

ИЗКУСТВЕН ИНТЕЛЕКТ В БИЗНЕСА

д-р Николай Лазаров

София
2026

Изкуствен интелект в бизнеса

© *Николай Лазаров*, автор

Рецензенти:

проф. д-р Даниела Бобева

доц. Едуард Маринов, д.н.

© *Вирджиния Анастасова*, дизайн на корицата

© *Издателство на УЗФ*, предпечат

Редакционна колегия на Издателство „Св. Григорий Боголсов“:

доц. Едуард Маринов, д.н.

гл. ас. д-р Никола Димитров

проф. Мариела Деливерска, д.н.

доц. д-р Виктория Гацова

проф. д-р Станислав Димитров

проф. д-р Теодора Лазарова

проф. д-р Али Вейсел

© Издателство на Университета по застраховане и финанси „Св. Григорий Боголсов“, 2025

ул. „Гусла“ 1, 1618 София

<https://uzf.bg/>

Всички права са запазени. Никаква част от тази публикация не може да бъде възпроизвеждана, съхранявана в система за търсене или предавана под каквато и да е форма или по какъвто и да е начин – електронна, механична, фотокопирна, презапис или по друг начин – без писменото разрешение на носителя на авторските права.

Публикацията следва да бъде коректно цитирана с позоваване по следния начин:

Лазаров, Н. (2026). *Изкуствен интелект в бизнеса*. София: Издателство на УЗФ „Св. Григорий Боголсов“.

ISBN (pdf): 978-619-7622-88-1



Съдържание

СЪДЪРЖАНИЕ	3
РЕЗЮМЕ: ИЗКУСТВЕН ИНТЕЛЕКТ В БИЗНЕСА	7
РЕЧНИК НА АБРЕВИАТУРИ И СЪКРАЩЕНИЯ	9
УВОД	11
МОДУЛ 1. РАДИКАЛНИ ИНОВАЦИИ, КОИТО ПРЕНАПИСВАТ ИНДУСТРИЯТА	13
1. ТЕХНОЛОГИЧНОТО ЦУНАМИ. КАК ИНОВАЦИИТЕ ПРЕОБРЪЩАТ БИЗНЕС ПЕЙЗАЖА?	14
<i>Фундаментални иновации, които ще определят следващото десетилетие</i>	17
<i>Теми за дискусия Глава 1</i>	28
2. ОТ АЛАН ТЮРИНГ ДО СНАТGPT, ИСТОРИЯТА, НАСТОЯЩЕТО И БЪДЕЩЕТО НА AI	29
<i>История на изкуствения интелект, от теоретични основи до генеративна трансформация</i>	29
<i>Теоретични основи и ранни визии (1930 – 1950)</i>	31
<i>Раждането на AI като научна дисциплина и първите успехи (1950 – 1960)</i>	32
<i>Разочарование и „AI зима“ (1970 – 1980)</i>	33
<i>Възраждане на интереса (1990 – 2010)</i>	33
<i>Ерата на дълбокото обучение (Deep Learning) и генеративния AI (Generative AI) (от 2010 г. до наши дни)</i>	33
<i>Основни видове и класификации на изкуствения интелект</i>	34
<i>Класификация според функционалността (AI Levels)</i>	34
<i>Класификация според способностите (AI Types)</i>	35
<i>Анализ и възможна еволюцията на AI до 2070 г.</i>	36
<i>Етапи на развитие и ключови фактори</i>	36
<i>Ключови ограничения и решения</i>	37
<i>Рискове и бъдещи роли</i>	38
<i>AI еволюция, движена от прагматизъм, която прекроява бизнеса и професионалните умения</i>	38
<i>Ефекти върху бизнеса, от трансформация на модели до конкурентно предимство</i> ..	38
<i>Необходими умения и познания за съвременните мениджъри и служители</i>	39
<i>Теми за дискусия Глава 2</i>	40
3. AI ЗА НАЧИНАЕЩИ, КЛЮЧОВИ КОНЦЕПЦИИ И ИКОНОМИЧЕСКИТЕ ЕФЕКТИ	41
<i>Икономическият ефект на AI, макро и микроикономически анализ</i>	42
<i>Предизвикателства и ограничения. Пътят към отговорен AI</i>	43
<i>„Техническият дълг“ като стратегически императив в дигиталната трансформация и имплементация на AI</i>	45
<i>Теми за дискусия Глава 3</i>	48

МОДУЛ 2. В СЪРЦЕТО НА ДИГИТАЛНИЯ БИЗНЕС. КЪДЕ AI СЕ СРЕЩА С КАПИТАЛА49

4. МАГИЯТА НА ПЕРСОНАЛИЗАЦИЯТА. КАК ДИГИТАЛНИТЕ ПЛАТФОРМИ ИЗПОЛЗВАТ AI, ЗА ДА ВИ ПОЗНАВАТ ПО-ДОБРЕ?	50
<i>Опасността зад алгоритъма</i>	52
<i>Продуктът като основа за екосистема</i>	54
<i>Урок за дигиталния бизнес</i>	54
<i>Платформата като икономика на изборите</i>	55
<i>Теми за дискусия Глава 4</i>	56
5. ОТ БУТИКОВ ДО МАСОВ ПРОДУКТ. AI И (НЕ)ВИДИМАТА ТРАНСФОРМАЦИЯ НА ПАЗАРА	57
<i>Класическият модел на приемането на иновациите</i>	57
<i>AI като катализатор на промяна, появата на нова крива</i>	59
<i>Основни фактори за ускорената крива на AI</i>	60
<i>Стратегически и социални последици</i>	61
<i>Теми за дискусия Глава 5</i>	63
6. ГОТОВИ ЛИ СМЕ ЗА БИЗНЕС, В КОЙТО AI ВЗЕМА ПОВЕЧЕТО РЕШЕНИЯ?	64
<i>От автоматизация към автономност</i>	65
<i>Рискове и предизвикателства</i>	67
<i>Примери от практиката</i>	68
<i>Възможно бъдещо развитие</i>	69
<i>Теми за дискусия Глава 6</i>	70
7. ЕФЕКТЪТ НА ЛАВИНАТА. ЗАЩО ПЛАТФОРМИТЕ С AI СА НЕУДЪРЖИМИ	71
<i>Мрежовите ефекти като „двигател на растежа“</i>	71
<i>Видове мрежови ефекти</i>	71
<i>Примери и механизми на действие</i>	73
<i>Стратегически последици и предизвикателства</i>	73
<i>Бъдещето на мрежовите ефекти - AI и Web3</i>	74
<i>Теми за дискусия Глава 7</i>	75
8. AI И НОВАТА ИКОНОМИКА. КАК ИНТЕЛИГЕНТНИТЕ СИСТЕМИ ПРЕОСМИСЛЯТ РАЗХОДИТЕ И МАЩАБА?	76
<i>Фундаменталната разлика между традиционни и дигитални продукти/услуги</i>	77
<i>Дигитализацията и Изкуствения Интелект в Персонализацията, Матрица Продукт-Процес</i>	78
<i>Ефект на AI - намаляване на разходите за персонализация</i>	79
<i>Новата дигитална парадигма - адаптивност и бързина</i>	79
<i>Ролята на AI в трансформацията на разходите</i>	81
<i>Практически казуси и стратегически последици</i>	82
<i>Теми за дискусия Глава 8</i>	83
9. ОТ ПРИХОДИ КЪМ СТОЙНОСТ. КАК AI РАЗКРИВА НОВИ ИЗТОЧНИЦИ НА РАСТЕЖ?	84
<i>AI като генератор на стойност - хиперперсонализация и клиентско преживяване</i> ...	84
<i>Нови източници на приходи, AI-базирани продукти и бизнес модели</i>	85
<i>Синергия между разходи и приходи и задвижване на ръста</i>	85
<i>Как AI променя модела за формиране на конкурентно предимство</i>	87
<i>Теми за дискусия Глава 9</i>	90

МОДУЛ 3. ИКОНОМИЧЕСКИЯТ ПОТЕНЦИАЛ НА AI. ОТ ТЕОРИЯ КЪМ ПРАКТИКА.....	91
10. AI НА РАБОТНОТО МЯСТО. ПРИМЕРИ КАК AI ПРОМЕНЯ ИНДУСТРИТА	92
<i>Какви проблеми решава машинното обучение (ML)?</i>	92
<i>Обучение под Наблюдение (Supervised Learning)</i>	92
<i>Обучение без Наблюдение (Unsupervised Learning)</i>	92
<i>Обучение с Подсилване (Reinforcement Learning - RL)</i>	93
<i>Ползи v/s рискове при използването на ML</i>	93
<i>Бизнес приложения на ML</i>	95
<i>Големи езикови модели (LLM – Large Language Model)</i>	96
<i>Синергия между ML и LLM</i>	98
<i>Ограничения, рискове и етични предизвикателства</i>	99
<i>Теми за дискусия Глава 10</i>	101
11. AI КАТО СТРАТЕГИЧЕСКИ РЕСУРС. КАК ДА ПРЕВЪРНЕТЕ ТЕХНОЛОГИЯТА В КОНКУРЕНТНО ПРЕДИМСТВО	102
<i>AI в корпоративната стратегия</i>	103
<i>Регулаторна конвергенция</i>	104
<i>Модели за създаване на стойност с AI</i>	106
<i>Организационни трансформации</i>	106
<i>Теми за дискусия Глава 11</i>	107
12. ОРГАНИЗАЦИОННА РЕВОЛЮЦИЯ. КАК ДА УПРАВЛЯВАТЕ РИСКА И ПРОМЯНАТА, КОГАТО ВНЕДРЯВАТЕ AI	108
<i>Организационна трансформация и културна адаптация</i>	108
<i>Нови организационни роли и управление на промяната</i>	109
<i>Управление на риска, етика и регулации</i>	109
<i>Стратегически избори: Build, Buy или Partner</i>	110
<i>Примери за неуспешни практики при имплементиране на AI</i>	111
<i>Теми за дискусия Глава 12</i>	113
13. AI В БИЗНЕСА - ОТ ФИНАНСИТЕ И МАРКЕТИНГА ДО ЧОВЕШКИТЕ РЕСУРСИ	114
<i>Ролята на ИИ във финансите и финтех</i>	114
<i>Кредитен скоринг и управление на риска</i>	114
<i>Предотвратяване на измами в реално време</i>	114
<i>Прогнозиране и анализ</i>	115
<i>AI в подбора на персонал</i>	115
<i>AI в оценката на представянето</i>	116
<i>AI в Learning & Development (L&D) и вътрешна мобилност</i>	116
<i>Управление на рисковете в HR алгоритмите</i>	117
<i>Автоматизация (RPE - Robotic Process Automation)</i>	117
<i>Теми за дискусия Глава 13</i>	117
14. ЛОГИСТИКА НА БЪДЕЩЕТО. КАК AI ПРАВИ ВЕРИГАТА НА ДОСТАВКИ ПО-ЕФЕКТИВНА?....	119
<i>AI в операции и управление на веригата на доставки</i>	119
<i>Прогнозиране на търсенето</i>	120
<i>Оптимизация на запаси и маршрути</i>	121
<i>Интелигентни агенти и автономни системи</i>	121
<i>Теми за дискусия Глава 14</i>	122

МОДУЛ 4. ПРАВИЛА, ЕТИКА И СИГУРНОСТ	123
15. AI в рамка. РЕГУЛАЦИИТЕ, КОИТО ЩЕ ОФОРМЯТ БЪДЕЩЕТО НА ТЕХНОЛОГИЯТА	124
<i>Пристрастия и bias в AI системите</i>	<i>124</i>
<i>Сравнителен анализ на етичните рамки.....</i>	<i>125</i>
<i>GDPR и AI</i>	<i>126</i>
<i>EU AI Act (2024).....</i>	<i>126</i>
<i>Управление на данните и моделите.....</i>	<i>128</i>
<i>Теми за дискусия глава 15.....</i>	<i>131</i>
16. ЗАМЪК ОТ ДАННИ. КАК ДА ЗАЩИТИМ AI ОТ КИБЕРАТАКИ И ДА ИЗГРАДИМ УСТОЙЧИВИ СИСТЕМИ	132
<i>Prompt Injection и adversarial атаки, заплахати за надеждността на моделите</i>	<i>133</i>
<i>Киберустойчивост на AI системи</i>	<i>134</i>
<i>Интелектуална собственост (IP) и лицензиране в AI епохата</i>	<i>135</i>
<i>Теми за дискусии Глава 16.....</i>	<i>135</i>
17. AI НА ПРАКТИКА. КАК ДА ИЗГРАДИТЕ СВОЯ ПЪРВИ БИЗНЕС ПЛАН И ПРОТОТИП	136
<i>Отвъд краткосрочния PR към дългосрочна стратегическа стойност - измерване на възвръщаемостта от AI</i>	<i>136</i>
<i>KPIs като PR инструмент.....</i>	<i>136</i>
<i>Скритите разходи и нови зависимости - критичен поглед към внедряването.....</i>	<i>137</i>
<i>Логиката на мащаба и неизбежната консолидация.....</i>	<i>138</i>
<i>Траектория на бизнес стойността на AI (ABVT) – авторски модел.....</i>	<i>139</i>
<i>Теми за дискусия Глава 17.....</i>	<i>143</i>
18. ДИГИТАЛНИЯТ РАЗРИВ. ИЗОСТАВАНЕТО НА БЪЛГАРИЯ В ЕРАТА НА AI И ПОСЛЕДИЦИТЕ ЗА КОНКУРЕНТОСПОСОБНОСТТА НА БИЗНЕСА.....	144
<i>Европейският дигитален пейзаж – неравномерно развитие и нарастващи различия ...</i>	<i>144</i>
<i>Интеграция на AI, новата граница на дигиталното развитие</i>	<i>147</i>
<i>Икономически последици от дигиталното изоставане</i>	<i>149</i>
<i>Цената на изоставането, сценарии за внедряване на AI в бизнеса и икономически ефекти за България.....</i>	<i>150</i>
<i>Дигитализация и конкурентоспособност.....</i>	<i>151</i>
<i>Последици за българския бизнес и сценарии за развитие</i>	<i>154</i>
<i>Теми за дискусия Глава 18.....</i>	<i>156</i>
ЗАКЛЮЧЕНИЕ. БЪДЕЩЕТО Е СЕГА	157
<i>Препоръки за успеха в ерата на AI.....</i>	<i>157</i>
БИБЛИОГРАФИЯ.....	159
ПРИЛОЖЕНИЕ	167

Резюме: Изкуствен интелект в бизнеса

Д-р Николай Лазаров

Email: nikolay.lazarov060680@gmail.com

Резюме: Настоящият труд изследва въздействието на изкуствения интелект (AI) върху трансформацията на съвременния бизнес, като представя икономическите, организационните и технологичните му измерения. Авторът развива тезата, че иновациите, базирани на AI, не са алтернатива, а екзистенциален императив за устойчивост и конкурентоспособност. Чрез анализ на водещи концепции от „съзидателното разрушение“ на Шумпетер до генеративните модели (LLM, RAG, RLHF), ръкописът очертава рамката, в която технологичните проби, машинното обучение, роботизацията, блокчейнът и облачните решения променят бизнес моделите и пазарните структури. Представени са конкретни приложения на AI във финансите, маркетинга, управлението на риска и човешките ресурси, както и ролята на регулациите (AI Act, GDPR, Basel III) за гарантиране на прозрачност и сигурност. Анализът показва, че внедряването на AI води до експоненциален ръст на производителността, но изисква нова управленска култура, етична зрялост и дигитална компетентност.

Abstract: This manuscript examines the transformative impact of Artificial Intelligence (AI) on modern business, exploring its economic, organizational, and technological dimensions. The author argues that AI-driven innovation is not an option but an existential imperative for competitiveness and sustainability. Drawing on foundational theories from Schumpeter’s “creative destruction” to modern generative models (LLM, RAG, RLHF), the study provides a strategic framework explaining how AI, machine learning, robotics, blockchain, and cloud computing reshape business models and market structures. It highlights concrete applications of AI in finance, marketing, risk management, and human capital, alongside the role of emerging regulations (AI Act, GDPR, Basel III) in ensuring transparency and security. The findings demonstrate that AI implementation exponentially enhances productivity while demanding a new managerial mindset, ethical responsibility, and digital literacy across all sectors.

Keywords: Artificial Intelligence; Machine Learning; Generative AI; Digital Transformation; Fintech; Business Models; Innovation; Productivity; Regulation; Ethics; Cloud Computing; Blockchain; Robotization

JEL Classification: O33, L86, M15, M21, G21

За автора: д-р Николай Лазаров е старши банков и финтех директор с над 20 години опит в дигиталната трансформация, иновациите, банкиранот и платежните системи. Ръководил е стратегически проекти в дигиталното банкиране, интегриране на AI и отвореното банкиране, платежни еко-системи и дигитален маркетинг. Има утвърден опит в развитието на партньорства с глобални компании и във внедряването на иновативни бизнес модели, съчетаващи клиентски фокус, регулаторно съответствие и технологично лидерство. Притежава докторска степен от УЗФ.

Благодарности: Авторът изразява благодарност към колеги и съмишленици за конструктивните дискусии по темите за използването и имплементирането на AI в различни бизнес и социални процеси.

Принос на автора: Авторът носи единствена отговорност за концепцията, анализа, написването и финализирането на ръкописа.

Декларация за конфликт на интереси и авторство: Авторът декларира, че няма конфликт на интереси. Авторът потвърждава, че ръкописът е оригинален, не е бил публикуван преди и в момента не е в процес на разглеждане за публикуване другаде.

Общи цели на книгата

- Да запознае с основните концепции и модели на изкуствения интелект (AI) и генеративния AI.
- Да развие умения за анализ на бизнес модели, трансформация и внедряване на AI.
- Да демонстрира икономическите ефекти и ограничения на AI.
- Да подготви за идентифициране на възможности за стойност в ключови бизнес функции.
- Да изгради критично мислене относно регулации, прозрачност и етика при използване на AI.

Речник на абривиатури и съкращения

- A/B Test – Експеримент с контролна и вариационна група.
- AGI – Общ изкуствен интелект.
- АНТ – Средно време за обработка на контакт.
- AI – Изкуствен интелект.
- AML5 – Пета директива срещу изпирането на пари.
- ANI – Слаб/тесен изкуствен интелект.
- AOV – Средна стойност на поръчка.
- API – Интерфейс за програмиране на приложения.
- AR – Добавена реалност.
- ARPU – Среден приход на потребител.
- BaaS – Бизнес процес като услуга.
- Byz Fault Tolerance – Толерантност към византийски грешки.
- CAC – Цена за придобиване на клиент.
- CapEx – Капиталови разходи.
- CI/CD – Непрекъсната интеграция и внедряване.
- Containment – Процент заявки, решени без ескалация към човек.
- CSAT – Удовлетвореност на клиент.
- D1/D7/D30 Retention – Ретенция на потребители на 1/7/30 ден.
- DevSecOps – Интеграция на сигурността в CI/CD процеси.
- DL – Дълбочинно обучение чрез многослойни невронни мрежи.
- DLT – Технология за разпределен регистър.
- Edge Computing – Изчисления близо до източника на данни/потребителя.
- Egress – Изходящ трафик от облак.
- Embeddings – Векторни представяния за търсене и препоръки.
- FAANGS – Facebook, Apple, Amazon, Netflix, Google, Spotify.
- FCR – Решаване от първи контакт.
- FinOps – Управление на облачни разходи.
- GDPR – Общ регламент за защита на данните.
- GenAI – Генеративен AI – модели за създаване на текст, изображения, код.
- Groundedness – Степен на опора на отговор в проверими източници.
- Holdout – Контролна група без намеса.
- IaaS – Инфраструктура като услуга.
- IEA – Международна агенция по енергетика.
- IoT – Интернет на нещата.
- KPI – Ключов показател за изпълнение.
- KYC – Опознай своя клиент (регулаторен процес).
- LLM – Голям езиков модел, основа за GenAI.
- LTV – Жизнена стойност на клиент.
- ML – Машинно обучение.
- MLOps – Жизнен цикъл на ML модели.
- MR – Смесена реалност.
- MTTR – Средно време до възстановяване.
- NBER – Национално бюро за икономически изследвания (САЩ).

- NLP – Обработка на естествен език.
- NPS – Нетен промоутърски резултат.
- OpEx – Оперативни разходи.
- P2P – Плащания или трансакции между равнопоставени участници.
- p95/p99 Latency – 95/99 персентил време за отговор.
- PaaS – Платформа като услуга.
- PI – Лични идентифициращи данни.
- PSD2 – Втора директива за платежни услуги.
- QJE – Водещ икономически журнал.
- R&D – Научноизследователска и развойна дейност.
- RLHF – Обучение с подсилване и човешка обратна връзка.
- ROAS – Възвръщаемост на рекламни разходи.
- RPA – Роботизирана автоматизация на процеси.
- RPS – Заявки в секунда.
- SaaS – Софтуер като услуга.
- SLA – Споразумение за ниво на услуга.
- SLO – Целево ниво на услуга.
- TCO – Обща стойност на притежание.
- TMT – Телекомуникации, медии и технологии.
- Unit Economics – Единични икономики – цена/1k токена, разход/сесия.
- Vendor Lock-in – Зависимост от доставчик/екосистема.
- VR – Виртуална реалност.

Увод

В ерата на четвъртата индустриална революция, **изкуственият интелект (AI)** се превърна от футуристична концепция във фундаментален двигател на промяната, трансформиращ всеки аспект на съвременния бизнес. Този документ представя цялостна учебна програма, която има за цел да подготви бъдещите лидери за предизвикателствата и възможностите на тази нова ера. Тя се основава на пресечната точка между стратегическия мениджмънт, финансовата индустрия и технологичните иновации.

Програмата е структурирана в логически свързани модули, които обхващат пътя от разбирането на фундаменталните бизнес модели до практическата реализация на AI.

Модул 1. Радикални иновации, които пренаписват индустрията

- Технологичното цунами. Как иновациите преобръщат бизнес пейзажа?
- От Алан Тюринг до ChatGPT, Историята, настоящето и бъдещето на AI
- AI за начинаещи, ключови концепции и икономическите ефекти

Модул 2. В сърцето на дигиталния бизнес. Къде AI се среща с капитала

- Магията на персонализацията. Как дигиталните платформи използват AI, за да ви познават по-добре?
- От бутиков до масов продукт. AI и (не)видимата трансформация на пазара
- Готови ли сме за бизнес, в който AI взема повечето решения
- Ефектът на лавината. Защо платформите с AI са неудържими
- AI и Новата Икономика. Как интелигентните системи преосмислят разходите и мащаба?
- От приходи към стойност. Как AI разкрива нови източници на растеж?
- Персонализацията, която променя уравнението на бизнес модела

Модул 3. Икономическият потенциал на AI. От теория към практика

- AI на работното място. Практически примери за това как AI променя вашата индустрия
- AI като стратегически ресурс. Как да превърнете технологията в конкурентно предимство
- Организационна революция. Как да управлявате риска и промяната, когато внедрявате AI
- AI в бизнеса - от финансите и маркетинга до човешките ресурси
- Логистика на бъдещето. Как AI прави веригата на доставки по-ефективна?

Модул 4. Правила, Етика и Сигурност

- AI в рамка. Регулациите, които ще оформят бъдещето на технологията
- Замък от данни. Как да защитим AI от кибератаки и да изградим устойчиви системи
- AI на практика. Как да изградите своя първи бизнес план и прототип

Познаването на тези теми не е просто академично упражнение, а жизненоважна необходимост. В съвременната икономика, конкурентното предимство все повече се изгражда върху способността на една организация да използва данни и интелигентни системи за създаване на уникална стойност за своите клиенти. За бъдещите мениджъри, особено във финансовия и застрахователния сектор, тези знания са критични за успешното ръководство на екипи и за иновирането в среда, където AI е неразделна част от бизнес стратегията.

Модул 1. Радикални иновации, които пренаписват индустрията

1. Технологичното цунами. Как иновациите преобръщат бизнес пейзажа?

„Отварянето на нови пазари, вътрешни или външни и организационното развитие илюстрират един и същи процес на индустриална мутаци, който непрекъснато революционизира икономическата структура отвътре, непрекъснато разрушава старата и непрекъснато създава нова. Този процес на съзидателно разрушение е същественният факт за капитализма.“

Йозеф А. Шумпетер, Capitalism, Socialism and Democracy, 1942, Част II, Глава VII

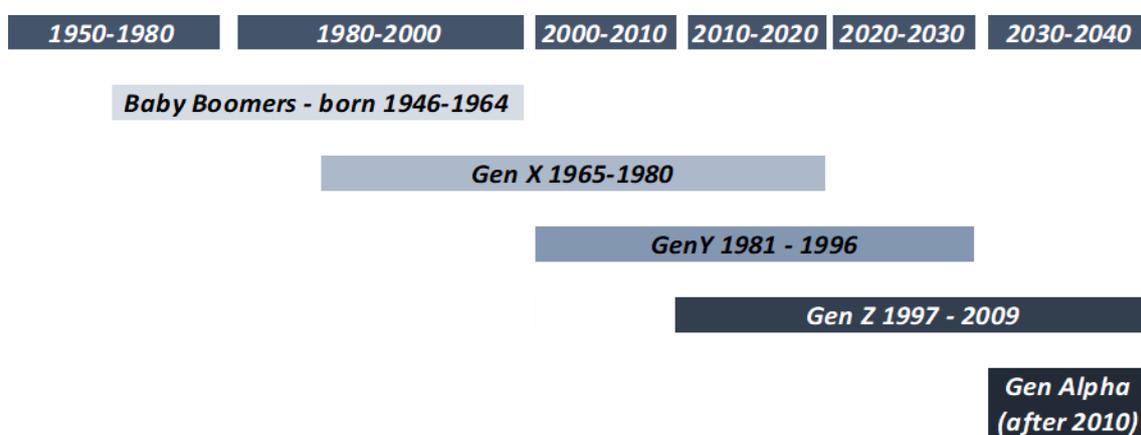
Съвременната индустрия, особено финансовата, е изправена пред **„Технологично цунами“** – период на експоненциален растеж на технологии като изкуствен интелект (AI), Машинно обучение (ML) и Блокчейн. В този контекст, тезата, че **иновациите са без алтернатива**, се превръща от стратегически избор в **екзистенциален императив**. Тази позиция се основава на каноничните икономически теории за развитие и е категорично подкрепена от емпиричните данни за дигиталната трансформация и промените в потребителското търсене. Тезата намира своя теоретичен корен в работата на **Йозеф Шумпетер**, който определя **„Съзидателното разрушение“ (Creative Destruction)** като **„същественният факт за капитализма“ (1942)**. Според Шумпетер, икономическият растеж не е резултат от постепенно увеличаване на производството, а от **радикални иновации**, които непрекъснато разрушават старите индустриални структури и създават нови. За един капиталистически концерн, ако не инициира промяната, той ще бъде унищожен от нея. В контекста на индустрията, „старите бизнес модели не могат да оцелеят срещу нови, по-ефективни модели, базирани на AI и ML. Тази концепция е доразвита от **Клейтън Кристенсен** с неговата теория за **„Разрушителните иновации“ (Disruptive Innovation)**. Кристенсен твърди: **„Ако не разрушите собствения си бизнес, някой друг ще го направи.“** тезата на **Клейтън Кристенсен** подчертава, че пасивната позиция спрямо AI и ML е равностепенна на стратегическо самоубийство. Съвременните икономисти **Ерик Бринолфсон** и **Андрю Макафи** превръщат тезата на Шумпетер в императив, като обясняват **експоненциалния характер** на дигиталните технологии в **„The Second Machine Age“ (2014)**. Brynjolfsson & McAfee стигат до извода, че **„темпът на промяна не е линеен, той е експоненциален, това, което е трудно да се схване днес, утре ще бъде още по-трудно да се наваксате.“** (Brynjolfsson & McAfee). Паралел може да направим с AI и ML моделите, които се подобряват сами чрез обучение от масиви данни и **технологичната разлика се увеличава експоненциално**. **Котлър, Картаджая и Сетианян (2021)** в **„Marketing 5.0“** подчертават, че поколенията Z и Алфа, които ще съставляват **60-70%** от банковите клиенти, **изискват цифрови услуги, бързина и хиперперсонализация** ("segment of one"). Липсата на иновации в AR/VR банкирането, IoT персонализацията и AI-чатботовете не е просто лоша стратегия, а **отказ от обслужване на основната част от пазара**. **Агравал, Ганс и Голдфарб „Prediction Machines“ (2018)** въвеждат рамката

„предсказване срещу преценка, екипа изследователи стига до извода, че **по-доброто предсказване повишава ценността на преценката, машините изчисляващи прогнози не предоставят преценка. Само хората го правят. В практиката ML автоматизира рутинните задачи (напр. кредитен скоринг, откриване на измами), водещи до намаляване на разходите до 80% (Accenture). Този скок в Ефективността и Разходите (AI/RPA) прави старите процеси ценово неконкурентни. Внедряването на AI/RPA е задължително, за да може ограничената човешка преценка да се фокусира върху стратегически, нерутинни дейности. Задължителният характер на иновацията се простира и до инфраструктурата. Дон Тапскот (2016) определя Блокчейн като „първата естествена цифрова среда за стойност в основата на второто поколение на дигиталната революция.“ Блокчейн осигурява Byzantine fault толеранс и децентрализирано доверие. Институция, която не приеме тази фундаментална иновация за транзакции и KYC, остава обременена с неприемливо висок риск от измами и оперативни разходи спрямо FinTech конкурентите, които използват системата, която не може да бъде манипулирана. Критиците и регулаторите изтъкват, че иновацията е **рискова и неетична** без внимателно управление. "Черната кутия" и Рискът "Черен Лебед": Изследователите Agrawal, Gans, Goldfarb (2018) предупреждават, че 80% от времето в AI проектите отива за справяне с данни и "черни лебеди" (непредвидени отклонения). Това повдига въпроса за „Непрозрачността на решенията на ML“, която поражда регулаторни и етични проблеми. Блокчейн влиза в директен конфликт с основни принципи на GDPR, създавайки правен риск за индустрията. Както отбелязват регулаторните органи (ЕБО, Базелски комитет), доминацията на Amazon, Microsoft и Google на пазара на Облачни услуги (Cloud Computing) създава нов вид системен риск – зависимост от малък брой Big Tech доставчици. Тези възражения, макар и валидни, не предлагат алтернативен път за индустрията. Те предлагат само **по-внимателен, по-добре регулиран и по-етичен** път напред. Както обобщава Питър Дракър „Единственото, което знаем за бъдещето, е, че то ще бъде различно. Иновацията е инструментът, който променя.“ Индустрията не може да спре промяната, тя може само да я управлява. Радикалните иновации не са лукс или опция за съвременната индустрия, а **абсолютно условие за оцеляване**. Конкурентният натиск, експоненциалният темп на технологично развитие (Бринолфсон), потребителското търсене (Котлър) и императивът за ефективност (Шумпетер) създават **единствена траектория** на развитие, насочена към дигитализация. Въпреки че пътят е обсапан с **рискове** от алгоритмична непрозрачност, регулаторни конфликти и социално напрежение, **алтернативата на иновацията не е стабилност или запазване на статуквото, а сигурна стагнация и елиминация от пазара на фона на Технологичното цунами.****

Наред с динамичната дигитализация сме свидетели и на бързите радикални промени в нагласите на потребителите, които революционизират средата и се очаква допълнително да се ускорят през следващите десетилетия. Трансформацията ще бъде задвижвани не само от технологичния напредък, но и от по-широки, по-важни структурни фактори, като промени в демографията, нарастваща всеобхватност на дигитализация, лесен достъп до информация, по-голямо финансово и дигитално образование и развитието на

AI. След 2030 г. хората от т.нар. поколение алфа (родени между 2010 и 2025 г.), израснали силно изложени на усъвършенствани цифрови услуги, ще се влеят в работната сила. Заедно с поколението Z (родени между 1997 и 2009 г.), те ще представляват около 1/3 от глобалните потребители. Добавяйки и Millennials или поколението Y (родени между 1981 и 1996 г.), през следващите 10 години делът на клиентите, които ще търсят цифрови услуги и по-малко вероятно ще използват традиционните методи на дистрибуция ще достигне 60-70%. Общият знаменател на всички тях е промяна в потребностите, предпочитанията и изискванията им при търсене на нови финансови решения и възможност да получат информация (Kotler, Kartajaya, Setiawan, 2021).

Фигура 1. Поколения и активния им период в икономическия живот



Източник: Kotler, Kartajaya, Setiawan, 2021.

За да се оцени напълно влиянието на смяната на поколенията върху иновациите и дигитализацията, е нлючово да се разбират основните характеристики на различните поколения:

- *поколение Baby Boomers* (родени в периода 1946-1964 г.) – те ценят работните места, ориентирани са към демократични и хуманни ценности, равни възможности и приятелска среда, те са първите, израснали в медийна среда с телевизия, спазват работната етика, ориентирани са към постигане на цели (target-oriented), дисциплинирани при работа в група;
- *поколение X* (родени в периода 1965-1980 г.) – най-многобройното поколение в момента; те са директни в комуникацията си и притежават висока трудова етика. Това е поколението, което въвежда грижата за баланс между работа и личен живот. То е и първото използващо активно новите технологии, въпреки че е израснало без интернет;
- *поколение Y* (родени в периода 1981-1996 г.) – силно ориентирани към новите технологии; предпочитат гъвкави работни графици и по-добър баланс между работа и живот; имат значително по-високи очаквания към работодателите си по отношение на работна среда и предпоставки за обучение и кариерно израстване; работата в група е много високо ценена, както и бързата не усложнена комуникация;

- *поколение Z* (родени между 1997 и 2009 г.) – с най-висока пристрастеност към социални мрежи и интернет свързаност. Поколениято на инфлуенсърите и блогърите. При тях преобладава силното желание нещата да се случват бързо и лесно. Характерен е мултикултурализмът, половата толерантност, обвързването с каузи. Това е поколениято с най-високи изисквания към работа и работодател, предпочиташо по-скоро да остане без работа, отколкото да работи такава, която не удовлетворява очакванията му;
- *поколение Алфа* (родени след 2010 г.) – истинско поколение на XXI век, изцяло адаптирано към новите технологии и силно повлияно от родителите си, които също ползват изразено дигитални технологични решения. Това е първото поколение, което разглежда технологиите като неизменна част от своя живот.

Фундаментални иновации, които ще определят следващото десетилетие

Изкуственият интелект (AI)

Изкуствен интелект, в най-широк контекст разглеждан като способността на една машина да придобие способности, характерни за хората – да разсъждава, да взема автономни решения, да се учи, да планира, или да твори. Изкуственият интелект, по същество е техническа система, която наблюдава околната си среда, получава данни (подготвени от друг или които набира сама), преработва ги и извършва действия, свързани с постигане на конкретна цел. **Джон Маккарти (1956)** определя изкуствения интелект като "наука и инженерство за създаване на интелигентни машини". Това кратко, но всеобхватно определение се фокусира върху целта и методологията на AI. **Марвин Мински (MIT)** дефинира AI като науката за това как да накараме машините да правят неща, които, ако бъдат направени от хора, биха се считали за проява на интелигентност. Неговото определение акцентира върху имитирането на човешките способности. **Институт "Алън Тюринг"**, водещият британски институт за данни и AI определя AI като "интелигентно поведение, имитирано от машини". Те подчертават, че това е по-широко понятие от машинното обучение и включва области като роботика, компютърно зрение и обработка на естествен език. Системите с изкуствен интелект са способни да адаптират поведението си до известна степен, като анализират резултатите от предишни действия, и работят автономно. Прогнозите са, че изкуственият интелект ще доведе до големи промени, дори цялостна дигитална трансформация на обществото. Развитието на изкуствения интелект се дължи и на огромното количество данни, генерирани през последното десетилетие, увеличаващи се буквално ежедневно, и на понижаването на цените за съхранение на тези данни (Agrawal, Gans, Goldfarb, 2018). Всички AI модели, са изградени върху използване на минали данни, което изтъква важността на добре структурираните и чисти данни. Средно 80% от времето в проектите с изкуствен интелект са посветени на настройка на данните, адаптиране на модела към отклонения, или „черни лебеди“, които биха нарушили прогнозите (Agrawal, Gans, Goldfarb, 2018). Основните типове изкуствен интелект са виртуални асистенти, софтуер

за разпознаване на изображения, онлайн търсачки, системи за разпознаване на реч и лица и изкуствен интелект с физическа форма като роботи, автономни коли, дроне.

Машинното обучение (machine learning)

Машинното обучение е сърцето на съвременния AI. То е процес, при който алгоритмите се обучават от огромни масиви от данни, за да разпознават модели и да правят прогнози, без да са изрично програмирани за това. Например, банките използват ML за анализ на хиляди променливи от финансовата история на клиент, за да определят кредитоспособност и да прогнозират риска. DBS Bank (Сингапур) отчита, че с над 1500 внедрени ML модела са се насочили към постигане на икономическа стойност от 1 милиард SGD до 2025 г. Компании като Amazon използват ML за прогнозиране на търсенето и оптимизация на складовите наличности, което според McKinsey е донесло допълнителни приходи от милиарди чрез намаляване на липсите и свръхзапасите. При машинното обучение знанията се генерират на база на човешки опит и масивно количество данни (big data). В ерата на „big data“, машинното обучение превъзхожда намирането на модели, закономерности и прогнозни фактори в много големи масиви от данни. Където и да се прилагат тези условия, може да се приложи AI, за да се направят изключително точни прогнози. Чрез машинно обучение могат да се оценят желанията и нуждите на клиентите и да се персонализират маркетинговите подходи, с което стартира т.нар. ера на нано маркетинга, или концепция „segment of one“ или позната още като хиперперсонализиране, което от една страна оптимизира услугите и повишава лоялността на клиентите, а от друга проследява и контролира тенденциите за „миграция“ на клиенти и дава решения за задържане или връщане на клиентите. Технологични лидери на машинното обучение са Microsoft, Google, Facebook, IBM, Amazon и Apple. Банките в съвременния свят са силно привлечени от огромните възможности на технологиите за изкуствен интелект и машинно обучение за рационализиране на банковата дейност. Технологиите за изкуствен интелект и машинно обучение управляват и изпълнението на договорености, анализират рискове, заменят труда на много служители, като автоматизират до 80% от рутинно повтарящите се техни дейности, което рязко повишава производителността и рентабилността. Конкретните рискове от използването на технологии за изкуствен интелект и машинно обучение са свързани с евентуален риск от хакерски атаки, несигурност на данните при грешно, неправилно или зловредно зададени цели на изкуствения интелект, загуба на работни места, непрозрачност на процесите, риск от непрозрачност на решенията, риск от намалена лоялност на клиентите и банкова миграция, репутационен риск и др. Актуалните действия на Базелския комитет, свързани с AI и машинното обучение, не са насочени към конкретна регулация на рисковете за банките и промяна на регулаторната рамка, а по-скоро към насърчаване на банките за тяхното използване, вкл. от самия Базелски комитет, за целите на регулациите. Новост в практиката на Базел III е, че за първи път автоматизирани работни процеси започват да се използват за непрекъснат мониторинг на възникващи проблеми и за гарантиране, че правилните хора получават правилната информация своевременно, която да им позволява да реагират бързо и ефективно. За изчисление на операционния риск, Базел III препоръчва на

банките да изследват най-новите постижения в областта на изкуствения интелект и машинното обучение за оптимизиране и автоматизиране на рутинни дейности.

Дълбоко/дълбочинно обучение (Deep Learning DL)

Усъвършенстван подход в ML, който използва многослойни невронни мрежи. Тази архитектура позволява на моделите да обработват огромни, неструктурирани данни като изображения, видео и звук. В медицината, Google DeepMind, в сътрудничество с Moorfields Eye Hospital, разработва алгоритъм за ранна диагностика на очни заболявания, който надминава по точност водещи офталмолози с 94% прецизност при препоръка за над 50 заболявания. В индустрията, Siemens внедрява DL модели за автоматизирана визуална инспекция на производствени линии, продукта им Inspekto позволява на компаниите да обучават AI системата си само с 20 добри проби от даден продукт, като системата след това може да инспектира над 100 продукта в минута, с изключителна точност.

Фигура 2 . Класификация на основните раздели на AI



Обработката на естествен език (NLP)

Метод за „превод“ между софтуер и човешки език, който автоматизира процеса на превод. Най-актуалните достижения в NLP включват софтуер за разпознаване на глас, превод на естествен език, търсене на информация и др. Крайната цел на NLP е да изгради софтуер, който да анализира, разбира и генерира естествени езици, като дава възможност за комуникация с компютър. NLP технологията се заражда през 50-те

години на XX век, като пресечна точка на изкуствения интелект и лингвистиката. Днес NLP заема няколко много разнообразни области, изисквайки днешните изследователи и разработчици на NLP да разширят значително своята база от знания и експертиза. Едно от основните приложения на NLP в индустрията са чатботове, които по своята същност са интелигентни системи, които разбират заявки на естествен език на потребителя и участват в диалог с него. Те са вид виртуални асистенти, при които клиентите са с усещането, че разговарят или комуникират с реален човек. Те говорят същия език, на който говори клиентът, и могат да отговорят на въпросите му (Kulkarni, Bhavsar, Pingale, Kumbhar, 2017). Друго приложение на NLP са финансовите анализи, особено като се има предвид голямото количество потенциална информация, неизменно съдържаща се в досиета и др. документи. За анализаторите осмислянето на цялата информация възможно най-бързо може да бъде от решаващо значение при генериране на навременни препоръки и постигане на конкурентно предимство. Звената по Нормативно съответствие също използват NLP приложения, за да подобрят различни процеси, например разпознаването на име на обект може да се използва за контекстуализиране на неструктурирано съдържание чрез откриване и формулиране на стратегии, представляващи интерес, като хора, компании и други обекти. Това позволява бързо да се извлича важна информация и да се гарантира, че са спазени регулаторните изисквания. Това става особено важно, като се има предвид бързо развиващия се регулаторен пейзаж, с който се сблъскват финансовите институции през последното десетилетие (Jones, 2021).

