматизирани продажби, могат да станат по-сложни за овладяване. За да се справят, пазарите биха внедрили **колективни ИИ "предпазни системи"**, напр., алгоритми, спиращи търговията при откриване на необичайни модели, или координиращи се помежду си борси, за да ограничат паниката. Също така, пазарната информация и анализ до 2040 г. ще се предават основно в машинно-четим формат, т.е. компаниите и правителствата ще публикуват данни и отчети във вид, директно усвоим от ИИ, за да се осигури мигновено и по-прозрачно отразяване. В дългосрочен план, ИИ би могъл да промени и **макроикономическата динамика**, ако например алгоритмите оптимизират разпределението на капитал глобално, това може да повлияе на потоците към различни страни, на формирането на балони и дори на паричната политика (централните банки може би ще трябва да реагират на действията на ИИ също толкова, колкото на тези на хората).

4. Банките, борсите, хедж фондовете и регулаторите на бъдещето

Традиционните банки през 2040 г. вероятно ще са претърпели радикална метаморфоза. Възможно е банката да се превърне по-скоро в платформа, която хоства множество ИИ приложения – за кредитиране, за управление на активи, за финансово планиране – разработени вътрешно или от трети страни, но предлагани на клиентите през екосистемата на банката. Това би приближило банките до модел подобен на ***App Store*** **за финансови услуги**. Хората биха били фокусирани основно върху управление на отношенията с ключови клиенти, комплексни казуси и надзор над ИИ системите. Някои банки може да изберат да **аутсорсват цели функции на ИИ** – например, оценка на кредити да се извършва изцяло от специализирана ИИ услуга, която обслужва множество банки (което повдига въпроси за концентрация и системен риск при външни доставчици). В най-крайния случай, възникват изцяло **цифрови банки**, където от откриването на сметка до отпускането на заем всичко се решава от алгоритми, а ролята на традиционната банка е само да осигури лиценз и капитал.

Фондовите и стоковите борси до 2040 г. вероятно ще работят на принципа на пълна автоматизация и взаимосвързаност. Борсите може да внедрят собствени „умни" ИИ, които да следят за пазарни злоупотреби и манипулации много по-ефективно от днешните системи. Например, ако напреднал ИИ забележи координирани аномални транзакции, той автоматично ще алармира регулаторите или дори временно ще спира определени видове поръчки. Също така, може да видим глобална консолидация на борсите, улеснена от ИИ – търговията с акции, деривати, суровини и криптоактиви може да се обединят в единни платформи, където ИИ агенти насочват потоците към най-подходящите места за изпълнение. Възможно е и възникването на **нови пазарни сегменти**, напр., **борси за данни или за модели**, където се търгуват комплекти от данни и модели като активи сами по себе си, с котировки, определяни от търсенето (ценни набори данни или успешни алгоритми могат да се лицензират и продават на търг).

До 2040 г. **хедж фондовете** може масово да са се превърнали в технологични фирми, където основното конкурентно предимство е качеството на ИИ моделите и достъпа до данни, а не толкова визията на отделен портфолио мениджър. Вероятно много фондове ще оперират с минимален човешки екип, например, основател и няколко инженери, поддържащи армия от алгоритми, които вършат анализа и търговията. **Тенденцията, започнала през 2010-2020 г., към пасивно инвестиране и индексни фондове може да се обърне благодарение на ИИ**: ако алгоритмите станат достатъчно добри в генериране на "*alpha*" (свръхдоходност над пазара),

активното управление може да преживее ренесанс, този път воден от машини. Същевременно, конкуренцията между ИИ на различни фондове може да доведе до куриозни **„войни на алгоритми"** – случаи, в които алгоритми открито се опитват да надхитрят или заблудят други алгоритми (напр., чрез генериране на подвеждащи пазарни сигнали). Това поставя въпроси за пазарния интегритет и нуждата от правила срещу манипулации от ИИ. Може да се стигне до ситуация, в която **победителят взима всичко**, т.е. най-съвършеният ИИ фонд привлича лъвския пай от капитали, оставяйки малко за конкурентите, което би концентрирало финансовата мощ. Като алтернатива, ИИ може да снижи толкова много разходите за управление, че дори малки бутикови фондове с добри алгоритми да просперират, предлагайки високотехнологични услуги на достъпна цена.

Регулаторните органи през 2040 г. ще трябва самите те да са изключително високотехнологични, за да бъдат адекватни на оборудваните с ИИ пазари. Очаква се регулаторите да разчитат на **надзорни ИИ агенти (SupTech)**, които в реално време да следят пазарите, да обменят данни между юрисдикции и да идентифицират системни рискове, преди да са ескалирали. Възможно е регулаторите дори да имат право да ограничават активността на определени алгоритми, ако ги счетат за твърде рискови - например, да забранят използването на необясними "черни кутии" за управлението на големи публични фондове, освен ако не минат специална сертификация. Международната координация вероятно ще се засили и форуми, подобни на FSB, МВФ, ОИСР (ако продължават да съществуват), ще изработят стандарти и протоколи за безопасен ИИ във финансите, тъй като действията на един неконтролиран ИИ на голям пазар могат да имат трансгранични последици. Възможно е и възникването на **нови регулаторни институции,** като например, *Агенция за надзор на алгоритмичните пазари*, специализирана в одит и тестване на използваните от финансовите фирми ИИ. Също така, регулаторите ще трябва да адресират етичните и социални ефекти: масовата автоматизация може да изисква политики за преквалификация на работната сила, за да не останат хиляди финансови анализатори и трейдъри без място в новата екосистема.

Приложение 7: Сценарии за киберсигурността в ерата на масов ИИ

За следващите 5-10 години могат да се очертаят няколко сценария, при които изкуственият интелект е навлязъл масово както в защитата, така и в атаката на киберпространството. По-долу представяме кратки сценарии, илюстриращи възможното развитие на ситуацията:

1. „Автономен SOC" – център за сигурност, управляван от ИИ: Представете си голяма финансова институция през 2030 г., чийто Security Operations Center (SOC) е почти изцяло автоматизиран. ИИ платформа денонощно следи системните журнали, мрежовия трафик и потребителските действия, откривайки аномалии за секунди и корелирайки индикатори на атака много по-бързо, отколкото някога е успявал екип от хора-аналитици. При засечен инцидент ИИ автоматично класифицира заплахата и задейства мигновени противодействия, изолира засегнатите хостове, прилага промени в конфигурацията и информира отговорните лица. Ролята на екипа от хора в този сценарий се измества към **наблюдение на ИИ системата и стратегически решения**. Те преглеждат обобщенията от ИИ (който използва обясними модели, за да посочи защо е предприел дадени действия) и се намесват само при нетипични сложни казуси. Това води до драматично намаляване на времето за реакция. Инциденти, които някога биха се развили в големи пробиви за часове и биха нанесли щети за стотици милиони долари, сега биват потушавани в зародиш минути след началото им. Подобен сценарий е реалистичен с оглед на настоящите тенденции. Takepoint посочва, че и към момента 60% от компаниите в индустрията прилагат ИИ за откриване на заплахи, а 36% – за автоматизация на отговора.[493], като тези проценти ще растат. До няколко години може да видим SOC, които е почти изцяло управляван от ИИ, където персоналът от хора е малък на брой, но висококвалифициран, надзиращ "умните" защити. Разбира се, ще останат процедури за човешка проверка при ключови решения (напр. изключване на сървъри), но оперативната дейност ще е предимно в ръцете на алгоритмите.