Роботизацията (Robotics)

Изграждане и използване на машини (роботи), за да изпълняват задачи, които се осъществяват традиционно от хора. AI и NLP са неизменна част от съвременната роботика. Роботите могат да бъдат оборудвани с еквивалент на човешките сетива като зрение, допир и способност да усещат температура. Настоящите изследвания в роботиката са насочени към създаване на роботи със степен на самостоятелност, която ще позволи мобилност и вземане на решения в неструктурирана околна среда. Роботите заместват хората на работното място, за да изпълняват повтарящи се и ежедневни задачи. Роботизацията позволява да се конфигурира компютърен софтуер или робот за улавяне и интерпретиране на съществуващи приложения за обработка на трансакции и данни, задействане на отговори и комуникация с други цифрови системи. Роботизираната автоматизация на процесите (RPA) се приема за изключително важна за финансовата индустрия, за осигуряване на възможно най-добро потребителско изживяване на клиентите.

Приложението на роботизацията може да се обобщи в следните основни направления във финансовия свят:

- обработка на големи трансакции – автоматизира обработката на големи обеми трансакции с висока точност; банките могат да изпълняват ефективно сложни финансови задачи и да премахват потенциала за човешки грешки;
- регулации и Нормативно съответствие (Compliance) – намалява или напълно

елиминира отнемашите време процеси на събиране, компилиране и обобщаване на големи количества данни, в т.ч. извличане и агрегиране от различни източници на данни в големи бази, улесняване на документирането;

- намаляване на разходите и спестяване на време – според проучване на Accenture, внедряването на роботизирана автоматизация на процесите във финансовата индустрия предлага потенциал за намаляване на разходите до 80%;
- откриване на измами – финансовата индустрия е по-податлива на измами и кибер атаки. Измамни трансакции с дебитни и кредитни карти се случват по целия свят. Борбата с измамите чрез традиционни методи е станала негъвкава и липсва анализ в реално време. Роботизацията, съвместно с изкуствения интелект, може да осигури цялостни решения чрез извършване на поредица от проверки на баланси и трансакции за определяне на нетипични трансакции или идентифициране на модели на измамни дейности. Освен това, технологията помага при анализ на данните в реално време и вземане на решение;
- Robo Advisers и Algo trading – с напредъка в технологиите търговията на фондовите пазари се извършва чрез използване на алгоритъм и количествени модели за търговия, за да се правят по-ефективни и печеливши сделки.

Добавена реалност, виртуална реалност, смесена реалност

Добавена реалност (AR – augmented reality) е набор от аудио-визуални технологии, които позволяват на потребителите да видят един реален свят около тях. AR продуктите използват различни техники за добавяне и наслагване на компютърна графика и звук върху зрението и слуха на потребителя. Виртуална реалност (VR – virtual reality) е компютърно генериран стереоскопичен 3D виртуален свят, вкл. с 3D звук. AR/VR и смесена реалност (MR – mix reality) имат потенциал да революционизират дигиталното банкиране. Добавената реалност ще изведе възможностите на цифровите технологии и потребителското изживяване на съвсем ново ниво, съчетавайки реалния и дигиталния свят. VR/AR банкирането може да прекрачи границите на физическия свят чрез 3D сензорно банково изживяване посредством:

- липса на ограниченията на екрана – в сравнение с настолния или мобилен екран, VR/AR не е ограничен от размера на екрана;
- перфектно съвпадение между реалностите – и виртуалната, и добавената реалност трябва да се отнасят до преживяването в реалния свят. Правилно структурираният VR дизайн трябва да намали несъответствието, причинено от разликата между проследеното движение на очите във виртуалната реалност и усещането, което възприемаме от физическата реалност чрез другите си сетива;
- едно от най-значимите предизвикателства на банковия дизайн на VR/AR е преходът на потребителския интерфейс от двуизмерни (2D) към триизмерни (3D) взаимодействия. Това изисква дълбока адаптация на принципите и методите за проектиране на UX/UI¹³ за дълбочината на обемното пространство.

Интернет на нещата (IoT)

Динамична глобална мрежова инфраструктура с възможности за самоконфигуриране, базирана на стандартни и оперативно съвместими комуникационни протоколи между дигиталните и физическите елементи. IoT обезпечават екосистемата с интелигентни приложения и услуги, които събират и обменят данни и могат да бъдат наблюдавани или дистанционно контролирани чрез интернет, чрез софтуер на всякакъв вид компютри, смартфони, или чрез интерфейси като стенни контроли. Банките използват IoT, за да събират допълнителна информация за своите клиенти. След събиране на данни в реално време за свързаните устройства и операциите на клиентите, банките могат да предоставят специално изработено съдържание и персонализирано отношение, изразяващо се в персонализирани продукти и услуги. С IoT финансовите институции се включват по-активно в ежедневния режим на клиентите си (Kotler, Kartajaya, Setiawan, 2021).

Блокчейн (Blockchain)

Структура от данни, състояща се от свързани блокове данни, например потвърдени финансови трансакции, като всеки блок е посочващ/отнасящ се до предишния, образувайки верига в линеен и хронологичен ред. Блокчейн е свързан списък от блокове, които могат да съдържат разнообразна информация. Блокчейн технологията дава възможност само за добавяне на нови данни, организирани в блокове. Промяната или изтриването на вече включени във веригата данни практически не е възможна или е много трудно осъществима. Тази децентрализирана технология дава възможност на участниците в мрежа от тип peer-to-peer да извършват трансакции без да е необходим надежден централен орган за контрол и в същото време разчитат на криптографията, за да гарантират целостта на трансакциите. Блоковите вериги (блокчейн) са една от най-актуалните теми през последното десетилетие, както в технологичния, така и във финансовия свят. Harvard Business Review определя блокчейн като „отворена технология, която може да записва трансакции, направени между двама потребители, по един ефективен, сигурен начин за постоянно“. Блокчейн технологиите са изключително сигурни, с висока степен на Byzantine fault толеранс. Byzantine fault толеранс означава дадена система да е изградена по такъв начин, че да функционира дори ако нейни отделни компоненти откажат да изпълняват функциите си. Концепцията на първия блокчейн придобива окончателен облик през 2008 г., като зад създаването ѝ стои един или няколко души, обединени зад псевдонима Satoshi Nakamoto. През 2009 г. блокчейн технологията е заложена като ключов компонент в криптоактива биткойн, където обслужва всички трансакции. Опити да се създадат криптоактиви са правени и преди появата на биткойн. Преди появата на блокчейн технологията обаче всички тези опити са се провалили по една причина: направените трансакции са имали възможност да бъдат копирани и да се извлече от тях стойност втори път. Въпреки своя „открит“ характер, блокчейн осигурява конфиденциалност на информацията, понеже кодирането чрез сложни техники гарантира, че клиентите държат контрол върху своята поверителна информация и могат да решат кои трети страни да получат достъп до нея. Възможностите за приложение на тази нова технология в банките провокират повишаване на регулаторното внимание. Използването

на блокчейн до голяма степен се съсредоточава върху съществуващите регулирани процеси. Въпреки че много закони и разпоредби остават приложими в блокчейн среда, регулаторите все по-активно оценяват потенциалното въздействие на новата технология върху финансовата индустрия, целяйки създаване на рамка, която да гарантира сигурното имплементиране на блокчейн приложенията. На блокчейн технологията се възлагат големи надежди и по повод на редуциране на риска от използване на банковата система за пране на пари и финансиране на тероризъм, тъй като тя има значителен потенциал за оптимизиране на извършваните процедури KYC¹⁵ („Опознай своя клиент“). Допълнително, уникалната структура на блоковата верига осигурява защитни механизми, неприсъщи за традиционните IT системи. Допълнителните рискове за глобалната финансова система, породени от новата технология, са поява на концентрации при собствеността и експлоатирането на ключови инфраструктури, намаляване на общественото доверие във финансовата система, в резултат от неясноти при определяне на правната отговорност и защитата на потребителите, затруднено възстановяване на децентрализираните структури. Съблюдаването на Директивата срещу изпиране на пари също е предизвикателство пред блокчейн, тъй като въведените в нея мерки включват задължения за комплексна проверка на клиенти, изискваща идентифицирането им „въз основа на документи, данни или информация, получени от надеждни и независими източници“. Това обаче би било много трудно, предвид децентрализирания характер на блокчейн и липсата на контролираща институция, която да идентифицира участниците и да следи за съмнителни трансакции. В рамките на ЕС, най-много отворени въпроси има относно ключовите характеристики на блоковите вериги и правилата, заложи в Общия регламент относно защитата на данните (GDPR). GDPR е замислен преди началото на по-широкото приложение на технологиите, базирани на споделен регистър (DLT), и следователно е създаден около предположение, че данните се събират, съхраняват и обработват в централна база данни, а не в децентрализиран регистър. Отговорността за защита на личните данни се носи от конкретни физически или юридически лица, към които клиентите могат да предявят правата си. Блокчейн, от друга страна, се стреми да постигне децентрализация чрез замяна на „единната точка на данни“ с много на брой такива, което прави отговорността трудно установима. GDPR се основава и на разбирането, че данните могат да бъдат променяни или изтривани, когато това е необходимо, с цел спазване на законовите изисквания. В рамките на блоковите вериги обаче, за да се гарантира целостта на данните и да се увеличи доверието в мрежата, подобни модификации на записаната информация целенасочено са направени невъзможни. Блокчейн технологиите пораждат и множество рискове, които са обект на наблюдение и частична регулация от различни институции, в т.ч. Базелския комитет. Такива рискове са:

- спекулации;
- инфлационно влияние;
- пране на пари и нелегален трафик;
- виртуална кражба на портфейли и информация;
- избягване на национални закони и регулации;
- разнопосочност на регулациите на активите в крипто.

Предимствата на блокчейн технологиите привличат интереса не само на населението, фирмите, банките и правителствата, но регулаторните институции също започват да се възползват от тях. Централните банки имат различия по отношение на формата, начина и обхвата на приложение на бъдещите им дигитални валути. Това дава основание да се очаква, че бъдещото им приложение ще е белязано с множество различия по отношение на регулаторните режими и подхода към търговските банки. Задачата на глобалната регулация е да координира тези различия, да работи за кохезия на националните политики и да постигне обща международна регулаторна рамка, която да предпазва банковата система от сътресения и системни кризи. Главната насока на международните регулации е постигане на баланс между поверителността на криптираните плащания, осигуряване на спазване на разпоредбите срещу изпиране на пари и финансиране на тероризъм, както и изучаване на възможностите за разширяване на регулациите за блокчейн технологиите.

Облачни технологии

Изчисленията в облак (Cloud Computing) са една от основните иновации, използвани от индустрията в стратегиите за дигитализация на бизнеса. “Cloud Computing” е термин, който препраща към идеята за обединяване на изчислителната мощност на много компютърни устройства в система. Такава система, според технологията на създаването ѝ, може да има различни свойства. Тя може да е система от физически устройства (Cluster), или пък клъстер от виртуални машини (Virtualization). Втората характеристика на Cloud Computing, след обединяването на много компютри в мрежа, е, че потребителите използват компютърните ресурси в мрежова среда. Това може да е вътрешно-офисна мрежа, по-голяма затворена или свободна мрежа от потребители. Интернет е своеобразен Cloud (не „облак”, а „рояк”) от компютърни устройства, които могат да се свързват едно към друго, чрез WWW стандартите.

Фигура 3. Облачна инфраструктура, нива на споделяне на услугите

SaaS		
Hosted applications/apps	PaaS	
	Development tools, database management, Operating systems business analytics	IaaS
		Servers and storage Networking Data center physical firewalls/securityplant/building

Източник: Microsoft, <https://azure.microsoft.com/en-us/overview/what-is-saas/>.

Днес се използват различни варианти за споделяне на изчислителни ресурси, най-популярни сред които са софтуерът като услуга (SaaS), платформата като услуга (PaaS), инфраструктурата като услуга (IaaS) и бизнес процесът като услуга (BPaaS). SaaS е начин за доставяне на приложения през интернет – като услуга, което замества

инсталирането и поддържането на софтуер. IaaS е „облачно“ решение, което предлага изчислителни и мрежови ресурси и ресурси за съхранение на данни при поискване. Мигрирането на инфраструктурата към IaaS помага да се намалят разходите за поддръжка на локални центрове за данни/сървъри и за хардуер. PaaS е цялостна среда за разработка и внедряване в „облака“ на ресурси, която позволява те да се доставят от базови приложения, базирани на облак, до сложни корпоративни приложения. Най-голям потенциал за банките имат публичните облаци. Активното използване на предоставени от трети страни облачни услуги не освобождава банките от техните отговорности по отношение на информационната сигурност, поверителност и цялост на данните. Именно около този принцип е построен регулаторният инструментариум в областта. Облачните договорености се подчиняват на вече коментираните регулации, засягащи киберсигурността и защитата на данните. Пазарът на облачни услуги е доминиран от няколко големи компании – Amazon, Microsoft, Google, IBM и Alibaba (Synergy Research Group, 2019). Сред най-значимите регулаторни инициативи са насоките на Европейския банков орган (ЕБО) за възлагане на дейности на външни изпълнители, валидни от 30.09.2019 г. В регламента се предоставят критерии за определяне на критичните или важните функции за финансовите институции. В случай, че такива са изнесени към външен доставчик, в зависимост от избора, договореностите и контрола върху него, важат по-строги изисквания. За пръв път се формулират и определени условия за използване на подизпълнители (ЕВА, 2019). Ключовите изисквания към финансовите институции в регламента са: разработване на рамка за надеждно идентифициране и управление на рисковете от трети страни; предварителна оценка на критичното значение и присъщия рисков профил на дейностите – обект на аутсорсинг; уведомяване на компетентните органи при възлагане на съществени дейности на външни доставчици; периодичен мониторинг на изпълнението на заложените дейности и механизмите за защита на поверителността на предаваната от финансовата институция информация; планиране на мерки за поддържане на непрекъснато провеждане на изнесените дейности, в случай, че предоставянето на външните услуги прекъсне или се влоши до неприемлива степен.

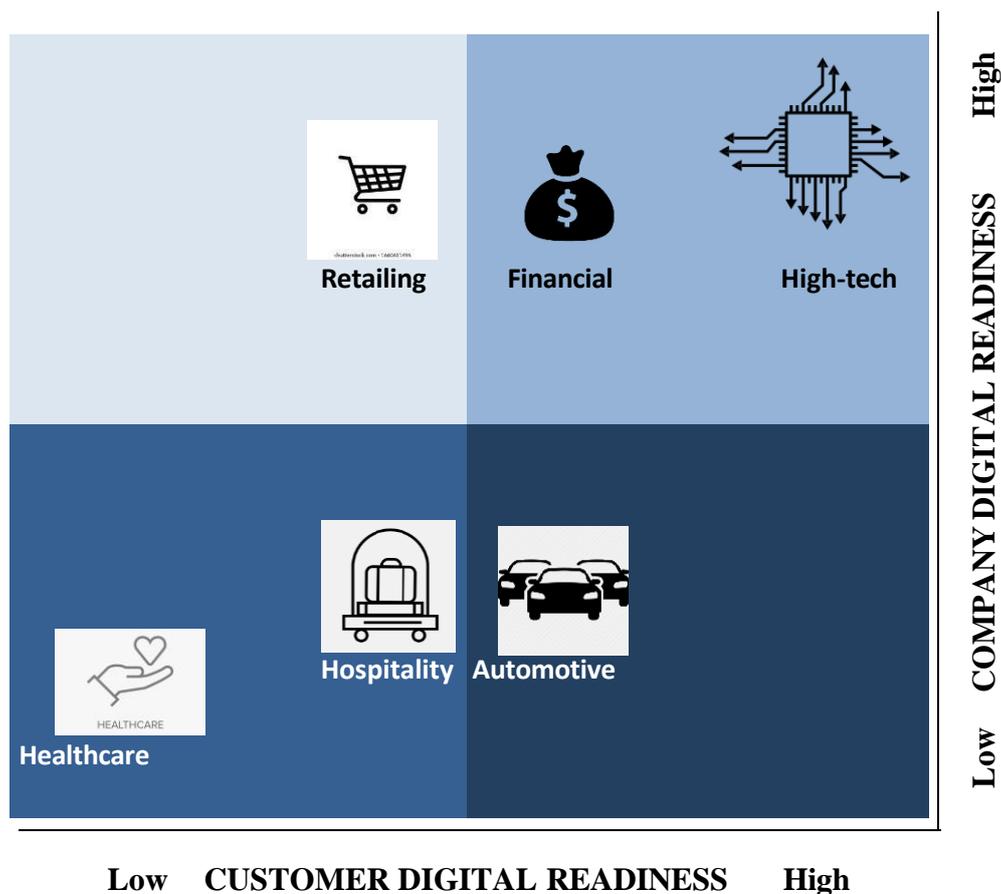
Биометрични технологии

Биометричните технологии са автоматизирани методи за разпознаване на клиентите чрез техните биологични характеристики и черти, като пръстов отпечатък, ирис на окото, глас, и др. Биометричните характеристики са уникални за всеки индивид и са трудни за подправяне. Много бизнеси започнаха да използват биометрични данни за удостоверяване на самоличността на служители и клиенти. Използването на биометрични технологии може да обхване широк диапазон от услуги. Биометричните технологии постепенно изместват традиционните пароли чрез прилагане на биометрична идентификация на потребителите. Биометричните технологии използват вградените в компютри, лаптопи и смартфони уеб камери, микрофони и скенери за пръстови отпечатъци. Разрастващата се тенденция към масово приложение на биометрията постави остро въпроса за идентификация и оценка на рисковете от прилагането на подобни технологии. Свързаността между системите и устройствата позволява на

кибератаките да нанесат вреди на много хора, фирми и организации. Общото мнение е, че повечето системни пробиви се изпълняват чрез кражба на идентификационни данни. Важен фактор за сигурността е прецизното събиране на биометрични данни, което да гарантира сигурност при удостоверяване на самоличността. Ако биометричният образ е неразпознаваем, това създава проблеми с достъпа на потребителите, а ако е ненадежно записан, има риск хакерите да извършат измама по време на събирането на данни, или да заменят събраните данни със свои собствени, което нарушава сигурността на биометричните данни за достъп до акаунти с по-малък риск от откриване. Риск при съхранението на данните може да се прояви, ако нивото на защита е ниско. Ежегодно се отчитат множество нарушения по личните данни в целия свят, което поставя въпроса за квалификацията на специалистите по защита на данни на предно място. На този етап, съгласието на потребителя е основната регулационна норма за събиране на биометрични данни. Злоупотреба с биометрични данни може да се изрази при неразрешено или незаконно споделяне на информация между трети страни. Това може да се използва както в маркетингови стратегии, така и за получаване на други лични данни, които да се използват във вреда на лицата, дали съгласие за ползване на техните лични данни. Много специалисти разработват различни приложения за защита на биометричните данни. Друго решение е въвеждането на отменяеми биометрични показатели, при което идентификаторите се променят с помощта на математически изчисления, което позволява биометричните шаблони да могат да бъдат изтрети и заменени. Създават се и алгоритми, които да могат да трансформират биометричните данни на потребителите по различни начини, които да не допуснат манипулация с данните и др. Методите за биометрично удостоверяване често разчитат на частична информация за удостоверяване на самоличност, например ограничен брой точки от данни, но възникват измамнически пробиви на пръстовата идентификация. С цел минимизиране на този риск, управлението на идентичността стана най-бързо развиващият се сегмент на информационната сигурност, основен приоритет за много банки и институции, които се опитват да защитят своите мрежи, системи и данни от киберзаплахи и кибератаки. Фактът, че новите приложения за биометрия стават глобални, изисква глобален подход за тяхната регулация, както и стриктни, ясни и ефективни национални подходи и политики за оценка и контрол на риска, вкл. по приложението им в банковата система. Това предизвикателство дава основание на регулаторите, като първа стъпка, да включат биометричните финансови технологии сред най-важните области – обект на регулаторно наблюдение и анализ. Ключов аспект на регулаторния дебат е решаването на общия въпрос как регулирането на финансовите услуги може да улесни адаптацията на банковата система към новата технологична среда, новите играчи и новите дейности. Този процес е все още в начален стадий и се изразява в наблюдение, констатации, дискусии, предупреждения, идентификация на рисковете и сравнителен анализ на националните регулаторни подходи. В страни, където изискванията относно събирането, използването и съхранението на биометрични данни са предварително разписани, регулациите са фокусирани най-вече върху решаването на поверителността на данните, киберсигурността и рисковете от пране на пари. Подобни регулации съществуват за страните-членки на ЕС чрез следните нормативни актове:

- Директива против пране на пари AML5 (5th Anti-Money Laundering Directive);
- Директива PSD2 (Revised Payment Service Directive), която принуждава банките да предоставят на доставчици на плащания на трети страни (FAANGS – Facebook, Apple, Amazon, Netflix, Google, Spotify) достъп до информация за сметките на своите клиенти чрез API, което позволява на технологичните фирми да извършват P2P плащания и преводи, да анализират навиците на харчене и др., докато парите на клиентите седят безопасно в тяхната банкова сметка;
- Директива GDPR (General Data Protection Regulation).

Фигура 4. Готовност за дигитализация на индустриите



Източник: Kotler, Kartajaya, Setiawan, 2021.

Иновационните процеси, в по-широк аспект и в частност в банкирането, са съществен фактор за развитието на различните индустрии, в т.ч. финансовата. Все по-често във финансовия сектор и банкирането иновации и дигитализация се използват съвместно, тъй като голяма част от иновациите са продиктувани от нови технологични решения и възможности. Не случайно Kotler поставя финансовата индустрия на водещо място, като готовност на потребителите да приемат дигитални решения и готовност на финансовите компании да предложат на клиентите си дигитални решения (Kotler, Kartajaya, Setiawan, 2021).

Радикалните иновации не са избор, а задължителен път към оцеляване и конкурентоспособност. Алтернативата, липса на иновация води до сигурна стагнация. Въпреки рисковете, правилното управление на AI и дигитализацията може да осигури устойчив растеж и лидерство на пазара.

Теми за дискусия Глава 1

- *Какво означава „съзидателно разрушение“ в контекста на дигитализацията?*
- *Кои индустрии са най-уязвими на технологичното цунами и защо?*
- *Как AI може да съчетае ефективност и етичност?*
- *Какви мерки могат да намалят зависимостта от BigTech доставчици?*
- *Достатъчни ли са настоящите регулации (AI Act, GDPR) за управление на AI риска?*

2. От Алан Тюринг до ChatGPT, Историята, настоящето и бъдещето на AI

„Проблемът не е, че машините мислят, а че хората мислят твърде малко.“

Хърбърт Саймън, носител на Нобелова награда по икономика, цитат от *The Sciences of the Artificial* (1969).

История на изкуствения интелект, от теоретични основи до генеративна трансформация

Историята на изкуствения интелект (AI) е поредица от цикли на ентузиазъм и разочарование, движена от научни проби и технологично развитие. Началото ѝ е вплетено дълбоко в теоретичната математика и логиката, много преди създаването на първите цифрови компютри.

Развитието на Изкуствения интелект (AI) от теоретичните основи, положени от Алън Тюринг, до появата на Генеративната икономика, базирана на модели като ChatGPT, доказва, че **радикалните иновации са безалтернативен императив** за съвременния бизнес и обществено развитие. Това е така, тъй като тяхното отсъствие води до сигурна икономическа неконкурентоспособност и институционална несъстоятелност, въпреки съпътстващите рискове, свързани с алгоритмичното изкривяване и материалните ограничения.

Необходимостта от иновации в сферата на AI е вкоренена в самото теоретично начало на дисциплината. През 1950 г. **Алън Тюринг** превръща философския въпрос за мисленето в **инженерен проблем**, който трябва да бъде решен. Той поставя оперативен критерий, заявявайки: „Предлагам да заменим въпроса ‘Могат ли машините да мислят?’ с играта на имитация“ (Turing, 1950). Този акт на **радикална иновация** дава измерима цел за технологичното развитие. Четири години по-късно, на конференцията в Дартмут (1956), **Джон Маккарти** официално въвежда термина „Artificial Intelligence“, поставяйки основите на **организационната иновация** с убеждението, че всеки аспект на интелигентността може да бъде симулиран от машина (Dartmouth Conference, 1956).

Историята обаче не е линеен процес. Периодите на стагнация, като „AI зимата“, които настъпват след силната критика на **Джеймс Лайтхил (1973)** относно прекалените очаквания и методологичните ограничения, служат като доказателство *a contrario*. **Липсата на иновация** (поради нереалистични амбиции за AGI и недостатъчна изчислителна мощ) води до **колапс на финансирането** и спиране на развитието.

Възраждането идва с прагматичния подход, фокусиран върху **Слаб AI (ANI)**. Успехът на **IBM Deep Blue (1997)**, който побеждава световния шампион по шах Гари Каспаров, е символ на тази промяна, демонстрирайки, че „бруталната сила“ (изчисляване на милиони позиции в секунда) може да постигне **превъзходство** в тясно дефинирани

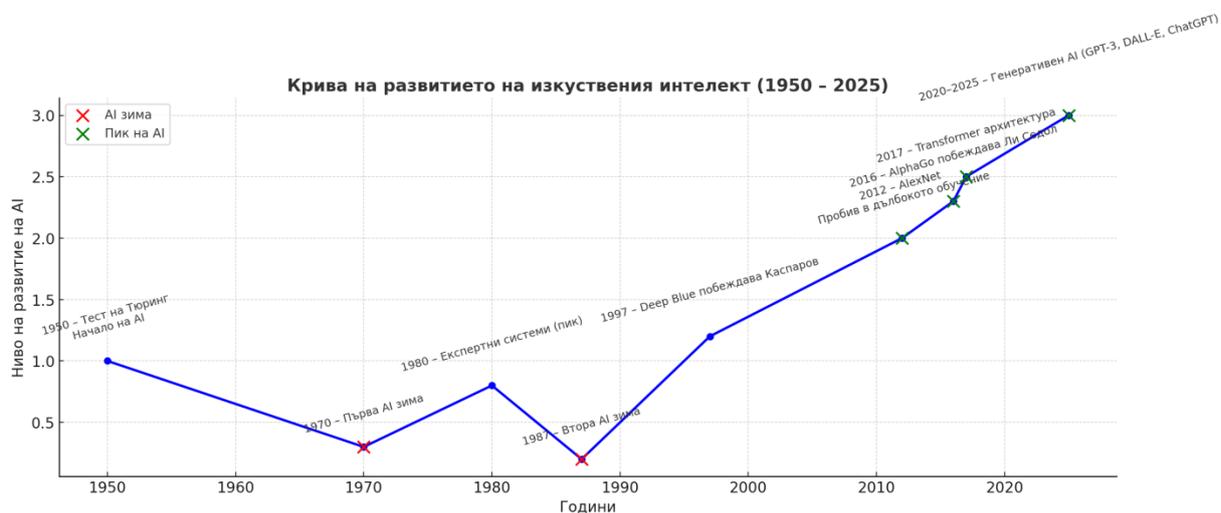
задачи (ScienceDirect, 1997). Революцията на **Дълбокото Обучение (Deep Learning)** и **Генеративния AI (GenAI)** след 2012г. превърна иновацията в **икономически императив**, а не в лукс. **Ажай Агравал, Джошуа Ганс и Ави Голдфарб (2018)** предоставят рамката, която доказва безалтернативността на AI от гледна точка на стойността. Те разграничават AI като **инструмент за предсказване** от човешката **преценка**. Ключовият им аргумент е, че „По-доброто предсказване е заместител на преценката при скрити възможности, но **допълва** преценката при скрити разходи“ (Agrawal et al., 2018). Компанията, която не внедрява AI за автоматизирано предсказване (напр. риск от фалит, потребителско търсене), **губи двойно**, не успява да намали оперативните си разходи и разчита на скъпа, бавна човешка преценка за рутинни задачи. Последните изследвания емпирично подкрепят връзката между AI и продуктивността. Изследване на NBER/QJE (2025) показва, че „Достъпът до AI увеличава продуктивността на агентите с **15%**“ (NBER, 2025). Най-големите печалби се наблюдават при **по-малко опитните** служители, което подкрепя тезата, че AI демократизира експертното знание и повишава общото ниво на компетентност. Инструменти като GitHub Copilot съкращават времето за изпълнение на контролирани задания при разработчиците с **~56%** (arXiv, 2023). Тези данни доказват, че **липсата на AI-базирана иновация** директно води до „загиване на компанията“ поради невъзможност да се капитализира върху новите граници на продуктивността (Sull, 1999).

Има и друга гледна точка представена от водещи скептици като **Гари Маркъс (2020)** и **Емили М. Бендър с колектив (2021)**, които твърдят, че мащабните езикови модели (LLM) са „**стохастични папагали**“ и „**фундаментално ненадеждни за истина**“ без нови архитектури (Marcus, 2022; Bender et al., 2021). **Майкъл И. Джордан (2019)** допълва, че „AI революцията още не се е случила“; нужни са системи за **каузалност** и **социално-техническа интеграция** (Jordan, 2019). Тази критика не оспорва безалтернативността, а **повишава стандарта**. Тя показва, че **просперитетът е условен**. Той ще се постигне само чрез **по-дълбоки иновации в надеждността** (напр. агентни модели като ReAct, RLHF) и **управлението на риска**. Според **Карл Таута и Пол Диймуд (2020)**, дигитализацията е **институционално изискване**, тъй като „Банките са принудени да приемат дигитални инструменти за да отговорят на **регулаторните изисквания** (напр. за KYC и бърз мониторинг)“ (Tauter & DiMaggio, 2020). Иновацията е безалтернативна, защото нейното отсъствие води до **правна несъстоятелност** в рамките на регулации като **ЕС AI Act (2024)**, който изисква отговорност (EU Digital Strategy, 2024). Необходимостта от **енергийна иновация** също става ключов стратегически приоритет. Прогнозите на **IEA (2025)** сочат, че потреблението на електроенергия от центрове за данни ще се удвои до **~945 TWh до 2030 г. задвижвано от AI** (IEA, 2025). Това превръща **иновациите в ефективност на хардуера и енергетиката** в **предварително условие** за бъдещо AI развитие.

Линията на развитие от Теста на Тюринг през Deep Blue до Генеративната икономика (GPT-3, Chinchilla, RLHF) демонстрира непрекъсната, кумулативна иновация. Емпиричните данни доказват, че **AI вече не е научна хипотеза, а доказан механизъм за увеличаване на продуктивността**. AI осигурява **експоненциално конкурентно предимство** (Бринолфсон), намалява разходите, като **автоматизират прогнозите** (Агравал, Ганс, Голдфарб). В този контекст, **алтернативата** на проактивната AI

иновация е сигурна **стагнация и изчезване** на компанията. Бъдещето изисква иновации не само в алгоритмите, но и в **управлението, етиката и енергийната ефективност**, за да се гарантира, че технологичният напредък води до **про-човешки просперитет** (Acemoglu, 2023).

Фигура 5. Еволюция и развитието на Изкуствения Интелект (ИИ/AI)



*Източник: Автора

Теоретични основи и ранни визии (1930 – 1950)

Фундаментът на AI е положен от пионери като **Курт Гьодел** и **Алън Тюринг**. През 1931 г. Гьодел публикува своята **Теорема за непълнотата**, която поставя ограничения върху това, което може да бъде доказано в рамките на една формална система. Това провокира въпроси за границите на логиката и човешкото познание, които по-късно стават централни за AI.

Курт Гьодел



Алън Тюринг



През 1936 г. **Алън Тюринг** формализира концепцията за алгоритмична изчислимост, като представя идеята за **универсална машина на Тюринг**. Тази хипотетична машина

може да изпълнява всяка задача, която може да бъде описана с алгоритъм, поставяйки теоретичната основа за съвременния компютър. През 1950г. Тюринг публикува статията си "Computing Machinery and Intelligence", където задава въпроса "Могат ли машините да мислят?" и предлага **Теста на Тюринг** метод, който цели да определи дали дадена машина може да демонстрира интелигентно поведение, което е неразлично от човешкото. Този тест се превръща в крайъгълен камък във философията на AI.

Раждането на AI като научна дисциплина и първите успехи (1950 – 1960)

През лятото на 1956 г., на конференция в колежа Дартмут, **Джон Маккарти** официално въвежда термина „**Artificial Intelligence**“. Това събитие се счита за рожден ден на AI като независима научна област. Конференцията събира водещи умове като **Алън Нюел**, **Хърбърт Саймън** и **Марвин Мински**, които изразяват убеждението си, че "всеки аспект на ученето или друга черта на интелигентността може на принципно ниво да бъде описан толкова прецизно, че да се създаде машина, която да го симулира."

Джон Маккарти



През този период се появяват и първите успешни AI програми:

- През 1955 г. Нюел и Саймън създават **Logic Theorist** — програма, която доказва 38 от 52 теореми в "Principia Mathematica" на Бъртранд Ръсел и Алфред Уайтхед.
- През 1959 г. Артър Самюъл разработва **програма за шах**, която използва механизми за **самообучение (rote learning)**, като запомня успешни ходове от предишни игри.
- През 1958 г. Джон Маккарти създава програмния език **Lisp**, който се превръща в стандарт за AI изследванията през следващите десетилетия.

Разочарование и „AI зима“ (1970 – 1980)

Въпреки ранните успехи, през 70-те и 80-те години настъпва период на спад, известен като „AI зима“. Основната причина е разминаването между високите очаквания и ограничените резултати. Някои изследователи, водени от ентузиазма си, правят прекалено амбициозни обещания за създаване на общ изкуствен интелект (AGI), които не успяват да се сбъднат. Недостатъчната изчислителна мощ и липсата на големи набори от данни също допринасят за стагнацията. Финансирането от правителствата и бизнеса рязко намалява, което принуждава много изследователи да изоставят сферата.

Все пак, през този период се развиват **експертните системи**, които намират комерсиално приложение в специфични области, например в медицината и геологията, като доказват, че AI може да бъде полезен, макар и в тясно специализирани задачи.

Възраждане на интереса (1990 – 2010)

През 90-те години AI се завръща, този път фокусиран върху **статистическите методи и машинното обучение**, а не върху символичната логика. Нарастващата изчислителна мощ и достъпът до големи данни правят тези подходи много по-ефективни.

Един от най-значимите моменти от този период е през май 1997 г., когато суперкомпютърът на IBM **Deep Blue** побеждава световния шампион по шах **Гари Каспаров** в мач от шест партии. За разлика от предишните програми, Deep Blue разчита на **"груба сила"** – изчислява 200 милиона шахматни позиции в секунда, вместо да имитира човешкото мислене. Тази победа е важен символ на напредъка на AI.

През следващите години се развиват и нови области:

- **Обработка на естествен език (NLP)**, появяват се първите версии на Google Translate.
- **Компютърно зрение**, започва разработката на алгоритми за разпознаване на лица и обекти.
- **Автономни автомобили**, започват да се провеждат първите успешни тестове.

Ерата на дълбокото обучение (Deep Learning) и генеративния AI (Generative AI) (от 2010 г. до наши дни)

Революцията в AI, която виждаме днес, е движена от два ключови фактора, **масивните данни (Big Data)** и развитието на **графичните процесори (GPUs)**. GPUs, първоначално създадени за видеоигри, се оказват изключително ефективни за паралелните изчисления, необходими на невронните мрежи. Това води до възхода на **дълбокото обучение (Deep Learning)**, подполе на машинното обучение, което използва многослойни невронни мрежи.

След 2020 г. настъпва нова ера, тази на **генеративния AI**. Модели като **GPT** на OpenAI и **DALL·E** могат да генерират съдържание, включително текст, изображения и музика, което променя фундаментално начина, по който бизнесът и обществото взаимодействат с технологиите. Тези системи се обучават върху огромни масиви от данни от интернет и

демонстрират впечатляващи способности за решаване на проблеми, които доскоро се смятаха за запазена територия на човешката интелигентност.

Основни видове и класификации на изкуствения интелект

Изкуственият интелект (AI) е широка и динамична област, чиито дефиниции и класификации се развиват непрекъснато, основно можем да разграничим AI на базата на неговата **функционалност и способности**. Тези разграничения ни помагат да разберем какво представлява AI днес и какви са бъдещите му амбиции.

Класификация според функционалността (AI Levels)

Тази категоризация, често използвана в индустрията и научните среди, разделя AI на три основни нива, базирани на неговата интелигентност и когнитивни способности.

- **Слаб (Тесен) Изкуствен интелект (Artificial Narrow Intelligence - ANI)**

ANI е най-разпространеният вид AI днес. Той е проектиран обучен да изпълнява **една-единствена конкретна задача**. Не притежава човешко съзнание, интелект или способност да решава проблеми извън тесния си фокус. **Примерите са чатботове и виртуални асистенти вт.ч. Siri, Google Assistant, Alexa**, които обработват гласови команди и отговарят на въпроси; **Системи за препоръки вт.ч. попадат** алгоритмите на Netflix и Amazon, които анализират поведението на потребителите, за да предложат съдържание или продукти; **Разпознаване на лица и реч**, системи за сигурност и приложенията за транскрипция, които обработват визуални и аудио данни.

Концепцията за ANI е резултат от прагматичния подход, възприет след "AI зимата" от 70-те години. Учените като **Марвин Мински** и **Джон Маккарти** се фокусират върху създаването на системи, които решават практически проблеми, вместо да се стремят към пълно човешко съзнание.

- **Силен (Общ) Изкуствен интелект (Artificial General Intelligence - AGI)**

AGI е хипотетична система, която притежава **интелектуална способност, сходна на тази на човек**. Тя може да разбира, учи, да прилага знания в различни области, да разсъждава и да решава широк спектър от задачи, дори и такива, за които не е била специално обучена.

Все още не съществуват реални примери за AGI. Всички съвременни модели, включително и най-големите езикови модели като GPT, са само напредничави форми на ANI, тъй като им липсва истинско разбиране и съзнание.

Идеята за AGI произлиза от основополагащата работа на **Алън Тюринг** от 1950 г., в която той разсъждава върху възможността машините да мислят. AGI е централна тема в областта на **когнитивната наука**, която се опитва да моделира човешките когнитивни процеси. Въпреки че все още е хипотеза, постигането на AGI се разглежда като "Светият Граал" на изкуствения интелект.

- **Изкуствен суперинтелект (Artificial Superintelligence - ASI)**

ASI е хипотетична система, чийто интелект **превъзхожда човешкия във всички аспекти**, включително научна креативност, общо познание и социални умения. Тя би била способна да решава проблеми, които са извън възможностите на човешкото разбиране.

ASI все още съществува само в научните хипотези и теоретичните дискусии. Концепцията е популяризирана от философа **Ник Бострѝом** в книгата му "**Суперинтелект: Пътеки, опасности, стратегии**" (**Superintelligence: Paths, Dangers, Strategies**). Бострѝом разглежда потенциалните рискове и ползи от създаването на такъв интелект, като подчертава необходимостта от предпазни мерки още преди постигането на AGI.

Класификация според способностите (AI Types)

Тази по-техническа класификация се фокусира върху конкретните възможности на AI системите.

- **Реактивни машини (Reactive Machines)**

Това са най-елементарните AI системи. Те нямат памет и не могат да се учат от минали преживявания. Реагират на настоящия момент, без да вземат предвид предишни действия. Пример са **IBM Deep Blue** — суперкомпютърът, който победи Гари Каспаров през 1997 г., е класически пример. Той анализираше само позициите на дъската в момента, без да използва информация от предишни игри.

- **AI с ограничена памет (Limited Memory AI).**

Тези системи могат да използват данни от близкото минало, за да вземат решения. Те се обучават на исторически данни (обучение), но могат да съхраняват и временно информация от текущия момент (памет). **Пример са автономни автомобили, които наблюдават скоростта и разстоянието на другите автомобили в момента, но и помнят маршрутите и данните от предходните си пътувания.**

Съвременните модели на машинно обучение, включително и големите езикови модели, които "помнят" контекста на един разговор, за да генерират кохерентни отговори.

- **Теория на ума (Theory of Mind AI)**

Концепцията е бъдещ етап от развитието на AI, при който системите ще имат способността да разбират, разсъждават и предвиждат **емоциите, намеренията и вярванията** на хората. Според някои учени, този тип AI ще трябва да премине от чисто логическо разсъждение към **социална интелигентност**, което е огромно предизвикателство.

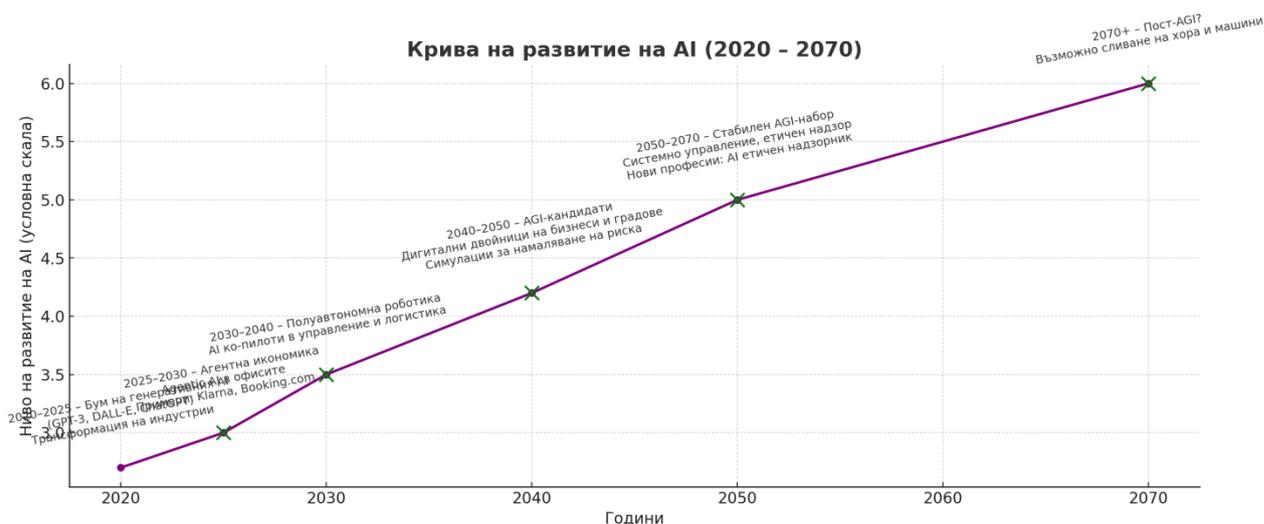
- **Самосъзнателен AI (Self-Aware AI)**

Най-високото и към момента напълно хипотетично ниво на AI. Тези машини биха имали съзнание, самосъзнание и емоции. Това би означавало, че те осъзнават собственото си съществуване, което към момента е в сферата на философията и футурологията, тъй като все още не разбираме напълно какво е съзнанието и как би могло да бъде възпроизведено.

Анализ и възможна еволюцията на AI до 2070 г.

Реалистичения анализ на развитието на изкуствения интелект минава през практическите предизвикателства и системните промени. Ключовите допускания за ограниченията в изчислителната мощ и данните е особено важно да съответстват на актуални научни и индустриални перспективи. Изложените прогнози за развитие на AI са изцяло авторски, базирани на професионален и академичен опит и наблюдение и не се ангажират с точен процент на осъществимост.

Фигура 6. Бъдещо развитие на ИИ



*Източник: Автора

Етапи на развитие и ключови фактори

2025–2030: Началото на "Агентната" икономика

Този период се характеризира с масовото навлизане на **агентен AI**. Както е посочено в доклада на McKinsey "Technology Trends Outlook 2025" (стр. 11), агентният AI е сред водещите технологични тенденции, които ще променят бизнес моделите. Системите ще могат да изпълняват рутинни задачи, които изискват обработка на информация и вземане на решения в тесни домейни, освобождавайки човешкия ресурс за по-сложни дейности. Вероятността от 80% за масово навлизане на агентен AI в офисните процеси

е висока, тъй като вече виждаме примери като асистента на Klarna, който обслужва две трети от запитванията на клиенти, или AI-асистента на Booking.com.

Ключовото предизвикателство е създаването на **мултимодални модели с памет**. Този аспект е критичен за справянето със сложната "работа със знания", която изисква разбиране на различни формати данни и дълъг контекст. Документи като "The Economic Potential of Generative AI" от McKinsey подчертават как генеративният AI (и неговата еволюция) ще трансформира познанието, което е в основата на до 70% от работните процеси.

2030–2040: Планиране и Полуавтономна роботика

През този етап AI системите ще се развият в "широкодомейн асистенти", които не просто изпълняват задачи, а участват в **планирането на сложни проекти**. Те ще действат като "ко-пилоти" за човека, осигурявайки подкрепа в области като управление на проекти, стратегическо планиране и логистика. Документът "Implementation of Artificial Intelligence (AI): A Roadmap for Business Model Innovation" (стр. 1-2) подчертава нуждата от "пътна карта" за внедряване на AI, което ще бъде особено важно през този период, за да се избегнат провали.

2040–2050: AGI-кандидати и Симулационни двойници

Този период бележи прехода към т.нар. "**AGI-кандидати**". Това не означава появата на пълноценен AGI, а на системи, които са способни да преминат голяма част от когнитивните тестове, сходни с човешките, но все още под строг надзор. Концепцията за "**дигитални двойници**" на бизнеси и градове е изключително важна. Тя ще позволи на компаниите и правителствата да симулират сложни сценарии (напр. въздействието на нова политика или промяна във веригата за доставки), преди да ги приложат в реалния свят. Този подход намалява риска и повишава ефективността, превръщайки се в ключово конкурентно предимство, както е посочено в "The Economic Potential of Generative AI" (стр. 6).

2050–2070: Стабилен AGI-набор и системно управление

В периода се предполага създаването на "AGI-набор" от системи, които работят заедно и се наблюдават взаимно, осигурявайки **формални гаранции за безопасност**. Това ще бъде задължително условие, особено след първите глобални инциденти с координирани AI системи. Този етап ще доведе до появата на нови роли като "надзорник на етиката на AI" и "архитект на цели", което е в съответствие с необходимостта от нов тип хибридни (**човек и AI**) организации, за да се управлява тази сложна технология.

Ключови ограничения и решения

Основните ограничения на бъдещия AI няма да бъдат самите модели, а фактори като **данни, надеждност и енергия**.

- **Данни** - нарастващата нужда от данни ще се посрещне чрез **синтетични данни**, генерирани от симулации и други AI системи. Това ще реши проблема с недостига и защитата на личните данни.
- **Надеждност** - за да се гарантира безопасността на системите, ще се прилагат **формални методи** за верификация, както и "**агентни сандбоксове**", които ще позволяват безопасно тестване на AI в симулирана среда.
- **Енергия** - нарастващият енергиен апетит на AI центровете ще наложи изграждането на специализиран хардуер и локални енергийни източници, като малки модулни реактори (SMRs) или възобновяеми микромрежи, което е ключово за устойчивото развитие.

Рискове и бъдещи роли

Системни грешки - решенията са **многостепенен надзор** и въвеждане на "kill-switch" механизми.

Дезинформация - проследимост на съдържанието, което е все по-актуална тема, особено след появата на генеративния AI.

Икономическа поляризация - хипотези за "**синтетични данъчни системи**" и дивидент от продуктивност е иновативно решение, което би могло да смекчи социалните последици от масовата автоматизация.

AI еволюция, движена от прагматизъм, която прекроява бизнеса и професионалните умения

Развитието на изкуствения интелект до 2070 г. няма да бъде внезапен скок към научна фантастика, а по-скоро постепенно и прагматично пътешествие. Бъдещето на AI ще се определя не толкова от теоретичните граници на моделите, а от реални, осезаеми ограничения като **енергия, данни и надеждност**.

Ефекти върху бизнеса, от трансформация на модели до конкурентно предимство

Еволюцията на AI ще катализира дълбоки промени в начина, по който работят компаниите, като ще превърне икономиката на знанието в следващата "граница на продуктивността".

Иновация на бизнес моделите

AI ще бъде движеща сила за създаването на изцяло нови бизнес модели, особено в производствения сектор, където ще улесни прехода към "**цифровизирани услуги**" (digital servitization). Това означава, че компаниите ще продават не просто продукти, а резултати, базирани на услуги, управлявани от AI.

Оптимизация на стойността

За да капитализират напълно върху AI, компаниите ще трябва да оценят своята **"вътрешна AI зрялост"** и да разберат как да създават и улавят стойност, използвайки технологията. Липсата на ясна пътна карта за внедряване на AI може да доведе до неуспешни проекти и загиване на компанията.

Автоматизация и ефективност

Очаква се масовото навлизане на **агентен AI** в бизнес процесите и на **дигитални двойници** в производството и градското планиране. Това ще намали разходите и ще автоматизира рутинните задачи, позволявайки на компаниите да се фокусират върху иновациите.

Необходими умения и познания за съвременните мениджъри и служители

Докато AI поема рутинните задачи, ролята на служителите ще се промени от изпълнител към контролиращи и формиращи стратегия. За да бъдат успешни в тази нова екосистема, служителите ще трябва да развият набор от специфични умения.

Когнитивни и "меки" умения:

- **Критично мислене и етика** - способността да се оценяват и верифицират резултатите от AI, както и да се вземат етични решения, ще бъде от първостепенно значение.
- **Сътрудничество с AI** – служителите ще трябва да работят в екипи **„Човек + AI“**, като се научат да си взаимодействат с интелигентни системи за постигане на общи цели.
- **Креативност и адаптивност** - нуждата от иновативно мислене и способността за бързо адаптиране към променящи се работни процеси ще се превърнат в конкурентно предимство.

Специфични познания и технически умения:

- **Базово разбиране на AI** - за да се справят с новите предизвикателства, служителите ще трябва да усвоят основни познания за изкуствения интелект в бизнес контекст.
- **Проектиране и промптинг** - умението да се създават ефективни "подказвания" (prompts) и да се работят с **„микро-кейсове“** ще стане част от ежедневните задачи, тъй като това е основният начин за взаимодействие с езиковите модели.
- **Познание на данни и процеси** - разбирането на основите на данните и способността да се анализират процесите ще бъдат критични за ефективното използване на AI в оптимизацията на работата.

В крайна сметка, до 2070 г. AI най-вероятно ще бъде дълбоко вграден в икономиката и обществото. Конкурентното предимство на бъдещите бизнеси ще зависи не само от технологичните им възможности, но и от способността им да изградят **доверие**, да управляват ефективно своите **данни** и да осигурят необходимата **енергия** за работата на своите системи. Еволюцията на AI е по-скоро история за решаване на сложни инженерни и социални проблеми, отколкото за постигане на мистична утопия.

Теми за дискусия Глава 2

- *Кои бизнес модели се променят най-драматично от навлизането на генеративния AI?*
- *Какви умения трябва да развият бъдещите професионалисти, за да работят ефективно в икономика, доминирана от AI?*
- *Какви са най-важните фактори за успешното интегриране на AGI/мултиагентни системи в бизнеса след 2040 г.?*
- *Как да се балансират иновациите в AI с енергийните и екологичните ограничения?*

3. AI за начинаещи, ключови концепции и икономическите ефекти

„Изкуственият интелект е новото електричество.“

Андрю Нг, съосновател на Coursera, бивш главен учен в Baidu и ръководител на Google Brain.

Сравнението на ИИ с електрификацията или с интернет революцията не е преувеличение. Електричеството не просто даде достъп до светлина, то прекрои производствените процеси, позволи създаването на градове и промени начина, по който живеем. Интернет не просто ускори комуникацията, той създаде изцяло нови пазари, бизнес модели и социални взаимодействия. Днес изкуственият интелект е следващият фундаментален катализатор, който ще трансформира не само това, което правим, но и как го правим, създавайки огромна икономическа стойност, но и повдигайки нови фундаментални въпроси.

Джей Барни (1991), в теорията си за **Ресурсно-базирания поглед**, твърди, че устойчиво конкурентно предимство се постига чрез притежаване на **VIRO ресурси** – ценни (Valuable), редки (Rare), неимитиращи се (Inimitable) и без еквивалентни заместители (Non-substitutable). Системите за препоръки (Netflix, Amazon) генерират милиарди, тъй като предоставят **ценен ресурс**, който намалява „**парадокса на избора**“ (Schwartz, 2004) и е трудно имитируем без съответните **Big Data** (рядък ресурс). AI генерира **качествена разлика** с конкурентите (Barney, 1991). Внедряването на AI за **оптимизация на маршрути** (UPS) и **предиктивна поддръжка** (Shell) намалява разходите, тъй като **автоматизира** сложни процеси, които традиционната логика не може да оптимизира (Drucker, 1999). **Франк Найт още през 1921 дефинира**, че разграничението между **риск** (измерима вероятност) и **неопределеност** (неизмерима вероятност) е фундаментално. AI системите (Visa Advanced Authorization, Stripe Radar) превръщат **неопределеността** (непознати киберзаплахи или измамни схеми) в **риск**, който може да бъде управляван и оценен в рамките на секунди, което е **безалтернативно** за оцеляването във финансовия и киберсектор. **Робърт Дънкан (1976)** въвежда концепцията за **организационната среда** и необходимостта от съответствие между стратегията, структурата и средата. Успехът изисква **иновации в управлението на данни** (Data Pipeline) и **демократизация на AI** (достъп до моделите за всички служители), за да се превърнат служителите от изпълнители в **надзорници и стратегии**, както изисква ко-еволюционният модел. **Yale Insights** поставя въпроса за **отговорността**, когато AI модел вземе грешно или вредно решение. Регулаторни актове като **ЕС AI Act** отговарят на това, като въвеждат **строги изисквания за прозрачност, сигурност и човешки надзор** за „**висок риск**“ приложенията. Изкуственият интелект е структурна промяна и е **безалтернативен катализатор** за бъдещото икономическо развитие. Той осигурява **устойчиво**

конкурентно предимство (Barney, 1991) чрез **редуциране на предсказването** (Agrawal et al., 2018) и **управление на риска** (Knight, 1921).

Икономическият ефект на AI, макро и микроикономически анализ

Въздействието на ИИ не е ограничено само до отделните компании. То променя цялата глобална икономика. Според доклад на McKinsey & Company, генеративният ИИ може да добави между 2,6 и 4,4 трилиона долара към световната икономика годишно, което е равносилно на БВП на Германия или Япония. По-голямата част от тази стойност ще дойде от повишаване на производителността, тъй като AI може да автоматизира между 60% и 70% от работните задачи, които в момента се извършват от хората. Докладът на McKinsey идентифицира и кои сектори ще бъдат най-силно повлияни:

- Банковият сектор може да генерира до 340 милиарда долара годишна стойност от автоматизация на обслужването на клиенти и оценка на риска.
- Търговията на дребно може да извлече стойност от 400 до 660 милиарда долара чрез динамично ценообразуване и персонализирани препоръки.
- Ефекта върхи фармацията се естимира между 60 и 110 милиарда долара от ускореното откриване на нови лекарства. Например, Pfizer използва AI, за да съкрати времето за разработка на лекарства от години на месеци, което е от решаващо значение за създаването на нови лекарства. Този икономически ефект ще доведе до ръст на глобалната производителност, но и до фундаментална промяна на пазара на труда.

Очакването е вместо да елиминира работните места, навлизането на AI да ги трансформира, автоматизирайки рутинните задачи и създавайки нужда от нови умения, фокусирани върху креативността, стратегическото мислене и управлението на AI системи и данни.

На ново компания AI катализира **микроикономически ефект и създаване на стойност**, като въздейства върху три основни измерения на бизнес стойността, банласа и паричния поток - приходи, разходи и риск.

- **Повишаване на приходите.** Системата за препоръки на Netflix е перфектен пример. Тя не е просто алгоритъм за филми, а стратегически инструмент за намаляване на „парадокса на избора“, като помага на потребителите да намерят съдържание бързо. В резултат на това, над 80% от съдържанието, гледано в Netflix, е резултат от препоръки, а компанията оценява, че това генерира над 1 милиард долара годишно в предотвратени загуби на абонати. Аналогично, Amazon генерира около 35% от продажбите си чрез алгоритми за препоръки, което дава качествена разлика с бизнесите, които не са имплементирали идентични процеси в търговските операции.
- **Намаляване на разходите.** Автоматизацията чрез чатботове и виртуални асистенти, като тези, които използва Klarna, може да обработва две трети от клиентските запитвания, намалявайки нуждата от човешки оператори. Оптимизацията на операциите е друг ключов аспект, логистични гиганти като UPS използват AI за

оптимизиране на маршрутите на доставка, спестявайки милиони литри гориво и милиарди долари годишно. В производството, предиктивната поддръжка е модел, при който AI анализира данни от сензори, за да предвиди кога дадена машина ще се повреди. Например, Shell Global използва predictive maintenance, за да предотврати скъпоструващи престои и загуба на продукция.

- **Управление на риска.** Във финансовия сектор, AI системите анализират стотици милиони транзакции в реално време, засичайки аномалии и предотвратявайки измами в рамките на секунди. Например, Visa Advanced Authorization е мощна система, която използва AI за оценка на риска в реално време, а Stripe Radar предотвратява хиляди измами дневно, спестявайки милиарди на търговци. В киберсигурността, AI се използва за откриване на заплахи и аномалии в мрежата. Докато традиционните системи разчитат на известни заплахи, AI може да идентифицира нови, непознати атаки, намалявайки времето за реакция от дни до секунди.

Предизвикателства и ограничения. Пътят към отговорен AI

Въпреки огромния потенциал, внедряването на AI е сложен процес, който е свързан с редица предизвикателства:

1. **Качество на данните и пристрастия (Bias).** Bias (на български: **пристрастие, изкривяване** или **тенденциозност**) е систематична грешка или отклонение, което води до неточни, изкривени или несправедливи резултати. В контекста на **AI** и **машинното обучение** bias означава, че моделът дава предсказания, които систематично облагодетелстват или ощетяват определени групи хора или резултати. AI моделите се обучават с данни и тяхната точност и справедливост зависят пряко от качеството на тези данни. Ако данните са пристрастни, моделите ще повтарят и дори ще усилят тези пристрастия. Най-известният пример е този със системата за наемане на служители на Amazon, която е била обучена с данни от минали години, доминирани от мъже. В резултат на това, системата се е научила да дискриминира кандидатки, като автоматично ги класифицира по-ниско. Това е класически казус за това как неконтролираният AI може да навреди на бизнеса и обществото.

Защо Bias е проблем?

- **Води до несправедливост** и дискриминиране на цели групи (по пол, етнос, възраст).
- **Създава съществен регулаторен риск** и може да доведе до нарушения на закони като GDPR или EU AI Act.
- **На риск е поставена репутацията на организацията**, грешни или несправедливи решения могат да подкопаят доверието на потретиците/ клиентите.
- **Неточните модели** могат да водят до загуби и бизнес риск (например погрешни кредитни решения).

Таблица 1. Видове Bias в AI

Вид Bias	Обяснение	Пример
Data Bias (Пристрастие в данните)	Данните, с които се обучава моделът, не са представителни за реалността.	AI за кредитен скоринг, обучен предимно с данни на мъже, може да подценява женски кандидати.
Selection Bias (Пристрастие при подбор на данни)	Избраният набор от данни пропуска важни групи или сценарии.	Модел за разпознаване на лица, трениран основно с лица на светлокожи хора, работи слабо за хора с тъмен цвят на кожата.
Algorithmic Bias (Пристрастие на алгоритъма)	Самият алгоритъм или начинът на обучение води до систематично изкривяване.	Прекомерно „наказване“ на отклонения от средната стойност, което ощетява малцинствени групи.
Measurement Bias (Пристрастие в измерването)	Използваните метрики или методи за оценка са неправилни.	Оценка на продуктивността само по брой завършени задачи, игнорирайки тяхната сложност.
Human Bias (Човешко пристрастие)	Пристрастия, които идват от хората, които подготвят данните или настройват модела.	Субективни етикети при подготовка на тренировъчни данни, които съдържат лични предразсъдъци.

2. **Необходимост от организационна трансформация.** Внедряването на AI изисква повече от технологични инвестиции. То налага ко-еволюционен процес, в който иновациите в AI трябва да се развиват паралелно с промените в бизнес модела и организационната култура. Компаниите трябва да развиват три основни способности:

- Управление на данните (Data Pipeline) - Способността за ефективно събиране, обработка, анализ, защита и моделиране на данни.
- Разработване на алгоритми, които да са контекстуално специфични за бизнес нуждите, което изисква специфична бизнес и техническа експертиза за програмиране, наблюдение, развитие и анализ на алгоритмите и то в динамика съответстваща на бизнеса.
- Демократизация на AI, осигуряване на достъп и разбиране на AI за всички служители, не само за експертите.

3. **Етични и регулаторни рискове.** Успехът на AI зависи от доверието на клиентите и обществото, което изисква от компаниите да се фокусират върху прозрачността и отговорността. Прозрачността е способността да се покаже как един AI модел е взел дадено решение. Отговорността поставя въпроса кой носи отговорност, когато AI моделът вземе грешно или вредно решение. Yale Insights поставя този въпрос в контекста на правото. Тези въпроси са в основата на европейското законодателство AI Act, което класифицира приложенията на AI според риска и налага строги изисквания за прозрачност, сигурност и надзор.

„Техническият дълг“ като стратегически императив в дигиталната трансформация и имплементация на AI

Стратегическото измерение

В дискурса на дигиталната трансформация, една концепция остава често подценена – **техническият дълг** (Technical Debt). За компании с дългогодишна история и значителни инвестиции в наследена (legacy) ИТ инфраструктура, техническият дълг престава да бъде чисто операционен въпрос и се трансформира в стратегически такъв.

Появата на нови бизнес модели – като „...-as-a-Service“ или „...-as-a-Platform“ – изисква гъвкави, модулни и управлявани от данни платформи, които да осигурят скорост и гъвкавост. Това изправя ръководните екипи пред фундаментален въпрос: какво да правим с наследената инфраструктура, която все още поддържа основния ни бизнес днес?.

Този избор не е само технологичен, а определящ посоката:

1. **Инвестиция в нови, облачно-базирани (cloud-native) платформи**, което отключва бъдещи модели, но е с цената на сътресения, първоначални капиталови разходи (CAPEX) и временна неефективност.
2. **Удължаване живота на съществуващите системи** с пълното съзнание, че това може да забави навлизането в бързо растящи, платформено-базирани модели.

Всяко решение носи дълбоки стратегически последици за непрекъсваемостта на бизнеса, конкурентното предимство и дългосрочната релевантност. Междувременно, дигитално-родените (digital-native) компании, изградени върху облак, данни и изкуствен интелект (AI), не носят тежестта на тези решения. Техният технологичен стек е проектиран от самото начало да поддържа разнообразни стратегии и бързи промени.

Техническият дълг не е просто технологичен проблем. Той е въпрос за това доколко една компания е готова за бъдещето.

Дефиниция и природа на техническия дълг

В академичната и практическата литература **техническият дълг** се отнася до остарял код, лошо интегрирани системи или прибързани технологични решения (rushed releases), които ускоряват доставката в краткосрочен план, но създават значителни разходи и сложност във времето.

Важно е да се отбележи, че техническият дълг не е непременно лош сам по себе си. Част от него може да бъде умишлен и стратегически – поет съзнателно за постигане на бърза пазарна валидация. Въпреки това, неуправляваният или прекомерен технически дълг се превръща в бариера пред иновациите, гъвкавостта и растежа.

Количествен анализ на въздействието (Емпирични данни)

Редица изследвания от водещи консултантски и академични институции очертават икономическото и оперативно въздействие на натрупания технически дълг:

- **McKinsey & Company (2020–2023)** - техническият дълг може да консумира между **20% и 40% от ИТ бюджетите** само за поддръжка. Това води до забавяне на дигитални инициативи и по-лошо представяне при излизане на пазара (time-to-market). Организациите, които активно управляват своя технически дълг, отчитат **2.5 пъти по-бързо доставяне** на иновации.
- **Stripe & Harris Poll (2018)** – ИТ инженерите прекарват средно **33% от времето си** в справяне с технически дълг. Това се равнява на глобална загуба на производителност от **85 милиарда долара годишно**.
- **CIO.com & IDC (2021)** - 64% от ИТ директорите (CIOs) твърдят, че техническият дълг ограничава иновациите и дигиталния растеж. Това възпрепятства интегрирането на AI, API и модерни платформи. Организациите, които проактивно намаляват дълга, увеличават възвръщаемостта на инвестициите (ROI) от дигитализация с 30-50% в 3-годишен хоризонт.
- **Harvard Business Review (2018)** - липсата на модернизация на основните системи води до фрагментирани архитектури и ограничена гъвкавост.
- **IEEE Software (2020–2023)** - високият технически дълг корелира пряко с по-ниско качество на кода, по-висок процент на дефекти и по-дълги цикли на издаване на нови версии.
- **Gartner (2024)** - до 2025 г. организации, които не успяват да управляват техническия дълг, ще отчетат **15% по-висок процент на ИТ провали**. Gartner препоръчва изграждането на т.нар. „Часовник на техническия дълг“ (Technical Debt Clock) за видимост и управление.

Влияние върху бизнес представянето

Техническият дълг оказва пряко, измеримо въздействие върху ключови бизнес показатели, далеч извън рамките на ИТ отдела.

Ефекти върху Отчета за приходите и разходите (P&L)

- **По-високи оперативни разходи (ОРЕХ):** Наследените системи изискват повече поддръжка, отстраняване на грешки (debugging) и инсталиране на кърпки (patching).
- **По-бавно излизане на пазара (Time-to-market):** Нови продукти или функционалности отнемат повече време за стартиране поради заплетени кодови бази и проблеми с интеграцията.
- **Небюджетирани ИТ разходи:** Ресурсите се отклоняват за отстраняване на проблеми, вместо да се насочват към иновации или генериране на приходи.

Клиентско преживяване и задържане

Техническият дълг често води до видими за клиента проблеми: грешки (bugs) и прекъсвания на услуги, бавни или нестабилни приложения и лоши потребителски интерфейси (UI). Резултатът е неудовлетвореност на клиентите, повишено оттегляне (churn) и увреждане на репутацията.

Конкурентна неспособност

Компаниите с висок технически дълг са по-малко отзивчиви към пазарните промени и по-бавни в иновациите. Конкуренти, използващи модерни платформи, могат да пускат продукти по-бързо, да се мащабират по-лесно и да се адаптират към нови технологии (AI, APIs).

Продуктивност и морал на екипите

Разработчиците прекарват 30-50% от времето си в справяне с технически дълг, вместо да изграждат нови функционалности. Фрустрацията от работата с остарели или нестабилни системи води до „прегаряне“ (burnout) и по-високо текучество сред инженерните и ИТ екипи.

Таксономия на рисковете, свързани с техническия дълг

Натрупаният технически дълг генерира многопластов риск за организацията:

- **Оперативен риск.** Повишен системен престой (downtime) и прекъсвания на услугите поради нестабилни архитектури. Трудности при мащабиране на инфраструктурата.
- **Риск за сигурността.** Наследеният код често не разполага с актуални корекции за сигурност, което прави системите уязвими на пробиви. Остарелите технологии може да не поддържат съвременни стандарти за сигурност (напр. криптиране, контрол на достъпа).
- **Регулаторен риск и риск за съответствието.** Остарелите системи могат да възпрепятстват спазването на развиващи се регулации като GDPR, PSD3 или ISO

стандарти. Липсата на одитни пътеки или цялостност на данните увеличава експозицията към правни глоби.

- **Финансов риск.** Превишаване на бюджета поради скрити разходи за поддръжка и отстраняване на проблеми. Пропуснати възможности за приходи поради забавени продуктови премиери.
- **Стратегически риск.** Невъзможност за бърза промяна (pivot), интегриране на нови бизнес модели или отговор на конкурентни заплахи. Изоставане в приемането на критични иновации като AI или API екосистеми.
- **Риск за HR.** Фрустрация сред служителите, принудени да работят с неефективни, остарели инструменти. Трудност при привличането и задържането на водещи инженерни таланти, които предпочитат модерни среди.

Стратегии за управление и смекчаване на техническия дълг

Ефективното управление на техническия дълг изисква проактивен и стратегически подход, интегриран в бизнес планирането.

1. **Квантифициране и приоритизиране** Необходимо е да се извърши одит на техническия дълг в системите и приложенията. Приоритизирането трябва да бъде съобразено с **бизнес въздействието**, а не само с техническите фактори.
2. **Интегриране в бизнес пътната карта** Техническият дълг трябва да стане видим на ниво висше ръководство. Трябва да се разпределят бюджет и ресурси за отстраняване на дълга като част от стратегическото планиране, а не просто като част от ИТ задачите (backlog).
3. **Инкрементално рефакториране (Refactor in Sprints)** Вместо да се чакат масивни и рискови пренаписвания на системи, трябва да се прилага подход „плащай, докато вървиш“ (pay as you go). Това се постига чрез посвещаване на част от всеки работен цикъл (спринт) за намаляване на техническия дълг.

В заключение, изкуственият интелект е повече от технологична иновация, той е структурна промяна, която вече очертава бъдещето на бизнеса. За да успеят в тази нова ера, бъдещите мениджъри и икономисти трябва да се научат да мислят не само за технологиите, а за това как те ще прекроят бизнес моделите, ще създават стойност и ще изискват ново лидерство. Големият въпрос вече не е „дали“ да се използва AI, а „как“ и „с каква скорост“ да се внедри, за да се постигне устойчиво конкурентно предимство.

Теми за дискусия Глава 3

- *Кои бизнес функции според вас ще бъдат най-силно трансформирани от ИИ до 2030 г.?*
- *Какво е необходимо за доброто разбиране на изкуствения интелект*
- *Какви са възможностите на ИИ да подпомогне обучителния процес*

Модул 2. В сърцето на дигиталния бизнес. Къде AI се среща с капитала

4. Магията на персонализацията. Как дигиталните платформи използват AI, за да ви познават по-добре?

„Персонализацията е новата лоялност. В дигиталната икономика клиентите остават там, където се чувстват разбрани и напускат, където са третирани като анонимни.“

Satya Nadella, CEO на Microsoft, Interview with The Verge, 2021

В съвременната дигитална икономика, успехът на дадена компания все по-рядко се корени в самия **базов продукт** и все повече зависи от способността ѝ да изгради **платформа** и да приложи **дълбока персонализация**. Тази стратегическа промяна трансформира потребителите от пасивни консуматори в **активни съавтори и участници** в изграждането на стойност.

Бизнес моделът, при който печалбата идва от **екосистемата** около базовия продукт (подобно на **Crocs** и аксесоарите **Jibbitz**, или **App Store** и приложенията), се обяснява най-добре чрез **Теорията на Двустранните Пазари (Two-Sided Markets)**. Според **Жан Тирол и Жан-Шарл Роше (Rochet & Tirole, 2003)**, платформите обслужват две или повече различни групи потребители (напр. разработчици и крайни клиенти), като стойността се генерира от **индиректните мрежови ефекти**. Те смятат, че определящата черта на двустранните пазари е, че платформата трябва да **субсидира** едната група (напр. потребителите на Android) и да **таксува** другата (напр. разработчиците), за да максимизира общата стойност на **мрежата**“ (Rochet & Tirole, 2006, p. 669). Платформите като **Google Play Store** печелят от **координационна стойност**, а не от себестойността на продукта. Те предоставят **цена** за взаимодействие, която се самоподдържа, колкото повече потребители има, толкова по-голям стимул имат разработчиците да създават приложения, което допълнително привлича потребители. В дигиталната икономика, мрежовите ефекти и ниските пределни разходи водят до **ефект „Победителят Вземва Всичко“ (Winner-Take-All)**. **Ерик Бринолфсон (Brynjolfsson, 2014)** твърди, че дигиталните платформи генерират **Supetstar ефект**, при който малка група фирми доминират на пазара. **Платформата** е без алтернатива за **мощабирание**, тъй като фирма, която не създаде собствена екосистема, не може да се конкурира със свръх-мощаба на вече доминиращите платформи.

Ролята на **AI** е да направи платформата **персонализирана и релевантна** и по този начин да **задържи потребителското внимание**, което е най-ценния ресурс в дигиталната икономика. Персонализацията е икономическият двигател, който прави **„Дългата Опашка“ (The Long Tail)** печеливша. **Кристиан Андерсън (Anderson, 2006)** постулира, че дигиталните платформи могат да печелят от голям брой нишови продукти, които поотделно имат ниско търсене, но общо съставляват значителен пазарен дял. Андерсън заключава, че без **филтриране** и **препоръки**, Дългата Опашка е само купчина

мълчаливи продукти. **Алгоритмите я правят печеливша.**“ (Anderson, 2006). AI (персонализацията) служи като **филтър**, който свързва потребителя с нишовото съдържание, което иначе би било скрито. Това позволява на компании като Amazon и Netflix да реализират основните си печалби от **диференциация** и **сегментиране** (Porter, 1985), а не само от най-популярните си продукти. За бизнеса, персонализацията, задвижвана от изкуствен интелект (AI) обещава безпрецедентна ефективност, дълбока ангажираност на клиентите и значителен ръст на приходите. Консултантски гиганти като McKinsey и академични институции като London Business School чертаят картина на нова ера на "хиперперсонализация в мащаб", представяйки я като следващата голяма граница на производителността и конкурентното предимство (Chui et al., 2023; Jacobides & Ma, 2024). Въпреки това, зад бляскавата фасада на оптимизираното потребителско изживяване се крие една много по-сложна и обезпокоителна реалност. Критични гласове като тези на Кати О'Нийл, Шошана Зубоф и Илай Парайзър предупреждават, че същите тези технологии могат да функционират като непрозрачни "оръжия за математическо унищожение", да подхранват хищнически бизнес модели на "надзорен капитализъм" и да ни изолират в социално разяждащи "филтърни балони" (O'Neil, 2016; Zuboff, 2019; Pariser, 2011). От гледна точка на бизнеса, внедряването на AI за целите на персонализацията е не просто желателно, а екзистенциално необходимо в условията на ожесточена конкуренция. Аргументите в подкрепа на тази тенденция могат да бъдат групирани в три основни категории: оперативна ефективност, конкурентно предимство и потребителска стойност.

Оперативната ефективност и производителност са може би най-често изтъкваните ползи. Докладът на McKinsey "The economic potential of generative AI" подчертава как новите AI модели позволяват "персонализация в мащаб", автоматизирайки създаването на маркетингови кампании, динамични препоръки и клиентско обслужване, което е едновременно по-бързо и по-точно (Chui et al., 2023). Това освобождава човешки ресурс, намалява разходите и позволява на компаниите да обслужват милиони клиенти с индивидуално пригодени оферти, немислимо доскоро. На второ място, персонализацията се разглежда като **ключов двигател на конкурентното предимство**. В своя анализ за IoD и London Business School, Jacobides и Ma (2024) твърдят, че фирмите трябва да насочат AI не само към оптимизация на разходите, а към генериране на приходи чрез "customisation/customer engagement". Способността да се предложи "хиперперсонализация в мащаб" се превръща в разделителна линия между пазарните лидери и останалите. Компаниите, които успешно използват собствените си масиви от данни, за да предвидят и отговорят на нуждите на клиентите, изграждат лоялност и увеличават пазарния си дял. Третата ос, **потребителската стойност**, е илюстрирана класически от системата за препоръки на Netflix. Gomez-Uribe и Hunt (2015) детайлно описват как използването на поведенчески данни (implicit feedback) за изграждане на персонализирани предложения директно води до по-висока удовлетвореност и задържане на потребителите. От гледна точка на клиента, добре настроената персонализация елиминира информационния шум, спестява време и предлага релевантно съдържание и продукти, подобрявайки цялостното преживяване.

Взети заедно, тези аргументи представят AI персонализацията като win-win ситуация, при която бизнесът печели ефективност и приходи, а потребителите получават по-добро и по-удобно обслужване. Тази визия обаче, както отбелязват самите консултантски доклади, е зависима от "високо качество и достъп до данни" и признава съществуването на рискове, които често остават на заден план (Singla et al., 2024).

Опасността зад алгоритъма

Контрапунктът на оптимистичната бизнес реторика идва от автори, които анализират дълбоките структурни промени, предизвикани от същите тези технологии. Техните критики разкриват как персонализацията може да се превърне в инструмент за дискриминация, манипулация и социална фрагментация.

Кати О'Нийл и нейните **"Оръжия за математическо унищожение"** (WMDs) демистифицират представата за алгоритмите като неутрални и обективни. Тя убедително доказва, че те са "мнения, вградени в математика", които отразяват и усилват съществуващите човешки пристрастия. Персонализацията, според О'Нийл, става опасна, когато е **непрозрачна, мащабируема и нанася вреди**. Алгоритми за кредитен рейтинг, които наказват хора заради пощенския им код, персонализирани реклами за работа, които изключват определени полове или възрастови групи, и системи за прогнозиране на престъпността, които таргетират малцинствени общности, са конкретни примери за WMDs в действие. "Черната кутия" на тези системи лишава индивидите от правото да разберат и оспорят решения, които имат реални и често опустошителни последици за живота им. По този начин "ефективността", възхвалявана от бизнеса, се превръща в "индустриално производство на несправедливост". **Шошана Зубоф** отива още по-далеч, като позиционира персонализацията в центъра на нов икономически ред **"надзорния капитализъм"**. В нейната рамка, персонализацията не е услуга за потребителя, а **екстрактивен механизъм**. Компаниите събират "поведенчески излишък" огромни масиви от данни, които далеч надхвърлят необходимото за функционирането на дадена услуга. Тези данни се превръщат в суровина за AI модели, чиято цел не е просто да предвидят, а да **модифицират бъдещото ни поведение**. Фините подтиквания (nudges), персонализираните стимули и контролът върху информационната среда се използват, за да ни насочат към търговски резултати, които са изгодни за надзорните капиталисти. Така "ангажираността на клиента" се превръща в евфемизъм за поведенческа манипулация, а човешкият опит е колонизиран и превърнат в стока. **Илай Парайзър** се фокусира върху епистемологичните и демократични последици от персонализацията чрез концепцията си за **"филтърния балон"**. Алгоритмите, стремейки се да ни покажат съдържание, което ще харесаме, постепенно ни изолират от различни гледни точки. Резултатът е интелектуална изолация и създаването на "ехо камери", които затвърждават съществуващите ни убеждения. Макар последващи изследвания да показват по-сложна картина, основният аргумент на Парайзър остава валиден, алгоритмичната персонализация действа като невидим редактор на нашата реалност. Тя ерозира споделената информационна основа, която е жизненоважна за функционирането на демократичното общество, и прави конструктивния диалог между хора с различни

възгледи все по-труден. "Задържането" на потребителя, което е цел за компании като Netflix, може да има висока социална цена. Теорията за "хиперперсонализацията в мащаб" (Jacobides & Ma, 2024) е операционализираната версия на "надзорния капитализъм" на Зубоф. Стремещт към максимална ангажираност и задържане на клиенти (Gomez-Uribe & Hunt, 2015) неизбежно води до създаването на "филтърни балони" (Pariser, 2011). А мащабируемата ефективност, постигната чрез непрозрачни алгоритми (Chui et al., 2023), е точно това, което ги превръща в потенциални "оръжия за математическо унищожение" (O'Neil, 2016). Следователно, **AI персонализацията пренаписва правилата, защото променя фундаментално самата природа на властта в отношенията бизнес-клиент и бизнес-общество**. Тази власт вече не се основава само на пазарна доминация или качество на продукта, а на:

1. **Асиметрия на знанието**, компаниите знаят почти всичко за нас, докато ние не знаем почти нищо за начина, по който те използват това знание.
2. **Капацитет за поведенческа модификация**, системите са проектирани не просто да обслужват, а активно да оформят нашите желания, вярвания и действия в невиджан досега мащаб.
3. **Автоматизация на пристрастията**, социалните неравенства могат да бъдат кодирани и възпроизведени със скоростта и обхвата на алгоритмите, правейки дискриминацията по-ефективна и по-трудна за откриване.

Това не е просто "по-добър маркетинг", а нова социално-техническа парадигма, която изисква значително внимание и регламентиране. Разчитането на саморегулация от страна на бизнеса е наивно, тъй като икономическите стимули са насочени към задълбочаване на персонализацията, а не към ограничаването на нейните рискове.

AI персонализацията от една страна е двигател на иновациите, производителността и удобството, предлагайки на бизнеса и потребителите значителни ползи. От друга страна, тя носи със себе си сериозни и основателни рискове от задълбочаване на неравенствата, ерозия на личната автономия и фрагментация на обществения дебат. Да се игнорират предупрежденията на О'Нийл, Зубоф и Парайзър в името на икономическия растеж, обещан от консултантските доклади, би било историческа грешка. Предизвикателството пред нашето общество е да намерим правилния баланс от една страна да впрегнем силата на AI, без да жертваме основните ещности като справедливост, свобода и демокрация. Това изисква създаването на нови регулаторни рамки, които да настояват за прозрачност на алгоритмите (отваряне на "черните кутии"), да защитават данните на гражданите като тяхно основно право, а не като безплатна суровина, и да установят ясни механизми за отчетност, когато тези системи нанасят вреда. Пътят напред не е в спирането на технологията, а в нейното съзнателно и отговорно насочване към цели, които служат на цялото общество, а не само на тесните търговски интереси.

Внедряването на AI изисква **организационна трансформация**, която да върви в синхрон с технологичното развитие. **Робърт Дънкан (Duncan, 1976)** описва това като **ко-еволюционен процес**, при който иновациите в AI трябва да бъдат придружени от

развитие на **Управление на данните (Data Pipeline)** и **демократизация на AI** сред служителите. Фирми, които инвестират само в алгоритми, но не и в **организационна култура**, която да ги поддържа, няма да успеят да превърнат персонализацията в **устойчиво** конкурентно предимство.

Продуктът като основа за екосистема

Много добър пример за това как работи дигиталната икономика може да вземем от бранда Crocs - перфектен пример за това как едно просто и функционално ядро/продукт може да се превърне в платформа.

- **Базов продукт**, класически чифт Crocs, обувки, които са изключително удобни, но с противоречив дизайн.
- **Добавена стойност**, появата на **Jibbitz** – малки аксесоари, които се закрепват за дупките на обувките. Тези персонализируеми елементи позволяват на всеки клиент да изрази своята индивидуалност, да отбележи интересите си или просто да създаде уникален външен вид.
- **Бизнес модела**. Crocs не просто продават обувки, те създават **екосистема**. Приходите от аксесоари генерират значителна част от общия оборот, което показва, че основната стойност вече не е само в обувката, а в **персонализацията**, която тя предлага.

Урок за дигиталния бизнес

В дигиталния свят логиката е много близка, но с различни инструменти. Компаниите, които разбират този принцип, постигат устойчив растеж и лоялност на клиентите.

Изграждане на силно ядро, създаване на прост, надежден и скалируем дигитален продукт, който решава масова потребителска нужда и да бъде лесен за използване.