2. „Надпревара във въоръжаването" – ИИ нападател срещу ИИ защитник: В този сценарий киберпространството се превръща в арена на дуел между вражески и отбранителни алгоритми. Например, голяма облачна услуга е под постоянно наблюдение от защитен ИИ, който непрекъснато сканира за прониквания и изучава модела на трафика. Срещу него, обаче, оперира разработен от престъпна група или вражеска държава ИИ агент, който самостоятелно търси слабости. Този атакуващ ИИ сменя тактиките си динамично. Ако една атака бъде засечена, той бързо генерира вариант, който да заобиколи защитния модел (прилагайки *adversarial* техники). Отбранителният ИИ пък на свой ред се адаптира, обновявайки своите детектори на база на неуспешните опити за пробив. Така двата изкуствени интелекта влизат в непрестанна игра на котка и мишка, ускорена до милисекунди: атака–защита–контраатака в цикъл. Екипите от хора на заден план наблюдават метриките и дообучават моделите периодично, но голяма част от сблъсъка протича автономно. Този сценарий всъщност започва да се очертава още днес, като Световният икономически форум отбелязва в своя статия относно киберсигурността и ИИ, че *"70% от ръководните служители по информационна сигурност вярват, че ИИ дава предимство на нападателите пред защитниците"*.[494] В бъдеще, когато и двете страни масово внедрят ИИ, ще се стигне до качество нова надпревара във въоръжаването в киберсигурността, където успехът ще зависи от това чий ИИ е по-умен, по-бърз и по-устойчив на противникови трикове. Възможно е да видим злонамерен червей с ИИ, който сам изменя кода си, за да избегне засичане, докато

[493] https://industrialcyber.co/ai/takepoint-research-80-of-cybersecurity-professionals-favor-ai-benefits-over-evolving-risks/#

[494] https://www.weforum.org/stories/2024/01/cybersecurity-ai-frontline-artificial-intelligence/#

защитна ИИ система предсказва следващата му мутация и я блокира предварително. В този сценарий човекът играе ролята на треньор и съдия, настройвайки правилата, но оставя "играта" на машините. **Икономическите ефекти от такъв сценарий са невъобразими и с широко отражение върху всички сектори на икономиката**. Фронтът на такива атаки от ИИ може да включва фини методи за **индустриален шпионаж, кражби на авторски права и самоличности, пробиви във финансови институции, пренасочване на финансови и търговски потоци, блокиране на борсовата търговия на глобалните пазари, спиране на производства, нанасяне на щети по ключова инфраструктура като енергийни мрежи и водоснабдяване и други**, което да генерира загуби от милиарди или дори трилиони долари. Ето защо индустрията за разработка и трениране на защитни модели на ИИ също ще бележи бум и ще профъцтява.

3. „Блато на автоматизацията" – провал заради липса на надзор: При този сценарий е реалистично да се случи значим инцидент, при който прекаленото доверие в ИИ довежда до катастрофален пробив – урок, който променя разбиранията в индустрията. Представете си голяма компания, която е поверила почти всичко на система за сигурност с ИИ, считана за най-доброто на пазара. В продължение на месеци ИИ ловко спира дребни атаки и компанията намалява екипа си, вярвайки че "умната" защита се справя. Атакуваща група обаче открива, че чрез минимални промени в своя малуеър могат да го направят "невидим" за модела (типичен *adversarial* подход). Техният зловреден софтуер прониква и бавно краде данни, като системата с ИИ погрешно го класифицира като легитимен процес. Тъй като няма достатъчно опитни аналитици, които да следят алармите (които ИИ дори не генерира), пробивът остава незабелязан с месеци. Когато най-накрая излиза наяве чрез източване на чувствителни данни към тъмната мрежа, щетите са огромни. Разследването показва, че пробивът е бил "скрит в сляпото петно" на ИИ защитата и нападателите умишлено са се прицелили в слабост на модела. Този сценарий подчертава риска, описан и от експерти: ИИ, макар и мощен, може да провали с гръм и трясък, ако не е подсигурен с човешки контрол и допълнителни проверки. [495] Последиците от подобен инцидент биха включвали преоценка на политиките и регулаторите биха въвели изисквания за участие на човек в цикъла (*human in the loop*) за критични системи, а компаниите отново биха инвестирали в комбинирани екипи, вместо да разчитат на 100% автоматизация. Този сценарий е предупреждение, че балансът между ИИ и човек трябва да се запази, дори когато технологията изглежда надеждна. **Икономическите измерения на такъв сценарий са свързани, не на последно място, и с нови умения на пазара на труда, свързани със способности за адекватна работа с ИИ и контрол.**

4. „Съюз човек–машина" – колаборативна киберсигурност: В този сценарий виждаме узряла екосистема, където ИИ и експертите по сигурност работят в синергия, всеки допълвайки слабостите на другия. Например, в една правителствена агенция за киберотбрана всеки анализатор е подпомаган от свой цифров асистент с ИИ, който пресява информацията, извлича релевантните данни от разузнаване за заплахи, симулира възможни ходове на нападателя и предлага опции за реакция. Човекът, от своя страна, внася стратегическа преценка, избира сред предложените варианти, отчитайки политически и етични фактори, които ИИ не разбира. При разследване на сложна атака ИИ системата може да визуализира *целия "Kill Chain"* и да маркира вероятните входни точки, но експертът решава къде да фокусира усилията и кои активи да защити с приоритет. В този режим ИИ, също така, се учи от решението на човека и чрез механизми за обяснимост и обратна връзка моделът коригира бъдещите си препоръки, приближавайки се до начина на мислене на хората-аналитици. Регулациите и стандартите в индустрията също са еволюирали до този момент и са налице

[495] https://www.isaca.org/resources/news-and-trends/industry-news/2024/overreliance-on-automated-tooling-a-big-cybersecurity-mistake#

международни норми за надеждни приложения на ИИ в киберсигурността, които гарантират прозрачност, поверителност и контрол. [496] Например, може да има сертификати за ИИ модели, доказващи че са устойчиви на подмяна и че решенията им могат да се обяснят на външен одитор в разбираем вид. Така бъдещето на масовия ИИ в киберсигурността се оформя не като пълно изместване на хората, а като тясно сътрудничество: машините осигуряват безпрецедентна бързина и дълбочина на анализа, докато хората осигуряват посока, мъдрост и отговорност. В резултат, киберотбраната става по-проактивна, адаптивна и всестранно обезпечена, а цифровият свят – по-сигурен в сравнение с предходните десетилетия на изцяло реактивна защита.

[496] https://pmc.ncbi.nlm.nih.gov/articles/PMC11656524/#

Източници

Maslej, N. et al., 2024. *Artificial Intelligence Index Report 2024*, Human-Centered Artificial Intelligence. United States of America. Retrieved from https://hai.stanford.edu/ai-index/2024-ai-index-report

SCImago Journal & Country Rank, "Artificial Intelligence; Eastern Europe, 1996-2023" 2024 https://www.scimagojr.com/countryrank.php?category=1702&area=1700&region=Eastern%20Europe 13.03.2025

SCImago Journal & Country Rank, "Artificial Intelligence; Eastern Europe, 1996-2023" 2024 https://www.scimagojr.com/countryrank.php?category=1702&area=1700&region=Eastern%20Europe 13.03.2025

Anthropic, "*Anthropic Economic Index*" 2025, Retrieved from https://www.anthropic.com/news/the-anthropic-economic-index , 13.03.2025

Oxford Insights, "Government AI Readiness Index 2024", 2024, Retrieved from https://oxfordinsights.com/ai-readiness/ai-readiness-index/

MIT Sloan, "Five Trends in AI and data science for 2025", 2025, https://sloanreview.mit.edu/article/five-trends-in-ai-and-data-science-for-2025/ 12.03.2025

"*GTC March 2025 Keynote with NVIDIA CEO Jensen Huang*", 2025, https://www.nvidia.com/gtc/keynote/ 24.03.2025

"*AI Statistics and trends: New Research for 2025*", 2025, https://www.hostinger.com/tutorials/ai-statistics 11.03.2025

First meeting of the International Network of AI Safety Institutes" , 2024

https://digital-strategy.ec.europa.eu/en/news/first-meeting-international-network-ai-safety-institutes 10.03.2025

"G7 Leaders' Statement on the Hiroshima AI Process" , 2023 https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document01_en.pdf 10.03.2025

"Hiroshima Process International Guiding Principles for Organizations Developing Advanced AI System ", 2023 https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document04_en.pdf 10.03.2025

"Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems", 2023, https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document05_en.pdf 1.03.2025

"Ministerial declaration", 2024, https://www.soumu.go.jp/hiroshimaaiprocess/pdf/document07_en.pdf , 14.03.2025

"The OECD Artificial Intelligence Policy Observatory" https://oecd.ai/en/ , 11.03.2025

"About the Global Partnership on AI (GPAI)" , 2024, https://oecd.ai/en/about/what-we-do , 11.03.2025

"AI Principles", 2019, https://www.oecd.org/en/topics/sub-issues/ai-principles.html 11.03.2025