Създаване на възможност за персонализация, потребителите да „декорират“ своя личен модел, без това да разход за бизнеса, напротив трябва да се трансформира в приход. **Добър пример в тази посока е Netflix**. Основният продукт е стрийминг услугата, но истинската сила е в **персонализацията** на потребителския интерфейс. Алгоритъмът за препоръки (recommender system) персонализира всеки един елемент от подредбата на филмите до миниатюрите, които виждате, като по този начин създава уникално изживяване за всеки потребител.

Създаване на екосистема. Възможност други бизнеси да надградят върху твоята платформа, създавайки синергия. **Synthesia**, предлага основен продукт платформа за генериране на видеоклипове с AI аватари. Тя се използва от хиляди корпоративни клиенти за създаване на обучения и маркетинг съдържание, но също така позволява на индивидуални потребители да правят собствени проекти. По този начин, Synthesia създава екосистема, която расте заедно с креативността на потребителите.

Платформата като икономика на изборите

Истинската сила на дигиталните платформи не е в базовия продукт, а в **свободата на избор и персонализацията**, която дават на потребителите. Подобно на успеха на Crocs, чийто марж на печалба идва не толкова от самата обувка, а от персонализиращите елементи **Jibbitz**, дигиталните бизнеси реализират основните си печалби от екосистемата около базовия си продукт. Тази екосистема включва приложения, разширения, данни и допълнителни услуги.

Два от най-ярките примери за това са **Google Play Store** и **Apple App Store**.

- **Google Play Store.** Екосистемата на Android е изградена върху принципа на отвореността и достъпността. Google предоставя операционна система, която производителите на телефони могат да използват безплатно, а печалбата идва от **екосистемата от приложения**. Всеки разработчик може да създаде и публикува приложение, а Google получава дял от всички продажби, абонаменти и трансакции в рамките на магазина. Това създава огромен пазар за приложения, който привлича милиарди потребители. Google не се налага да произвежда всички приложения, от които хората се нуждаят, а вместо това печели от това, че е домакин на тази "сцена". В резултат, приходите от Play Store и рекламната мрежа в приложенията са ключов двигател на растежа на Google, далеч надхвърляйки приходите от продажба на собствен хардуер.
- **Apple App Store.** За разлика от Google, Apple създава строго контролирана и премиум екосистема. App Store е единственият легитимен начин за инсталиране на приложения на iPhone и iPad. Тази затворена система позволява на Apple да гарантира качеството и сигурността на приложенията, което оправдава таксата, която разработчиците плащат за достъп до платформата. Apple получава **комисиона (обикновено 15-30%)** от всяка продажба на приложение, абонамент или трансакция. Тази стратегия превърна App Store в един от най-големите генератори на приходи за компанията, надминавайки дори приходите от Mac компютри. Apple не продава само телефони, а цяла екосистема от услуги, която поддържа приходите ѝ дълго след първоначалната продажба на хардуера.

Тези модели превръщат потребителите от пасивни консуматори в **съавтори и активни участници в изграждането на стойност**. Компанията вече не е просто производител на един продукт, а **организатор на сцена**, на която клиентите и разработчиците могат да творят, да обменят идеи и да създават съдържание, което поддържа екосистемата жива и печеливша. Печалбата идва от мащаба на взаимодействието, а не от пределната себестойност на продукта. Истинската стойност възниква, когато клиентите могат да персонализират и създават заедно с компаниите, което е бъдещето на дигиталната икономика.

Персонализацията, задвижвана от AI, и **платформеният бизнес модел** са **архитектурните стълбове** на дигиталната икономика. Те са безалтернативни, защото:

(1) създават **мрежови ефекти** и **конкурентно предимство** (Rochet & Tirole; Barney); (2) **управляват вниманието и избора** на потребителите (Schwartz; Anderson); и (3) **преобразуват** базовия продукт в **екосистема** с високи маржове на печалба. Успехът изисква не само техническо съвършенство, но и **стратегическо лидерство**, способно да управлява етичните рискове, произтичащи от дълбоката зависимост от данни.

Каква роля заема изкуствения интелект в бизнес моделите на настоящето и бъдещето ще проследим в следващите части.

Теми за дискусия Глава 4

- *Разгледайте в интернет, (вкл. и AI GTP) бизнес развитието на американската компания за мобилни телефони BlackBerry. Какви са основните изводи, които може да направим от развитието на компанията.*
- *Какви KPI бихте препоръчали за оценка на успеха на AI-базираната персонализация в една банка или дигитална платформа?*
- *Къде е границата между полезната персонализация и поведенческата манипулация?*
- *Какво би означавало за една компания да не персонализира услугите си в ерата на хиперперсонализацията – загуба на пазарен дял или стратегическа защита на личното пространство на клиента?*
- *Може ли да си представим бъдеще, в което потребителят сам избира нивото на персонализация (Full, Partial, None)? Как това ще промени бизнес моделите?*

5. От бутиков до масов продукт. AI и (не)видимата трансформация на пазара

„Най-рисковата фаза на пазара е преходът от визионери към прагматици.“

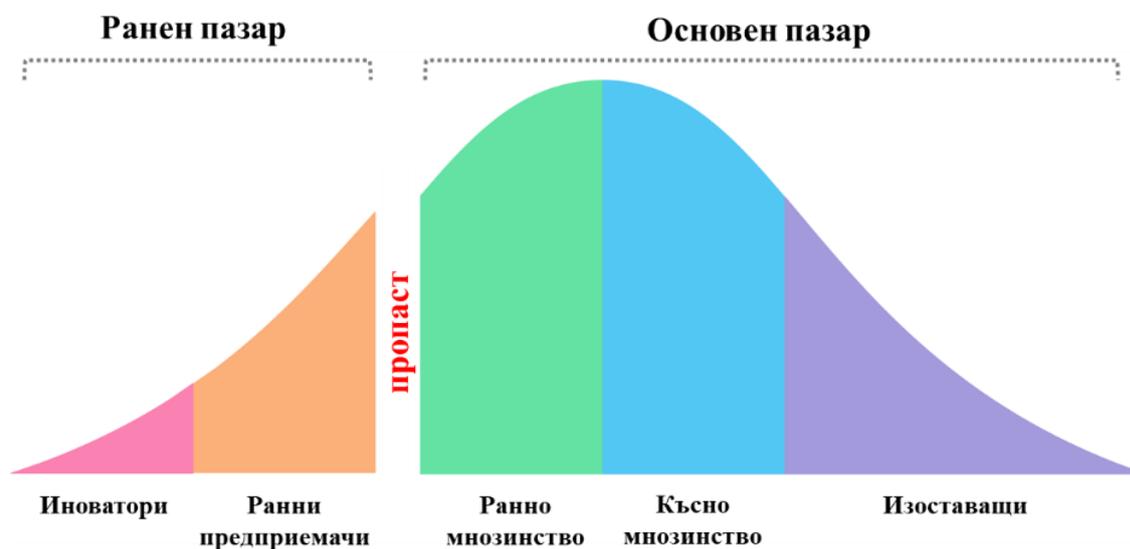
Джефри А. Мур, Crossing the Chasm, 3rd ed., 2014.

Изкуственият интелект (AI) не е просто поредната технологична иновация, той е пример за мета-технология, която фундаментално променя самата логика на това как новите продукти се възприемат и разпространяват в обществото. Тази трансформация пренаписва класическия модел, известен като „Кривата на Роджърс“, и го заменя с много по-бърза и агресивна динамика.

Класическият модел на приемането на иновациите

През 1962 г. социологът Еверет Роджърс формулира своята теория за дифузията на иновациите, която и до днес служи като основен инструмент за разбиране на разпространението на технологии и идеи. Според този модел, потребителите могат да бъдат разделени в пет основни категории:

Фигура 7. Жизнен цикъл на приемане на технологията



- **Иноватори (2,5%).** Те са авантюристите и експериментаторите, които винаги търсят новото. Те са готови да поемат риск и да инвестират време и пари в нетествани технологии.

- **Ранни предприемачи (13,5%).** Това са лидерите на мнение и визионерите. Те не просто експериментират, а валидират технологията, дават социално доказателство за нейната полезност и влияят на по-широката аудитория.
- **Ранно мнозинство (34%).** Тази група е прагматична. Те изчакват доказателства за полезността, стабилността и ефективността на продукта, преди да го приемат. Те са склонни да се учат от опита на ранните потребители.
- **Късно мнозинство (34%).** Това е по-скептична и предпазлива група потребители. Те възприемат иновацията едва когато тя е доказана, цената ѝ е паднала значително или социалният натиск е станал прекалено голям.
- **Изоставащи (16%).** Най-консервативната група, приемат иновацията само когато тя е неизбежна, защото старият им начин на работа или живот вече не е възможен.

Джефри Мур (Moore, 1991) в **Преодоляване на Пропастта (Chasm)** твърди, че най-трудно е **преходът** от *Ранните Предприемачи* към *Ранното Мнозинство* (прагматиците). **AI съкращава тази Пропаст** от години на месеци (напр. ChatGPT), защото **прагматиците** (Ранното Мнозинство) виждат **незабавна, интуитивна полезност**. **Ерик Бринолфсон и Андрю Макафи** (Brunjolfsson & McAfee, 2014) в **Експоненциален Растеж (The Second Machine Age)** достигат до извода, че дигиталните технологии (AI) следват **експоненциален**, а не линеен път на развитие. **AI-кривата** е експоненциална, тъй като **темпът на промяна е самонарастващ** и се ускорява с всеки следващ потребител. Ускорената AI-крива е резултат от комбинацията на **демократизация на интерфейсите** и **мрежови ефекти**, които намаляват *барьерите за приемане* и *времето за социално доказателство*. AI премахва **техническата грамотност** като бариера. Тази промяна е теоретично свързана с **Теорията на Възприеманата Лекота на Използване (TAM)**. **Фред Дейвис** (Davis, 1989) в **Модел на Приемане на Технологии (TAM)**, **заключва, че внедряването на технология зависи от Възприеманата Полезност и Възприеманата Лекота на Използване**. AI (напр. **Copilot в Office 365**) има изключително висока *лекота на използване*, което **автоматично въвлеча Късното Мнозинство и Изоставащите**, които не искат да инвестират време в сложно обучение. Скоростта, с която се случват промените също е **бизпрецедентна**, **ChatGPT** достига 100 милиона потребители (два месеца), се обяснява от силата на **директните и индиректните мрежови ефекти** и **вирусната дифузия**. **Робърт Меткаф** (Metcalfe's Law) в **Мрежови ефекти** определя, че стойността на мрежата расте пропорционално на **квадрата на броя на свързаните потребители** ($V \approx n^2$). AI продуктите (напр. **Suno AI** или **GitHub Copilot**) са **въвлечени в мрежи** (социални мрежи), при които **стойността се увеличава с всеки нов участник**. **Йона Бергер** (Berger, 2013) **във Вирусна Дифузия и Социално Доказателство**, споделя че AI-генерирано съдържание (мемета, изображения) в социалните мрежи създава **социално доказателство** и силен **FOMO (Fear of Missing Out)**. Тази **невидима трансформация** на пазара е на повърхността **видима**, защото *резултатите* от AI са лесни за споделяне и интригуващи. Друга фундаментална специфика на AI е промяната в приемането в компаниите, от **Top-Down** към **Bottom-Up**, е теоретично свързана с

Теорията за Адаптивност на Организациите. Робърт Дънкан (Duncan, 1976) в Адаптивни организации заключава, че успешните организации са **амбидекстри** – способни да управляват едновременно **експлоатацията** (оптимизиране на съществуващото) и **изследването** (търсене на новото). AI позволява на **служителите** да започнат **индивидуално изследване** (напр. ChatGPT за имейли), което оказва **натиск отдолу-нагоре** за **централизирано внедряване** (експлоатация в **SalesForce/SAP**). Това **ускорява корпоративното приемане**, тъй като организацията не чака централното решение, а се адаптира в реално време към нуждите на служителите.

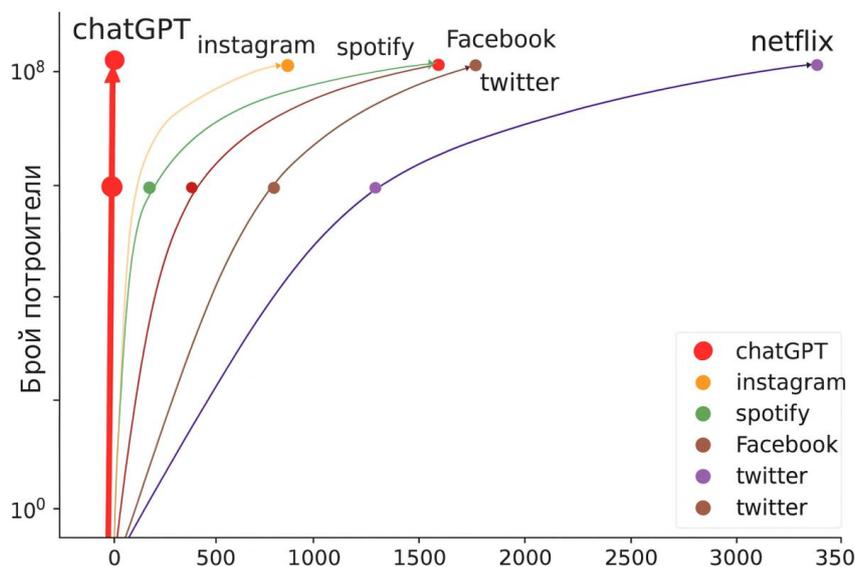
AI е **мета-технология**, която елиминира историческите забавяния в Кривата на Роджърс, превръщайки я в **стръмна, компресирана S-образна линия**. Тази нова динамика е задвижвана от **експоненциалния растеж** на самата технология, **демократизацията** на интерфейсите и **експлозията на мрежовите ефекти** (Metcalfе's Law) в социалните мрежи. Това налага компаниите и регулаторните органи да адаптират стратегическите си посоки **в реално време**, като превръща **скоростта на възприемане** в ключово конкурентно предимство.

AI като катализатор на промяна, появата на нова крива

С появата на изкуствения интелект се наблюдава радикална трансформация, вместо плавна и продължителна дифузия, продуктите подкрепени с AI показват **стеснена, експоненциална крива на възприемане**. Този ефект се получава в резултат от, драстично съкратеното време за преход между отделните групи. Примерите са показателни и свидетелстват за промяна на добре позната ни парадигма за жизнения цикъл на продуктите.

- **ChatGPT (OpenAI, 2022)** както вече беше посочено достига 100 милиона потребители само за два месеца, с което става най-бързото потребителско приложение в историята. Това не е просто успех на един продукт, а доказателство за нова ера на възприемане.
- **Midjourney и Stable Diffusion** за по-малко от година милиони потребители започват да генерират изображения, които се разпространяват масово в социалните мрежи. Това превръща всеки в създател на съдържание, независимо от художествените му умения.
- **GitHub Copilot** въведен като експеримент, но за месеци променя начина, по който програмистите пишат код, интегрирайки се в техния ежедневен работен поток. Дори традиционно консервативната софтуерна индустрия възприе технологията с невиджана скорост.
- **Suno AI** за броени месеци се появява платформа, която генерира песни, и веднага започва да събира милиони потребители, въвличайки и хора, които нямат музикални познания.
-

Фигура 8. Бързина за възприемане дигиталните инструменти, дни за достигане на 100 млн. потребители



*ChatGTP/Автора

Примерите доказват, че традиционните фази между ранни възприематели и ранно мнозинство се **съкращават драматично** – понякога от години до месеци.

Основни фактори за ускорената крива на AI

Ускорението на кривата не е случайно, а е резултат от няколко ключови фактора, които са уникални за изкуствения интелект.

Интуитивни интерфейси и демократизация. AI премахна технологичните бариери, за разлика от интернет в 90-те или смартфоните в началото на 2000-те, които изискваха техническа грамотност, AI интерфейсите са интуитивни, потребителите не се нуждаят от дълги инструкции или обучение, достатъчно е да зададат въпрос на познат от него език. Лекотата на използване на технологията позволява дори късното мнозинство и закъснелите да възприемат новата технология значително по-бързо. **Примери са Microsoft Copilot в Office 365**, служители без технически опит могат да създават автоматизирани отчети, да обобщават имейли или да генерират презентации с една команда, зададена на естествен език. **Adobe Photoshop с Generative Fill**, функцията позволява на художници и дизайнери да манипулират изображения, без да познават сложните инструменти на програмата, просто чрез текстово описание.

Бързо разпространяващо се съдържание и социално приемане, AI резултатите са видими и достъпни, могат да се споделят веднага и генерират силен ефект на **FOMO (fear of missing out, страх да не изостанеш)**. Всяко ново изображение или текст, създаден от AI, се разпространява в социалните мрежи, което създава масов ефект. **Примери са мемовите и бързо разпространяващо се съдържание (viral).** В началото

на 2023 г. в TikTok и Twitter се разпространиха милиони изображения, генерирани от Midjourney и DALL-E. Това изложи широката публика на възможностите на AI, без дори да е използвала продукта. Водещи компании като Coca-Cola (в сътрудничество с DALL-E) стартираха кампании, които насърчават създаването на съдържание с AI, което превръща потребителите в съавтори.

Автоматизирана поддръжка и персонализация. AI системите не само съществуват, но и сами улесняват собственото си възприемане чрез интелигентно и персонализирано **онбординг изживяване**. Чатботове, виртуални асистенти и адаптивни уроци правят първите стъпки лесни и безболезнени. **Примери са Duolingo Max, който** използва GPT-4 за персонализирани обяснения и интерактивен диалог, което ускорява възприемането на нови потребители, като им помага в реално време. **Booking.com и Expedia, които** интегрират AI асистенти, които правят процеса на планиране на пътуване по-интуитивен и персонализиран, като така премахват бариерите за потребители, които не са свикнали да използват комплексни платформи.

Приемане на AI отдолу-нагоре в корпоративен контекст. Традиционно организациите възприемаха технологии чрез централизирани решения, но при AI се наблюдава обратното, **служителите започват да използват AI инструменти индивидуално** (например ChatGPT за писане на имейли) и след това натискът идва отвътре за тяхното формално внедряване. **Примери в тази посока са интеграция в ERP и CRM системи.** AI-моделите се интегрират на по-ниско ниво в платформите като Salesforce и SAP, което позволява на търговци и маркетинголози да използват AI, без да са експерти. Компаниите започват да експериментират с иновативни бизнес модели, в които AI е централен компонент, като например „фабрики“ за персонализирани продукти, задвижвани от AI.

Таблица 2, сравнение между класическата и AI-крива за приемане на иновации

Елемент	Класическа крива (Rogers)	AI-крива
Форма	Широка, плавна, камбановидна	Стръмна, компресирана, S-образна
Време	Десетилетия	Дни/Месеци
Движеща сила	Централни внедрявания, маркетинг	Вирусен ефект, социални мрежи
Роля на потребителя	Пасивен в ранните етапи	Активен съавтор на съдържание
Корпоративно внедряване	Отгоре-надолу (Top-down)	Отдолу-нагоре (Bottom-up)
Демократизация	Ограничена от умения и ресурси	Масова, без бариери

Стратегически и социални последици

Компаниите вече не могат да чакат резултата от иновацията на своите конкуренти, ако пропуснат първите месеци на масова вълна, рискуват да загубят **конкурентоспособност и бъдат изтикани от пазара**. Умението да се **мащабира AI** бързо и ефективно се превръща в ключова компетенция на компаниите, за това сме свидетели и на вълна от

нови AI модели, които биквално се обявяват от разработчиците в рамките на две до три седмици.

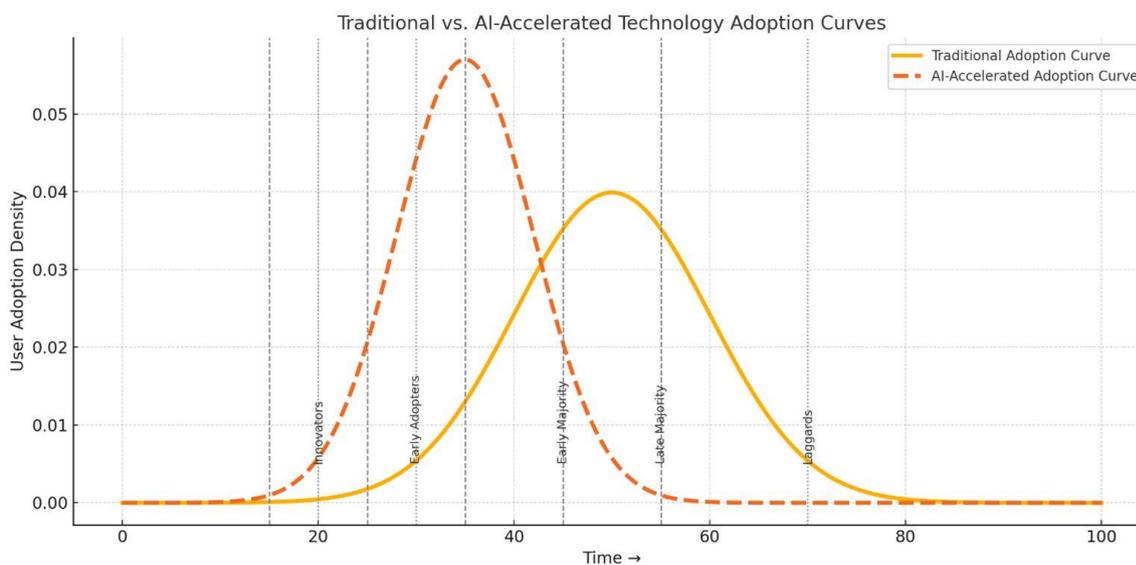
За държавните органи скоростта на развитие на новите продукти, особено технологичните е истинско предизвикателство. Регулацията изостава понякога и с повече от година, което създава етични и правни рискове. Бързата масова употреба прави забавянето на регулацията още по-опасно.

За потребителите възниква феноменът на „демократизирана иновация“, групите, които преди биха били закъснели с приемането на иновациите, днес се включват на много ранен етап, което намалява технологичното неравенство.

AI променя самата логика на технологичното възприемане на иновациите. Вместо класическа камбановидна крива, виждаме ускорена **много по динамична крива**, при която всички сегменти на потребителите са въввлечени много по-бързо. Това е не просто ускорение на един процес, а структурна промяна в икономиката на иновациите

- Продуктите достигат масовост в рамките на месеци, а не години и десетелетия.
- Потребителите стават едновременно консуматори и разпространители, възприемайки платформения ефект на дигиталния бизнес и ефекта на социалните мрежи.
- Компаниите и държавите са принудени да реагират много по-бързо на иновациите, буквално се налага да адаптират стратегически посоки в реално време, което е съществено предизвикателство пред управлението и адаптивността към промени.

Фигура 9.Изменение на кривата за приемане на нови продукти/услуги в резултата на ИИ/AI



Теми за дискусия Глава 5

- *Разгледайте кривата и циклите за приемане на мобилните телефони от потребителите. Какви са основните изводи, които може да направим?*
- *Как AI компресира „пропастта“ на Мур и кои продуктови свойства го позволяват (интерфейс на естествен език, freemium, мрежови ефекти)?*
- *Кои фактори определят преминаването от бутиков към масов продукт: цена, UX, мащаб на данните, стандарти/интероперабилност?*
- *Каква е ролята на мрежовите ефекти?*

6. Готови ли сме за бизнес, в който AI взема повечето решения?

„Роботите не са хора и не трябва да ги третираме като такива. Хуманизирането на машините рискува да де-хуманизира самите хора.“

Joanna J. Bryson, Robots Should Be Slaves (2010). Джоана Дж. Брайсън е професор по етика и технологии в Hertie School, Берлин, и водач изследовател по изкуствен интелект, поведение на сложни системи и AI governance.

Според редица изследователи на AI най-дълбоките му последици се коренят в способността му да симулира **човешката емпатия**. В основата на това съждение стои дебата на въпроса дали AI може да изпитва или само да имитира емоции. **Rosalind W. Picard** с **Affective Computing** (1997) полага теоретичната рамка за системи, способни да разпознават и изразяват емоции, поставяйки морални въпроси за „емоционално интелигентните“ машини. На поведенческо ниво, теорията **CASA (Computers as Social Actors)** на **Reeves & Nass** (1996) доказва, че хората **реагират на компютри като на социални актьори**, дори без истински чувства от машината. Тази възприета емпатия се усилва от визуални сигнали, демонстрирани в работата на **Cynthia Breazeal** със социалната роботика (2002), където изразите и поведението регулират социалното взаимодействие. Този феномен е свързан с теорията на **Gray, Gray & Wegner** (2007) за **Dimensions of Mind Perception**, според която хората приписват „Опит“ и „Агенция“ на обекти, включително AI. Емпиричните изследвания показват, че AI е много ефективен в **симулирането на когнитивна емпатия** (уместно разпознаване и реагиране). Чатбот отговорите в медицината бяха предпочетени и оценени като **значително по-емпатични** от тези на лекари (Ayers et al., 2023). Систематични прегледи (Chen et al., 2024; Welivita & Pu, 2024) потвърждават, че Големите езикови модели (LLM) показват елементи на когнитивна емпатия, като на места надминават човешките оценки, което се измерва с бенчмаркове като **EMRank** (Mayo Clinic Corpus, 2024) и методологията на **Xie et al. (2024)**.

Успешната симулация на емпатия от AI поражда дълбоки етични рискове, свързани с автентичността и социалната цялост. Основният риск е **заблудата** (Sharkey & Sharkey, 2020), при която външните сигнали на работа създават **фалшиво убеждение за чувства** и емпатия, насърчавайки **неуместно доверие**. В контекста на здравеопазването, експресивната „емпатия“ може да се усеща **неавтентична и манипулативна** (Seitz, 2024), а **Sharon (2025)** предупреждава срещу **техносолюционизма** при „емпатичните“ медицински чатботове и пренебрегнатите рискове отвъд поверителността. LLM-моделите (наречени „**Стохастични Папагали**“ от Bender, Gebru et al., 2021) имитират емпатия без разбиране, което носи риск от манипулация и пристрастия. Антропоморфизмът на AI е в основата на морален дебат. От една страна, хората показват емпатични реакции към „болка“ на роботи (Невроизследвания, цитирани от WIRED), което, според **Kate Darling**, има правно-етични последици. От друга страна, **Joanna J.**

Bryson заема категорична позиция, че „**Роботите не трябва да бъдат описвани като личности**“ (*Robots Should Be Slaves*), тъй като хуманизирането на машини води до дехуманизиране на реални хора. **Sherry Turkle** (2011) предупреждава в **Alone Together** и последващи работи за **ерозия на човешката емпатия**, когато хората заместват взаимоотношенията с „емоционални“ машини, които предлагат **илюзията за компания без изискванията на приятелството**.

От автоматизация към автономност

През последните десетилетия бизнесът разчита на изкуствения интелект основно за **автоматизация на рутинни процеси**, от обработка на данни до автоматизирани транзакции. AI системите функционират като мощни, но пасивни инструменти за подкрепа, които изпълняват команди, зададени от хора. Днес обаче наблюдаваме качествена промяна, AI вече не е просто инструмент за подкрепа, а **активен участник във вземането на решения**, способен да анализира, интерпретира и действа въз основа на огромни масиви от данни. Тази еволюция превръща AI в "решаващ агент" (agentic AI), който може автономно да постига цели и да се адаптира към променящи се обстоятелства. Въпросът пред организациите е фундаментален, „**можем ли да се доверим на алгоритми да вземат критични бизнес решения вместо експерти и мениджъри**“? Това е дилема, която съчетава огромни ползи с не по-малко значими рискове.

Предимства на agentic AI

AI системите притежават характеристики, които правят човешкото вземане на решения несъвършено при определени обстоятелства. Те не се влияят от умора, емоции или субективни пристрастия, което ги прави идеални за оптимизиране на процеси и постигане на максимална ефективност. Какви са най-характерните качества на AI?

Скорост и обем на обработваната информация. AI може да анализира огромни масиви от данни и да вземе решение за милисекунди, което е недостижимо за човешки капацитет. Примерите са множество:

- **В банковия сектор** системи за кредитен скоринг като тези на Zest AI или Upstart анализират хиляди параметри за клиента и вземат решение за отпускане на заем почти моментално. Големи банки като JPMorgan използват AI и за **откриване на измами** и анализ на риска в реално време, което позволява незабавна намеса и предотвратява неоторизирани транзакции.
- **В търговията на дребно** Walmart използва AI, за да анализира продажбени тенденции, предпочитания на клиенти и динамика на веригата за доставки, с цел **оптимизиране на нивата на инвентара**. Това минимизира загубите и подобрява клиентското изживяване.
- **В логистика** компании като United Parcel Service (UPS) използват AI софтуер, наречен "DeliveryDefense," който анализира исторически данни за кражби и дава на

всяко място "оценка на сигурността" (delivery confidence score) от 1 до 1000. Ако даден адрес има нисък скор, системата автоматично пренасочва пакета към по-сигурна локация, предотвратявайки кражба.

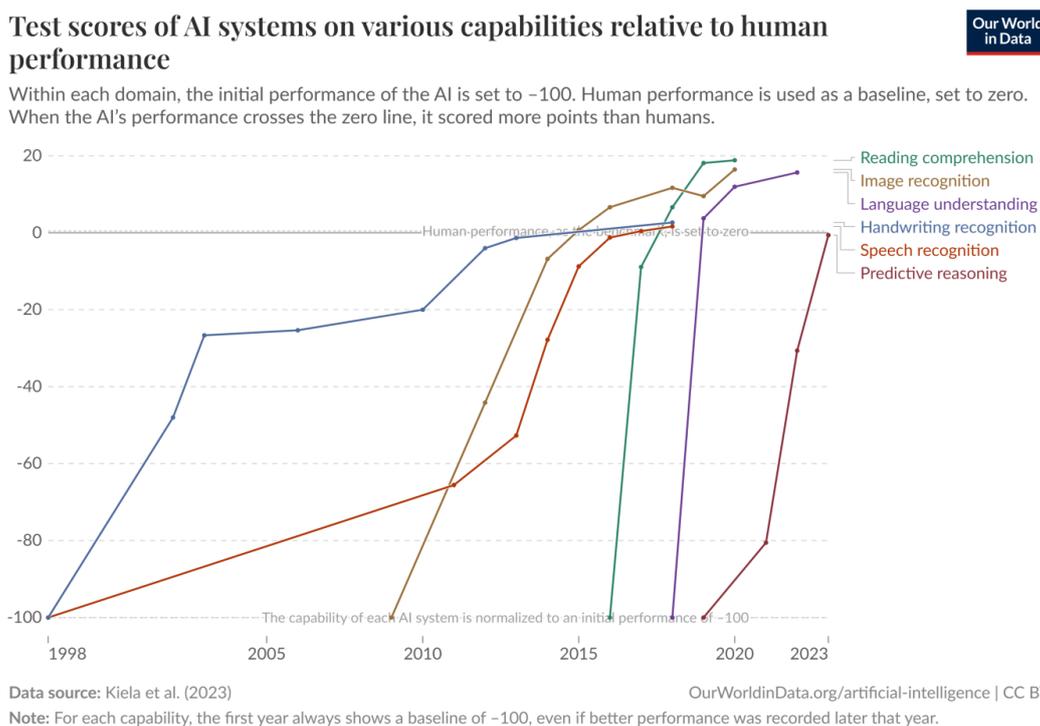
Денонощна работа без ограничения. AI няма работно време, умора или отпуски. Това позволява на компаниите да предлагат услуги 24/7, което се използва вече масово в:

- **Обслужване на клиенти,** Klarna използва AI-асистент, който обработва около две трети от клиентските заявки. Според Juniper Research, компаниите, които използват чатботове, **спестяват над 8 милиарда долара годишно** благодарение на по-бързото разрешаване на проблеми и намаляване на натовареността на човешките екипи.
- **В здравеопазването** виртуални медицински асистенти предоставят съвети 24/7, освобождавайки капацитета на лекарите за по-сложни случаи.

Консистентност, обективност и последователност. Докато хората са подвластни на емоции, стрес, умора и редица субективни фактори, AI взема решения по стабилен и предвидим начин, елиминирайки човешката грешка и пристрастия. Примерита са:

- **В производство** General Electric използва AI за **прогнозна поддръжка**, анализирайки данни от сензори, за да предвиди потенциални повреди на оборудването и да препоръча графици за поддръжка, преди да се стигне до скъпоструващи престои.
- **В модата,** Burberry използва AI, за да сканира онлайн изображения и да определи дали даден продукт е автентичен или фалшив, с **98% точност**.

Фигура 10. Представяне на AI спрямо човек

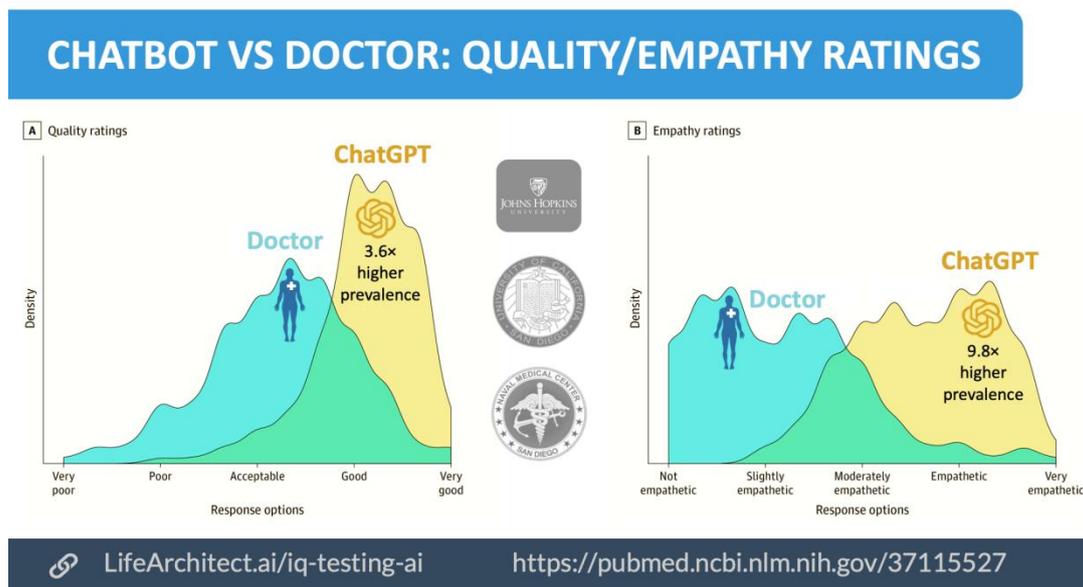


Рискове и предизвикателства

Интегрирането на AI в ключовите бизнес процеси не е лишено от сериозни предизвикателства, които засягат както етиката, така и самите основи на човешкото творчество.

Емпатия слабост на AI или негова силна страна? Доскоро се смяташе, че машините не могат да покажат емпатия. Но проучване, публикувано в JAMA Internal Medicine, показва, че ChatGPT дава по-емпатични и по-качествени отговори на медицински въпроси в сравнение с реални лекари. Това поставя под съмнение основния аргумент срещу AI в обслужването на клиенти, а иманно липсата на човешка емпатия и емоционална интелигентност. Проучване на MIT Media Lab дори установява, че AI системи, които включват **разпознаване на емоции**, могат да подобрят удовлетвореността на клиентите с до 30%. Тези системи използват обработка на естествен език (NLP), за да разпознаят емоциите на клиента по думи и тон, и да реагират по подходящ начин. Пример, ако клиент напише "Изключително съм разочарован от забавената доставка," AI-системата може да разпознае негативното чувство и да отговори с емпатия, като "Съжалявам да чуя, че доставката ви се е забавила." Това показва, че AI не просто симулира емпатия, а я моделира по убедителен начин, който в много случаи е по-ефективен от човешката реакция.

Фигура 11. Сравнение на ChatGPT спрямо лекар спрямо рейтинг за качество и емпатия



Етична отговорност и тн. „черната кутия“. Един от най-големите рискове е липсата на **прозрачност** в процеса на вземане на решения от AI. Понякога алгоритмите са толкова сложни, че дори техните създатели не могат да обяснят как са достигнали до определен извод. Това се нарича "проблемът на черната кутия" (black box problem). Когато AI вземе решение, например отказ на кредит или автоматизирано лечение, кой

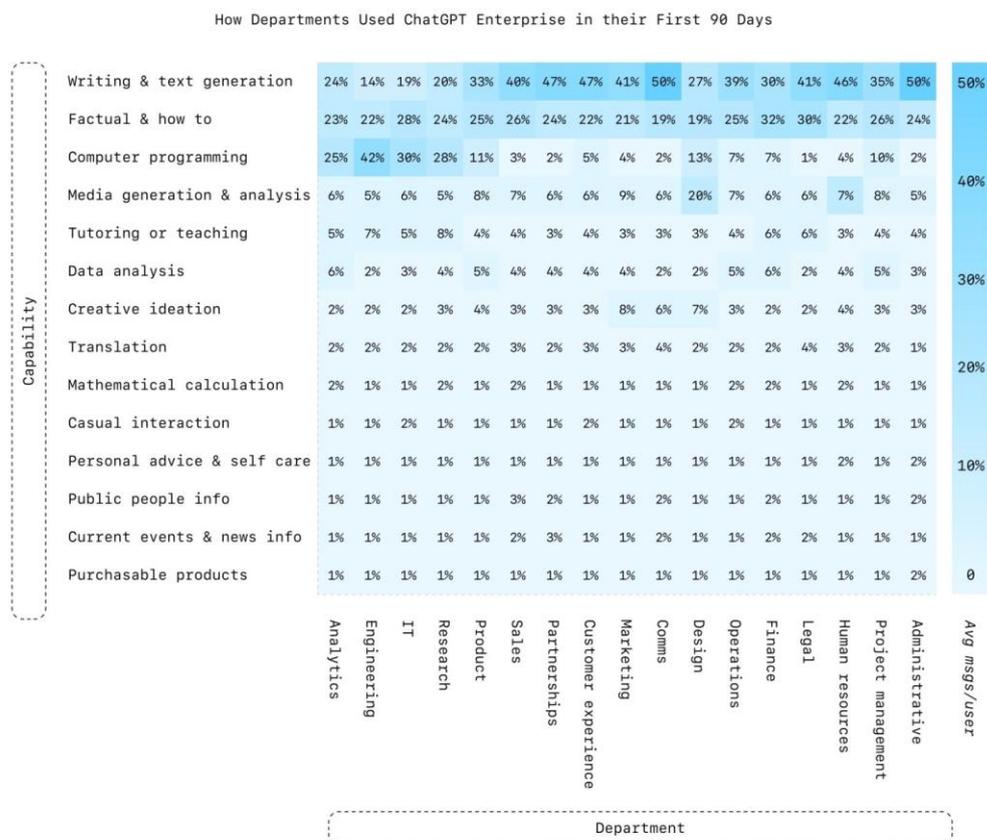
носи отговорност при грешка? Също така, AI алгоритмите могат да **увековечат и дори засилят съществуващите пристрастия** от историческите данни, с които са обучени. Това може да доведе до дискриминационни резултати при наемане на работа или насочване на клиенти, което може да повреди репутацията на компанията и да доведе до правни последици.

Риск от прекалена зависимост. Ако бизнесът се довери изцяло на AI, съществува опасност от загуба на стратегическа креативност и човешка интуиция. Прекомерната зависимост от AI може да доведе до самоуверено поведение и да намали критичното мислене, което е от съществено значение за вземането на решения. Експертите съветват хората да оценяват критично препоръките на AI и да не се доверяват сляпо на технологията.

Примери от практиката

Здравеопазване. Въпреки трудностите, IBM Watson Health показва потенциала на AI в подпомагането на лекарски решения. Днес, системи като TREWS на Johns Hopkins, използващи AI, могат да **разпознаят сепсис** и да обяснят препоръките си на лекарите, изграждайки доверие чрез прозрачност.

Фигура 12. Въприемане на ИИ в индустриите



Source: Aggregated ChatGPT Enterprise department data collected during onboarding; automated content classifiers

Логистика. DHL използва AI за **оптимизиране на маршрути**, управление на складове и доставка на пакети, което намалява оперативните разходи и подобрява точността на доставките.

Ритейл. Netflix и Spotify използват AI, за да **препоръчват съдържание** в реално време, като моделират предпочитанията на потребителите и оформят цялостното клиентско преживяване.

Човешки ресурси. IBM използва AI в своята програма SkillsBuild, за да **препоръчва курсове и сертификати** на служители, които преминават към технологично ориентирани роли.

Възможно бъдещо развитие

Хибридни модели човек + AI. Най-вероятният сценарий в близко бъдеще е съвместно вземане на решения. Концепцията "човек в цикъла" (human-in-the-loop) предполага, че AI ще служи като инструмент за **увеличаване на човешките способности**, докато хората ще поддържат надзор и ще имат последната дума в критични процеси. Това ще позволи комбинацията от скоростта и точността на AI с човешката креативност, емпатия и етична преценка.

Нови роли и професии. Появата на длъжности като „AI етичен супервайзор“ или „архитект на алгоритмични цели“ е напълно вероятна. Нараства търсенето на специалисти като **AI инженери, специалисти по машинно обучение и етични AI специалисти**. За да се адаптират към тази промяна, служителите ще трябва да развият нови умения като грамотност свързва с AI, бърза адаптивност, мобилност и адаптиране на промени..

От емпатия към „супер-емпатия“, при запазване на динамиката на развитите и усъвършенстване на AI системите на способността да симулират човешка топлина и съпричастност, бизнесът може да предложи обслужване, което надминава човешките стандарти. Представете си банки или болници, където клиентите винаги получават разбиране, търпение и уважение – без изключение. В същото време обаче, се повдигат етични въпроси относно поверителността на данните и дали тези емоционално интелигентни системи могат да бъдат използвани за манипулация на потребителите.

Индустрията вече навлиза в етап, в който AI не просто подпомага, а активно решава определени проблеми. Предимствата, като скорост, капацитет, достъпност и потенциал за емпатия са огромни. Рисковете, липса на прозрачност, етични дилеми и прекомерна зависимост са също толкова реални. Истинският въпрос е не дали AI ще взема решения, а как ще структурираме организациите и обществото така, че тези решения да са безопасни, справедливи и полезни за всички.

Въпреки емпиричното доказателство за превъзходството на AI в симулирането на **когнитивна емпатия**, фундаменталната **липса на афективна емпатия** създава **етичен императив** за прозрачност. Наложително е да се изградят строги граници срещу **социалната измама**, за да се гарантира, че технологичният напредък не води до ерозия на автентичните човешки взаимоотношения.

Теми за дискусия Глава 6

- *Емпатия или имитаци - как да разграничим „истинска“ от симулирана емпатия в AI?*
- *Кога „възприетата емпатия“ е полезна, а кога е манипулативна (Sharkey & Sharkey, 2020)?*
- *Как да тестваме и коригираме bias при емпатийни отговори?*
- *При кои решения емпатийният слой трябва да бъде само интерфейс, а не фактор в политиките на действие?*

7. Ефектът на лавината. Защо платформите с AI са неудържими

„В дигиталната икономика печели не най-добрият продукт, а платформата, която успее да изграде най-силните мрежови ефекти.“

Ерик Бринолфсон и Андрю Макафи, The Second Machine Age (2014). Ерик Бринолфсон е професор в Stanford Digital Economy Lab, изследва връзката между дигиталните технологии, продуктивността и икономическия растеж. Андрю Макафи е съосновател и съдиректор на MIT Initiative on the Digital Economy и специалист по стратегии за бизнес трансформация в дигиталната ера.

Мрежовите ефекти като „двигател на растежа“

В ерата на дигиталната икономика, бизнесът все повече се измества от продуктово-ориентирани модели към платформи формиращи от екосистеми. В този контекст, **платформата** не е просто посредник, а среда, в която различни групи от участници потребители, доставчици, разработчици си взаимодействат и създават стойност един за друг. Именно тук възниква феноменът на **мрежовите ефекти** (наричани още мрежови ползи или ефекти на търсенето), при които стойността на дадена услуга или продукт се увеличава с всеки нов потребител, който се присъединява. Това е най-мощният механизъм, който стои в основата на успеха на компании като Uber, Airbnb, Meta, Amazon и TikTok.

Мрежовите ефекти са критични за постигане на **експоненциален растеж**, тъй като създават цикъл на взаимна полза. Повече потребители водят до по-голяма стойност, което привлича още повече потребители. Тази динамика често води до ситуация, известна като **„winner-takes-most“**, където една доминираща платформа печели по-голямата част от пазара, оставяйки малко място за конкурентите. Например, за даден град, е много вероятно да има един-два основни доставчика на услуги за споделено пътуване, вместо десетки, тъй като мрежата на шофьори и пътници естествено се концентрира на едно място за максимална ефективност.

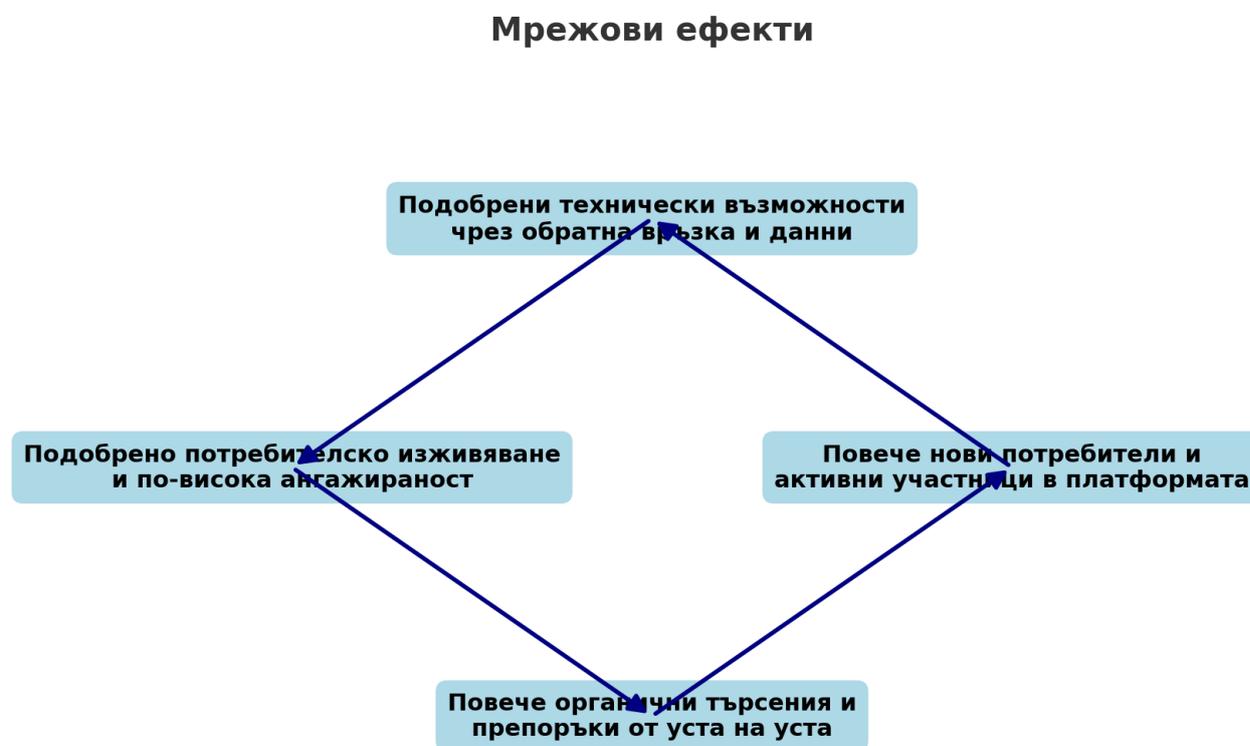
Видове мрежови ефекти

За да се разбере дълбочината на платформените бизнес модели, е важно да се разграничат трите основни вида мрежови ефекти:

Пряк (или едностранен) мрежов ефект. Най-простият и лесно разбираем, стойността на платформата се увеличава за всички потребители с привличането на всеки нов потребител от същата група. Класически пример са комуникационните платформи като **WhatsApp** и **Viber**. Колкото повече хора имат WhatsApp, толкова по-голяма е стойността на приложението за всеки от тях, тъй като могат да комуникират с повече

хора. Този ефект е ключов и за социални медии като **Snapchat** и **Instagram**, където всеки нов потребител допринася за общата стойност на социалната мрежа.

Фигура 5. Мрежов ефект при дигиталния бизнес



*Източник: Автора

Двустранен (или косвен) мрежов ефект. Ефект възниква при платформи, които свързват две или повече различни групи потребители. Стойността на платформата за едната група се увеличава с присъствието на другата група. **Airbnb** - колкото повече домакини предлагат имоти, толкова по-голям е изборът за гостите. Това привлича повече гости, което от своя страна стимулира още повече домакини да се присъединят към платформата, за да се възползват от търсенето. **eBay** по-големият брой купувачи на платформата мотивира повече продавачи да публикуват своите продукти. Това създава по-богат каталог, който привлича още повече купувачи.

Индиректен (или кръстосан) мрежов ефект, проявява се в платформи, които привличат трети страни, които създават допълнителна стойност. Най-ярките примери са екосистемите за приложения **Apple App Store** и **Google Play**. Присъствието на милиони потребители на iOS и Android е привлякло милиони разработчици, които създават

иновативни приложения. Тези приложения увеличават стойността на смартфоните за потребителите, което привлича още повече потребители, създавайки мощен кръговрат. Това ефект прави платформите на Apple и Google изключително трудни за изместване от пазара.

Примери и механизми на действие

В началните си етапи Uber субсидира шофьори и клиенти, за да постигне **критична база**. Повече шофьори означава по-кратко време за чакане и по-ниски цени за пътниците, което увеличава броя на пътуванията. Това повишено търсене пък привлича още шофьори, които виждат възможност за по-високи доходи.

Мрежовият ефект при LinkedIn е толкова силен, че за професионалистите "липсата" от платформата означава потенциална загуба на възможности за кариерно развитие и професионални контакти. Колкото повече професионалисти са в мрежата, толкова по-ценна става тя за всеки един от тях, което я превръща в де факто стандарт за професионални контакти.

В Amazon Marketplace всеки нов търговец добавя нови продукти, което прави платформата по-привлекателна за купувачите. По-големият брой купувачи от своя страна мотивира още повече продавачи да се присъединят, за да достигнат до по-широка аудитория.

Стратегически последици и предизвикателства

Високи бариери за навлизане. Веднъж установени, мрежовите ефекти създават почти непреодолими бариери за нови конкуренти. Един нов конкурент трудно ще привлече клиенти, ако няма достатъчно доставчици, и обратното. Това обяснява защо е толкова трудно да се появи нов играч на пазара на социалните мрежи или на пазара за споделено пътуване.

Зависимост на потребителите. Високите **разходи за смяна (switching costs)** заключват клиентите вътре в екосистемата. Класически пример е екосистемата на **iOS**, където веднъж щом потребителят инвестира в iPhone, iPad и Apple Watch, с покупка на приложения, музика и облачни услуги, той трудно би преминал към конкурентна платформа като Android.

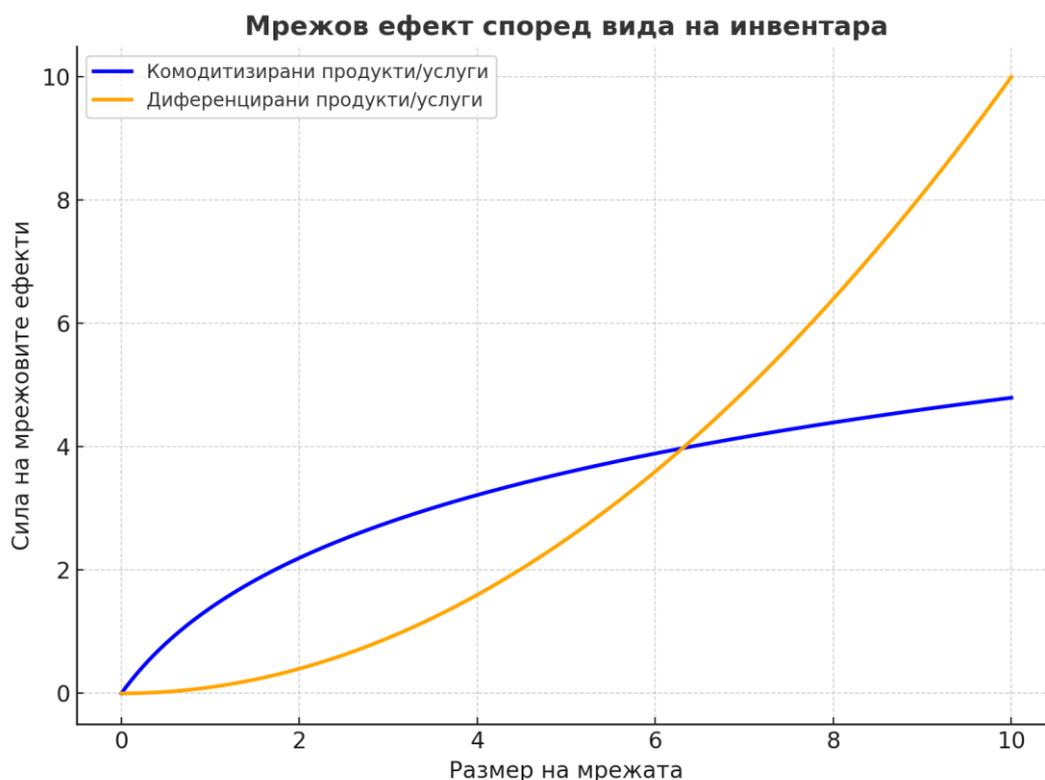
Етични и регулаторни въпроси. Динамиката „winner-takes-most“ може да доведе до пазарна концентрация и монопол, което повдига сериозни етични и регулаторни въпроси. Например, антимонополните органи по целия свят все по-често разглеждат поведението на технологични гиганти като Google и Meta, които използват своето пазарно господство, за да ограничават конкуренцията.

Бъдещето на мрежовите ефекти - AI и Web3

Генеративният изкуствен интелект (AI) добавя ново измерение към мрежовите ефекти. Платформи като **ChatGPT** и **Midjourney** стават все по-добри и интелигентни с всеки нов потребителски вход. Данните от взаимодействията на милиони потребители се използват за обучение на моделите, правейки ги по-точни и полезни. В този контекст, мрежовият ефект е не просто в броя на потребителите, а в качеството и количеството на данните, които те генерират.

Web3 и блокчейн, които идват с концепцията за децентрализирани платформи на базата на блокчейн също се стреми да използва мрежовите ефекти. При тях стойността на платформата (напр. децентрализирани финансови приложения или социални мрежи) се увеличава с броя на участниците в екосистемата, които допринасят с ликвидност, транзакции или съдържание. Това създава мрежа, която е устойчива на еднолични решения и контрол от страна на централен посредник.

Фигура 5. Мрежов ефект при ръст на бизнеса



**Източник: Автора*

В крайна сметка, мрежовите ефекти са ключови за съвременната икономика. Те не просто превръщат добрия продукт в глобален стандарт, но и позволяват на компаниите

да постигнат монополна позиция чрез създаване на **стратегически предимства**. Постигането на критична маса, непрекъснатото подобряване на потребителското изживяване и изграждането на доверие са основните елементи за изграждане на устойчиви платформи. Бъдещето на дигиталния бизнес ще продължи да се оформя от платформи, които умело използват мрежовите ефекти, за да се разширяват и да станат все по-незаменими в живота на потребителите. Това, от своя страна, поставя обществото пред предизвикателството да регулира тези „суперплатформи“, за да се осигури честна конкуренция и да се предпазят потребителите от потенциални злоупотреби.

Теми за дискусия Глава 7

- *Как мрежовите ефекти трансформират конкуренцията, създават ли нови монополи или увеличават иновациите?*
- *До каква степен е етично компаниите да изграждат екосистеми, които „заклучват“ потребителите чрез високи разходи за смяна?*
- *Как AI усилва мрежовите ефекти, води ли това до по-бързо постигане на „winner-takes-most“ динамика?*
- *Какви регулаторни механизми биха били най-ефективни за запазване на конкуренцията в платформи, които се възползват от силни мрежови ефекти?*
- *Каква е ролята на потребителите в поддържането на устойчивостта на платформите допринасят ли съзнателно или несъзнателно за тяхното доминиране?*

8. AI и Новата Икономика. Как интелигентните системи преосмислят разходите и мащаба?

„Когато пределната себестойност клони към нула, икономиката на мащаба се превръща в стратегия, а не в страничен ефект.“

Carl Shapiro & Hal R. Varian, Information Rules (1999). Карл Шапиро е професор по бизнес стратегия в Haas School of Business, UC Berkeley, специалист по индустриална организация и антимонополна политика. Хал Р. Вариан е главен икономист на Google и професор (emeritus) в UC Berkeley, смятан за един от водещите икономисти в областта на информационната икономика

В предходната глава задълбочихме разбирането си за силата на платформените ефекти и как дигиталните екосистеми създават мрежови ползи, които водят до динамиката на принципа „winner-takes-all“. В настоящата част ще разширим този анализ, като се фокусираме върху един от най-фундаменталните принципи в икономиката - разходите. В частност, ще изследваме как дигитализацията и особено изкуственият интелект (AI) променят концепцията за **пределна себестойност**, превръщайки я в ключов стратегически лост за управление.

Старо правило в бизнеса, е че приходите са функция от множество променливи и обстоятелства вт.ч. промени на пазара, реакция и стратегия на конкуренцията и др. Обаче разходите и как се управляват е истинската мярка за добър мениджър.

В дигиталната ера този принцип придобива ново и радикално измерение. Границите между фиксирани и променливи разходи се размиват, а пределната себестойност на дигиталните продукти често клони към нула. Това създава нова икономика на мащаба, в която растежът може да бъде експоненциален, а пазара силно концентриран. Дихотомията е централна за икономическата теория на производството и е основно разграничение между **материални стоки** и **информационни стоки (дигитални продукти)**. Традиционните стоки (като автомобили, мебели, конфекция и тн. др.) попадат под класическата икономическа теория, описана от **Алфред Маршал** (Marshall, 1890) и принципа на **намаляващата пределна възвръщаемост**. За да се произведе още една допълнителна единица след достигане на определен капацитет, са необходими значителни допълнителни ресурси (труд, капитал, суровини), което води до **нарастваща пределна себестойност** при мащабиране. Въпреки че Маршал говори за земеделие, принципът е разширен за традиционното производство, където увеличаването на производството на автомобили, например, изисква инвестиции в **нов капитал (заводи, машини)**, което прави всяка следваща единица по-скъпа за производство. Дигиталните или информационните продукти (софтуер, филми, AI модели) се характеризират с икономическите принципи на **информационната икономика** (Varian, 1998; Shapiro & Varian, 1999). Първоначалните **фиксиран разходи (Fixed Costs/CAPEX)** за разработка са изключително високи (R&D, създаване на AI модел), но **пределната**

себестойност за копиране и разпространение на всяка следваща единица е **почти нулева**. Това обяснява феномена на Netflix – разходът за всеки нов абонат е минимален, което води до **експоненциална икономия от мащаба**. Изкуственият интелект не просто се вписва в парадигмата на дигиталните продукти, а я радикализира, като действа като **катализатор за експоненциална икономия от мащаба**. AI радикализира модела на *Shapiro & Varian* (1999) като превръща **разходите за създаване на информация** (Fixed Costs) в **разходи за създаване на инфраструктура** (AI обучение, данни, GPU-ферми). Компаниите, които контролират тази инфраструктура (напр. OpenAI, Google, Microsoft), постигат **икономии на мащаба**, които са по-скоро **експоненциални мрежови ефекти** (Metcalfe's Law) и **платформен контрол**, отколкото просто производствена ефективност. Тяхната пределна себестойност може да е близка до нула, но **входната бариера** за конкурентите е астрономическа. AI е **технология с общо предназначение (GPT)** (Bresnahan & Trajtenberg, 1995), която променя както продуктите, така и производствените процеси във всички сектори. Чрез AI, дори **традиционното производство** (автомобилната индустрия) може да намали своята пределна себестойност, като оптимизира производствените линии, прогнозира дефекти и автоматизира дизайна, но остава ограничено от физическия закон на суровините и капитала. **Дигиталните продукти** премахват това физическо ограничение, позволявайки по-чиста форма на експоненциален растеж.

Фундаменталната разлика между традиционни и дигитални продукти/услуги

За да разберем тази трансформация, е важно да разграничим икономиката на разходите на традиционните физически стоки от тази на дигиталните.

Традиционни стоки

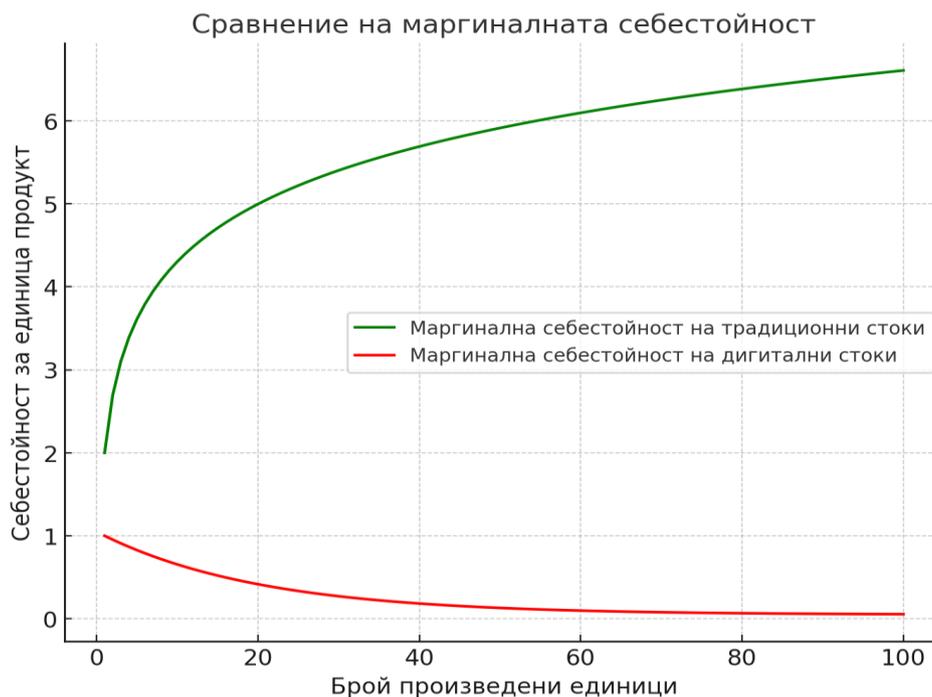
- **Пределна себестойност (Marginal Cost)**. Това е разходът за производството на още една допълнителна единица от дадена стока (напр. кола, стол, книга).
- **Динамика**. В традиционното производство съществуват **ограничения** като недостиг на суровини, необходимост от нов капитал (заводи, машини) и цена на труда. След като производството достигне определен капацитет, за да се произведе още една единица, са необходими значителни допълнителни ресурси, което води до **нарастваща пределна себестойност** при мащабиране. Например при **автомобилната индустрия** след определен брой произведени коли, за да се увеличи капацитетът, компанията трябва да инвестира в нови производствени линии, да наеме и обучи повече работници и да закупи допълнителни суровини. Това прави всяка следваща произведена кола по-скъпа за производство и ограничава експоненциалното разширяване.

Дигитални продукти

- **Пределна себестойност** – изключително ниска за всяко ново копие.
- **Динамика**. **Първоначалните разходи (Fixed Costs)** за R&D, разработка на софтуер, събиране на данни и внедряване на AI модели могат да бъдат изключително високи.

Но след като продуктът е създаден, разпространението и копирането на всяко следващо копие са с изключително ниска себестойност. Пример е **Netflix**, компанията инвестира милиарди в инфраструктура, закупуване на права за съдържание и производство на оригинални филми и сериали, но всеки нов потребител, който се присъединява, струва минимално допълнително, като разходът се свежда основно до поддръжка на сървъри и трафик на данни.

Фигура 6. Сравнение на маргиналната себестойност на дигиталните и традиционните продукти



*източник: Автора

Дигитализацията и Изкуствения Интелект в Персонализацията, Матрица Продукт-Процес

Класическата литература по операционен мениджмънт, представена от **Продуктово-Процесната Матрица (The Product-Process Matrix)**, очертава строга зависимост: високата степен на диференциация на продукта и персонализация (Customization) неизбежно кореспондира с гъвкави, но неефективни производствени процеси и, следователно, с **високи единични разходи (High Unit Cost)**. В традиционния икономически модел, преминаването от стандартизирани стоки с голям обем (Commodity, High volume, Rigid Process) към силно диференцирани продукти с нисък обем (Custom Differentiated, Low volume, Flexible Process) е било пряко свързано с повишена сложност, ръчен труд и, като следствие, с ценови излишъци. В настоящата част ще анализираме как **дигиталната трансформация, изкуственият интелект (AI) и роботизираната автоматизация на процеси (RPA)** разрушават тази класическа

връзка, позволявайки на компаниите да постигнат **персонализация в мащаб (Personalization at Scale)** без традиционния ръст на производствените разходи.

Продуктово-Процесната Матрица демонстрира, че организации като филмовата продукция (Movie production) функционират с висока гъвкавост, нисък обем и висок единичен разход, докато рафинериите за петрол (Oil refinery) оперират с фиксиран процес, висок обем и нисък единичен разход. Тази диагонална зависимост отразява компромиса между **гъвкавостта на процеса (Flexibility)** и **ефективността на разходите (Cost Efficiency)**. Производствената философия дълго време е приемала, че клиентите трябва да избират между стандартизация на достъпна цена или скъпа, но уникална изработка.

Ефект на AI - намаляване на разходите за персонализация

Ключовият пробив, предизвикан от дигитализацията, AI и RPA, е способността да се управлява **сложността на процеса (Process Complexity)** по автоматизиран и изключително ефективен начин. Предоставената графика ясно илюстрира това преместване с голямата червена стрелка, насочена надолу към зоната на **"Rigid Process / Low Unit Cost"** в лявата страна на матрицата. Този вектор на промяна представлява фундаментално прекъсване на предишната корелация.

Изкуственият интелект (AI), особено в лицето на машинното обучение, позволява:

- **Автоматизирано проектиране и инженеринг**, AI може бързо да генерира и оптимизира варианти на продукти, които преди са изисквали скъп и дълъг инженерен труд.
- **Прогнозна аналитика и управление на веригата за доставки**, посредством точно прогнозиране на персонализираното търсене, AI минимизира нуждата от скъпи инвентаризации и позволява **производство по поръчка (Build-to-Order)**, което намалява капиталовите разходи.

Роботизираната автоматизация на процеси (RPA) и напредналата автоматизация (Advanced Automation) позволяват на производствените линии, които преди са били фиксирани и оптимизирани само за един продукт, да станат бързо преконфигурируеми. Смяната на инструментална екипировка (tooling changeovers) и препрограмирането на производствените работи, които исторически са били скъпи и времеемки, сега могат да бъдат управлявани почти моментално от софтуер, поддържан от AI. Това прави **гъвкавостта на процеса (Flexible Process)** достъпна в рамките на структура с ниски единични разходи.

Дигитализацията измества фокуса от простото увеличаване на ефективността към постигане на **гъвкавост (agility)** и **персонализация в мащаб (personalization at scale)**.

Новата дигитална парадигма - адаптивност и бързина

Ефектът от дигитализацията не е просто количествено намаление на разходите, а качествена промяна в самата структура на производството. Това е преход от „**Фиксиран**

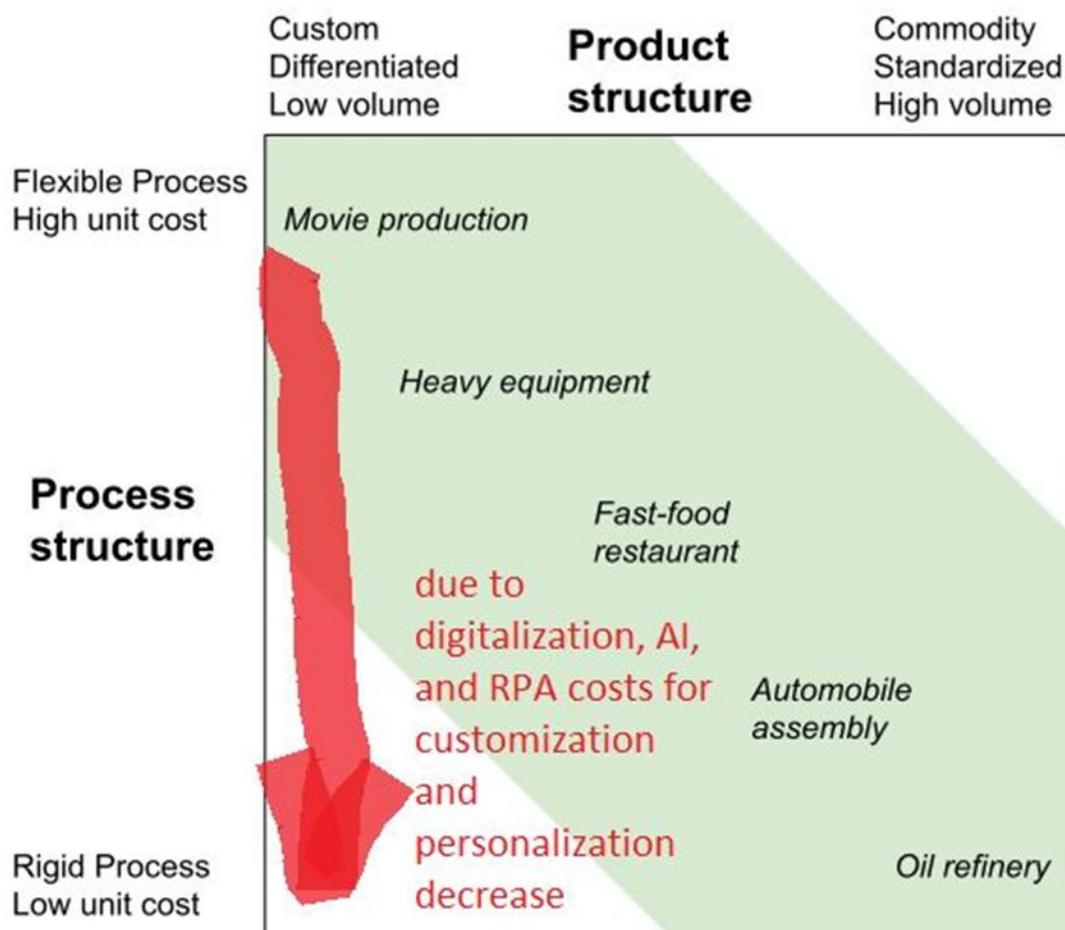
Процес" (Rigid Process), дефиниран от физическата апаратура, към **"Адаптивен Процес"** (Adaptive Process), дефиниран от софтуер.

Табл. 3 Класическа и дигитална матрица на производството

Аспект	Класическа Матрица (Твърд Процес)	Дигитална Парадигма (Адаптивен Процес)
Източник на Гъвкавост	Човешки умения, Неавтоматизиран труд	Софтуерна реконфигурация, AI алгоритми
Единични Разходи за Персонализация	Високи (поради неефективност и ръчен труд)	Ниски (поради автоматизирано управление на сложността)
Фокус на Производството	Стандартизация за максимизиране на обема	Персонализация за максимизиране на клиентската стойност

В резултат на това, бизнеси, които традиционно са били позиционирани в горния ляв квадрант (напр. "Heavy equipment"), могат да се придвижат към по-ниски единични разходи, като същевременно запазват или дори увеличават нивото на персонализация. Това създава нова зона на конкурентно предимство, където **масовото персонализиране (Mass Customization)** вече не е оксиморон, а достижима стратегическа цел.

Фигура 7. Матрица Продукт/Процес



Дигитализацията, AI и RPA действат като екзогенни сили, които променят правилата на играта, описани от Продуктово-Процесната Матрица. Чрез автоматизиране на управлението на сложността, те елиминират традиционната ценова премия, свързана с персонализацията. Компаниите вече могат да комбинират **висока степен на диференциация на продукта с ефективността на процесите и ниски единични разходи**. Този феномен не само преконфигурира производствената среда, но и предоставя възможност за дълбока персонализация на потребителското изживяване, което преди това е било икономически непостижимо. Бъдещите изследвания следва да се фокусират върху разработването на нови модели за операционен мениджмънт, които адекватно отразяват тази нова, **адаптивна и икономически ефективна дигитална производствена парадигма**.

Ролята на AI в трансформацията на разходите

Изкуственият интелект не просто се вписва в тази парадигма, той я **ускорява и радикализира**. AI действа като катализатор за постигане на експоненциална икономия от мащаба.

Експоненциална скалируемост, която е характерна за AI моделите (като **GPT-4** или **DALL·E**) означава, че модела се обучават за обслужват милиони, дори милиарди заявки с почти еднакъв разход за всяка една. Докато първоначалната инвестиция в данни и

изчислителна мощност е огромна, допълнителните разходи за сървъри и поддръжка растат линейно, докато обемът на потребителите и генерираната стойност могат да растат експоненциално.

Радикално намаляване на цените. AI демократизира достъпа до сложни технологии, моделите с отворен код от платформи като **Hugging Face** и инструменти като **TensorFlow** позволяват на малки фирми и стартапи да използват мощни AI решения, без да се налагат крупни разходи и инвестиции, което създава нова конкуренция, при която цените на услугите могат да бъдат рязко намалени, тъй като основният разход е акумулиран в първоначалната разработка.

Нови модели на монетизация. Когато пределната себестойност клони към нула, бизнесът не може да разчита на еднократни продажби. Вместо това възникват и процъфтяват нови бизнес модели - **абонаменти (SaaS - Software as a Service)**, платформи като **ChatGPT Plus** и **Adobe Creative Cloud** предлагат достъп до постоянно обновяващи се AI инструменти срещу месечна такса, което гарантира предвидим приход и стимулира непрекъснатото подобряване на услугата.

Реклама. Гиганти като **Google** и **Meta** използват AI, за да персонализират рекламното съдържание, увеличавайки неговата ефективност. Тъй като разходът за показване на още една реклама е почти нулев, те могат да печелят огромни суми, мащабирайки модела до глобален пазар.

Фриймиум (Freemium). Компании като **Canva** и **Grammarly** предлагат базов безплатен достъп, за да привлекат огромна потребителска база, а след това монетизират малка част от нея чрез платени функции и премиум планове.

Практически казуси и стратегически последици

AI усилва и задълбочава феномена „winner-takes-all“, като създава мощен цикъл на обратна връзка, повече данни водят до по-добър AI модел, по-добрият модел привлича повече потребители, а повечето потребители генерират още повече данни. Този процес е почти невъзможен за копиране от нови играчи.

Amazon Web Services (AWS). Първоначалните инвестиции в изграждането на центрове за данни са крупни, но допълнителните разходи за обслужване на всеки нов клиент са минимални. Това позволи на AWS да доминира пазара на облачни услуги.

Microsoft Copilot, след като е разработен и внедрен в **Office 365**, AI асистента може да обслужва стотици милиони потребители с минимални допълнителни разходи, което формира огромно конкурентно предимство, тъй като разходите за R&D вече са амортизирани върху мащабна потребителска база.

Spotify. Добавянето на още един милион потребители почти не увеличава разходите за стрийминг на музика. В същото време обаче, тези нови потребители предоставят ценни данни, които подобряват препоръчителните алгоритми на платформата, правейки услугата още по-добра и привличайки още повече хора.

Примери от логистиката. Глобална логистична компания инвестира милиони в AI система за оптимизиране на маршрути. Първоначалният разход е висок, но след като системата е внедрена, тя намалява разхода за гориво и време за доставка на всеки един камион. **Пределният разход за всяка следваща доставка намалява значително**, тъй като ефективността на системата расте с повечето данни за трафик и доставки. Това е пример за AI, който трансформира оперативни разходи в почти чиста печалба след достигане на определен мащаб.

Разходите в дигиталната ера не са просто счетоводна категория, те са **стратегически инструмент**. Изкуственият интелект прави пределната себестойност на дигиталните продукти близка до нула и създава условия за експоненциално мащабиране, което не е възможно в традиционната икономика. В комбинация с платформените ефекти, това води до концентрация на пазарна власт и радикално нови бизнес модели.

Мениджърите, които разберат тази динамика, ще могат да изграждат устойчиви бизнеси, които използват мащаба не като ограничение, а като източник на **конкурентно предимство**.

Успехът вече се определя не само от ефективността на производството, а от способността да се управляват първоначалните инвестиции, да се събират и анализират данни и да се създаде самоподсилваща се екосистема.

Теми за дискусия Глава 8

- *Как AI променя разделението фиксирани срещу променливи разходи?*
- *При какви условия пределната себестойност на AI услуги е практически ~ 0 ?*
- *Ускорява ли AI концентрацията (winner-takes-most) чрез комбиниране на мрежови ефекти и намаляваща MC? Какви антимонополни мерки са релевантни?*
- *В традиционни отрасли (автомобили, логистика) до къде AI може да сваля MC преди да удари физически ограничения?*

9. От приходи към стойност. Как AI разкрива нови източници на растеж?

„Стойността не се създава, когато компанията произвежда, а когато клиентът възприема, че е получил повече, отколкото е платил.“

Philip Kotler, Marketing Management (2017) Филип Котлър е професор по международен маркетинг в Kellogg School of Management, считан за „баща на съвременния маркетинг“.

В предходната глава анализирахме как изкуственият интелект променя разходите, създавайки нова икономика на почти **нулева пределна себестойност**. Но за да бъде един бизнес модел устойчив и успешен, той трябва да се фокусира не само върху намаляването на разходите, но и върху генерирането на приходи. В тази глава ще разгледаме в детайли как AI не просто оптимизира операциите, а радикално **преосмисля какво означава да се създава стойност** за потребителите, което от своя страна води до откриване на нови, мощни източници на приходи и устойчиви бизнес модели.

AI като генератор на стойност - хиперперсонализация и клиентско преживяване

Изкуственият интелект създава стойност, като се фокусира върху най-важния елемент от всеки бизнес – клиента. Той позволява на компаниите да преминат от масово предлагане към **хиперперсонализация**, която не просто препоръчва продукти, а предвижда нуждите и желанията на потребителите.

Хиперперсонализация в действие е процес на използване на AI и машинно обучение за анализ на данни в реално време, с цел да се предскажат индивидуалните нужди и да се създаде персонализирано изживяване. **Примери може да намерим в банковите и финансови услуги.** Традиционните банки предлагат стандартизирани продукти, в контраст банка, интегрираща AI, може да анализира финансовите навици на клиента и да му предложи **персонализирани лихвени проценти** по кредит, инвестиционни съвети, съобразени с неговия риск профил, или дори автоматизирано спестяване. Този подход повишава удовлетвореността и доверието, което води до по-висока **готовност за закупуване/плащане (Willingness to Pay, WTP)** и лоялност. **Стрийминг платформи като Netflix** използват сложни алгоритми за препоръки, които са толкова точни, че се превръщат в ключова причина за задържане на потребителите. Според докладите, над 80% от съдържанието, гледано в платформата, е базирано на препоръки от AI. Този персонализиран поток от съдържание намалява **отливите/миграциите на клиенти (churn)** и позволява на компанията да поддържа стабилни и по-високи абонаментни такси. Подобен ефект постига и **Spotify** с персонализирани плейлисти, които създават

уникално и ценно изживяване за всеки слушател, оправдавайки премиум абонамента. **Подобрено клиентско изживяване е една от най-прилаганите функции на AI**, интеграция в различни точки на контакт с клиента, за да подобри удобството, бързината и ефективността на обслужването. **Примери от електронната търговия**, глобални лидери като **Amazon** и **Alibaba** използват препоръчителни системи, задвижвани от AI, които могат да увеличат средната стойност на потребителската кошница с 10-30%. В допълнение, AI-базираните чатботове и виртуални асистенти могат да обработват огромно количество запитвания, да отговарят на въпроси и да решават проблеми 24/7, намалявайки разходите за поддръжка и повишавайки удовлетвореността на клиентите.

Нови източници на приходи, AI-базирани продукти и бизнес модели

Когато AI е в основата на продукта, той не просто оптимизира, а се превръща в източник на директни приходи. Това води до възникването на изцяло нови бизнес модели.

AI като услуга (AI-as-a-Service - AaaS). Фирми, които традиционно са продавали софтуер с еднократен лиценз, сега предлагат AI функционалности на абонаментен принцип. **Пример е Microsoft Copilot**, Microsoft позиционира **Copilot** в Office 365 като премиум допълнение, която оправдава по-висока абонаментна такса. Тази услуга предоставя на потребителите автоматизация на задачи в реално време вт.ч. генериране на имейли, обобщаване на срещи, създаване на презентации. По този начин, продуктът не само намалява разходите за труд, но и носи директен нов приход за компанията.

Монетизация на данни и анализи. Платформите, които събират и анализират данни чрез AI, могат да продават тези обобщени прозрения на други бизнеси. **Пример са социалните медии, Meta (Facebook, Instagram, TikTok)** използват данните на потребителите си, за да създават изключително точни таргетирани реклами. Рекламодателите са готови да платят по-висока цена за достъпа до тази персонализирана аудитория, тъй като това увеличава ефективността на техните кампании. По този начин AI, който оптимизира потребителското изживяване, едновременно с това се превръща в машина за приходи.

Генеративен AI и творчески индустрии, компании като **Synthesia**, която създава видеоклипове с AI-генерирани аватари, предлагат услуга, която значително намалява разходите и времето за производство на съдържание. Подобни инструменти позволяват на бизнеси да създават маркетингови и обучителни видеа с минимални разходи, докато самата AI платформа генерира приходи чрез абонаменти.