"G7 reporting framework – Hiroshima AI Process (HAIP) international code of conduct for organizations developing advanced AI systems", 2024, https://transparency.oecd.ai/ , 11.03.2025

"The Bletchley Declaration by Countries Attending the AI Safety Summit, 1-2 November 2023", 2023, https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration/the-bletchley-declaration-by-countries-attending-the-ai-safety-summit-1-2-november-2023 , 11.03.2025

"Seoul Declaration for Safe, Innovative and Inclusive AI by Participants Attending the Leaders' Session of the AI Seoul Summit, 21st May, 2024", 2024, https://www.mofa.go.kr/eng/brd/m_5674/view.do?page=1&seq=321007 , 11.03.2025

"Pledge for a Trustworthy AI in the World of Work", 2025, https://www.elysee.fr/emmanuel-macron/2025/02/11/pledge-for-a-trustworthy-ai-in-the-world-of-work 11.03.2025

OB L, 2024/1689, 12.7.2024, ELI: http://data.europa.eu/eli/reg/2024/1689/oj

European Commission, "Third Draft General-Purpose AI Code of Practice", 2025, Retrieved from https://digital-strategy.ec.europa.eu/en/library/third-draft-general-purpose-ai-code-practice-published-written-independent-experts

European Commission, "Second Draft General-Purpose AI Code of Practice " 2024, Retrieved from https://digital-strategy.ec.europa.eu/en/library/second-draft-general-purpose-ai-code-practice-published-written-independent-experts

European Commission, "1- Commitments -Third Draft General-Purpose AI Code of Practice", 2025, Retrieved from https://digital-strategy.ec.europa.eu/en/library/third-draft-general-purpose-ai-code-practice-published-written-independent-experts

European Commission, "3- Copyright -Third Draft General-Purpose AI Code of Practice", 2025, Retrieved from https://digital-strategy.ec.europa.eu/en/library/third-draft-general-purpose-ai-code-practice-published-written-independent-experts

European Commission, "2- Transparency -Third Draft General-Purpose AI Code of Practice", 2025, Retrieved from https://digital-strategy.ec.europa.eu/en/library/third-draft-general-purpose-ai-code-practice-published-written-independent-experts

European Commission, "4- Safety and Security -Third Draft General-Purpose AI Code of Practice", 2025, Retrieved from https://digital-strategy.ec.europa.eu/en/library/third-draft-general-purpose-ai-code-practice-published-written-independent-experts

Browne, R. "Google, Meta execs blast Europe over strict AI regulation as Big Tech ups the ante", 2025 https://www.cnbc.com/2025/02/21/google-meta-execs-blast-europe-over-strict-ai-regulation.html 19.03.2025

Haeck, P. "EU rules for advanced AI are step in wrong direction, Google says", 2025, https://www.politico.eu/article/google-eu-rules-advanced-ai-artificial-intelligence-step-in-wrong-direction/ 19.03.2025

Federal Register, "Framework for Artificial Intelligence Diffusion", 2025,

https://www.federalregister.gov/documents/2025/01/15/2025-00636/framework-for-artificial-intelligence-diffusion 17.03.2025

"Notable AI Models", updated 2025, https://epoch.ai/data/notable-ai-models, 25.03.2025.

OJ L, 2024/2831, 11.11.2024, ELI: http://data.europa.eu/eli/dir/2024/2831/oj

OJ L 277, 27.10.2022, ELI: http://data.europa.eu/eli/reg/2022/2065/oj

OJ L 265, 12.10.2022, ELI: http://data.europa.eu/eli/reg/2022/1925/oj

OJ L, 2024/2847, 20.11.2024, ELI: http://data.europa.eu/eli/reg/2024/2847/oj

OJ L 119, 4.5.2016, ELI: http://data.europa.eu/eli/reg/2016/679/oj

Motoki, F., & Pinto, J."Regulating data: Evidence from corporate America" . Journal of Business Finance & Accounting, 52(1), 541-568. https://doi.org/10.1111/jbfa.12820

Moazed, Alex, "How GDPR is Helping Big Tech and Hurting the Competition", https://www.applicoinc.com/blog/how-gdpr-is-helping-big-tech-and-hurting-the-competition/#:~:text=Giants%20can%20thrive%20in%20the%20face%20of%20GDPR%20and%20other%20privacy%20laws&text=Google%2C%20for%20example%2C%20has%20seen,America%20and%20Europe%20lost%20ground 23.03.2025

European Commission, "Competitiveness Compass", 2025, retrieved from https://commission.europa.eu/document/download/10017eb1-4722-4333-add2-e0ed18105a34_en

EESC, "Generative AI and foundation models in the EU: Uptake, opportunities, challenges, and a way forward", QE-01-25-014-EN-N, 2025

EESC, "Pro-worker AI: levers for harnessing the potential and mitigating the risks of AI in connection with employment and labour market policies", SOC/803 - EESC-2024-01024-01-00-TCD-TRA - EN, 2025

EESC, "Counter Opinion on "Pro-worker AI: levers for harnessing the potential and mitigating the risks of AI in connection with employment and labour market policies"", SOC/803 - EESC-2024-01024-00-02-AMP-TRA, 2025

DigWatch Team, "Apple Intelligence expands to the EU amid regulatory changes", 2024, https://dig.watch/updates/apple-intelligence-expands-to-the-eu-amid-regulatory-changes 19.03.2025

Lee D., "EU's Hard Line on Tech Keeps Users Off the Cutting Edge", 2024, https://www.bloomberg.com/opinion/articles/2024-07-03/apple-and-meta-users-lose-out-in-eu-s-hard-line-on-tech 19.03.2025

Madiega T.and Ilnicki R. "AI investment: EU and global indicators", , European Parliament Research Service, 2024, retrieved from https://www.europarl.europa.eu/RegData/etudes/ATAG/2024/760392/EPRS_ATA(2024)760392_EN.pdf

Bratton, L. "Big Tech set to invest $325 billion this year as hefty AI bills come under scrutiny", 2025, https://finance.yahoo.com/news/big-tech-set-to-invest-325-billion-this-year-as-hefty-ai-bills-come-under-scrutiny-182329236.html 19.03.2025

Smith, K. et al, "AI power surge: Growth Scenarios for GenAI Datacenters Through 2030", Center for Strategic and International Studies, 2025, retrieved from https://www.csis.org/analysis/ai-power-surge-growth-scenarios-genai-datacenters-through-2030

"Announcing the Stargate Project", OpenAI, 2025 https://openai.com/index/announcing-the-stargate-project/ 11.03.2025

Speech by President von der Leyen at the Artificial Intelligence Action Summit, France, 11-12 February 2025 - https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_25_471

https://aichampions.eu/, прессъобщение https://files.elfsightcdn.com/eafe4a4d-3436-495d-b748-5bdce62d911d/9dc25764-5c0a-4c04-b55e-6be1d26f625f/EU-AI-Champions-Initiative-Media-Release.pdf

Министерски съвет, "Концепция за развитие на изкуствения интелект в България до 2030 г." , 2020, retrieved from https://egov.government.bg/wps/portal/ministry-meu/strategies-policies/digital.transformation/itis-national-strategic-documents/ai.development.concept.2030

"25 recommandations pour l'IA en France", 2024 https://www.info.gouv.fr/actualite/25-recommandations-pour-lia-en-france , 17.03.2025

"La France se dote d'un Institut national pour l'évaluation et la sécurité de l'intelligence artificielle (INESIA)

",2025https://www.economie.gouv.fr/actualites/la-france-se-dote-dun-institut-national-pour-levaluation-et-la-securite-de-lintelligence 17.03.2025

Executive Order 14110 "Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence", 2023

https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence 09.03.2025

Executive Order 14179, "Removing Barriers to American Leadership in Artificial Intelligence

", 2025, https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence/ 09.03.2025

NSPM "America First Investment Policy", 2025, https://www.whitehouse.gov/presidential-actions/2025/02/america-first-investment-policy/ 19.03.2025

Executive Order 14177, "Presidential Council on Science and Technology", 2025, https://www.whitehouse.gov/presidential-actions/2025/01/presidents-council-of-advisors-on-science-and-technology/ 19.03.2025