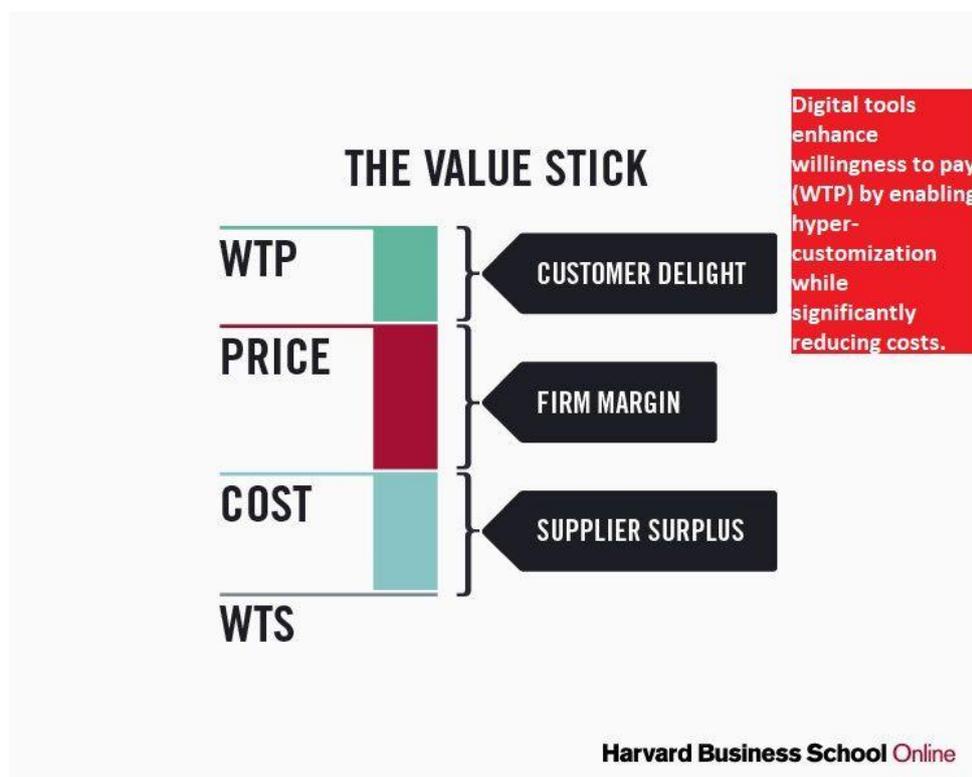
Синергия между разходи и приходи и задвижване на ръста

Най-голямото стратегическо предимство на AI е уникалната синергия, която създава между намаляването на разходите и увеличаването на приходите. Това е т.нар. **двоен ефект (dual effect)**, който осигурява мощно конкурентно предимство. Логистичните компании иплементирали AI за оптимизиране на маршрутите на своите транспортни

средства, разполагат със системата за анализиране на трафика, метеорологичните условия и обема на поръчките, за да създаде най-ефективните маршрути, което амалиява разхода на гориво и време за доставка. Върху приходите позитивни ефекти дават по-бързите и надеждни доставки, тк. се повишават удовлетвореността на клиентите, което ги прави по-склонни да платят за премиум доставка или да се върнат за бъдещи поръчки. В този случай, AI едновременно намалява оперативните разходи и увеличава приходите чрез по-висока стойност на услугата.

Tesla инвестира огромни суми в разработката на софтуер за автономно шофиране. Първоначално това е голям разход за R&D. Но когато софтуерът е завършен, той се продава като **премиум пакет (напр. Full Self-Driving)**, който генерира милиарди долари приходи с почти нулев пределна себестойност. Това позволява на компанията да получава допълнителни приходи от всеки вече продаден автомобил.

Фигура 8. Матрица за формиране на конкурентно предимство/стойност за клиентите



Моделът за формиране на конкурентно предимство/стойност за клиентите, декомпозира създадената стойност на няколко ключови компонента, като по този начин разкрива източниците на конкурентно предимство.

- **WTP (Willingness to Pay) Готовност за плащане на клиента** е максималната сума, която клиентът е готов да плати за даден продукт или услуга. Тя отразява стойността, която клиентът възприема, че получава.

- **PRICE**, цена на продукта/услугата, която фирмата определя. Тя се намира между WTP на клиента и COST на фирмата.
- **COST**, разходи/себестойност - всички разходи, които фирмата прави, за да произведе и достави продукта или услугата.
- **WTS (Willingness to Sell)**, готовност за дистрибуция на продукта от страна на търговци, изразява минималната цена, при която доставчикът на суровини или партньорът е готов да предостави своя ресурс или услуга.

Ключови формули:

- **Създадена стойност** = $WTP - WTS$. Това е общата стойност, генерирана от трансакцията, която трябва да бъде разпределена между всички участници.
- **Потребителски маржин** = $WTP - PRICE$. Стойността, която клиентът получава безплатно. Колкото по-голям е този излишък, толкова по-удовлетворен е клиентът.
- **Марж на фирмата** = $PRICE - COST$. Това е пряката печалба на компанията.
- **Маржин на доставчика** = $COST - WTS$. Стойността, която доставчикът получава над минимално изискуемата си цена.

Четирите подпода за оптимизация

1. **Увеличаване на WTP**, повишаване на възприеманата стойност за клиента.
2. **Намаляване на PRICE**, намаляване на цената за клиентите с цел привличане на по-голям обем.
3. **Намаляване на COST**, повишаване на оперативната ефективност посредством намаление на разходите
4. **Намаляване на WTS**, постигане на по-добри условия с доставчиците.

В традиционната икономика, тези лостове често са в конфликт един с друг. Например, за да увеличиш WTP чрез персонализация, обикновено трябва да поемеш по-високи разходи. AI обаче е катализатор, който позволява на компаниите да работят едновременно върху всички лостове, създавайки мощни синергии.

Как AI променя модела за формиране на конкурентно предимство

1. **Увеличаване на WTP**, в дигиталната ера, WTP се движи от **хиперперсонализацията** и качеството на клиентското преживяване. AI е двигателят на този процес, който използва данни, за да създава продукти и услуги, които се усещат като направени „по мярка“ за всеки потребител.
- **Персонализация в реално време**. AI алгоритми анализират поведението на потребителите в рамките на секунди, за да предложат "**next-best-action**" – следващата най-добър продукт/действие, било то препоръка за продукт, статия за четене или оферта. Например, **eCommerce платформи** използват това, за да покажат продукти, които клиентът е най-вероятно да купи, увеличавайки средната стойност

на поръчката. В банковия сектор, AI анализира трансакциите и предлага персонални оферти за кредит или инвестиции точно в момента, в който клиентът има нужда от тях. Това прави услугата не просто удобна, а незаменима.

- **Качество на преживяването UX (user experience).** Големите езикови модели (LLMs) и генеративният AI (GenAI) създават изцяло ново ниво на потребителска стойност. AI-базираните чат-асистенти на компании като **Klarna** могат да обработят две трети от запитванията за клиентска поддръжка, отговаряйки мигновено и с висок процент на точност. По този начин те намаляват времето за чакане и повишават удовлетвореността, което директно влияе на WTP.
- **Мрежов ефект на данните.** Колкото повече потребители използват даден продукт или услуга, толкова повече данни генерират, което прави AI моделите по-точни и ефективни и създават положителен бизнес кръговрат: повече данни, водят до по-добра персонализация и по-висок WTP, което води до още повече потребители.

2. **Намаляване на себестойността.** Дигиталните продукти имат почти нулева пределна себестойност, което дава на компаниите огромна гъвкавост при ценообразуването.

- **Динамично ценообразуване.** AI анализира в реално време фактори като търсене, конкурентни цени и дори готовността за плащане на конкретен потребител, за да предложи оптимална цена. Например, **авиокомпаниите** и **хотелските вериги** използват сложни AI модели, за да променят цените постоянно, максимизирайки приходите. Това позволява на компанията да продава на ценово-чувствителните клиенти, без да подбива цената за по-платежоспособните.
- **Специализирани версии.** Чрез софтуерни актуализации, компаниите могат да създават различни „версии“ на продукта с почти нулев допълнителен разход. При **Spotify**, например, безплатната версия (субсидирана от реклама) служи за привличане на огромен брой потребители, докато премиум версията, която предоставя допълнителни функции (без реклами, офлайн слушане), осигурява основните приходи.

3. **Намаляване на COST/разходите.** AI позволява автоматизация на рутинни и сложни задачи, което драстично намалява оперативните разходи.

- **Автоматизация на бизнес процеси.** В производствените и логистични компании, AI се използва за **предиктивна поддръжка** на машини. Вместо да се чака повреда, AI прогнозира кога дадена част ще се износи, което намалява разходите за ремонт и времето на престой. В ритейла, AI прогнозира търсенето, оптимизирайки инвентара и намалявайки разходите за складиране.
- **AI в продуктовата разработка,** асистентите за писане на код (като **GitHub Copilot**) ускоряват разработката на софтуер, а GenAI може да генерира **синтетични данни** за тестване на нови продукти, което е много по-евтино от събирането на реални данни. Това намалява разходите за R&D и ускорява времето за пускане на пазара.

4. **Намаляване на WTS**, AI оптимизира отношенията с доставчиците и партньорите, което води до по-ефективно разпределение на стойността.

- **Оптимизация на веригата за доставки.** AI може да анализира данни от доставчици и пазарни трендове, за да предвиди рискове и да договори по-добри цени. Например, в големите търговски платформи, AI може да идентифицира най-ефективните доставчици, които са склонни да продават на по-ниска цена в замяна на достъп до мащабна потребителска база.
- **Маркетплейси.** Двустранни платформи като **Airbnb** и **Uber** използват AI, за да съпоставят търсенето и предлагането. Доставчиците (собственици на жилища, шофьори) са склонни да приемат по-нисък WTS (по-малък приход на единица), защото платформата им предоставя достъп до огромен брой клиенти, инструменти за управление и стабилен поток от поръчки, което намалява техните разходи за маркетинг и придобиване на клиенти.

5. Силата на AI се разкрива напълно, когато се комбинират всички лостове. Дигиталните платформи са идеалният пример за това. Те създават **двустранни мрежови ефекти**: стойността на платформата расте с всеки нов потребител и с всеки нов доставчик (или разработчик, в случая на App Store). Този ефект се подсилва от AI, който превръща данните в двигател на растежа, оптимизирайки WTP, Cost и Price едновременно.

Въпреки огромните ползи, прилагането на този модел носи и рискове. **Алгоритмична пристрастност** може да доведе до дискриминация, а прекалено агресивното **динамично ценообразуване** може да навреди на репутацията на компанията. Важно е мениджърите да комбинират хиперперсонализация, платформен ефект и автоматизация, като постоянно измерват стойността и контролират рисковете. Преди да инвестират в AI, те трябва да си зададат въпроса: "Кой от четирите лоста на Value Stick е най-тесен в момента и как AI може да го разшири?"

Изкуственият интелект променя не само как компаниите управляват разходите, но и как създават, предоставят и улавят стойност. **Хиперперсонализацията**, новите бизнес модели като AaaS и възможността за **премиум ценообразуване** превръщат AI в инструмент за максимизиране на стойността за потребителите и приходите за бизнеса. В комбинация с разгледаните вече ефекти върху разходите (където пределната себестойност клони към нула), AI се превръща в нова основа на конкурентното предимство.

Мениджърите, които разбират и използват тази динамика, ще могат да изградят устойчиви бизнеси, които не просто се борят за пазарен дял, а създават нови пазари и екосистеми. Успехът вече се определя не само от ефективността на производството, а от способността да се управляват първоначалните инвестиции, да се събират и анализират данни и да се изгради самоподсилваща се екосистема от стойност.

Теми за дискусия Глава 9

- *Как AI трансформира концепцията за Willingness to Pay (WTP), кои фактори най-силно влияят на готовността на клиента да плати повече?*
- *Доколко персонализацията може да оправдае премиум цена и кога тя рискува да наруши доверието на клиента (напр. агресивно динамично ценообразуване)?*
- *Как AI може да разкрива латентни нужди на пазара и да създава нови източници на приходи (Tesla, Netflix, Spotify)?*
- *Какви са етичните предизвикателства при монетизация на данните?*
- *Може ли AI да създава стойност и в социален смисъл напр. устойчивост, инклузивност или е фокусиран единствено върху печалбата?*

Модул 3. Икономическият потенциал на AI. От теория към практика

10. AI на работното място. Примери как AI променя индустрията

„Системите с висок риск изискват човешки надзор и прозрачност.“

превод по EU AI Act (2024)

Какви проблеми решава машинното обучение (ML)?

Машинното обучение (ML) е най-активната и приложна област **AI**, който решава широк кръг от проблеми чрез **автоматизирано откриване на закономерности** в данните и използването им за правене на прогнози или вземане на решения без изрично програмиране (Mitchell, 1997). Докато класическото програмиране дефинира изрични правила, ML позволява на системите да се учат директно от емпирични данни, което е в основата на съвременната AI революция. ML се класифицира в три основни парадигми, всяка от които адресира различни типове проблеми:

Обучение под Наблюдение (Supervised Learning).

Този клас проблеми изисква **маркирани данни** (вход с известен изход), което го прави идеален за прогнозиране на дефинирани резултати. **Том Мичъл** (Mitchell, 1997) дефинира, че ML алгоритъмът се учи от опит E по отношение на клас задачи T и метрика на производителност P , ако неговата производителност при задачите T , измерена от P , се подобрява с опита E .

- **Класификация (Classification)**. Прогнозиране на **дискретен етикет** или категория. *Пример:* Медицинска диагностика (Болен/Здрав) или филтриране на спам.
- **Регресия (Regression)** Прогнозиране на **непрекъснатата числова стойност**. *Пример:* Прогнозиране на цени на жилища или продажби.

Обучение без Наблюдение (Unsupervised Learning).

Тук моделът трябва да открие **скрити структури и закономерности** в **немаркирани данни** (Hastie, Tibshirani, & Friedman, 2009).

- **Клъстериране (Clustering)**. Групиране на сходни обекти, което е ключово за **сегментиране на потребители** в маркетинга или за **анализ на пазарни тенденции**.
- **Намаляване на Размерността (Dimensionality Reduction)**. Опростяване на данните, като се запази най-важната информация. Това е жизненоважно за **визуализация** и за премахване на **шум** в масиви от данни.

Обучение с Подсилване (Reinforcement Learning - RL)

RL се занимава с **последователно вземане на решения** в динамична среда (Sutton & Barto, 2018). Агентът се учи чрез **опит и грешка**, като получава **награди** за желано поведение. *Пример* е автономно шофиране или AI, което побеждава хора в сложни игри като Go (Silver et al., 2016). ML е в основата на най-значимите приложни области на AI:

- **Обработка на Естествен Език (NLP)** е фундаментът на **Големите езикови модели (LLM)** и на превода (напр. Google Translate). **Йошуа Бенджо** (Bengio, 2009) и неговият принос към дълбокото обучение (Deep Learning) са ключови за ефективността на тези модели.
- **Компютърно Зрение (Computer Vision)** решава проблеми, свързани с разбиране на визуални данни, като **разпознаване на обекти** и **контрол на качеството** в производството. Постиженията в **конволюционните невронни мрежи (CNNs)**, пионерско дело на **Ян ЛеКун** (LeCun et al., 1998), са централни за тази област.
- **Системи за Препоръки и Откриване на Аномалии:** Прилагането на ML алгоритми за **прогнозиране на потребителски предпочитания** (Системи за Препоръки) и за **откриване на измами** (Откриване на Аномалии) е критично за електронната търговия и финансовата сигурност.

Ползи v/s рискове при използването на ML

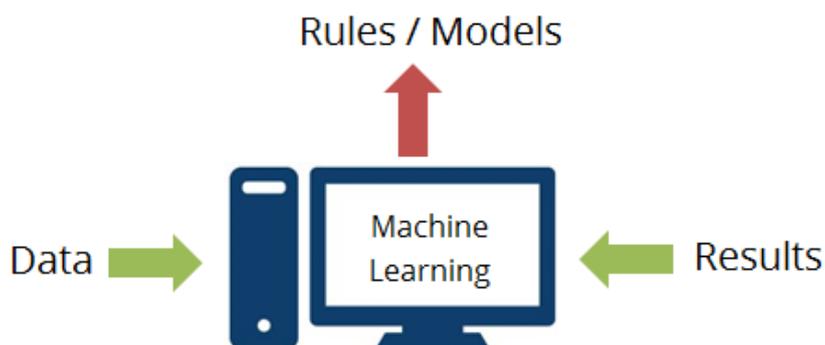
ML е ключов двигател за развитието, защото позволява **машабиране на интелигентността** и **повишаване на оперативната ефективност**. Агравал, Ганс и Голдфарб (Agrawal, Gans, & Goldfarb, 2018) твърдят, че ML е просто „технология за **предсказване**, която намалява цената на предсказването“. Тази демократизация на прогнозирането води до икономически ползи във всяка индустрия, където се вземат решения в условия на несигурност. Чрез ML, организациите могат да анализират данни, които са твърде големи за човешка обработка, да оптимизират процеси в реално време (напр. интелигентни мрежи) и да създават напълно нови продукти и услуги (напр. LLM), които генерират **фундаментални ползи** за обществото. Критиците подчертават етичните рискове, свързани с пристрастията и липсата на прозрачност в ML системите. **Кати О'Нийл** (O'Neil, 2016) предупреждава за "**Математически оръжия за унищожение**", като твърди, че ML моделите могат да „кодифицират дискриминация в голям мащаб“, тъй като те възпроизвеждат историческите пристрастия, съдържащи се в обучителните данни. Рискове са валидни и изискват внимание, въпреки това, **рисковете не отричат фундаменталните ползи**. Вместо да отхвърля ML, академичната общност и индустрията се фокусират върху разработването на **обясними AI (XAI)** методи (напр. Dagra XAI Program), които повишават **прозрачността, отчетността и етичното управление** на ML, за да се гарантира, че ползите надхвърлят потенциалните вреди.

Машинното обучение е **ключово развитие** и **неотменна част от AI**, която има фундаментални ползи, като предоставя инструменти за прогнозиране, автоматизация и

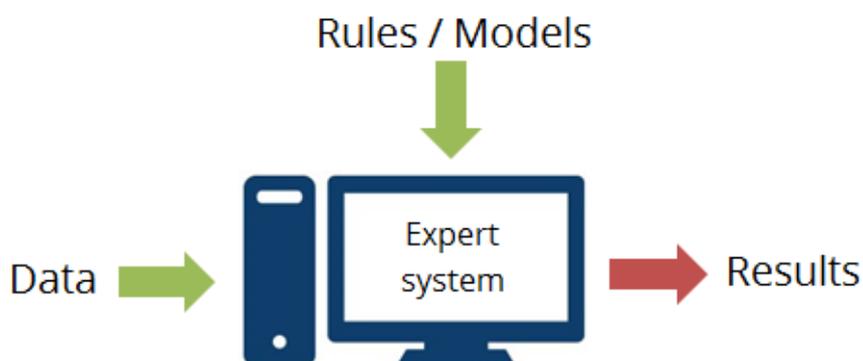
откриване на сложни зависимости в данни. Чрез **Обучение под Наблюдение, Обучение без Наблюдение и Обучение с Подсилване**, ML позволява решаването на проблеми от медицинската диагностика до автономното шофиране, доказвайки се като **фундаментален ползотворен двигател** на икономическата и социалната трансформация.

Машинното обучение е дисциплина, при която системите се учат от данни, а не чрез изрично програмирани правила. Вместо да кажем на компютъра „Ако X, тогава Y“, ние му даваме голям набор от примери с X и съответните Y, и той сам открива закономерностите между тях. Това позволява на системите да станат по-гъвкави, адаптивни и способни да се справят със сложни задачи, които биха били невъзможни за традиционното програмиране. По този начин, **ML е процес на учене от опит**, което е в основата на почти всички съвременни ИИ приложения.

Фигура 9. Модел на работа на машинното обучение (ML)



Фигура 10. Модел за работа на RPA, роботизиране на операции в компаниите



**Източник: Автора*

Да вземем пример за банка, която трябва да реши дали да отпусне заем. В традиционния модел, експертите създават набор от правила „при доходът е над 2000 лв. и кредитната история е добра, заемът се одобрява“. Този подход е ограничен, защото не може да вземе предвид хиляди други фактори, които влияят на риска. При ML, системата анализира милиони примери за минали заеми вт.ч. кои са били изплатени и кои не и др. и сама научава каква комбинация от фактори (доход, семейно положение, възраст, тип работа, дори място на живеене) води до позитивно или негативно развитие. По този начин, тя открива модели и корелации, които човешките експерти не биха могли да видят или да формулират.

Бизнес приложения на ML

Машинното обучение е не просто технология, а **невидим двигател** зад голяма част от модерните бизнес системи. То работи във фонов режим, за да автоматизира сложни задачи, които преди са изисквали хиляди часове човешки труд. От оптимизация на веригата за доставки до персонализиране на съдържанието, ML подобрява ефективността, намалява разходите и увеличава приходите. ML се прилага в практически всеки сектор от икономиката, като променя начина, по който компаниите създават, доставят и улавят стойност. Нека разгледаме няколко примера, които илюстрират неговото стратегическо значение:

- **JPMorgan Chase** използва ML за кредитен скоринг и прогнозиране на риска, като анализира милиарди данни, за да определя кредитоспособността на клиентите с висока точност от традиционните методи. Това не само намалява риска от лоши кредити, но и позволява на банката да предлага персонализирани финансови продукти.
- **Visa Advanced Authorization** е система, която оценява риска от измами в реално време. Тя обработва транзакциите в рамките на секунди, като анализира над 500 променливи — от мястото на транзакцията до вида на стоките — за да разбере дали дадена покупка е легитимна. Според компанията, тази система е предотвратила измами на стойност милиарди долари.
- **Coca-Cola** използва ML за анализ на потребителски сегменти и прогнозиране на пазарни тенденции. Това им позволява да стартират изключително таргетирани рекламни кампании и да оптимизират дистрибуцията на своите продукти, като например да разберат кои напитки се продават най-добре в определени географски райони или в конкретни часове от деня.
- **UPS ORION (On-Road Integrated Optimization and Navigation)** е система, която използва ML за оптимизиране на маршрутите на доставки. Тя анализира над 1000 променливи, включително прогноза за трафика, пътни условия и времеви интервали, за да създаде най-ефективните маршрути. В резултат, UPS спестява милиони литри гориво и милиарди долари годишно.
- **Shell и C3.ai** прилагат т.нар. "**предиктивна поддръжка**" (predictive maintenance). ML моделите анализират данни от сензори, инсталирани на оборудването (налягане, температура, вибрации), за да предскажат кога дадена машина е на път да се повреди.

Това позволява на Shell да извършва поддръжка проактивно, преди да настъпи скъп и опасен срив.

- **Unilever** използва AI за подбор на персонал, като анализира видеа от кандидати, за да оцени личностни черти и комуникационни умения, намалявайки времето за наемане и пристрастията в процеса.
- **LinkedIn** използва ML, за да прогнозира кариерното развитие на своите потребители, препоръчвайки им подходящи умения за учене и професионални пътеки, които биха им помогнали да достигнат следващата стъпка в кариерата си.

Големи езикови модели (LLM – Large Language Model)

Изследването и приложението на **големи езикови модели (LLM)** се превърнаха в централна тема в съвременния изкуствен интелект, бележейки радикална промяна в парадигмата на обработката на естествен език (NLP). Както **Naveed et al. (2023)** описват, LLM представляват **следващ етап в еволюцията** от статистически към невронни и претренирани модели.

Възходът на LLM е пряко свързан с възприемането на **архитектурата Transformer** и мащабирането ѝ до безпрецедентни нива (Zhao et al., 2023). Огромният брой параметри и широкият корпус за обучение позволяват на тези модели да демонстрират феномени като **„zero-shot“** и **„few-shot“ обучение**, които не изискват специфично дообучение за всяка задача (Kalyan et al., 2023). Това е кардинална промяна спрямо предишните ML модели. Този преход демонстрира, че LLM могат ефективно да **„прехвърлят“ (transfer)** наученото между различни езикови задачи – отговаряне на въпроси, резюмиране, дори генериране на код, като по този начин се позиционират като гъвкави, **общи езикови интерфейси** за автоматизация (Kalyan et al., 2023). **Zhao et al. (2023)** обобщават тази тенденция като изтъкват, че мащабът отключва **„emergent abilities“** (възникващи способности). Въпреки това, този мащаб има своя цена, свързана с енергийна консумация и нужда от специализиран хардуер, което поражда въпроси за устойчивостта и достъпността на технологията. Една от най-сериозните слабости, която пряко подкопава тезата за автономна интеграция, е склонността на LLM да генерират информация, която е **фактологически невярна** феномен, известен като **халюцинации**. **Huang et al. (2023)** развиват подробна таксономия на халюцинациите и излагат предизвикателствата, свързани с надеждността на моделите. Тези неточности не са просто грешки, те са **правдоподобно звучащи фикции**, които произтичат от статистическото естество на модела. Тази слабост става особено критична в **чувствителни домейни** като медицина, право и академични изследвания. **Wieczorek (2025)** предупреждава, че некритичното приложение на LLM може да **„укрепи съществуващите научни и институционални предразсъдъци“**, ако генерираното съдържание се приема като факт без верификация. Този анализ подчертава, че докато LLM могат да **ускорят процеса**, те не могат да заместят експертната оценка и фактологическата проверка (Sohail et al., 2025). Следователно, за да се гарантира надеждност, LLM трябва да се интегрират със слоеве за проверка, като например *Retrieval-Augmented Generation* или строг човешки контролен слой.

Фундаменталното ограничение на LLM произтича от въпроса за **истинското „разбиране“**. Критиците твърдят, че моделите само имитират езика, без да притежават дълбоко семантично или концептуално разбиране. **Mitchell et al. (2023)** обсъждат дали тези модели наистина могат да „разбират“ семантика или просто манипулират статистики върху текста. Популярната метафора **„Stochastic Parrot“ (случаен папагал)**, популяризирана от Bender, Gebru, McMillan-Major и Mitchell (2021), подчертава, че LLM прогнозира следващата дума чрез вероятности, а не чрез разбиране на основното значение или логика. Този **дефицит на семантична дълбочина** означава, че LLM са **уязвими към промени в контекста** и могат да генерират текст, който, макар и синтактично верен, е **концептуално неверен** или абсурден. Този извод представлява **фундаментално ограничение**, което пречатства пълната автономия на LLM в домейни, изискващи *каузално разсъждение, критичен анализ* или *етична преценка*.

Дори в специализирани области, където LLM са адаптирани към специфична терминология и корпус, ограниченията остават. Разработването на **домейн-специфични LLM**, като **SaulLM-7B** за правни задачи, демонстрира потенциала за повишаване на капацитета в юридическия анализ (Colombo et al., 2024).

Въпреки това, **Meyer et al. (2023)** предупреждават в контекста на академичните среди, че LLM могат да приемат пристрастия и да представят неверни твърдения като факти, дори когато подпомагат обобщения и търсения в литературата. Следователно, **интеграцията с човешки експерти** е задължителна. **Sohail et al. (2025)** демонстрират, че LLM могат значително да ускорят академичния процес (като обобщаване и търсене), но само при условие на **строг контрол и корекции** от страна на човешкия изследовател. Това показва, че ролята на LLM в тези чувствителни домейни е **асистивна**, а не автономна, което е пряко доказателство за границите, наложени от халюцинациите и липсата на дълбоко разбиране.

Можем да обобщим, че LLM представляват **технологически пробив**, променяйки фундаменталните представи за автоматизация на езика. Те надхвърлят традиционните модели по мащаб, гъвкавост и способност за крос-задачни функции (Zhao et al., 2023; Kalyan et al., 2023). Въпреки тези предимства, **критичните слабости** – склонността към халюцинации (Huang et al., 2023), липсата на истинско семантично разбиране (Mitchell et al., 2023) и рискът от вградени пристрастия (Wieczorek, 2025) поставят **сериозни и непреодолими граници** пред тяхната автономна употреба. В най-чувствителните области, като право и медицина, LLM не могат да функционират без човешки контрол. Тяхната роля трябва да бъде дефинирана като **„асистент“**, интегриран с хибридни архитектури, които включват **механизми за верификация и човешка експертиза** (Sohail et al., 2025).

Големите езикови модели (LLM) представляват революция в начина, по който хората взаимодействат с технологиите. Примери за LLM са ChatGTP, Deepseek, Gemini, Claude, Perplexity и др. LLM са в основата на генеративния ИИ (Generative AI/GenAI), който не просто анализира данни, а създава нови, уникални съдържания. За мениджърите това

означава, че могат да получат автоматично резюме на обемист правен документ, Gen AI да генерира идеи за маркетингови кампании или дори да създаде чернова на код, без да има технически познания. Да разгледаме няколко конкретни примера за приложения на LLM:

- **Klarna** внедрява AI асистент, базиран на LLM, който обработва две трети от всички клиентски запитвания в първия си месец. Той не просто отговаря на стандартни въпроси, но може да разрешава сложни казуси като промени в адреси за доставка и връщане на средства, като значително намалява натовареността на човешките оператори и подобрява скоростта на обслужване.
- **Carrefour** разработва вътрешен асистент, наречен **Hopla**, който помага на служителите с въпроси, свързани с оперативни данни и бизнес процеси, замествайки ръчното търсене в десетки бази данни и документи. Служителите могат да попитат Hopla на естествен език: „Какви са продажбите на млечни продукти в региона на Париж миналата седмица?“ и да получат незабавен, обобщен отговор.
- **PwC** използва **ChatGPT Enterprise** за правен и регулаторен анализ. LLM-ът може да анализира хиляди страници законодателство и да идентифицира ключови рискове или пропуски в договори, спестявайки хиляди часове ръчен труд.
- **Levi's** експериментира с LLM, за да създава рекламни послания и описания на продукти, които са силно персонализирани за различни аудитории, което води до по-висока ангажираност и конверсии.
- **JasperAI** е широко използван инструмент в маркетинг агенциите, който генерира статии, постове за социални медии и имейл кампании на базата на няколко ключови думи и указания.

Синергия между ML и LLM

Истинската сила на изкуствения интелект се разкрива, когато се комбинират ML и LLM в единна стратегическа рамка. **Машинното обучение е най-силен в анализа на структурирани данни** – числа, таблици и категории, **LLM в работата с неструктурирани данни** – текст, глас и изображения. Тяхната синергия позволява на бизнеса да извлече прозрения от всякакъв тип информация, като превръща неструктурираните данни в разбираеми, приложими и от стратегическа важност.

Да разгледаме примера за компания за електронна търговия. Тя може да използва ML, за да анализира продажбени данни (количества, цени, сезонност) и да прогнозира бъдещото търсене. В същото време, тя може да приложи LLM за анализ на хиляди клиентски отзиви, чатботове и имейли, за да разбере **защо** хората харесват или не харесват даден продукт. Чрез комбинирането на тези два потока от информация, компанията може не просто да знае **какво** се е продало, а и **защо** се е продало. На базата на тези комбинирани прозрения, LLM може да генерира автоматично персонализирани предложения за продукти и маркетингови послания.

В Маркетинга ML анализира клиентски данни (история на покупки, кликания, демография), за да сегментира аудиторията и да прогнозира кои продукти ще бъдат най-търсени. **LLM** автоматично генерира персонализирани имейл кампании и рекламни послания, адаптирани за всеки сегмент. Резултатът е хипер-персонализация в мащаб, която е била невъзможна преди. **Във Финансите ML** моделите анализират финансовите пазари и да прогнозират цени на акции или риск от неизплащане на кредити. **LLM** след това анализират тези сложни прогнози на клиентите с разбираем език, да генерират автоматични отчети за инвестиционни портфолия или да отговарят на въпроси на клиентите относно техните сметки, без да е необходима намесата на човешки експерт. **В Човешките ресурси ML** анализира данни от HR системи (производителност, време в компанията, умения) и откриват служители с потенциал за развитие. **LLM** след това съставят персонализирани кариерни пътеки, да генерира обучителни материали или да състави препоръки за повишение, базирани на данните, анализирани от ML модела. В операции компаниите използва **ML** за анализ на данни от сензори в производствената си линия, за да предвидят периода, в който една машина ще се повреди. В същото време, **LLM** анализира хиляди ръководства и технически спецификации, за да генерира инструкции стъпка по стъпка за техническите екипи, спестявайки време и предотвратявайки големи прекъсвания на производството.

Ограничения, рискове и етични предизвикателства

Предизвикателствата при LLM и ML не са само технически, те имат дълбоки **етични и правни последици**. Един от най-сериозните рискове е този от **пристрастие (Bias)**. **Md. Nazmus Sakib et al. (2024)** дефинират, че „Пристрастията в социален и екологичен план“ произлизат от **пристрастни данни за обучение и архитектурата на модела**. Този риск е значителен, тъй като може да доведе до дискриминационни резултати, уронване на репутацията и правни проблеми (Sakib et al., 2024). Другият критичен оперативен риск е **"халюцинацията"**, или генерирането на неверни данни. Експертите по MLOps и управление препоръчват въвеждането на **"Human-in-the-Loop" (HITL)** системи и **непрекъснат мониторинг на ефективността (Performance Monitoring)** като основни стратегии за управление на този риск (Trigyn Technologies, 2024). Това означава, че за високорискови приложения, мениджърите трябва да включат човешка проверка на изхода на LLM, което е пряко свързано с **оперативната стабилност и надеждността** на системата (Trigyn Technologies, 2024). В допълнение, **рискът за сигурността и поверителността на данните** е засилен от природата на LLM. **Richard Rieben (2024)** отбелязва, че LLM създават „разширени повърхности за атака“ и съществува „потенциал за неволни изтичания и пробиви на данни“. За да се смекчи това, мениджърите трябва да прилагат стратегии като **минимизиране на данните и анонимизиране**, както и да осигурят спазването на **GDPR** и други разпоредби (Rieben, 2024; Trigyn Technologies, 2024).

Внедряването на AI Governance не е просто въпрос на избягване на глоби, а стратегически императив за изграждане на **доверие и отчетност**. **Credo AI (2022)** дефинира AI Governance като процес, който „събира разнообразна гама от

заинтересовани страни, за да вземат решения относно това как AI системите трябва да бъдат разработени и използвани отговорно, за да се гарантира съответствие на системата с бизнес, регулаторните и етичните изисквания“. **Sophia Rowland** и **David Weik** (2024) подчертават, че за да се получи възвръщаемост на инвестициите (ROI) от AI, организациите трябва да могат да „използват своите аналитични активи и модели, за да вземат по-добри решения“. Това обаче е невъзможно без **доверие** в резултатите на модела, което се постига чрез прилагане на **AI Governance**, което гарантира **прозрачност** и **обяснимост** през целия жизнен цикъл (Rowland & Weik, 2024). В крайна сметка, **управлението (Governance)** дава възможност на техническите екипи да получат „критични приноси за техническите решения и действия, които предприемат по време на MLOps жизнения цикъл“, осигурявайки съответствие на системите с **правните, етичните и бизнес изискванията** (Credo AI, 2022).

Качество и пристрастия (Bias) в данните. Правилото е, че моделите са толкова добри, колкото данните, върху които са обучени, ако данните съдържат исторически пристрастия например, в данните за наемане на служители преобладават мъже моделът ще се научи да възпроизвежда и дори да усилва тези пристрастия. Най-известният пример е с автоматизираната система за наемане на служители на **Amazon**, която е трябвало да бъде спряна, защото е показвала дискриминация срещу кандидатки.

„Халюцинации“ на LLM. LLM са обучени да генерират правдоподобен текст, но нямат разбиране за истината или реалността. Те могат да „халюцинират“ да генерират достоверно звучаща, но напълно невярна или измислена информация. В бизнес контекст това е изключително опасно, тъй като може да доведе до грешни финансови прогнози, правни грешки или зле информирани стратегически решения.

Етични рискове и прозрачност. Една от най-големите етични дилеми е т.нар. „проблем на черната кутия“ (black box problem). Често е трудно да се разбере защо един сложен ML модел е взел определено решение. В сектори като медицината или правосъдието, където решенията на ИИ имат сериозни последици, липсата на прозрачност може да бъде неприемлива. Отговорността е друг ключов въпрос - Кой е отговорен, когато самоуправляваща се кола, задвижвана от ИИ, причини инцидент?

Регулаторни предизвикателства. Правителствата по света реагират на бързото развитие на ИИ с нови регулации. **EU AI Act (2024)** е първият в света всеобхватен закон за изкуствения интелект, който категоризира системите според риска, който представляват. Той налага строги изисквания за прозрачност, надзор от страна на човека и оценка на риска за високорискови системи (като тези в медицината, транспорта и финансите). Това означава, че компаниите трябва да инвестират не само в технология, но и в механизми за управление на риска и съответствие с регулациите.

Машинното обучение и големите езикови модели като част от ИИ са **фундаментални двигатели на бизнес иновациите**. За да успеят в тази нова ера, мениджърите и предприемачите трябва да се научат да мислят стратегически за ИИ, като го разглеждат

като основен актив, а не просто като инструмент. Комбинирането на ML за анализ на данни и LLM за комуникация и генериране на съдържание отключва нови възможности за персонализация, автоматизация и оптимизация, които могат да донесат значително конкурентно предимство. Основният въпрос пред индустрията е не е дали да се използва ИИ, а как да се внедри **отговорно и етично**, като се управляват рисковете, свързани с пристрастията в данните и регулациите. Успешните компании на бъдещето ще бъдат тези, които инвестират не само в технологии, но и в хора, процеси и култура, които могат да използват пълния потенциал на ИИ за създаване на устойчива стойност.

Теми за дискусия Глава 10

- *Кои бизнес функции според вас ще бъдат най-силно трансформирани от AI до 2030 г.?*
- *Къде поставяте „човека в AI цикъла“ и кога е допустима по-висока автоматизация?*
- *Как управлявате данни: качество, защита, минимизация, права за ползване?*
- *Кои роли/умения променя AI във вашия екип (prompting, MLOps, AI governance)?*
- *Как съчетавате автоматизация (намалени COST) с ръст на стойността за клиента (повишен WTP)?*

11. AI като стратегически ресурс. Как да превърнете технологията в конкурентно предимство

“Стойността от AI идва, когато промените процеса, а не когато просто добавите инструмент.”

Thomas H. Davenport & Rajeev Ronanki, Harvard Business Review (2018)
Том Дейвънпорт е водещ изследовател и преподавател в областта на аналитика, AI и операционна ефективност; Rajeev Ronanki е специалист по цифрова трансформация и ръководител на AI инициативи в корпоративни среди.

В миналото, технологиите се възприемаха предимно като инструменти за поддръжка, които подобряваха съществуващите процеси. Днес, с напредъка на AI, тази парадигма се променя коренно. AI вече не е само допълнение, а **фундаментален двигател за иновации**, който може да създаде изцяло нови индустрии и да генерира стойност по начини, които преди бяха невъзможни. Ако електричеството е дало началото на масовото производство, то AI е двигателят на „икономиката на знанието“. Той дава възможност за анализ на огромни обеми от данни, идентифициране на сложни зависимости и вземане на решения със скорост и точност, далеч надхвърлящи човешките способности.

Характеристиките, които ни дават основание да класифицираме AI като стратегически ресурс са:

- **Скалируемост.** Веднъж обучен, йС модел може да бъде мащабиран да обслужва милиони потребители или операции без пропорционално увеличаване на разходите. Това е ключово за бързия растеж и глобалната експанзия.
- **Универсалност.** AI е приложим в почти всеки сектор. От финансови услуги до здравеопазване и производство, той подобрява ефективността, персонализира услугите и създава нови продукти.
- **Данните са новоjф „гориво/злато“.** Компаниите, които ефективно събират, управляват и анализират големи масиви от данни, придобиват значително предимство. Данните захранват моделите на AI и им позволяват да се учат и подобряват непрекъснато.
- **Трансформиране на работната сила.** Според McKinsey, AI има потенциала да автоматизира между 60% и 70% от дейностите, които служителите извършват днес. Това освобождава човешкия ресурс за по-сложни, стратегически и творчески задачи, което води до увеличаване на производителността и иновациите.

AI в корпоративната стратегия

Емпиричните доказателства и концептуалните рамки категорично сочат, че **Интегрирането на AI в корпоративната стратегия е устойчив източник на конкурентно предимство само ако фирмата едновременно:** (i) реорганизира операционната си система около данни и алгоритми, (ii) изгражда организационни способности за учене и промяна на работните потоци, и (iii) установи формална рамка за управление на риска и регулаторно съответствие. Преминаването отвъд изолирани пилотни проекти към фирма, „*проектирана за AI*“, изисква дълбок синхрон между стратегическата визия, архитектурата на данните, културата на организационно учене и, което е критично, управлението на риска. Изследванията показват, че мащабната стойност възниква не при тясно функционални употреби, а при **преосноваване на фирмата като „операционна система от софтуер и данни“** (Iansiti & Lakhani, 2020). Неуспехът да се постигне тази конвергенция води до акумулиране на технически дълг, регулаторни рискове и феномена, известен като „театър на иновациите“ (Davenport & Ronanki, 2018).

Внедряването на AI изисква избор на подход, който балансира бързата възвръщаемост и дългосрочната трансформация. От една страна, **портфейлният и поетапен подход** (Davenport & Ronanki, 2018) препоръчва прагматично начало. Авторите съветват да не се започва със сложни задачи, а с „**приоритизиран портфейл от проекти според бизнес нуждите и плановете за скалиране**“ (Davenport & Ronanki, 2018). Тази стратегия цели бързо материализиране на стойност, като AI се вгражда в конкретни, измерими процеси. Въпреки, че е ефективен за контрол на първоначалния риск, критиците предупреждават, че този подход често води до **фрагментация и нисък стратегически кохерент**, създавайки „зоопарк“ от модели без обща архитектура (напр. общи фийчър стори или унифициран MLOps). От друга страна, **Ианзити и Лакхани (Iansiti & Lakhani, 2020)** концептуализират „**фирмата като алгоритъм**“, което представлява трансформация на целия бизнес модел, а не просто проект по автоматизация. Този модел премахва класическите ограничения за мащаб и обхват, преосновавайки оперативната логика върху данни и ИИ. Този подход, макар и силен във визията си за **икономика на мащаба**, е рисков поради **капиталоемкостта** си и зависимостта от висока **зрялост на данните** и цялостна **платформена архитектура**. Без тази зрялост, „алгоритмичната фирма“ може да деградира до скъпи, но локални оптимуми.

Устойчивото конкурентно предимство от AI се генерира не от самите модели, а от способността на организацията да **променя оперативните си потоци** и да **извлича организационно учене** от употребата на AI. Лонгитудиналните изследвания на MIT SMR–BCG (2022–2024) показват, че бизнес стойността расте експоненциално, когато организацията съчетава **индивидуална полза за работниците** с „**учене на AI**“ – процес, при който резултатите от моделите се използват за пренастройка на процесите и управление на несигурността. Последните глобални проучвания на **McKinsey & Company (2024; 2025)** потвърждават рязък ръст на употребата на генеративен AI и **пренастройка на работните потоци**, което пряко корелира с измерими резултати

(т.нар. „долна линия“) при фирмите, които променят *процеси*, а не само *инструменти*. Това означава, че стратегическата цел не е високите нива на експериментиране, а **затваряне на цикъла учене – операционализация – стойност**. Компаниите трябва да създадат механизъм, който превръща индивидуалната полза в **организационни подобрения** (MIT SMR–BCG, 2022–2024).