National Science Foundation, "Request for Information on the Development of an Artificial Intelligence (AI) Action Plan",2025 https://www.federalregister.gov/documents/2025/02/06/2025-02305/request-for-information-on-the-development-of-an-artificial-intelligence-ai-action-plan 15.03.2025

"Google's comments on the U.S. AI Action Plan", 2025

https://blog.google/outreach-initiatives/public-policy/google-us-ai-action-plan-comments/ 17.03.2025

"OpenAI's Proposals for the U.S AI Action Plan", 2025, https://openai.com/global-affairs/openai-proposals-for-the-us-ai-action-plan/ 17.03.2025

"Anthropic's Recommendations to OSTP for the U.S. AI Action Plan", 2025, https://www.anthropic.com/news/anthropic-s-recommendations-ostp-u-s-ai-action-plan 17.03.2025

"a16z's Recommendations for the National AI Action Plan", 2025, https://a16z.com/a16zs-recommendations-for-the-national-ai-action-plan/ 17.03.2025

"Lockheed Martin and Google Cloud Announce Collaboration to Advance Generative AI For National Security", 2025,https://news.lockheedmartin.com/2025-03-27-Lockheed-Martin-and-Google-Cloud-Announce-Collaboration-to-Advance-Generative-AI-For-National-Security 28.03.2025

"Anduril Partners with OpenAI to Advance U.S. Artificial Intelligence Leadership and Protect U.S. and Allied Forces", 2024, https://www.anduril.com/article/anduril-partners-with-openai-to-advance-u-s-artificial-intelligence-leadership-and-protect-u-s/ 19/03.2025

"Azure for US Department of Defense", https://azure.microsoft.com/en-us/explore/global-infrastructure/government/dod 19.03.2025

National Security Memorandum, "Memorandum on Advancing the United States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence", 2024, https://bidenwhitehouse.archives.gov/briefing-room/presidential-actions/2024/10/24/memorandum-on-advancing-the-united-states-leadership-in-artificial-intelligence-harnessing-artificial-intelligence-to-fulfill-national-security-objectives-and-fostering-the-safety-security/ 20.03.2025

"What Trump's New AI and Crypto Czar David Sacks Means For the Tech Industry

" 2024 https://time.com/7200518/david-sacks-new-white-house-ai-crypto-czar-trump-administration/ 20.03.2025

"新一代人工智能伦理规范", 2021, https://www.most.gov.cn/kjbgz/202109/t20210926_177063.html 21.03.2025

"互联网信息服务算法推荐管理规定", 2022, http://www.cac.gov.cn/2022-01/04/c_1642894606364259.htm 21.03.2025

"生成式人工智能服务管理暂行办法", 2023, https://www.cac.gov.cn/2023-07/13/c_1690898327029107.htm 21.03.2025

"AI Safety Governance Framework", 2024, https://www.tc260.org.cn/upload/2024-09-09/1725849192841090989.pdf

"International AI Safety Report 2025", 2025, retrieved from https://www.gov.uk/government/publications/international-ai-safety-report-2025

"Ethics guidelines for Trustworthy AI", 2019, retrieved from https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Ahmed, Abdullah & Khan, Mudassar, "AI and Content Moderation: Legal and Ethical Approaches to Protecting Free Speech and Privacy." 2024, 10.13140/RG.2.2.23619.41767; DOI:10.13140/RG.2.2.23619.41767

"Building a responsible AI: How to manage the AI ethics debate", https://www.iso.org/artificial-intelligence/responsible-ai-ethics 20.03.2025

Governing AI for Humanity, September 2024; https://www.un-ilibrary.org/content/books/9789211067873

UN Resolution on Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development, 11.03.2024 - https://docs.un.org/en/A/78/L.49

Ethics of Artificial Intelligence - The Recommendationhttps://www.unesco.org/en/artificial-intelligence/recommendation-ethics Retrived 3.4.2025

ILO's Observatory on AI and Work in the Digital Economy - https://www.ilo.org/artificial-intelligence-and-work-digital-economy

"What is Responsible AI?", 2025, https://learn.microsoft.com/en-us/azure/machine-learning/concept-responsible-ai?view=azureml-api-2 09.03.2025

"Responsible AI Transparency Report", retrieved from https://www.microsoft.com/en-us/ai/principles-and-approach, 09.03.2025

"Transform responsible AI from theory into practice", https://aws.amazon.com/ai/responsible-ai 09.03.2025

"Responsible AI", https://cloud.google.com/responsible-ai?hl=en 09.03.2025

"Global Index on Responsible AI", 2024, retrieved from https://girai-report-2024-corrected-edition.tiiny.site/

Byman, D. et al, "Government Use of Deepfakes: Questions to Ask", Center for Strategic and International Studies, 2024, retrieved from https://csis-website-prod.s3.amazonaws.com/s3fs-public/2024-03/240312_Byman_Government_Deepfakes.pdf?VersionId=OAFKvrOi9ojseIjOKhzqi672claqU.if

Funk, A. et al., "The Repressive Power of Artificial Intelligence", Freedom House, 2023 https://freedomhouse.org/report/freedom-net/2023/repressive-power-artificial-intelligence

"Ethics guidelines for Trustworthy AI", 2019, retrieved from https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Ben Buchanan, "The AI Triad and What It Means for National Security Strategy", Center for Security and Emerging Technology, 2020. https://doi.org/10.51593/20200021

Department of Commerce, BIS, "Framework for Artificial Intelligence Diffusion". 2025, retrieved from https://www.federalregister.gov/documents/2025/01/15/2025-00636/framework-for-artificial-intelligence-diffusion

Department of Commerce, BIS, "Commerce Implements New Export Controls on Advanced Computing and Semiconductor Manufacturing Items to the People's Republic of China (PRC)", 2022, https://www.bis.doc.gov/index.php/documents/about-bis/newsroom/press-releases/3158-2022-10-07-bis-press-release-advanced-computing-and-semiconductor-manufacturing-controls-final/file

Allen, Gregory C., "DeepSeek, Huawei, Export Controls, and the Future of of the U.S.- China AI Race", Center for Strategic and International Studies, 2025, retrieved from https://csis-website-prod.s3.amazonaws.com/s3fs-public/2025-03/250307_Allen_AI_Race.pdf?VersionId=8Qpp5OWwGj9OKB25qzQX4tQ_DzIZ9NS_

Baskaran, G. and Schwartz, M., "China Imposes Its Most Stringent Critical Minerals Export Restrictions Yet Amidst Escalating U.S.-China Tech War", Center for Strategic and International Studies, 2024, https://www.csis.org/analysis/china-imposes-its-most-stringent-critical-minerals-export-restrictions-yet-amidst

"商务部 工业和信息化部 海关总署 国家密码局公告 2024 年第 51 号 关于发布《中华人民共和国两用物项出口管制清单》的公告", 2024, https://aqygzj.mofcom.gov.cn/qdml/art/2024/art_8bf88382af1143168e2b11beadf000ba.html 14.03.2025

"中华人民共和国 两用物项出口管制清单", 2024, retrieved from https://aqygzj.mofcom.gov.cn/qdml/art/2024/art_8bf88382af1143168e2b11beadf000ba.html

OJ L 206, 11.6.2021, ELI: http://data.europa.eu/eli/reg/2021/821/oj

"Exporting dual-use items", https://policy.trade.ec.europa.eu/help-exporters-and-importers/exporting-dual-use-items_en 21.03.2025

"Statement regarding partial revocation export license", 2024, https://www.asml.com/en/news/press-releases/2023/statement-regarding-partial-revocation-export-license 20.03.2025

"AI Talent Report", 2025, https://bidenwhitehouse.archives.gov/cea/written-materials/2025/01/14/ai-talent-report/#:~:text=The%20United%20States%20could%20increase,and%203 22.03.2025

Perault, M. "A Policy Blueprint for US Investment in AI Talent and Infrastructure", 2025, https://a16z.com/a-policy-blueprint-for-us-investment-in-ai-talent-and-infrastructure/ 21.03.2025

Zwetsloot, R., "Keeping Top AI Talent in the United States", 2019 https://cset.georgetown.edu/publication/keeping-top-ai-talent-in-the-united-states/ 21.03.2025

"AI Excellence: Ensuring that AI works for people", https://digital-strategy.ec.europa.eu/en/policies/ai-people 21.03.2025

"2024 State of the Digital Decade package", https://digital-strategy.ec.europa.eu/en/policies/2024-state-digital-decade-package 21.03.2025

"AI Skills Strategy for Europe", 2023, retrieved from https://aiskills.eu/wp-content/uploads/2024/01/AI-Skills-Strategy-for-Europe.pdf

Yang, Z.,"Four things you need to know about China's AI talent pool ", 2024, https://www.technologyreview.com/2024/03/27/1090182/ai-talent-global-china-us/ 21.03.2025

Paterson, D. et al., "Assessing China's AI Workforce Regional, Military, and Surveillance Geographic Clusters", 2023, https://cset.georgetown.edu/publication/assessing-chinas-ai-workforce/ 21.03.2025

"China's AI talent gap", 2023, https://www.mckinsey.com/featured-insights/sustainable-inclusive-growth/charts/chinas-ai-talent-gap 21.03.2025

"China expands university enrolment to boost AI talent", 2025, https://dig.watch/updates/china-expands-university-enrolment-to-boost-ai-talent 21.03.2025

"National AI Strategy Policy Directions", 2024, retrieved from https://www.msit.go.kr/bbs/view.do?sCode=eng&mId=4&mPid=2&bbsSeqNo=42&nttSeqNo=1040

"Seoul vows to nurture 10,000 AI talents annually", 2025, https://www.koreaherald.com/article/10417202#:~:text=The%20Seoul%20city%20government%20on,living%20and%20 pleasure%20are%20 combined 22.03.2025

"AI and energy: Will AI help reduce emissions or increase demand? Here's what to know", 2024, https://www.weforum.org/stories/2024/07/generative-ai-energy-emissions/ 22.03.2025

"AI has high data center energy costs — but there are solutions", 2025, https://mitsloan.mit.edu/ideas-made-to-matter/ai-has-high-data-center-energy-costs-there-are-solutions - 22.03.2025

"AI's power binge", 2024, https://www.mckinsey.com/featured-insights/sustainable-inclusive-growth/charts/ais-power-binge 22.03.2025

"AI has high data center energy costs — but there are solutions", 2025, https://mitsloan.mit.edu/ideas-made-to-matter/ai-has-high-data-center-energy-costs-there-are-solutions - 22.03.2025

Bennett, S. and Spencer, T., "How will artificial intelligence transform energy innovation?", 2024,https://www.iea.org/commentaries/how-will-artificial-intelligence-transform-energy-innovation 22.03.2025

"Is nuclear energy the answer to AI data centers' power consumption?", 2025, https://www.goldmansachs.com/insights/articles/is-nuclear-energy-the-answer-to-ai-data-centers-power-consumption - 22.03.2025

"European Nuclear Business Alliance", 2025, https://www.medef.com/uploads/media/default/0020/05/16382-european-business-nuclear-alliance-fully-signed.pdf 22.03.2025

Sizing the prize: PwC's Global Artificial Intelligence Study: Exploiting the AI Revolution https://www.pwc.com/gx/en/issues/artificial-intelligence/publications/artificial-intelligence-study.html

The next big arenas of competition, McKinsey Global Institute, 2024, https://www.mckinsey.com/~/media/mckinsey/mckinsey%20global%20institute/our%20research/the%20next%20big%20arenas%20of%20competition/the-next-big-arenas-of-competition.pdf

The global economy will be $16 trillion bigger by 2030 thanks to AI, World Economic Forum, 2017, https://www.weforum.org/stories/2017/06/the-global-economy-will-be-14-bigger-in-2030-because-of-ai/, Retrieved 24.03.2025

Europe in the Intelligent Age: From Ideas to Action, World Economic Forum, 2025 https://reports.weforum.org/docs/WEF_Europe_in_the_Intelligent_Age_2025.pdf, Retrieved 24.03.2025

The future of European competitiveness 2024, European Commission, https://commission.europa.eu/document/download/97e481fd-2dc3-412d-be4c-f152a8232961_en , Retrieved 25.03.2025

Industrial AI gives people 'superpowers' in advanced manufacturing. Here's how, World Economic Forum, January 2024 https://www.weforum.org/stories/2024/01/industrial-ai-superpowers-advanced-manufacturing/ , Retrieved 25.03.2025

Sizing the prize. What's the real value of AI for your business and how can you capitalise?, PwC, 2017, https://www.pwc.com/gx/en/issues/analytics/assets/pwc-ai-analysis-sizing-the-prize-report.pdf, 26.03.2025

Industrial AI gives people 'superpowers' in advanced manufacturing. Here's how, World Economic Forum, January 2024, https://www.weforum.org/stories/2024/01/industrial-ai-superpowers-advanced-manufacturing/, 26.03.2025

Why AI's role in advancing sustainability is underestimated , World Economic Forum, March 2025, https://www.weforum.org/stories/2025/03/can-ai-foster-sustainability/, 26.03.2025

From bytes to bushels: How gen AI can shape the future of agriculture, McKinsey and Company, June 2024, https://www.mckinsey.com/industries/agriculture/our-insights/from-bytes-to-bushels-how-gen-ai-can-shape-the-future-of-agriculture, 26.03.2025

Scientific AI: Unlocking the next frontier of R&D productivity, McKinsey Digital, January 2025, https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/tech-forward/scientific-ai-unlocking-the-next-frontier-of-r-and-d-productivity, 26.03.2025

Maynard, Andrew, "Is AI poised to suck the soul out of science?", November 2024, The Future of Being Human https://futureofbeinghuman.com/p/is-ai-poised-to-suck-the-soul-out-of-science, 26.03.2025

AI investment: EU and global indicators, European Parliament, Directorate-General for Parliamentary Research Services, https://epthinktank.eu/2024/04/04/ai-investment-eu-and-global-indicators/#, 26.03.2026

Alphabet plans massive capex hike, reports cloud revenue growth slowed, Reuters, February 2025, https://www.reuters.com/technology/google-parent-alphabet-misses-quarterly-revenue-estimates-2025-02-04/#, 26.03.2025

How Artificial Intelligence is Transforming the Financial Services Industry, Deloitte, https://www.deloitte.com/ng/en/services/risk-advisory/services/how-artificial-intelligence-is-transforming-the-financial-services-industry.html, 26.03.2025

What is artificial intelligence (AI) in finance?, IMB, December 2023, https://www.ibm.com/think/topics/artificial-intelligence-finance#, 27.03.2025

Artificial Intelligence and its Impact on Financial Markets and Financial Stability, International Monetary Fund, September 2024, https://www.imf.org/en/News/Articles/2024/09/06/sp090624-artificial-intelligence-and-its-impact-on-financial-markets-and-financial-stability#

How Agentic AI will transform financial services with autonomy, efficiency and inclusion, World Economic Forum, December 2024, https://www.weforum.org/stories/2024/12/agentic-ai-financial-services-autonomy-efficiency-and-inclusion/#, 27.03.2025

Artificial Intelligence in Financial Markets: Systemic Risk and Market Abuse Concerns, Sidley, December 2024, https://www.sidley.com/en/insights/newsupdates/2024/12/artificial-intelligence-in-financial-markets-systemic-risk-and-market-abuse-concerns#, 27.03.2025

The Financial Stability Implications of Artificial Intelligence, Financial Stability Board, November 2024, https://www.fsb.org/2024/11/the-financial-stability-implications-of-artificial-intelligence/, 27.03.2025

Vanian, Leswing, 2023, ChatGPT and generative AI are booming, but the costs can be extraordinary, CNBC, https://www.cnbc.com/2023/03/13/chatgpt-and-generative-ai-are-booming-but-at-a-very-expensive-price.html, 27.03.2025

AI Model Training Cost Have Skyrocketed by More than 4,300% Since 2020, Edge AI Vision, September 2024 https://www.edge-ai-vision.com/2024/09/ai-model-training-cost-have-skyrocketed-by-more-than-4300-since-2020/#, 27.03.2025

"Visualizing the Training Costs of AI Models Over Time", Visual Capitalist, June 2024, https://www.visualcapitalist.com/training-costs-of-ai-models-over-time/#, 27.03.2025

Data Center GPU Market Valuation – 2025-2032, Verified Market Research, https://www.verifiedmarketresearch.com/product/data-center-gpu-market/#:~:text=high,9%20Billion%20by%202032, 27.03.2025