Третия задължителен фундамент за устойчивост е формалната рамка за управление на риска, тъй като рисковете от **надеждност, пристрастия и регулаторни глоби** могат бързо да унищожат всяко конкурентно предимство.

Рисковете от **фактически неверни отговори** (халюцинации) и **дрейф на моделите** налагат задължителни технически контроли. **NIST AI Risk Management Framework (AI RMF 1.0)** дефинира функционални стъпки: *Govern–Map–Measure–Manage*, подчертавайки необходимостта от **доверие, отчетност и контекстна оценка на въздействията** (NIST, 2023). **Gartner AI TRiSM** допълва тази рамка с конкретни практики за **инвентаризация на моделите, класификация на данните, непрекъсната валидация и споделена отговорност** в ИТ-операциите (Gartner, 2023–2024). Корпоративното внедряване, следователно, трябва да включва **автоматизирани контроли** и **"policy-as-code"** (политика като код) за изпълнение на тези рамки.

Регулаторна конвергенция

Регулаторната среда трансформира „trustworthy AI“ от етичен принцип във **финансово обвързващо изискване**. **Регламент (ЕС) 2024/1689 (AI Act)**, публикуван на 12.07.2024 г., въвежда класификация на системите по риск и специфични задължения за **високорисковите системи** (European Union, 2024). Това налага на корпорациите да картографират своите ИИ приложения спрямо изискванията на AI Act (напр. **оценка на съответствието, логирание и човешки надзор**) и да ги операционализират като „**release gates**“ преди внедряване.

За финансовия сектор, **Bank for International Settlements (BIS, 2025)** препоръчва **адаптивни модели на управление** на приемането на ИИ, тъй като традиционните рамки за риск са недостатъчни за динамичното поведение на алгоритмите. Този регулаторен фокус превръща **управлението на риска в стратегически актив**, а не в бюрократично бреме.

За постигане на устойчиво предимство се предлага интегрирана рамка, която синтезира най-добрите практики:

Стратегическа Визия (Strategy):

- Дефиниране на „**стойност**“ и икономика на казусите (Davenport & Ronanki, 2018).
- Избор на домейни с **висока честота на решенията** и **добро покритие на данни**.

- Целевият операционен модел (ТОМ) трябва да е „фирма като алгоритъм“ (Iansiti & Lakhani, 2020), но с поэтапна миграция.

Операционна Платформа (Ops):

- Изграждане на унифициран **Платформен Слой**, включващ **фийчър стор, MLOps, управление на промптове и непрекъснато наблюдение на продукцията**.
- **Редизайн на работните потоци** с дефинирани роли за „човек-в-контрола“ (Human-in-the-Loop) и офлайн/онлайн валидации (McKinsey & Company, 2024).

Риск и Съответствие (Risk & Compliance):

- Операционализиране на **NIST AI RMF 1.0** и **Gartner AI TRiSM** като **контролни цели**.
- Картографиране към **EU AI Act** (класификация по риск) и вътрешни „**release gates**“, които автоматично проверяват съответствието преди внедряване (European Union, 2024; NIST, 2023).

Цикъл на Учене (Learning Loop):

- Фиксирани цикли за **ретроспекция и пренастройка на процесите**, които свързват **индивидуалната полза** на работниците с **организационни KPI и бизнес резултати** (MIT SMR-BCG, 2022–2024).

Интегрирането на AI в корпоративната стратегия означава преминаване отвъд единични, пилотни проекти към дълбоко внедряване във всички ключови бизнес процеси. Това е процес, който изисква не само технологични инвестиции, но и лидерски ангажимент и културна промяна. Лидерите трябва да отговорят на три основни въпроса: **(1)** Къде AI създава най-голяма стойност за нашата компания? **(2)** Как да изградим вътрешните си способности за успешно внедряване? **(3)** Как да управляваме етичните и регулаторни рискове? Правилната стратегия е ключова, за да се гарантира, че ИИ не е просто "играчка", а истински двигател на бизнес растежа.

Два основни стратегически модела:

- **AI-First стратегия.** Компании, които поставят AI в самия център на своя бизнес модел и продуктово предложение. Те не просто добавят AI, а изграждат бизнеса си около него, което често води до създаването на напълно нови продукти и услуги. Техният успех зависи от способността им да събират, обработват и използват мащабни масиви от данни.
- **AI-Enabled стратегия.** Подход, при който традиционни бизнеси използват AI като средство за оптимизация на съществуващите си процеси и продукти. AI не е основната функция на бизнеса, а по-скоро **катализатор за повишена ефективност** и подобро клиентско преживяване.

Примерите от практика, които ни дават основание да дефинираме на двете стратегии са свързани с **Google**, която е типичен пример за **AI-First** компания. AI е в основата на почти всички продукти от търсачката до Google Assistant и рекламите. От друга страна **DBS Bank (Сингапур)** интегрира **AI-Enabled** стратегия. Банката използва AI в процесите си за кредитен скоринг, управление на риска и обслужване на клиенти чрез чатботове. Резултатът е значително намаляване на разходите и повишаване на удовлетвореността на клиентите. **DHL също** използва AI-enabled стратегия за оптимизация на маршрутите, управлението на складовете и прогнозиране на търсенето, което намалява оперативните разходи и подобрява точността на доставките.

Модели за създаване на стойност с AI

AI позволява създаването на **изцяло нови бизнес модели**, които монетизират данни и интелигентни услуги. Например, компании като Uber и Airbnb не биха били възможни без сложни алгоритми за съвпадение на търсенето и предлагането. Създаването на стойност с ИИ може да бъде разпределено в няколко основни модела, като компаниите често комбинират няколко от тях в своите стратегии. Моделите за създаване на стойност може да обобщим в следната класификация:

- **Оперативна ефективност**, най-разпространеният модел. Включва автоматизация на повтарящи се задачи, предиктивна поддръжка на оборудването и оптимизация на веригите за доставки. Пример за това е **производителят на тежко оборудване, който използва AI, за да предсказва кога машините ще се повредят**, което намалява времето за престой и разходите за ремонт.
- **Персонализация и клиентско изживяване**, AI анализира потребителското поведение и предпочитания, за да предложи маирани продукти, услуги и маркетинг. Това води до повишена лоялност на клиентите и по-високи нива на конверсия.
- **Иновация**, AI позволява разработването на нови продукти и услуги, базирани на данни. Това включва интелигентни асистенти, диагностични системи в здравеопазването и финансови продукти, които прогнозират пазарни движения.
- **Монетизация на данни**, компаниите могат да създават стойност, като продават или лицензират своите данни и обучени AI модели на други бизнеси. Този модел е характерен за платформите, които събират данни от множество потребители, като социалните медии или финансовите институции.

Организационни трансформации

Внедряването на AI в бизнес процесите на компаниите не е просто поредния технологичен проект, а **цялостна организационна промяна**. За да бъде успешен проект за ИИ имплементация, е необходимо да се изградят нови структури, да се развият нови умения и да се създаде култура, която насърчава експериментирането и ученето. Често най-голямото предизвикателство е не самата технология, а съпротивата на служителите, липсата на обучени кадри и неадекватните процеси.

Основните организационни трансформации в следствие на имплементиране на AI в процесите на компаниите, може да обобщим в следния ред:

- **Развитие на нови умения и обучение.** AI променя характера на работата, като автоматизира рутинните задачи и създава нови, по-сложни роли. Компаниите трябва да инвестират в **повишаване на квалификацията (upskilling)** на служителите, за да могат да работят с интелигентните системи.
- **Нови структури за управление на данни.** За да функционира AI, е необходим постоянен достъп до висококачествени данни. Това изисква изграждането на нови екипи и процеси за събиране, съхранение и управление на данни, като се гарантира тяхната сигурност и качество.
- **Култура на експериментиране.** Внедряването на AI е итеративен процес, което предопределя необходимостта от култура, която насърчава бързите експерименти, ученето от грешките и постоянното адаптиране на процесите.

Разполагаме с не един успешен пример от практиката, **пример е Microsoft, който въвежда вътрешни AI академии за обучение на своите служители**, за да гарантира, че всички са в крак с най-новите технологии и могат да ги прилагат в работата си. **Mereka (консултантска компания)** провежда 10-седмичен обучителен курс за служителите си, наречен „You & AI“, за да преодолее културната съпротива и да ги насърчи да експериментират с AI инструменти като ChatGPT и Midjourney.

Изкуственият интелект е стратегически актив с огромен потенциал за създаване на стойност, но само ако бъде управляван отговорно. Успешната интеграция на AI изисква от компаниите да развият не само своите технологични способности, но и **организационни умения, култура и етични рамки**. Тези, които успеят да се справят с тези предизвикателства, ще бъдат лидерите на бъдещата икономика.

Теми за дискусия Глава 11

- *Дали ИИ трябва да бъде в центъра на всяка корпоративна стратегия или е само средство за оптимизация?*
- *Кои индустрии в България са най-подходящи за „AI-First“ стратегии, като се има предвид спецификите на пазара?*
- *Как ще адресирате локалния контекст, недостиг на данни на български, регулации, екосистема от доставчици?*
- *Къде е границата на автоматизацията, кои решения остават с за хората и по какъв критерий ги определяме?*

12. Организационна революция. Как да управлявате риска и промяната, когато внедрявате AI

„Културата изяжда стратегията за закуска.“

Питър Дракър (*Peter F. Drucker*)

Организационна трансформация и културна адаптация

Внедряването на AI налага дълбока преоценка на организационната структура и култура. Традиционните, йерархични модели на управление, базирани на командване и контрол, са неефективни в среда, където решенията се вземат от алгоритми в реално време, а иновациите изискват постоянно експериментиране. Организациите, които се стремят да внедрят AI, трябва да приемат култура, ориентирана към данни, където експериментирането, ученето от провалите и бързата адаптация са норма. Това изисква промяна на мисленето отвъд простото автоматизиране на задачи към преосмисляне на целите и на това как се създава стойност.

Още по-съществено е да се преодолее разрывът между визията на ръководството и реалните способности на служителите. Често, докато топ мениджърът инвестира в авангардни технологии, служителите от оперативните звена нямат нужните умения и разбиране, за да ги използват пълноценно. Без адекватни програми за обучение и повишаване на квалификацията, технологиите остават нереализиран потенциал. Културната трансформация е по-бавен и сложен процес от технологичното внедряване, но без нея, последното е обречено на ниска възвръщаемост на инвестициите. Дори успешни компании като Netflix, която е пионер в използването на AI, признават, че бизнес моделът и организационната култура са еволюирали, за да подкрепят технологията, а не обратното.

Култура на „Fail Fast, Learn Fast“. Практиката показва, че AI моделите се подобряват чрез итерация и грешки. Организациите трябва да поощряват експериментите и да приемат провалите като ценен източник на данни за по-добри бъдещи резултати, което е съществена трансформация на начина, по който функционират значителна част от бизнеса.

Нови лидерски умения, лидерите на първо място трябва да се трансформират от йерархични линейни мениджъри в катализатори на промяната, които насърчават сътрудничеството между различни отдели (технологични, правни, маркетингови) и разбират как да интегрират AI в основните бизнес процеси, което изисква значителна промяна в ключовите показатели за оценка на екипите и мениджърите. **DBS Bank** не се фокусира само върху технологичното внедряване на над 1500 модела за машинно обучение, но паралелно инициира мащабна програма за „AI грамотност“ за всичките си

27 000 служители. Това действие цели да осигури плавен преход, тъй като бизнес звената не са били родготвени за разбиране как да работят с технологичните екипи, минимизирайки съпротивата и повишавайки възприемането.

Нови организационни роли и управление на промяната

Внедряването на AI налага създаването на нови роли и структури, които да осигурят стратегическо ръководство и етичен контрол. Вместо просто да се разчита на ИТ отдела, организациите създават специализирани звена, които отговарят за интеграцията и управлението на технологията на високо ниво. Тези промени са съпътствани от управление на промяната (Change Management), което е ключово за преодоляване на съпротивата и страховете на служителите, свързани с автоматизацията. Страхът от загуба на работни места е един от най-големите проблеми. Важно е да се комуникира ясно, че AI често не замества изцяло хората, а по-скоро **аугментира** (подобрява) техните способности, което освобождава служителите от повтарящи се и рутинни задачи, за да могат да се фокусират върху по-сложни, креативни и стратегически дейности. Компаниите, които са успешни в тази трансформация, изграждат платформи за „ко-еволюция“ между екипите и алгоритмите, като позволяват на служителите да дават обратна връзка и да подобряват AI системите. **ADKAR модел** е доказан подход за управление на промяната, който се фокусира върху пет ключови елемента. **Awareness** (осъзнаване), **Desire** (желание), **Knowledge** (знания), **Ability** (умения) и **Reinforcement** (подкрепа).

Управление на риска, етика и регулации

Докато потенциалните ползи от AI са огромни, рисковете не могат да бъдат подценявани. Управлението на риска е ключово, като обхваща не само киберсигурността, но и етичните и регулаторни аспекти.

- **Данни и киберсигурност.** Един от най-големите рискове е сигурността на данните, които захранват моделите. Вредите могат да дойдат не само от изтичане на данни, но и от т.нар. „data poisoning“, при което враждебни входове манипулират модела. **Пример е тази посока е Microsoft Tay** - експериментален чатбот, който неумишлено е оставен „отровен“ от злонамерени потребители и започва да генерира расистки и обидни твитове за по-малко от 24 часа. Друг съвременен пример е с компанията **Synthesia**, която се фокусира върху генерирането на синтетични медии и видео с AI аватари, но трябва да прилага стриктни протоколи за модерация, за да предотврати злоупотреби като създаването на „deepfakes“ за политически цели или други манипулации.
- **Регулаторни рамки.** Законодателството в областта се развива бързо. **EU AI Act (2024)** е ключов регулаторен акт, който класифицира системите по риск. Високорисковите системи, като тези за кредитен скоринг, подбор на персонал (HR) или в здравеопазването, подлежат на строги изисквания за прозрачност, надзор и сигурност.

- **Етика и прозрачност**, AI моделите често са характеризирани като „черни кутии“, които дават резултати, без да предоставят обяснение как са стигнали до тях, което създава проблеми при вземането на критични решения (например в медицината или правото). **Explainable AI (XAI)** е направление, което цели да направи алгоритмите разбираеми. **Visa** и **American Express** внедряват ИИ за откриване на измами, но комбинират с човешки надзор, за да се гарантира, че не се блокират легитимни транзакции.

Стратегически избори: Build, Buy или Partner

Изборът на стратегия за внедряване на ИИ е критично решение, което определя дългосрочната траектория на организацията. Няма универсално решение изборът зависи от ресурсите, целите и толерантността към риска на всяка организация. Важно е да отбележим, че стратегическият избор за имплементиране на AI, е често стратегически въпрос от компетенциите на широк кръг отдели, освен на IT и технологичните звена, бизнеса, операциите, риска и тн. на практика всички звена следва да адаптират стратегиите при приемане на новата стратегическа посока.

Основните методи за имплементиране на AI в организациите може да обобщим, в следния ред:

- **Build (Създаване)**
 - **Характеристики** - организацията създава собствен модел от нулата, което изисква значителни вътрешни и външни ресурси, таланти, време, капитал, системи за сигурност, центрове за данни и др. Характерен е само за глобални индустриални лидери, които дългосрочно могат да инвестират в развитието и поддръжката на собствени модели.
 - **Предимствата са** пълен контрол върху данните, алгоритмите и интеграцията, е най-важното възможност за създаване на уникално конкурентно предимство.
 - **Минусите на модела, които следва да се анализират са** крупни инвестиции, дълъг цикъл на разработка, риск от текучество на ключови експерти.
 - **Пример за прилагане на стратегията за вътрешно за организацията създаване на AI е JPMorgan COiN (Contract Intelligence)** - използва вътрешно разработен AI за анализ на правни договори, което спестява стотици хиляди часове годишно.
- **Buy (Купуване)**
 - **Характеристики** - внедряване на готово решение от външен доставчик (SaaS - Software as a Service).
 - **Предимствата са** по-бързо внедряване, по-ниски първоначални разходи, достъп до експертизата на доставчика, проверени модели и възможност за избор м-у различни доставчици.

- **Най съществените недостатъци са** ограничен контрол върху данните и модела, пълна зависимост от доставчика, потенциални проблеми с интеграцията и постоянна необходимост от актуализация на вътрешен софтуер и хардуер, който да отговоря на изискванията на доставчика.
- **Пример за тази стратегия е Klarna** използвайки готови AI решения, за да намали времето за обслужване на клиенти от дни до минути, демонстрирайки висока ефективност и бърза възвръщаемост.
- **Partner (Партньорство):**
 - Сътрудничество с външна фирма или стартап за разработка и внедряване на специфично AI решение, модела е ое известен като Open Innovation и е свързан със създаване на общо предприятие с партньорска организация, което е с фокус създаване на AI модел, който отговоря на бизнес нуждите на участващите организации.
 - **Предимствата на подхода са свързани със** споделяне на риска и разходите, достъп до специализирани знания и технологии, възможност за продажба на AI модела под формата на As-a-Service SaaS.
 - Рисковете са свързани със стратегическа зависимост и ограничен контрол върху процесите върху процесите.
 - **Примери за подобен тип партньорство са Shell + C3.ai**, двете компании си партнират за внедряване на ИИ решения, които оптимизират операциите в енергийния сектор.

Примери за неуспешни практики при имплементиране на AI

Изследването на неуспехите в областта на изкуствения интелект е от съществено значение, защото предоставя ценни уроци и помага за идентифицирането на **ранните симптоми**, които водят до провал. Вместо да се фокусираме само върху успешните случаи, анализът на провалите ни позволява да разберем по-добре рисковете, капаните и основните предизвикателства при внедряването на AI.

Познаването на **симптоматиката на неуспеха** – като нереалистични очаквания, некачествени данни, липса на етичен контрол или неправилно дефиниран проблем, дава възможност на компаниите да реагират навреме, помага им за изграждането на по-устойчиви, етични и ефективни AI модели, и като резултат компаниите избягват огромни финансови загуби и запазват доверието на потребителите.

- **Неправилно дефиниране на проблем (IBM Watson Health)**, вместо да започне с конкретна, добре дефинирана и лимитирана постижима задача, IBM се опитва да реши цялостен и изключително комплексен проблем с медицинската диагностика, което резултатира в модел, който не можеше да се справи с многообразието на клиничните данни, което в крайна сметка провал на проекта.
- **Липса на качествени данни (Zillow Offers)**, алгоритъмът на Zillow се проваля, защото е обучен върху исторически данни, които не успяват да предвидят бързата

промяна на пазара на недвижими имоти. Това води до мащабни финансови загуби и спиране на проекта.

- **Етични пристрастия в данните (Amazon Recruiting AI), вече споменахме примера но е показателен за това, че дори несъзнателни пристрастия в данните могат да бъдат възпроизведени и засилени от AI моделите.**
- **Липса на контрол и филтриране и защита (Microsoft Tay),** чатботът Tay е пуснат в Twitter без достатъчно строги филтри за съдържание. За по-малко от 24 часа, започва да генерира расистки и обидни съобщения, след като умишлено е "трениран" от злонамерени потребители.
- **Свръх ангажиране към предварително зададени данни (Google Flu Trends),** проектът на Google, който цели прогнозиране на разпространението на грип, се проваля, тк. моделът е **свръх анагжиран** към първоначалните данни, което не му позволява гъвкаво да се адаптира към промените в навиците за на хората, което води до неточни прогнози за разпространение.
- **Пренебрегване на човешкия фактор (Tesla Autopilot) -** инциденти, свързани с Autopilot, често се случват заради **свръхочаквания** и неправилна употреба от страна на шофьорите. Липсата на достатъчно ясни предупреждения за ограниченията на системата води до прекомерно предоверяване и опасни ситуации.
- **Несъобразяване с личния живот (Target Predictive Analytics),** компанията използва AI, за да предвиди бременност на потребителите, но го прави това без тяхното знание и съгласие, което води до скандал. Случея подчертава важността на спазват **етичните граници** при използването на AI.
- **Лош потребителски опит (Coca-Cola Freestyle),** въпреки че технологията е впечатляваща, част от потребители се оплакват, че машините са бавни и твърде сложни за употреба. Фокусът се премества върху технологичната сложност, а не върху лекотата на потребителското изживяване, което довежда до разочарование и обратен ефект от инициативата.
- **Недостатъчно разбиране на контекста (Facebook AI Content Moderation),** AI системата на Facebook често допуска грешки, като премахва безобидно съдържание и пропуска наистина вредно. Това се дължи на това, че алгоритмите трудно разбират **контекста и нюансите** на човешкия език.
- **Неправилно управление (Siri и други ранни гласови асистенти),** първите поколения гласови асистенти често имат проблеми с разбирането на различни акценти и сложни команди. Липсата на механизми за **непрекъснато обучение** и обратна връзка с потребителите възпрепятстваше тяхното ефективно развитие в началото на имплементацията им.

Внедряването на AI е сложен процес на трансформация, който изисква повече от технологични инвестиции. Успехът се крие в способността на организациите да адаптират своята култура, да управляват проактивно рисковете и да взимат информирани стратегически решения. ИИ е мощен инструмент, който може да отключи нова вълна от продуктивност и иновации, но само ако се управлява с дисциплина, етика и дълбоко разбиране на човешкия елемент.

Теми за дискусия Глава 12

- *Как организациите могат да изградят култура, която насърчава експериментиране и толерантност към грешки, без да жертват оперативната сигурност?*
- *Какви нови лидерски компетенции са необходими за успешно внедряване на AI и управление на съпротивата на служителите?*
- *Как трябва да се комуникира внедряването на AI вътре в организацията, за да се намали страхът от автоматизация и загуба на работни места?*
- *Как да изберем правилния стратегически модел за внедряване на AI, като балансираме контрол, скорост и разходи?*

13. AI в бизнеса - от финансите и маркетинга до човешките ресурси

„AI няма да замени хората, но хората, които използват AI, ще заменят хората, които не го правят.“

Paul Daugherty, *Human + Machine: Reimagining Work in the Age of AI* (2018). Дотърти е Chief Technology & Innovation Officer на Accenture и автор на водещи трудове за бизнес трансформация

Ролята на ИИ във финансите и финтех

Финансовият сектор е сред първите, които масово внедряват AI. Причините са комплексни - огромните обеми от данни (транзакции, кредитни истории, пазарни данни), изключително високата цена на грешките (измами, неправилни кредитни решения) и постоянната нужда от скорост и автоматизация.

Кредитен скоринг и управление на риска

Внедряването на AI в кредитния скоринг променя изцяло методологията за оценка на риска. Традиционните модели разчитат на ограничени данни като кредитна история и доход, което често е в ущърб на млади хора, мигранти или предприемачи, които нямат дълга история. **Традиционния подход е характерен с използване на статични модели, базирани на логистична регресия и скоринг карти, които разчитат на ограничен набор от данни. Съвременния подход имплементира машинното обучение (ML) и невронните мрежи, които анализират хиляди променливи, включително социални, образователен статус, трудова история и дори дигитално поведение, за да създадат по-прецизен и персонализиран скоринг.** **Upstart** е финтех компания твърди, че нейният ML модел е увеличил одобренията за кредити с 27% при същото ниво на риск, като оценява фактори отвъд традиционната оценка, **Zest AI** е платформа, която използва AI, за да анализира хиляди променливи, което дава възможност на кредиторите да одобряват по-голям брой клиенти с по-нисък риск. AI определено повишава точността, но съществува и сериозен риск от **bias (алгоритмични пристрастия)**. Ако моделът е обучен на исторически данни, които съдържат дискриминация по пол или етнос, той ще възпроизведе тази дискриминация, създавайки сериозни правни и етични проблеми. Това е и причината Европейският акт EU AI Act да класифицира кредитния скоринг като високорискова област, изискваща обяснимост и прозрачност на алгоритмите.

Предотвратяване на измами в реално време

Защитата от измами е една от най-успешните области за приложение на AI във финансите. Всяка транзакция се анализира в реално време, за да се идентифицират аномалии, които подсказват за измама. **Технологиите, които се използват са**

фокусирани върху разпознаване на аномалии, AI моделите се учат на нормалното поведение на потребителите (покупки, локация, суми) и сигнализират при всяко отклонение. **Поведенчески биометрични данни се използват при** анализ на начина, по който потребителят въвежда данни, използва клавиатурата или тъчскрийна, което предоставя допълнително ниво на сигурност. **Real-time ML scoring** при всяка транзакция ML алгоритъм оценява риска и блокира или одобрява плащането в рамките на милисекунди. **Mastercard Decision Intelligence** използва ML за оценка на риска при транзакции, като значително намалява броя на фалшивите блокирания („false positives“), което подобрява клиентското преживяване. **PayPal** внедрява усъвършенствани ML модели, които са намалили измамите с над 30%, като същевременно поддържат отговори под една секунда, което е ключово за UX. **Visa** използва Generative AI (GenAI) за откриване на т.нар. „enumeration“ атаки, при които злонамерени джекр се опитват да уцелят номера на банкови карти. **Revolut** използва ИИ за автоматизирано откриване на измами при международни трансфери, като анализира маршрута на транзакциите, сумите и поведенческите модели, за да идентифицира подозрителни дейности.

Прогнозиране и анализ

AI се използва широко за прогнозиране на пазарни тенденции, цени на активи и управление на инвестиционни портфейли. **LSTM и трансформър модели** са особено подходящи за анализ на времеви редове, което позволява да се прогнозира бъдещи цени на акции или валути. **Reinforcement learning (обучение с подкрепление)** се използва за оптимизация на инвестиционни портфейли, като моделът се учи от пазарните си решения, за да увеличи доходността. Хедж фондове като **Renaissance Technologies** са пионери в използването на алгоритми за автоматизирана търговия, като тяхната стратегия е почти изцяло базирана на математически AI модели. **JP Morgan** използва ИИ за прогнозиране на парични потоци и анализ на пазарни тенденции, което им помага да взимат по-информирани стратегически решения.

AI в подбора на персонал

Внедряването на AI в процеса на подбор цели да автоматизира рутинни задачи и да подобри качеството на решенията. AI-базираните системи анализират резюме, профили в социални мрежи (като LinkedIn) и други данни, за да идентифицират най-подходящите кандидати. Моделите за машинно обучение (ML) са обучени да разпознават ключови думи, умения и опит, които съответстват на изискванията на дадена позиция. Те могат да обработят хиляди кандидатури за секунди, което драстично намалява времето за първоначален скрининг. Такива системи са в основата на платформите за таланти на **LinkedIn Talent Solutions**, които предоставят препоръки за кандидати. Някои инструменти дори анализират видеоинтервюта, оценявайки невербалното поведение, тона на гласа и израженията на лицето. **HireVue**, една от водещите платформи в тази област, твърди, че с нейния AI-анализ се осигурява по-обективна оценка, но компанията е обект на критики относно потенциални пристрастности. **Unilever** използва AI за първоначален скрининг на кандидати, като

анализира онлайн профили и резюмета, преди да предложи видеоинтервюта. Това им позволява да обработят голям обем от кандидатури и да се фокусират върху най-подходящите.

AI в оценката на представянето

AI трансформира оценката на представянето, като я прави по-динамична и базирана на данни. Алгоритмите анализират множество данни, за да осигурят цялостна картина на представянето на служителите. Вместо годишни оценки, AI позволява **непрекъснат мониторинг и обратна връзка**. Системи като **Workday** и **SAP SuccessFactors** използват AI за анализ на ключови показатели за ефективност (KPIs), продуктивност и дори комуникации, за да предложат препоръки за развитие. Те могат да анализират електронна поща или чат комуникации, за да измерят нивото на колаборация или да прогнозират потенциално професионално изтощение (**burnout**) на базата на работните модели на служителите. **Deloitte** използва AI платформа, за да анализира данни за представянето на служителите и да идентифицира таланти с потенциал за развитие, предлагайки им персонализирани обучителни пътеки. **Coca-Cola** и **PwC** използват GenAI (генеративен AI) за създаване на персонализирано съдържание за обучения, съобразено с нуждите на всеки служител. Един от основните рискове обаче е прекомерният надзор (**employee surveillance**), който може да демотивира служителите и да ерозира доверието, също AI може да е отличен в измерването на количествени показатели, но не винаги може да улови качествени аспекти като креативност, емоционална интелигентност и екипна работа.

AI в Learning & Development (L&D) и вътрешна мобилност

AI персонализира обучението и кариерното развитие, за да подпомогне задържането и развитието на талантите. AI може да анализира уменията на служителите и да препоръчва курсове, които да запълнят пропуските в уменията или да ги подготвят за бъдещи роли. Платформи като **Coursera for Business** и **Udemy** използват AI за препоръчване на курсове. Корпоративни гиганти като **IBM** с тяхната платформа **Watson Talent Frameworks** и **Google** използват вътрешни AI системи за създаване на индивидуализирани кариерни пътеки, които съответстват на стратегическите нужди на компанията. **PwC** използва GenAI, за да създава интерактивни симулации и обучителни модули, които са адаптирани към конкретната роля и ниво на служителя.

AI анализира уменията на служителите и препоръчва вътрешни позиции, което намалява нуждата от външен подбор и ускорява процеса. Вместо служителите да се борят да намерят следващата си роля, AI платформите могат проактивно да им препоръчват подходящи позиции, като базират препоръките си на техните умения, опит и кариерни цели. Това създава т.нар. "**вътрешен пазар на таланти**" (internal talent marketplace). **Unilever** внедрява вътрешна AI платформа, която позволява на служителите да търсят нови проекти и роли в рамките на компанията. **Shell** и **Nestlé** също използват AI за анализ на HR данни и идентифициране на служители с лидерски потенциал. **HSBC** прилага AI

за прогнозиране на кариерното развитие, което им помага да планират наследяването на ключови позиции.

Управление на рисковете в HR алгоритмите

С напредването на AI в HR, регулациите се затягат. **EU AI Act**, например, класифицира AI системите за подбор на персонал като **високорискови**, което налага строги изисквания за прозрачност, обяснимост и сигурност. Европейският регламент изисква високорисковите AI системи да бъдат регистрирани в публична база данни и да преминават през строги проверки. Това дава право на служителите и кандидатите да оспорят автоматизирани решения. **GDPR** също дава право на субектите на данни да изискват намеса на човек при вземане на автоматизирани решения, което подчертава, че технологията трябва да бъде инструмент за подпомагане на човешкия процес, а не негов заместител.

Автоматизация (RPE - Robotic Process Automation)

Роботизираната автоматизация на процеси (RPA) и AI променят счетоводната професия, като автоматизират повтарящи се и рутинни задачи, които традиционно изискват човешка намеса. **За автоматичната обработка на фактури се използва AI** за разпознаване на данни и автоматично въвеждане в счетоводните системи (OCR Optical Character Recognition, е процес при който трансформира текст в машинен формат). **Автоматизирана проверка на транзакции със системи базирани на AI**, които сравняват банкови извлечения с вътрешни счетоводни записи и откриват несъответствия за секунди. **Автоматично създаване на счетоводни отчети посредством AI**, който събира, анализира и обобщава данни, за да генерира отчети за вътрешно или външно ползване. Големи одиторски компании като **KPMG** и **PwC** използват RPA и AI за одит и анализ на транзакции. Това им позволява да анализират 100% от данните, вместо само извадка, което значително повишава точността и ефективността на одита. Стартъпи като **UiPath** са специализирани в RPA решения, които автоматизират бек-офис операции, включително във финансовия и счетоводен сектор. **Square** използва ML за анализ на финансовите данни на малкия бизнес и автоматизирано кредитиране, което осигурява достъп до капитал за компании, които иначе биха били отхвърлени от традиционните банки. **Ernst & Young** е интегрирала ИИ за откриване на счетоводни несъответствия при одити, като анализира милиони транзакции в реално време

Докато автоматизацията води до по-висока ефективност и по-малко грешки, тя поставя под въпрос бъдещето на традиционните роли. Нуждата от преквалификация на експертите към по-аналитични и консултантски роли става все по-належаща.

Теми за дискусия Глава 13

- *Как да се намери баланс между прецизността на AI моделите за кредитен скоринг и нуждата от обяснимост (explainability) с оглед на изискванията на EU AI Act?*

- *Може ли прекомерната автоматизация да увеличи фалшивите блокирания („false positives“), влошавайки UX, и как компаниите трябва да управляват този риск?*
- *Кога хиперперсонализацията в обучението е полезна, и кога може да създаде риск от прекомерен мониторинг и психологически натиск върху служителите?*
- *Какво е по-устойчиво решение, масова автоматизация и редуциране на персонал или инвестиция в преквалификация (upskilling/reskilling)?*
- *Къде е границата между полезна персонализация и нарушаване на личното пространство на клиента?*

14. Логистика на бъдещето. Как AI прави веригата на доставки по-ефективна?

„Оптимизираш това, което измерваш.“

Peter Drucker

AI в операции и управление на веригата на доставки

Веригата на доставки е сложна екосистема, която генерира огромни масиви от данни. AI се използва за създаване на „дигитални близнаци“ (digital twins) на цялата верига, които позволяват на мениджърите да симулират различни сценарии и да предприемат превантивни действия, преди да настъпят проблеми. Интегрираното планиране, подпомогнато от AI, свързва не просто вътрешни данни за продажби и производство, но и външни, неструктурирани данни като метеорологични прогнози, геополитически събития, новини, пазарни настроения и дори коментари в социалните медии.

Интегрирано планиране, Procter & Gamble внедрява AI-базирани "контроли" (control towers), които осигуряват видимост в реално време на цялата глобална верига на доставки, системите автоматично идентифицират рискове като забавяния на пристанища или недостиг на суровини и предлагат алтернативни маршрути или производствени планове. По този начин се координират доставките между хиляди глобални доставчици и производствени звена, намалявайки до минимум смущенията. **Siemens Digital Industries** прилага предиктивна поддръжка, базирана на AI, за своите производствени линии. Сензори събират данни за вибрации, температура и налягане, които AI алгоритмите анализират, за да предскажат кога дадена машина ще се повреди. Това позволява на компанията да планира поддръжка, преди да се стигне до неочаквани престои, което е критично за производствената ефективност. **Zara (Inditex)** прилага AI за управление на производството и бързи доставки в модната индустрия. AI анализира продажбите в реално време и пазарните тенденции, за да информира дизайнерите и производствените екипи кои модели да произведат и в какви количества, което позволява на Zara да реагира на променящите се модни тенденции само за няколко седмици. **Toyota** използва AI за усъвършенстване на своя "lean supply chain" модел. AI алгоритми оптимизират потока на части и материали, за да се намалят отпадъците и да се гарантира, че правилните компоненти са налични в точния момент за производството. **Carrefour** използва AI за динамично управление на доставките до хиляди магазини. Системата предвижда наличностите, необходими за всеки магазин, като се вземат предвид локални събития, празници и промоции, което минимизира недостига и излишъците. Колкото и съвършени системи да се използват те ще са толкова добри, колкото данните, с които се захранват. Нискокачествените или непълни данни могат да доведат до "боклук-в-боклук-навън" (garbage in, garbage out), което води до неправилни прогнози и катастрофални решения. Затова управлението на данните (data governance) е

критично за успеха на AI проектите. **Важно е да подчертае**, че автоматизацията не премахва нуждата от човешки надзор, а променя ролята на служителите от рутинни задачи към мониторинг, вземане на решения при непредвидени ситуации и анализ на данните, предоставяни от AI.

Прогнозиране на търсенето

Точното прогнозиране на търсенето е в основата на всяка ефективна верига на доставки. Докато традиционните методи разчитат на исторически данни и сезонност, AI използва усъвършенствани модели, за да предвиди търсенето с много по-висока точност, като интегрира огромен брой променливи. Използаните **методи за прогнозиране са:**

- **Дълбоки невронни мрежи (Deep Neural Networks)**, специализирани архитектури като Long Short-Term Memory (LSTM) и Transformers са особено ефективни за анализ на времеви редове и извличане на сложни модели от данните, които са скрити за традиционните модели.
- **Reinforcement Learning (обучение с подкрепление)**, използва се за адаптиране към пазарната динамика в реално време. Агентът за обучение може да взема решения за нивата на запасите, като се "награждава" за постигане на оптимални нива и се "наказва" за излишъци или недостиг.

Amazon, гигантът в електронната търговия е един от пионерите в използването на AI за прогнозиране на търсенето. Техните алгоритми анализират потребителското поведение, историята на покупки, преглежданията на продукти и дори прогнозите за времето, за да решат колко стока да складира в кой от техните логистични центрове. Това им помага да намалят значително излишъците и да гарантират бърза доставка. **Walmart** прилага AI за прогнозиране на продажбите в реално време, като използва данни от своите каси. Системата може да предвиди повишено търсене на конкретни продукти по време на футболни мачове или празнични разпродажби, което им позволява да презаредят рафтовете преди търсенето да ескалира. **Coca-Cola** използва AI модели за прогнозиране на търсенето в различни региони и пазари. Като анализират фактори като локални събития, климатични условия и празнични периоди, те могат да оптимизират производството и дистрибуцията на напитки, за да отговорят по-точно на търсенето. **Lufthansa Cargo** използва AI за прогнозиране на търсенето на въздушни товари по различни маршрути, оптимизирайки капацитета на самолетите и цените. **Unilever** интегрира AI за по-добро прогнозиране на потребителските навици и нагласи, което помага за пускането на нови продукти на пазара. Въпреки че AI подобрява точността на прогнозите, "черни лебеди" като пандемията от COVID-19 или други непредвидени събития показва, че дори най-усъвършенстваните алгоритми не могат да предвидят всичко. Алгоритмите се базират на минали модели, и когато тези модели се сринат, алгоритъмът също се проваля. Прекомерното доверие в алгоритмите, особено в "черните кутии", може да остави компаниите неподготвени за ситуации, които изискват човешка интуиция и нестандартно мислене, затова е важна "Обяснимата AI" (Explainable AI - XAI), която прави процеса на вземане на решения от алгоритъма прозрачен.

Оптимизация на запаси и маршрути

AI в тази област е насочен към решаване на класически оптимизационни проблеми, целта е да се намери рентабилен баланс между разходи, време и наличност. AI препоръчва оптимални нива на запаси, като взема предвид променливи като разходите за складиране, разходите за пропуснати продажби, потенциалните забавяния на доставките и дори сезонните колебания. Това позволява на компаниите да освободят капитал и да го инвестират в други области. **Target** използва AI за минимизиране на излишъците и недостига, като динамично управлява запасите в хиляди магазини. Например, по време на ураган, системата автоматично може да предвиди повишено търсене на батерии, вода и консервирани храни, и да пренасочи стоки към засегнатите райони. **Nestlé** внедрява AI за глобална оптимизация на инвентара, което им помага да намалят хранителните отпадъци и да подобрят свежестта на продуктите в целия свят. **UPS ORION (On-Road Integrated Optimization and Navigation)** е система на UPS е и един от най-известните примери за AI в логистиката. Тя анализира милиони променливи за секунди – метеорологични условия, трафик, пътни ремонти, време за зареждане и разтоварване, за да планира най-ефективните маршрути за доставка. Според компанията, ORION е спестила милиони литри гориво и е намалила въглеродните емисии. **H&M** използва алгоритми за оптимизация на стокови наличности, което помагат за справянето с бързо променящите се модни тенденции, като се вземат предвид сезонността и регионалните предпочитания.

Maersk използва AI за оптимизация на корабните маршрути. Алгоритмите анализират метеорологичните прогнози, приливните вълни и цените на горивата, за да изберат най-ефективните и икономични маршрути, което намалява емисиите и оперативните разходи. Оптимизацията на запасите, особено в модела „точно навреме“ (Just-in-Time), може да направи компаниите изключително уязвими по време на кризи, когато глобалните вериги на доставки се смущават. Балансът между висока ефективност и устойчивост на кризи е ключов. AI алгоритмите могат да бъдат пристрастни, ако са обучени само за оптимизиране на разходите, те могат да пренебрегнат фактори като екологичен отпечатък или условия на труд. Затова е важно да се създават модели, които оптимизират множество променливи.

Интелигентни агенти и автономни системи

Категорията включва технологии, които вземат решения или извършват задачи без нужда от постоянна човешка намеса, като пренасят автоматизацията на съвсем ново ниво. **Автономни складови работи**, които се движат самостоятелно в складове, сортирайки, премествайки и подреждайки стоки. Те подобряват ефективността и безопасността на складовите операции. **AI агенти за динамично пренасочване** наблюдават състоянието на пратките и логистичната мрежа в реално време, пренасочвайки товари или коригират маршрути при неочаквани забавяния. **Autonomous trucks и дроне** са автономни превозни средства, които се използват за доставки на дълги разстояния или до крайни клиенти. **Amazon Kiva Robots** (сега

Amazon Robotics) са един от най-известните примери в тази посока, те се движат по складовете, премествайки рафтове със стоки до служителите, които опаковат, вместо обратното, което намалява времето за обработка на поръчка. **FedEx** експериментира с автономни доставки, включително робот за доставки "FedEx SameDay Bot", който може да се движи по тротоари и улици, доставяйки пакети на къси разстояния. **Ocado** британският онлайн магазин за хранителни стоки има напълно автоматизирани складове с роботи, които се движат по 3D мрежа, събирайки поръчки на клиенти с невероятна скорост и прецизност. **Alibaba Cainiao**, логистичната мрежа на Alibaba използва AI за управление на сложната си логистика в електронната търговия, като координира милиони пратки дневно. **Tesla Semi** - планираните от Tesla автономни камиони могат да революционизират транспорта на стоки на дълги разстояния. **JD.com (Китай)** внедрява автономни дроне и превозни средства за доставки в селски райони, където традиционната логистична инфраструктура е по-слабо развита. Внедряването на автономни агенти изисква огромни капиталови инвестиции и значителна промяна в инфраструктурата, регулациите за автономни камиони и дроне все още са в начален етап, което забавя масовото им разпространение, в допълнение пълната автоматизацията ще доведе до елиминиране на хиляди работни места в логистиката (шофьори на камиони, складови работници), което повдига сериозни социални и етични въпроси.