Artificial Intelligence Market, Lucintel, https://www.lucintel.com/artificial-intelligence-market.aspx, 27.03.2025

Reddit pricing: API charge explained, TechTarget, July 2023, https://www.techtarget.com/whatis/feature/Reddit-pricing-API-charge-explained#, 27.03.2025

AI and efficiency, OpenAI, May 2020, https://openai.com/index/ai-and-efficiency/, 27.03.2025

The Cost of AI Talent: Who's Hurting in the Search for AI Stars?, Information Week, February 2025, https://www.informationweek.com/it-leadership/the-cost-of-ai-talent-who-s-hurting-in-the-search-for-ai-stars-# , 27.03.2025

AI Talent Report, Онлайн архивна страница на Белия дом, Януари 2025, https://bidenwhitehouse.archives.gov/cea/written-materials/2025/01/14/ai-talent-report/#

KPMG GenAI Survey 2024, август 2024, https://kpmg.com/kpmg-us/content/dam/kpmg/corporate-communications/pdf/2024/kpmg-genai-survey-august-2024.pdf, 27.03.2025

Artificial Intelligence. Creating The Assembly Line For Knowledge Workers, ARK Invest, 2023, https://www.ark-invest.com/big-ideas-2023/artificial-intelligence#

Sparsity – определение, https://www.sciencedirect.com/topics/computer-science/sparsity, 27.03.2025

Cost of a Data Breach Report 2024, IBM, https://www.ibm.com/reports/data-breach , 27.03.2025

Salem, A.H., Azzam, S.M., Emam, O.E. et al. Advancing cybersecurity: a comprehensive review of AI-driven detection techniques. J Big Data 11, 105 (2024). https://doi.org/10.1186/s40537-024-00957-y

Stanbridge, Ryan, Overreliance on Automated Tooling: A Big Cybersecurity Mistake, Information Systems Audit and Control Association (ISACA), 2024, https://www.isaca.org/resources/news-and-trends/industry-news/2024/overreliance-on-automated-tooling-a-big-cybersecurity-mistake#

Takepoint Research Survey 2024, обобщено в Industrial Cyber, https://industrialcyber.co/ai/takepoint-research-80-of-cybersecurity-professionals-favor-ai-benefits-over-evolving-risks/#, 27.03.2025

Artificial Intelligence: Adversarial Machine Learning, National Cybersecurity Center of Excellence, https://www.nccoe.nist.gov/ai/adversarial-machine-learning, 27.03.2025

Adversarial Machine Learning and Cybersecurity: Risks, Challenges, and Legal Implications, 2023 Center for Security and Emerging Technology, Stanford Cyber Policy Center https://cset.georgetown.edu/publication/adversarial-machine-learning-and-cybersecurity/# , 28.03.2025

Economic impact of automation and artificial intelligence, WatchGuard, 2023, https://www.watchguard.com/wgrd-news/blog/economic-impact-automation-and-artificial-intelligence, 28.03.2025

Prepare for AI Hackers, Harvard Magazine, 2023, https://www.harvardmagazine.com/2023/02/right-now-ai-hacking# , 29.03.2025

AI Phishing Revolution: Study Reveals Bots Now Out-Scheme Humans in Cyber Deception, February 2025, https://thedebrief.org/ai-phishing-revolution-study-reveals-bots-now-out-scheme-humans-in-cyber-deception/#

WormGPT and FraudGPT – The Rise of Malicious LLMs , Trustwave, August 2023, https://www.trustwave.com/en-us/resources/blogs/spiderlabs-blog/wormgpt-and-fraudgpt-the-rise-of-malicious-llms/#

EU Raw Materials Information System https://rmis.jrc.ec.europa.eu/ , Retrieved 26.03.2025

Cornelis P. Baldé, Ruediger Kuehr, Tales Yamamoto, Rosie McDonald, Elena D'Angelo, Shahana Althaf, Garam Bel, Otmar Deubzer, Elena Fernandez-Cubillo, Vanessa Forti, Vanessa Gray, Sunil Herat, Shunichi Honda, Giulia Iattoni, Deepali S. Khetriwal, Vittoria Luda di Cortemiglia, Yuliya Lobuntsova, Innocent Nnorom, Noémie Pralat, Michelle Wagner (2024). International Telecommunication Union (ITU) and United Nations Institute for Training and Research (UNITAR). 2024. Global E-waste Monitor 2024. Geneva/Bonn. Pdf version: 978-92-61-38781-5 https://www.itu.int/en/ITU-D/Environment/Pages/Publications/The-Global-E-waste-Monitor-2024.aspx

Анекс към Препоръката на Комисията относно критичните технологични области за икономическата сигурност на ЕС за по-нататъшна оценка на риска съвместно с държавите членки, Страсбург 03.10.2023, C(2023) 6689 final

Анекс към изявлението на Консултативния съвет към EIC Pilot при старта на Европейския иновационен съвет, https://eic.ec.europa.eu/document/download/61d52ef5-5b28-4c00-bfb8-a67e9c22666f_en?filename=Statement%20on%20technological%20sovereignty.pdf , 27.03.2025

The future of European competitiveness, Марио Драги 2024, https://commission.europa.eu/document/download/97e481fd-2dc3-412d-be4c-f152a8232961_en?filename=The%20future%20of%20European%20competitiveness%20_%20A%20competitiveness%20strategy%20for%20Europe.pdf

DRAFT REPORT on European technological sovereignty and digital infrastructure (2025/2007(INI)), PE768.180v01-00, 25.02.2025

Supply chain analysis and material demand forecast in strategic technologies and sectors in the EU – A foresight study, Joint Research Council, ISSN 1831-9424

Global Data Center Trends 2024, https://www.cbre.com/insights/reports/global-data-center-trends-2024, 29.03.2025

World Energy Outlook 2024 Report https://www.iea.org/reports/world-energy-outlook-2024, 30.03.2025

Patterson, David et al., Carbon Emissions and Large Neural Network Training, https://arxiv.org/ftp/arxiv/papers/2104/2104.10350.pdf достъпено през https://news.climate.columbia.edu/2023/06/09/ais-growing-carbon-footprint/#, 29.03.2025

Green AI Explained: Fueling Innovation with a Smaller Carbon Footprint, Carbon Credits, December 2024, https://carboncredits.com/green-ai-explained-fueling-innovation-with-a-smaller-carbon-footprint/, 30.03.2025

Champion, Zachary, Optimization could cut the carbon footprint of AI training by up to 75%, University of Michigan, https://news.umich.edu/optimization-could-cut-the-carbon-footprint-of-ai-training-by-up-to-75/#, 30.03.2025

Davis, Jacqueline, Large data centers are mostly more efficient, analysis confirms, Uptime Institute, February 2024, https://journal.uptimeinstitute.com/large-data-centers-are-mostly-more-efficient-analysis-confirms/

Carbon-free energy, Amazon, https://sustainability.aboutamazon.com/climate-solutions/carbon-free-energy Retrieved 03.04.2025

Accelerating the Next Wave of Nuclear to Power AI Innovation, 3 December 2024, https://sustainability.atmeta.com/blog/2024/12/03/accelerating-the-next-wave-of-nuclear-to-power-ai-innovation/

New nuclear clean energy agreement with Kairos Power, Google Blog 2024, https://blog.google/outreach-initiatives/sustainability/google-kairos-power-nuclear-energy-agreement/ Retrieved on 3 April 2025

Accelerating the addition of carbon-free energy: An update on progress, Microsoft September 2024, https://www.microsoft.com/en-us/microsoft-cloud/blog/2024/09/20/accelerating-the-addition-of-carbon-free-energy-an-update-on-progress/?msockid=323e7dd3d8f462c42425688cd96d6336 Retrieved on 3 April 2025

Accelerating a Carbon-Free Future Microsoft policy brief on advanced nuclear and fusion energy, December 2023, https://assets-global.website-files.com/6115b8dddcfc8904acfa3478/656f712cf82ecbc5ac1feab6_Accelerating%20a%20Carbon-Free%20Future%20-%20Microsoft%20Policy%20Brief%20on%20advanced%20nuclear%20and%20fusion%20energy.pdf

Goldman Sachs, 23 January 2025, https://www.goldmansachs.com/insights/articles/is-nuclear-energy-the-answer-to-ai-data-centers-power-consumption

The WIPO Conversation on IP and AI, https://www.wipo.int/about-ip/en/artificial_intelligence/policy.html, 30.03.2025

WIPO/IP/AI/2/GE/20/1 EUIPO Comments, https://www.wipo.int/export/sites/www/about-ip/en/artificial_intelligence/call_for_comments/pdf/org_european_union_euipo.pdf. 30.03.2025

Notes from the AI Frontier. Modeling the Impact Of AI on the World Economy, McKinsey Global Institute, September 2018,

https://www.mckinsey.com/~/media/McKinsey/Featured%20Insights/Artificial%20Intelligence/Notes%20from%20the%20frontier%20Modeling%20the%20impact%20of%20AI%20on%20the%20world%20economy/MGI-Notes-from-the-AI-frontier-Modeling-the-impact-of-AI-on-the-world-economy-September-2018

Emerging Trends and Future Outlook: The Data Labeling Industry in 2024-2030, Remote Labeler, 2025 https://www.remotelabeler.com/data-labeling-trends-in-2024-2030/. 29.03.2025

Generative AI Market, Markets and Markets, April 2024, https://www.marketsandmarkets.com/Market-Reports/generative-ai-market-142870584.html, 29.03.2025

Roadmap: The rise of synthetic media, Bessemer Venture Partners, July 2022, https://www.bvp.com/atlas/roadmap-the-rise-of-synthetic-media#, 29.03.2025

Autonomous Vehicle Market To Reach $214.32 Billion By 2030, Grand View Research, December 2023 https://www.grandviewresearch.com/press-release/global-autonomous-vehicles-market#, 29.03.2025

Global AI Chip Market Worth $305 Billion by 2030, Tech Strong. AI, https://techstrong.ai/articles/global-ai-chip-market-worth-305-billion-by-2030/#, 29.03.2025

ASI - Artificial Superintelligence Alliance https://singularitynet.io/ , 29.03.2025

Latest Data Annotation Trends: What's Transforming AI Models in 2024, data Science Society, October 2024, https://www.datasciencesociety.net/latest-data-annotation-trends-whats-transforming-ai-models-in-2024/#

새로운 비즈니스 AI 모델 10 가지, https://www.hankyung.com/article/202105256916i# , 29.03.2025

Rabby., F. Chimhundu., R. & Hassan, R. (2021). Artificial intelligence in digital marketing influences consumer behaviour: a review and theoretical foundation for future research. Academy of Marketing Studies Journal, 25(5), 1-7, https://www.abacademies.org/articles/artificial-intelligence-in-digital-marketing-influences-consumer-behaviour-a-review-and-theoretical-foundation-for-future-research-11892.html#

Generative AI and Consumer Behaviour: A New Era of Personalisation and Efficiency, Horton International, https://hortoninternational.com/generative-ai-and-consumer-behaviour-a-new-era-of-personalisation-and-efficiency/# , 29.03.2025

Livadiotis, Chris, The Impact of Artificial Intelligence on Consumer Behavior Analysis in Digital Marketing, June 2024, Updated 25 March 2025, https://www.visionfactory.org/post/the-impact-of-artificial-intelligence-on-consumer-behavior-analysis-in-digital-marketing#, 29.03.2025

The Future of Marketing: Predicting Consumer Behavior with AI, MarTech Health Technology, 2023, https://martech.health/articles/the-future-of-marketing-predicting-consumer-behavior-with-ai# , 29.03.2025

GANs vs. VAEs: What is the best generative AI approach?, TechTarget, March 2025, https://www.techtarget.com/searchenterpriseai/feature/GANs-vs-VAEs-What-is-the-best-generative-AI-approach , 29.03.2025

Sanju, Maharjan, Artificial Intelligence in Online Shopping: Impact on Consumer Behaviour, LAB University of Applied Sciences, 2024 https://www.theseus.fi/bitstream/handle/10024/865395/Maharjan_Sanju.pdf?sequence=2&isAllowed=y# , 29.03.2025

OJ L, 2024/1689, 12.7.2024, ELI: http://data.europa.eu/eli/reg/2024/1689/oj

ISO,, Artificial Intelligence (AI) — Assessment of the robustness of neural networks — Part 1: Overview, ISO/IEC TR 24029-1:2021, 18.03.2025

OECD, "Explanatory memorandum on the updated OECD definition of an AI system", *OECD Artificial Intelligence Papers*, 2024, No. 8, OECD Publishing, Paris, https://doi.org/10.1787/623da898-en.

Stanford HAI, "AI Key Terms Glossary Definition", 2022, https://hai-production.s3.amazonaws.com/files/2023-03/AI-Key-Terms-Glossary-Definition.pdf , 18. 03.2025

Google Cloud, "What is Artificial Intelligence (AI)?"https://cloud.google.com/learn/what-is-artificial-intelligence, 18.03.2025

NVIDIA, "Artificial Intelligence" https://www.nvidia.com/en-us/glossary/artificial-intelligence/ , 18.03. 2025

IBM, "What is artificial intelligence (AI)?" 2024, https://www.ibm.com/think/topics/artificial-intelligence, 18.03.2025

Eaton Business School, Artificial Intelligence (AI): What it is and why it matters", https://ebsedu.org/blog/artificial-intelligence/ , 18.03.2025

IBM, "Understanding the different types of Artificial Intelligence", 2023, https://www.ibm.com/think/topics/artificial-intelligence-types, 18.03. 2025

Eaton Business School, "Understanding 7 types of AI (artificial intelligence) with examples" https://ebsedu.org/blog/7-types-of-artificial-intelligence, 18.03. 2025

Avast, "Different Types of Artificial Intelligence (AI) You Need To Know", 2025 https://www.avast.com/c-types-of-ai-explained 18.03.2025

Lantern Studios, "History of AI: From Rule-based Algorithms to Generative Models, 2024, "https://lanternstudios.com/insights/blog/the-history-of-ai-from-rules-based-algorithms-to-generative-models/ , 19.03.2025

Microsoft, "What are the different types of AI?", 2023, https://www.microsoft.com/en-us/microsoft-365-life-hacks/writing/what-are-the-different-types-of-ai, 18.03.2025

IBM,"AI vs. machine learning vs. deep learning vs. neural networks: What's the difference?", 2023, https://www.ibm.com/think/topics/ai-vs-machine-learning-vs-deep-learning-vs-neural-networks 19.03.2025

Microsoft, "What is deep learning?", https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-deep-learning#:~:text=Deep%20learning%20is%20a%20type,%2C%20make%20recommendations%2C%20and%20adapt. , 19.03.2025

Google Cloud, "What is deep learning?" https://cloud.google.com/discover/what-is-deep-learning?hl=en 19.03. 2025

IBM, "What is deep learning", 2024, https://www.ibm.com/think/topics/deep-learning ,19.03.2025

IBM, "What is a neural network?", https://www.ibm.com/think/topics/neural-networks, 19.03.2025

Oracle, "What is AI model training & why is it important?" , 2023 https://www.oracle.com/sa/artificial-intelligence/ai-model-training/ , 19.03.2025

IBM, "What is an AI model?" https://www.ibm.com/think/topics/ai-model 19.03.2025

Oracle, "What is AI model training & why is it important?" , 2023 https://www.oracle.com/sa/artificial-intelligence/ai-model-training/ , 19.03.2025

Huang, Ken, "Test Time Compute", 2024, https://cloudsecurityalliance.org/blog/2024/12/13/test-time-compute , 19.03.2025

OJ L, 2024/1689, 12.7.2024, ELI: http://data.europa.eu/eli/reg/2024/1689/oj 4

OECD, "Explanatory memorandum on the updated OECD definition of an AI system", *OECD Artificial Intelligence Papers*, 2024, No. 8, OECD Publishing, Paris, https://doi.org/10.1787/623da898-en. 9

ISO, "Information technology — Vocabulary — Part 28: Artificial intelligence — Basic concepts and expert systems", **ISO/IEC 2382-28:1995** , 20.03.2025

"National Strategy for Artificial Intelligence", 2018 https://www.niti.gov.in/sites/default/files/2023-03/National-Strategy-for-Artificial-Intelligence.pdf

"Approach Document for India Part 1 – Principles for Responsible AI ",2021, https://www.niti.gov.in/sites/default/files/2021-02/Responsible-AI-22022021.pdf

"Approach Document for India Part 2 – Operationalizing Principles for Responsible AI ",2021, https://www.niti.gov.in/sites/default/files/2021-08/Part2-Responsible-AI-12082021.pdf

"IndiaAI Portal", https://indiaai.gov.in/ 18.03.2025

"IndiaAI - Compute Capacity", https://indiaai.gov.in/hub/indiaai-compute-capacity 18.03.2025

"IndiaAI Innovation Center", https://indiaai.gov.in/hub/indiaai-innovation-centre 18.03.2025

"IndiaAI Datasets Platform", https://indiaai.gov.in/hub/indiaai-datasets-platform 18.03.2025

"IndiaAI Application Development Initiative", https://indiaai.gov.in/hub/indiaai-application-development-initiative ,18.03.2025

"IndiaAI FutureSkills", https://indiaai.gov.in/hub/indiaai-futureskills, 18.03.2025

"IndiaAI Startup Financing", https://indiaai.gov.in/hub/indiaai-startup-financing , 18.03.2025

"IndiaAI Safe & Trusted AI", https://indiaai.gov.in/hub/safe-trusted-ai , 18.03.2025

"National Strategy for Artificial Intelligence", 2019, https://www.msit.go.kr/bbs/documentView.do?atchFileNo=30011&fileOrdr=1 , 19.03.2025

"인공지능 윤리기준", https://ai.kisdi.re.kr/aieth/main/contents.do?menuNo=400029 19.03.2025

"인공지능 발전과 신뢰 기반 조성 등에 관한 기본법", 2025, https://www.law.go.kr/%EB%B2%95%EB%A0%B9/%EC%9D%B8%EA%B3%B5%EC%A7%80%EB%8A%A5%20%EB%B0%9C%EC%A0%84%EA%B3%BC%20%EC%8B%A0%EB%A2%B0%20%EA%B8%B0%EB%B0%98%20%EC%A1%B0%EC%84%B1%20%EB%93%B1%EC%97%90%20%EA%B4%80%ED%95%9C%20%EA%B8%B0%EB%B3%B8%EB%B2%95/(20676,20250121)

"AI Safety Institute", https://www.aisi.re.kr/kor 20.03.2025

"AI 안전정책", https://www.aisi.re.kr/kor/contents/10 20.03.2025

"AI Opportunities Action Plan", 2025, https://assets.publishing.service.gov.uk/media/67851771f0528401055d2329/ai_opportunities_action_plan.pdf 14.03.2025

"UK AI Security Institute", https://www.aisi.gov.uk/ 19.03.2025

"Artificial intelligence: ethics, governance and regulation", 2024, https://post.parliament.uk/artificial-intelligence-ethics-governance-and-regulation/ 19.03.2025

"Algorithmic Transparency Recording Standard Hub", 2024, https://www.gov.uk/government/collections/algorithmic-transparency-recording-standard-hub 21.03.2025

"Data and digital policy landscape", https://www.dataanddigital.gov.au/about/landscape 14.03.2025

"The Data and Digital Government Strategy", https://www.dataanddigital.gov.au/strategy 14.03.2025

"Mission: Delivering for All people and business", https://www.dataanddigital.gov.au/strategy/missions/delivering-for-all-people-and-business 14.03.2025

"Mission: Simple and seamless services", https://www.dataanddigital.gov.au/strategy/missions/simple-and-seamless-services 14.03.2025

"Mission: Government for the future", https://www.dataanddigital.gov.au/strategy/missions/government-for-the-future 14.03.2025

"Mission: Trusted and Secure", https://www.dataanddigital.gov.au/strategy/missions/trusted-and-secure 14.03.2025

"Mission: Data and digital foundations", https://www.dataanddigital.gov.au/strategy/missions/data-and-digital-foundations 14.03.2025

"Engaging with Artificial Intelligence", 2024, https://www.cyber.gov.au/resources-business-and-government/governance-and-user-education/artificial-intelligence/engaging-with-artificial-intelligence 13.03.2025

"Australia's AI Ethics Principles", https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-principles/australias-ai-ethics-principles 13.03.2025

"Policy for the Responsible Use of AI in Government", ver. 1.1, 2024, https://www.digital.gov.au/sites/default/files/documents/2024-10/Policy%20for%20the%20responsible%20use%20of%20AI%20in%20government%201.1.pdf 13.03.2025

"Artificial Intelligence in Government", https://www.digital.gov.au/policy/ai 13.03.2025

"Engaging with Artificial Intelligence", 2024, https://www.cyber.gov.au/resources-business-and-government/governance-and-user-education/artificial-intelligence/engaging-with-artificial-intelligence 13.03.2025

"Developing a National AI Capability Plan", 2024, https://www.industry.gov.au/news/developing-national-ai-capability-plan 13.03.2025

"Artificial Intelligence Ecosystem", https://ised-isde.canada.ca/site/ised/en/artificial-intelligence-ecosystem 11.03.2025

"Pan-Canadian Artificial Intelligence Strategy", https://ised-isde.canada.ca/site/ai-strategy/en 11.03.2025

"Canadian Sovereign AI Compute Strategy", https://ised-isde.canada.ca/site/ised/en/canadian-sovereign-ai-compute-strategy 11.03.2025

"Canadian Artificial Intelligence Safety Institute", https://ised-isde.canada.ca/site/ised/en/canadian-artificial-intelligence-safety-institute 11.03.2025

"Artificial Intelligence and Data Act", 2022, https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act - 11.03.2025

"Guide on the use of generative AI", https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/guide-use-generative-ai.html 17.03.2025

"NAIS 2.0 Singapore National AI Strategy", 2023, https://file.go.gov.sg/nais2023.pdf / 19.03.2025

"Model Artificial Intelligence Governance Framework" ed.2, 2020, https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf 19.03.2025

"Trusted Data Sharing Framework", 2019, https://www.imda.gov.sg/-/media/imda/files/programme/ai-data-innovation/trusted-data-sharing-framework.pdf

"Research", https://sgaisi.sg/our-works/ 19.03.2025

"Singapore Announces New AI Safety Initiatives at the Global AI Action Summit in France", 2025, https://www.mddi.gov.sg/media-centre/press-releases/singapore-announces-new-ai-safety-initiatives/ 13.03.2025

"Social Principles of Human-Centric AI ", https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/humancentricai.pdf

"Japan AI Safety Institute", https://aisi.go.jp/ 20.03.2025

"Overview of the Japan AI Safety Institute (J-AISI)", 2025, https://aisi.go.jp/assets/pdf/20250106_AISI_en.pdf

Habuka, H., "New Government Policy Shows Japan Favors a Light Touch for AI Regulaton", Center for Strategic and International Studies, 2025, retrieved from https://csis-website-prod.s3.amazonaws.com/s3fs-public/2025-02/250225_Habuka_Japan_AI_0.pdf?VersionId=aD5DJCIrllfgGup3x_ojtmE_vn4Y8NTP

"Microsoft and G42 announce $1 billion comprehensive digital ecosystem initiative for Kenya", 2024, https://news.microsoft.com/2024/05/22/microsoft-and-g42-announce-1-billion-comprehensive-digital-ecosystem-initiative-for-kenya/ 20.03.2025

"Kenya National Artificial Inteliligence (AI) Strategy 2025 – 2030 (draft)", 2025 https://ict.go.ke/sites/default/files/2025-01/Kenya%20National%20AI%20Strategy%20%28Draft%29%20for%20Public%20Validation%20%20%5B14-01-2025%5D.pdf

Cybersecurity is on the frontline of our AI future. Here's why, World Economic Forum 2024, https://www.weforum.org/stories/2024/01/cybersecurity-ai-frontline-artificial-intelligence/#

Achuthan K, Ramanathan S, Srinivas S, Raman R. Advancing cybersecurity and privacy with artificial intelligence: current trends and future research directions. Front Big Data. 2024 Dec 5;7:1497535. doi: 10.3389/fdata.2024.1497535. PMID: 39703783; PMCID: PMC11656524. https://pmc.ncbi.nlm.nih.gov/articles/PMC11656524/#, 29.03.2025

1527 София, ул. Чаталджа 76
www.bia-bg.com