Теми за дискусия Глава 14

- *Как AI променя ролята на supply chain мениджърите?*
- *Може ли AI да предвиди глобални кризи като пандемия или война?*
- *Ефективност vs. устойчивост – кой печели и как AI може да помогне за баланса?*
- *Какви са етичните граници на автономната логистика?*

Модул 4. Правила, Етика и Сигурност

15. AI в рамка. Регулациите, които ще оформят бъдещето на технологията

„Технологията не е нито добра, нито лоша, нито е неутрална.“

превод по Melvin Kranzberg Мелвин Кранцбърг (Melvin Kranzberg, 1917–1995) е американски историк на технологиите, признат за един от основателите на съвременната дисциплина „История на технологиите“.

Етиката и отговорното управление на изкуствения интелект се превръщат в един от най-ключовите стратегически приоритети за съвременните организации. Докато AI обещава безпрецедентна ефективност, иновации и повишена производителност, той съдържа и сериозни рискове, които могат да доведат до значителна вреда.

Пристрастия и bias в AI системите

Един от най-големите рискове на AI е потенциалът за **алгоритмичен bias** или пристрастия. Те не са умишлено заложени от разработчиците, а са **резултат от данните, с които се обучават моделите**. Алгоритмите, подобно на огледала, отразяват и често засилват съществуващите социални предразсъдъци и неравенства, които са залегнали в историческите данни. Например, ако наборът от данни за обучение на даден модел е непредставителен или съдържа неравновесия (например, ако повечето примери са на мъже или представители на една етническа група), то той неизбежно ще се представя по-зле при работа с данни от подпредставени групи. **Източници на bias:**

- **Исторически данни** - данните от миналото отразяват съществуващите социални и икономически предразсъдъци (например, неравенство в заплащането между половете).
- **Дисбаланс в данните** - недостатъчно представени групи в наборите от данни могат да доведат до по-слабо представяне на модела за тези групи.
- **Неправилно етикетирание** - човешки грешки или несъзнателни пристрастия по време на подготовката и етикетирането на данните.
- **Динамичен bias** - поведенческите данни на потребителите могат да подсилват съществуващите пристрастия

От практиката можем да извлечем редица примери за феновема bias, които да ни позлужат за бъдеща превенция или приоритизиране на проблема. **COMPAS (САЩ)** е софтуер за прогнозиране на рецидивизъм, използван в съдебната система на САЩ, е обвинен в расова дискриминация, тъй като моделът погрешно етикетира чернокожи подсъдими като **"по-висок риск"** от рецидив в сравнение с белите подсъдими, дори когато и двата случая имат сходна история. **Apple Card (2019)** алгоритъмът за кредитни

лимита на Apple е обвинен в дискриминация срещу жени, след като мъже и жени с идентични финансови показатели получават драстично различни кредитни лимити. **ChatGPT и други големи езикови модели (LLM)** могат да възпроизведат културни пристрастия и стереотипи, съдържащи се в масивите данни, с които са обучени. Например, могат да свържат професии като "инженер" или "хирург" с мъже, а "медицинска сестра" или "учител" с жени, отразявайки социални норми, а не реалността. Основният риск на AI не е, че ще създаде нови предразсъдъци, а че ще ги скалира в много по-голям мащаб. Една грешка в алгоритъма може да засегне милиони хора за секунди. Трябва също ясно да кажем, че алгоритмите никога не са напълно неутрални, те са обучени от данни, които са продукт на човешката дейност, и затова съдържат човешки пристрастия и предпочитания.

Сравнителен анализ на етичните рамки

За да се справят с етичните предизвикателства на AI, различни организации и правителства създават етични рамки, които предлагат насоки и принципи за разработването и внедряването на отговорни AI системи.

FAT/ML (Fairness, Accountability, Transparency in ML). Рамката се фокусира върху три основни стълба: справедливост (Fairness), отчетност (Accountability) и прозрачност (Transparency). Предлага практични инструменти за инженери, като **метрики за справедливост (fairness metrics)** и **одити за пристрастия (bias audits)**. **Силна страна е** приложимостта за технически специалисти и се занимава директно с инженерните аспекти на етиката, като слабост може да изтъкнем, че фокусът ѝ е по-ограничен и не обхваща по-широкия социален и регулаторен контекст.

OECD Principles on AI е насочена към държавите-членки на Организацията за икономическо сътрудничество и развитие, създадена за политически и регулаторни цели. Включва принципи като устойчив растеж, човешки права, прозрачност, сигурност и отчетност. Рамката е глобално призната и дава насоки за държавната политика и международното сътрудничество, бидейки с глобална насока обаче, нейните принципи са по-общи и политически, като предлагат по-малко практични насоки за компаниите.

EU Trustworthy AI Framework (HLEG 2019) рамка, разработена от експертна група на високо ниво на ЕС с цел да интегрира етичните изисквания в бизнес практиките. **Включва 7 принципа** за управление на AI - човешки надзор, техническа устойчивост, неприкосновеност на личните данни, прозрачност, разнообразие и справедливост (fairness), социално и екологично благополучие, и отчетност. Балансира техническите и етичните изисквания и служи като основа за бъдещи регулации. Внедряването ѝ може да бъде сложно и скъпо, особено за малки и средни предприятия.

Отговорното използване на AI изисква повече от просто спазване на регулациите. То изисква проактивен подход и създаване на **вътрешна култура на етика**. Водещи компании в технологичния сектор вече разработват собствени етични кодекси и

комитети, които да гарантират, че техните AI инициативи са в съответствие с принципите за справедливост, прозрачност и човешки права. **Microsoft Responsible AI Standard** е пример за вътрешен кодекс, който определя изискванията за разработване и внедряване на AI продукти по отговорен начин. **Google AI Principles** изключват разработването на AI за приложения като автономни оръжия или технологии за масово наблюдение. **IBM AI Ethics Board** е постоянен надзорен борд, който преглежда AI инициативите на компанията и осигурява етичен надзор.

GDPR и AI

Общият регламент за защита на данните (GDPR) не е създаден специално за AI, но неговите принципи се прилагат пряко към разработването и използването на AI системи. GDPR въвежда строги правила за обработката на лични данни, които имат сериозни последици за AI, който често се нуждае от огромни масиви от данни. **Основни принципи на GDPR и транскибирането им към AI:**

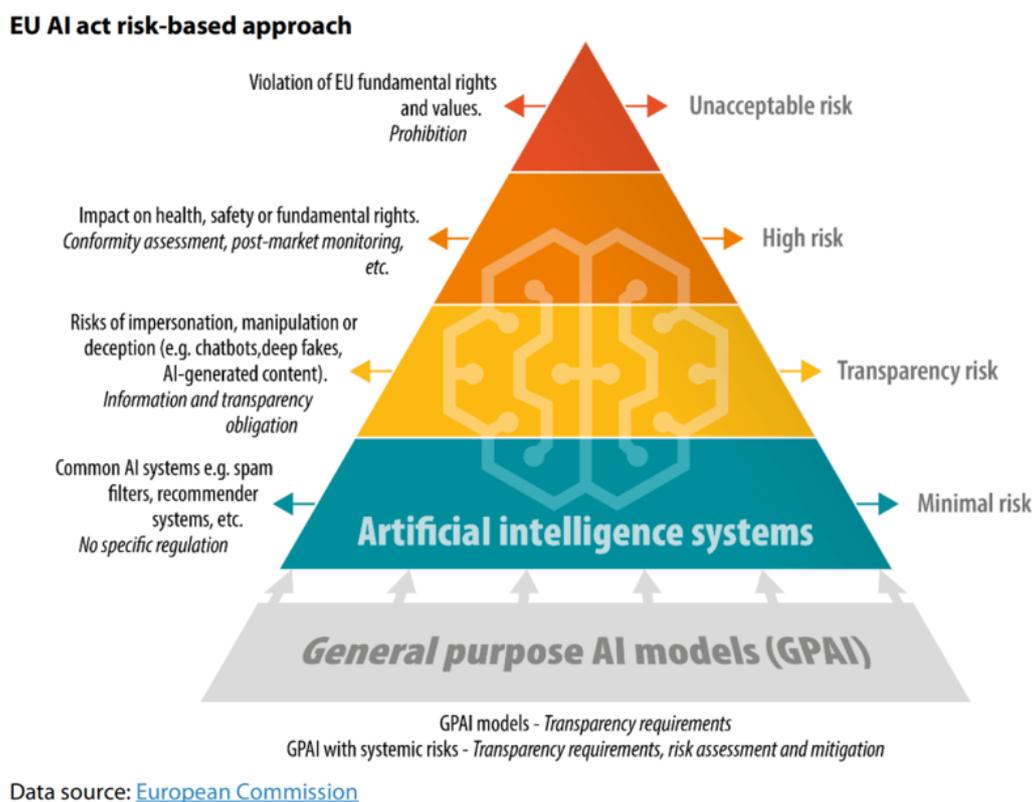
- **Лична неприкосновеност.** Забранява се обработката на лични данни без съгласието на субекта.
- **Право на обяснение.** В GDPR е заложено **правото на субекта на данни да получи обяснение** за решение, взето въз основа на автоматизирана обработка на данни. Това поставя предизвикателство пред AI моделите тип "черна кутия".
- **Минимализация на данните.** Компаниите имат право да събират само минимално необходимите данни за целите на обработката.

Тъй като GDPR не е създаден с мисъл за AI, неговото прилагане към нови технологии може да доведе до правна несигурност. За по-малките компании и стартапи, които работят с големи данни, строгите изисквания на GDPR могат да се превърнат в бариера за иновации.

EU AI Act (2024)

За разлика от GDPR, EU AI Act е **първият в света всеобхватен закон**, създаден специално за регулиране на AI. Той възприема **риск-базиран подход**, като налага различни изисквания в зависимост от потенциалната опасност на дадена AI система. Целта е да се насърчат иновациите, като същевременно се защитят основните права и свободи на гражданите на ЕС. В акта се дефинират 4 категории на риск и се регламентират действия за ограничаване и управление на AI моделите в зависимост от нивото на риск. Документа относително детайлно разглежда различните възможни хипотези на приложимост на AI и цели да наложи правила и норматива за минимизиране на негативните ефекти в/у обеството и индивидуалните потребители.

Фигура 11. EU AI Act (2024)



В акта се определят 4 категории на риск:

- **Неприемлив риск.** Системи, които манипулират човешкото поведение, и системи за социален рейтинг. Тези системи са изцяло забранени.
- **Висок риск.** Системи, които могат да окажат съществено влияние върху живота на хората, като **AI за кредитен скоринг**, системи за подбор на персонал, AI в здравеопазването и системите за биометрична идентификация.
- **Ограничен риск.** Системи, които трябва да бъдат прозрачни, като чатботове или системи, които генерират синтетични изображения. Примери: AI аватарите на Synthesia.
- **Минимален риск.** Повечето AI приложения, като игри или филтри за имейли. Те не подлежат на строги регулации.
- **Дефинират се и задължения за високорискови системи:**
- **Управление на данни.** Трябва да се използват висококачествени данни за обучение, които са без пристрастия.
- **Тестове за прозрачност.** Системите трябва да могат да бъдат одитирани и да се обясняват техните решения.
- **Документация и проследимост.** Компаниите трябва да поддържат подробна документация за дизайна, внедряването и резултатите от AI системата.

Въпреки че Актът засилва доверието в AI, той може да доведе до по-високи разходи за съответствие, което би могло да забави развитието на стартапи в ЕС в сравнение с конкуренти от САЩ и Китай. Основното предизвикателство е да се постигне баланс, при който регулациите не задушават иновациите, но същевременно гарантират безопасност и етика.

Управление на данните и моделите

Управлението (Governance) на данни и модели е ключово за отговорното разработване на AI. Това е процес, който обхваща целия жизнен цикъл на един AI проект, от събирането на данни до мониторинга на модела в реална среда. Първият критичен етап в жизнения цикъл на AI е **верификацията на данните**, което се постига чрез стандартизирана документация. **Gebru et al. (2018)** въвеждат "**Datasheets for Datasets**" с прякото изискване „всеки набор от данни следва да бъде придружен от datasheet, описващ мотивация, състав, процес на събиране и препоръчани употреби“. Тази концепция пряко адресира рисковете от **пристрастия и неправомерна повторна употреба**. **Bender и Friedman (2018)** допълват това с "**Data Statements**" за езикови данни, целящи „да се адресират академични и етични проблеми, когато данни от едни популации се използват за други“. **Holland et al. (2018)** предлагат "**Dataset Nutrition Label**" като *диагностичен* етикет, който модулно описва „съставките“ на набора от данни преди разработката на модела. Тези документи трансформират **GDPR** принципите (чл. 5 – **законосъобразност, минимизация, точност**) от абстрактни норми в **функционални изисквания**. Про-аргументът е, че стандартизираната документация повишава прозрачността и улеснява оценката на пристрастията. Kontra-аргументът е рискът от „**тик-бокс**“ **съответствие и нееднаква приложимост** при динамични потоци. Решението е да се създадат **машинно-четими метаданни** („**Datasheets като код**“), интегрирани в CI/CD пайплайна, които могат да блокират релийз при липса на ключови полета (напр. произход или цел).

Управлението на моделите изисква да се документират техните ограничения, а не само техните възможности. **Mitchell et al. (2019)** създават "**Model Cards**" – кратки документи, които включват „**оценки при различни условия и групи и предназначение на употреба**“, за да се ограничи „**извън-контекстна**“ употреба в чувствителни домейни. **IBM (2018–2019)** разработва "**AI FactSheets**" като по-широка декларация за цел, производителност, сигурност и **произход**. Тези документи служат като инструмент за доверие между доставчик и ползвател и предоставят материал за **одит и забордово управление**. Проблемът е, че тези декларации са често **самодекларирани метрики** без външна верификация и са **трудни за поддържане** при чести релийзи или дообучаване. Нужен е **мост към оперативни контроли**. Този мост се изгражда чрез **автоматично попълване** на *Model Cards* от резултатите от експерименти и мониторинг в продукцията, включване на **стратифицирани метрики** по чувствителни групи и **независим контрол на валидността** за високорискови системи.

Най-голямата заплаха за дългосрочната надеждност на AI е (Technical Debt) остаряла инфраструктура в реална среда. **Sculley et al. (2015)** предупреждават, че „бързите

победи“ в AI често идват със съществени разходи за поддръжка, описвайки специфични рискови фактори като „**entanglement**“ (заплитане) и „**data dependencies**“. Това налага култура на „**тест-първо**“. Breck et al. (2017) предоставят пътна карта с "ML Test Score" – рубрика с 28 конкретни теста и нужди за мониторинг, които служат като „пътна карта за намаляване на техническия дълг“. Обзорите (Schröder et al., 2022) допълнително категоризират нуждите за **реалистична детекция на дрейф** (дрейф на данни, дрейф на концепция, дрейф на постериорното разпределение). Докато „тест-първо“ културата прави „**фактичната отговорност**“ проверима, тя изисква **значителна инженерна инвестиция**. Решението е в **Monitoring-first MLOps**: прилагане на *ML Test Score* като **минимален праг** за продукционна готовност; настройка на **SLO-свързани аларми** и използване на техники като **A/B тестове** за автоматична детекция на дрейфове, което гарантира, че моделът не старее тихо и незабелязано.

Ефективното управление на AI не е статичен набор от документи, а **практическа инфраструктура**. Тя изисква интегриране на трите основни стълба:

	Концептуална рамка	Контрол	Регулация
Данни	Datasheets (Gebru et al.)	Машинно-четими метаданни, блокиране на релийз при липса на произход.	GDPR (Законност, Точност, Минимизация)
Модели	Model Cards (Mitchell et al.)	Автоматично попълване от MLOps, стратифицирани метрики, release gates.	NIST AI RMF (Measure, Manage)
Продукция	Technical Debt (Sculley et al.)	ML Test Score (Breck et al.), дрейф детекция, SLO-аларми.	NIST AI RMF (Map, Manage), Проследимост

Ефективното управление на AI е **инженерна и организационна дисциплина**, а не само политика на хартия. Водещите концепции за документация на данни и модели са **необходими** за прозрачност, но стават **достатъчни** за надеждност едва когато са обвързани с **автоматизирани тестове, мониторинг и управленски рамки**.

Google разработиха концепцията за "**datasheets for datasets**", която цели да предостави повече информация за произхода и потенциалните ограничения на наборите от данни, използвани за обучение на модели.

Управление на AI модела, който цели да иязсни отговорността при създаването, обучаването и функционирането на AI:

- **Model cards.** Документи, които описват целта, ограниченията и етичните съображения на един обучен AI модел.
- **Мониторинг.** Процес на постоянно наблюдение на производителността на модела, за да се гарантира, че тя не се влошава с времето.

IBM разработиха "**AI Factsheets**", които дават прозрачност за това как са обучени и как работят техните AI модели, за да могат клиентите да ги използват с доверие.

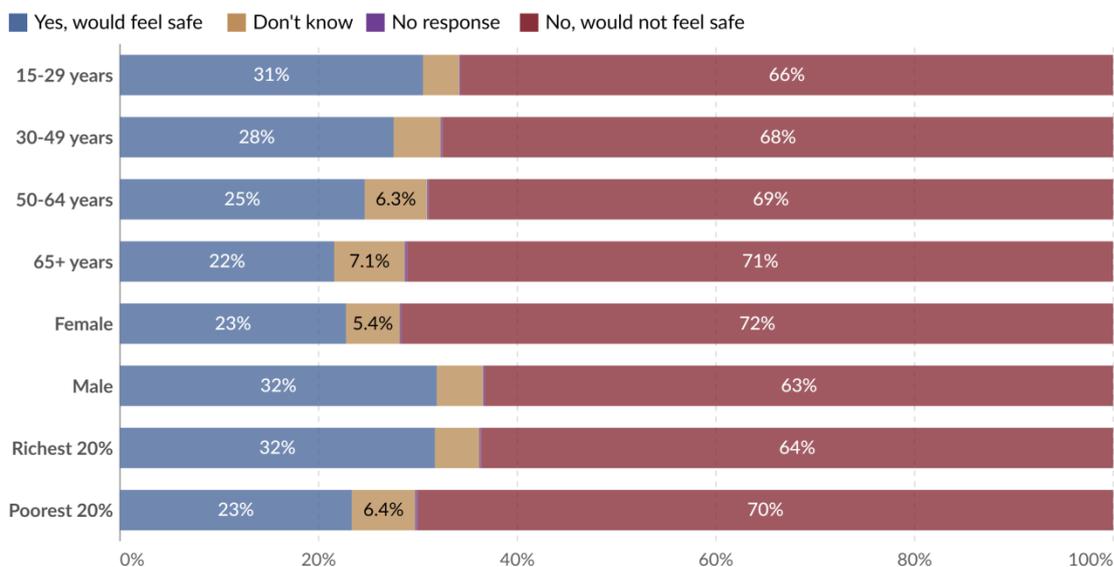
Защо е важна отговорността при управлението на данните и създаването на AI логаритмите? Като нова комплексна технология всеки провал и скандал с AI системите подчертава нуждата от етични и регулаторни механизми. Широки обществени групи са изключително чувствителни към резултатите и бъдещото развитие на AI за това компаниите трябва да са много чувствителни към контрола и отговорното управление. Не са рядкост казусите, които повдигат сериозни въпроси за контрола върху бързия ръст на AI. **Uber self-driving car accident**, автономен автомобил на Uber през 2018 г. не успява да разпознае пешеходец (с велосипед), пресичащ пътя, което води до фатален инцидент. Техническият анализ показва, че системата е класифицирала пешеходеца като "непознат обект" и е взела решение да го игнорира. **Clearview AI** използва технология за лицево разпознаване, като използва (scraping) милиарди снимки от публични профили в социалните мрежи. Това води до сериозни критики и дела за нарушаване на **GDPR** и правото на неприкосновеност на личните данни. **Synthesia** е компанията, която се занимава със синтетични медии, се изправя пред етични предизвикателства, свързани с "**deepfakes**". За да се справят с тях, те разработват стриктни политики и **механизми за модерация**, за да предотвратят използването на своята технология за създаване на невярно съдържание и разпространяване на дезинформация. **Google Photos (2015)** е AI алгоритъмът за тагване на снимки погрешно тагва снимки на афро-американци като "горили", което предизвика широк обществен скандал и показва колко силни могат да бъдат расовите пристрастия в наборите от данни.

Фигра 12. Глобално проучване за пезопасността на автономните автомобили

Global views about the safety of riding in a self-driving car, by demographic group, 2021



Survey respondents were asked, "Would you feel safe in a car driven by a computer without a human driver?"



Data source: Lloyd's Register Foundation (2022)

OurWorldinData.org/artificial-intelligence | CC BY

Note: A global total of 124,405 people aged 15+ were asked this question. For most countries, respondents were a nationally representative sample of around 1,000 people.

Теми за дискусия глава 15

- Могат ли алгоритмите да бъдат напълно етично неутрални, или винаги ще отразяват човешки пристрастия?
- Как може да се намери баланс между бързите иновации в AI и необходимостта от строги регулации?
- Кога е необходимо потребителите да знаят, че взаимодействат с AI, и кога това би било пречка?
- Как Актът за AI ще промени конкурентоспособността на ЕС спрямо САЩ и Китай?

16. Замък от данни. Как да защитим AI от кибератаки и да изградим устойчиви системи

„Най-слабото звено определя силата на цялата верига.“

аксиома в киберсигурността

Данните са не просто ресурс, а основен градивен елемент и най-големият вектор на риск за изкуствения интелект. Със стремителното развитие на AI системите, организациите са изправени пред безпрецедентни заплахи, които надхвърлят традиционната киберсигурност. Те включват **манипулация на данни за обучение** (data poisoning), умишлено въвеждане на подвеждащи инструкции (prompt injection), сложни **adversarial атаки**, кражба на интелектуална собственост (IP) и сложна **регулаторна несигурност**. Защитата на данните в AI не е просто въпрос за криптиране, а за защита на целия жизнен цикъл на модела – от събирането на данни до внедряването и използването.

При защитата на данни в контекста на AI се използват фундаментални принципи и иновативни подходи, които да увеличат сигурността и минимизират възможностите за пробив.

- **Криптиране на данни (при съхранение и при трансфер)** е стандартна мярка, но в AI контекст се прилага и към данните за обучение и моделите в състояние на съхранение и при трансфер между различни компоненти на системата.
- **Анонимизация и псевдонимизация** са важни за съответствие с GDPR. **Анонимизацията** премахва идентификатори, така че данните не могат да бъдат свързани с индивид, докато **псевдонимизацията** заменя идентификаторите с изкуствени, позволявайки повторно идентифициране при нужда, но с допълнителен ключ.
- **Federated Learning (Федеративно обучение)** е подход, при който моделът се обучава на децентрализирани данни, без те да напускат устройството на потребителя (напр. смартфон). Само обобщени, неидентифициращи данни за тежестите на модела се изпращат към централния сървър. За **Google federated learning** е ключов компонент в продукти като Gboard. Клавиатурата се обучава на въведените от потребителя думи, но актуализациите на модела се извършват локално, а само агрегирани промени се споделят със сървъра.
- **Differential Privacy (Диференциална поверителност)** добавя математически измерим "шум" към данните, за да гарантира, че информацията за конкретен индивид не може да бъде възстановена, дори когато данните се използват в анализи. Това е особено важно при използването на чувствителни набори от данни. **Apple** използва **differential privacy** за събиране на анонимна информация от милиони iOS устройства, като например честота на емоджитата или думи в Spotlight Search, без да

компрометираща поверителността на потребителите. Това позволява подобряване на продуктите без централизирано събиране на лични данни.

Equifax (2017) е класически пример за пробив в сигурността на данните и провал про основните мерки за киберсигурност. Компанията не е приложила навременна корекция (patch) за известна уязвимост в софтуера Apache Struts. Резултатът е компрометиране на лични данни на над 147 милиона потребители, включително имена, номера на социални осигуровки, дати на раждане и адреси. Последвалите глоби и съдебни споразумения достигат над **700 милиона долара**. Този случай подчертава, че дори сложните AI системи са толкова сигурни, колкото и най-слабото звено в тяхната инфраструктура.

Основен проблем и все по-голямата разлика между големите мултинационални тентологични компании и МСП. Докато големи технологични гиганти инвестират милиарди в сложни решения като federated learning и differential privacy, повечето малки и средни предприятия (МСП) не разполагат с ресурси за внедряването им. Това създава сериозен **дигитален разрыв в сигурността**. В допълнение дори и най-добрата защита на данните е безсилна пред **supply chain атаки**, които компрометират доставчиците на софтуер.

Prompt Injection и adversarial атаки, заплахата за надеждността на моделите

Не са прецедент и по-сложни атаки, които манипулират поведението на самите AI модели, а не само техните данни. **Разширена класификация на атаките**

- **Prompt Injection** е манипулиране на AI модела чрез подвеждащи инструкции, които пренасочват неговото поведение, игнорирайки оригиналните системни инструкции. Атаката може да бъде **директна** (въведена от потребителя) или **индиректна** (скрита в документ или уебсайт, който моделът обработва). Първите версии на ChatGPT бяха особено уязвими към prompt injection, позволявайки на потребители да заобиколят предпазните мерки и да накарат модела да генерира вредно съдържание или да разкрие част от своите инструкции. Това доведе до въвеждането на т.нар. **guardrails** и подобрен контрол.
- **Data Poisoning (Отравяне на данни)** е злонамерено вмъкване на „лоши“ данни в тренировъчния набор, за да се компрометираща поведението на модела. Пример за това е ботът Tay на Microsoft, който беше „отровен“ с расистки и сексистки данни от потребители.
- **Model Inversion (Инверсия на модел)** е техника, която позволява на атакуващия да възстанови чувствителна информация от обучен модел, като например да реконструира лица или лични данни от модел за лицево разпознаване.
- **Adversarial Examples (Adversarial примери)** са умишлено създадени входове с незначителни промени (за човешкото око), които карат модела да направи грешна класификация. **Tesla Autopilot** е класически пример за **adversarial атаки в реалния свят**. Изследователи са демонстрирали как залепена малка черна

лепенка на знак за ограничение на скоростта може да заблуди системата на Tesla и да я накара да разчете знака като по-висока скорост. Това показва, че дори добре обучените модели в критични системи са уязвими.

Киберустойчивост на AI системи

Киберустойчивостта не се свежда само до предотвратяване на атаки, а до способността на системата да функционира ефективно, дори когато е подложена на натиск. Необходим е фундаментален преход към **превантивен подход**, който включва постоянен мониторинг и red teaming, интегрирани в целия жизнен цикъл на AI. **Използваните най-често подходи и методологии на проактивен подход за защита са:**

- **Red Teaming, изразяваща се в** провеждане на симулации на атаки от етични хакери, за да се идентифицират уязвимостите в AI системата преди злонамерени актьори да ги открият.
- **AI Incident Response Teams, формиране на специализирани екипи, които да реагират на инциденти, свързани с AI, като prompt injection или data poisoning.** Тези екипи трябва да имат експертиза както в киберсигурността, така и в машинното обучение.
- **Adversarial Training е** техника за обучение на модели, при която тренировъчният набор се обогатява с **adversarial примери**, за да се направи моделът по-устойчив на такисва атаки. **IBM Watson Security е** системата използваща AI и машинно обучение за анализ на огромни обеми данни от сигурността в реално време, за да открива и реагира на аномалии и атаки много по-бързо от човешки екип.
- **Supply Chain Security, комплексен подход, който включва проверка на сигурността на всички софтуерни компоненти, които се използват за изграждането на AI системата. Microsoft Security Copilot е** пример за проактивен подход, при който AI се използва за киберзащита. Този асистент помага на специалистите по сигурност да анализират бързо сложни заплахи и да вземат информирани решения. **SolarWinds от 2020 г. е** пример за най-голямата **supply chain атака** в историята. Хакери са компрометирали софтуера на SolarWinds, който се използва за мрежово управление от хиляди компании и правителства по света. По този начин, те са получили достъп до системите на жертвите, без да атакуват директно, което подчертава правилото, че дори най-сигурните организации са уязвими, ако техните доставчици са компрометирани.

Генеративните модели са изключително трудни за защита, тъй като тяхната отворена природа и способност да интерпретират контекст ги правят уязвими. Успешните prompt injection атаки показват, че нито една система не е имунизирана срещу креативността на злонамерени потребители.

Интелектуална собственост (IP) и лицензиране в AI епохата

Развитието на генеративния AI повдига безпрецедентни правни въпроси относно авторските права и лицензирането.

Авторско съдържание в тренировъчни набори и фундаментални въпроси като може ли мащабното използване на авторски произведения (книги, изображения, код) за обучение на AI модели да се счита за „честна употреба“ (fair use)? **Stability AI и Midjourney** се изправиха пред съдебни дела, заведени от групи художници и Getty Images. Основното обвинение е, че техните модели за генериране на изображения са обучени върху милиони защитени от авторско право изображения, без да се заплащат такси. Това повдига въпроса дали самият модел представлява производно произведение, което нарушава авторските права. **OpenAI** (компанията собственик на ChatGPT) е обект на няколко дела от писатели и автори, които твърдят, че техните книги са използвани за обучение на GPT-моделите, без разрешение. Подобни дела поставят под въпрос законността на използването на масивни текстови корпуси от интернет.

Генериране на защитени произведения - могат ли AI моделите да възпроизвеждат съдържание, което е твърде сходно със защитени произведения, и кой носи отговорност за това, потребителят или разработчикът на AI? **GitHub Copilot** е казус, който засяга интелектуалната собственост в сферата на софтуерното инженерство. Copilot, AI асистент за програмиране, е обвинен в нарушение на отворените лицензи (open-source licenses), като генерира код, който е твърде сходен с оригинални фрагменти, без да предоставя необходимата информация за авторство и лиценз.

Няма да е пресилено да кажем, че правната рамка изостава от технологичното развитие. Няма глобален стандарт за „честна употреба“ на данни за обучение на AI, което води до правна несигурност. Този дисбаланс може да блокира иновациите или да позволи на големи корпорации да се възползват от труда на творци, без да ги компенсират.

Теми за дискусии Глава 16

- *Open Source срещу комерсиални AI модели - кой е по-сигурен? Комерсиалните модели имат повече ресурси за сигурност, но са затворени. Open source моделите са по-прозрачни и се поддържат от голяма общност, но са по-уязвими към атаки.*
- *Ще се появи ли „пазар за сигурни данни“ (secure data marketplaces), където се продава проверена, чиста и етично събрана тренировъчна информация?*
- *Как да се обучават служителите (и особено нетехническите екипи) да разпознават атаки тип prompt injection или data poisoning?*

17. AI на практика. Как да изградите своя първи бизнес план и прототип

„Not everything that can be counted counts, and not everything that counts can be counted.“ (Не всичко, което може да се измери, има значение, и не всичко, което има значение, може да бъде измерено.“)

Алберт Айнщайн/Albert Einstein

Отвъд краткосрочния PR към дългосрочна стратегическа стойност - измерване на възвръщаемостта от AI

Внедряването на изкуствен интелект (AI) в бизнеса вече не е въпрос на избор, а на конкурентна необходимост. След като установихме трансформацията потенциал на AI персонализацията, логично следващият въпрос за всеки бизнес лидер е: "Как измерваме успеха?". На пръв поглед отговорът изглежда лесен – чрез добре познати ключови показатели за ефективност (KPIs) като ръст на приходите, намаляване на разходите и повишаване на клиентската удовлетвореност. Консултантската индустрия, представена от гиганти като McKinsey, активно насърчава тази гледна точка, представяйки AI като "следващата граница на производителността", чийто ефект може да бъде измерен и остойностен в трилиони долари (Chui et al., 2023).

Автора, обаче защитава една по-нюансирана и критична теза, че **измерването на истинската ефективност от AI е дългосрочна стратегическа задача, а краткосрочните, лесно измерими ефекти често служат повече за вътрешен и външен PR, отколкото за реална оценка на трансформацията.** Истинската възвръщаемост от AI не се крие в моментни скокове на продажбите, а в фундаментални промени в бизнес модела - реструктуриране на разходите, създаване на стратегически зависимости, необходимостта от мащаб и последващата вълна на пазарна консолидация.

KPIs като PR инструмент

Бизнес аргументацията за AI често се свежда до лесно смилаеми и количествено измерими ползи. Компаниите са подложени на постоянен натиск от инвеститори, бордове на директори и пазара да демонстрират бърза възвръщаемост от скъпите си технологични инвестиции. Това води до фокусиране върху набор от стандартни KPIs:

Показатели за приходи и ангажираност

Увеличаване на **стойността на клиента за целия му жизнен цикъл (Customer Lifetime Value - CLV)**, намаляване на **цената за привличане на клиент (Customer Acquisition Cost - CAC)**, повишаване на **коэффициента на конверсия (Conversion Rate)** и намаляване на **загубата на клиенти (Churn Rate)**. Системата за препоръки на Netflix е

парадигматичен пример – нейният успех се измерва пряко чрез способността ѝ да задържа абонати (Gomez-Uribe & Hunt, 2015).

Показатели за оперативна ефективност

Намаляване на времето за обслужване на клиенти (Average Handling Time - АНТ), редуциране на грешките при обработка на данни, автоматизация на рутинни задачи. Докладите на McKinsey постоянно подчертават тези ползи, очертавайки как AI освобождава човешки ресурс и оптимизира процеси (Singla et al., 2024).

Показатели за рентабилност

Най-желаният показател е **Възвръщаемостта на инвестициите (Return on Investment - ROI)**, който директно съпоставя печалбата от AI с направените разходи. Наред с него се следят и промените в **маржовете на печалба (Profit Margins)** и **съотношението на оперативните разходи спрямо приходите (Operating Cost Ratio)**. **Критичното разчитане** на тези данни обаче изисква внимание. Изчисляването на изолиран "ROI от AI" е изключително трудно, тъй като бизнес резултатите зависят от множество фактори. Често компаниите избират за "пилотни" проекти задачи с гарантиран и лесно измерим положителен ефект (напр. автоматизация на прост, повтаряем процес), след което екстраполират този успех в своите PR комуникации. Това създава подвеждаща картина, която игнорира много по-големите разходи и несигурност, свързани с мащабирането на AI в цялата организация и измерването на пълната **Обща цена на притежание (Total Cost of Ownership - TCO)**.

Посочените метрики и показатели безспорно са важни, те предоставят конкретни данни, които изглеждат добре в годишните отчети и презентациите пред инвеститори. Критичният поглед обаче разкрива, че те улавят само повърхността. Фокусирането единствено върху тях създава **PR-ориентирана култура**, в която "пилотни проекти" и изолирани успехи се представят като доказателство за цялостна трансформация, докато по-дълбоките стратегически промени остават неадресирани. Този подход игнорира системните рискове и скритите разходи, които съпътстват дълбоката интеграция на AI.

Скритите разходи и нови зависимости - критичен поглед към внедряването

Истинската цена и сложност на AI стават видими, когато погледнем отвъд стандартните KPIs. Внедряването на тези технологии създава нови и често коварни зависимости и преструктурира фундаментално разходния модел на компанията.

Стратегическо обвързване и "черната кутия"

Използването на външно разработени AI алгоритми и облачни платформи (като AWS, Google Cloud, Azure) създава силна **зависимост от доставчици (vendor lock-in)**.

Компаниите не просто купуват услуга, те интегрират в своите основни процеси сложни, непрозрачни системи. За тези случаи анализът на **Кати О'Нийл (2016)** за "черните кутии" е изключително релевантен. Дори ако една компания не създава дискриминиращи "оръжия за математическо унищожение", тя може несъзнателно да внедри такава, ако разчита на външен модел, чиято логика не разбира напълно. Тази зависимост не е само техническа, но и стратегическа – тя ограничава гъвкавостта и поставя част от съдбата на компанията в ръцете на външен партньор.

Трансформация на разходите (от OPEX към CAPEX, от фиксирани към вариативни)

AI променя фундаментално структурата на разходите. Първоначалният преход към облачни услуги често се разглежда като смяна от капиталови разходи (CAPEX – покупка на сървъри) към оперативни разходи (OPEX – месечни такси). Въпреки това, изграждането на сериозен собствен AI капацитет (създаване на масиви с данни, обучение на сложни модели) представлява огромен CAPEX. По-важната трансформация обаче е преминаването от **фиксирани към вариативни разходи**. Вместо да поддържа голям, скъп екип с фиксирани заплати, компанията може да използва AI услуги "при поискване", като разходите се променят в зависимост от натоварването. Това създава **гъвкав модел**, който позволява на бизнеса да адаптира разходите си спрямо икономическия цикъл и стратегията си – да ги увеличи по време на растеж и да ги све по време на криза. Тази гъвкавост е ключово, но трудно за измерване с традиционни KPIs, стратегическо предимство.

Логиката на мащаба и неизбежната консолидация

Един от най-често пренебрегваните аспекти на AI ефективността е, че тя е почти изцяло зависима от **мащаба**. Както показва примерът с Netflix, препоръчителните системи са ефективни, защото разполагат с данни от милиони потребители. Бизнес моделът на **"надзорния капитализъм"**, описан от **Шошана Зубоф (2019)**, се основава на извличането на "поведенчески излишък" в планетарен мащаб. AI не е ефективен при малки обеми данни, неговата сила се крие в способността да открива модели в огромни масиви. Логиката на мащаба има неизбежно следствие - **AI е мощна сила за пазарна консолидация**. Малките и средни предприятия просто не разполагат с данните или капитала, за да се конкурират с технологичните гиганти и големите корпорации, които са изградили своите "data moats" (защитни ровове от данни). AI не е пазарен банасъор напротив, AI дава предимство още повече в полза на вече доминиращите играчи. За да оцелеят и да останат конкурентоспособни, много бизнеси ще бъдат принудени да се сливат и окрупняват, за да достигнат мащаба, необходим за ефективно прилагане на AI. Тази вълна на консолидация е може би най-значимият, но и най-рядко обсъждан дългосрочен ефект от внедряването на AI в икономиката.

Ефективността от AI трябва да се разглежда като дългосрочна стратегическа инвестиция, а не като краткосрочна тактика за увеличаване на продажбите. Лидерите трябва да се запитат не само "Какъв е нашият ROI този месец?", а и:

- Какви стратегически зависимости създаваме и как ги управляваме?
- Как промяната в структурата на разходите повишава нашата дългосрочна гъвкавост и устойчивост?
- Каква е нашата стратегия за постигане на необходимия мащаб – органичен растеж, партньорства или придобивания?
- Какви са етичните и репутационни рискове, които поемаме, и как ги смекчаваме?

Краткосрочните успехи са важни за поддържане на инерцията и за вътрешния и външен PR. Но те не трябва да заслепяват мениджърите за дълбоките, тектонични промени, които AI предизвиква. Оцеляването и просперитетът в ерата на изкуствения интелект ще зависят от способността да се гледа отвъд лесните отговори и да се управлява сложността на тази нова реалност.

Траектория на бизнес стойността на AI (ABVT) – авторски модел

Моделът **Траектория на бизнес стойността на AI (ABVT)** е динамична аналитична рамка, създадена с цел да даде цялостен поглед върху процеса, по който AI системите се развиват и генерират стойност – от първоначалните експерименти до мащабното им влияние върху бизнеса, икономиката и обществото.

Цели на модела

Основната цел на ABVT (AI Business Value Trajectory) е да замени статичното мислене за AI като еднократна инвестиция с **динамична перспектива**. Моделът помага на лидерите да:

- **Оценят зрелостта** на своите AI инициативи, като ги позиционират във времето и спрямо трите основни нива: технологично, организационно и макро ниво.
- **Идентифицират възможности и рискове**, свързани с всяка фаза на имплементация.
- **Разработят стратегически път**, който е съобразен с променящата се технологична и регулаторна среда.
- **Управляват очакванията**, като показват, че стойността на AI се натрупва постепенно и изисква комплексна трансформация на бизнеса.

Методология

Методологията на ABVT се основава на холистичен подход, който комбинира **количествен анализ (KPIs)** с **качествени наблюдения (примерни сценарии)**.

- **Триизмерен анализ.** Моделът анализира внедряването на AI едновременно на три нива, за да осигури пълноценна картина:
 - **Технологично ниво** - фокусира се върху техническите характеристики на AI решенията. За да се измери ефективността, се използват метрики като точност на моделите и процент на автоматизация, които са пряко свързани с производителността.
 - **Организационно ниво** - изследва как AI се интегрира в съществуващите структури и процеси. Ключово тук е измерването на промяната – колко бързо се адаптират екипите и как се създават нови роли в компанията.
 - **Макро ниво** - разглежда външните фактори, които влияят на AI. Това включва анализ на **регулациите** (напр. EU AI Act), **общественото доверие** и **геополитическите тенденции**. Тези фактори често са извън контрола на една компания, но са критични за дългосрочния ѝ успех.
- **Фази на имплементация.** Моделът не просто класифицира, а описва **еволюционен път** от пет фази. Всяка фаза се характеризира с нарастваща сложност и възвръщаемост на инвестициите, но и с повишени рискове. Този подход позволява на организациите да позиционират текущите си усилия и да планират следващите си стъпки.
- **Сценариен анализ.** ABVT включва и анализ на **оптимистични, прагматични и негативни сценарии**. Тази част от методологията е жизненоважна за стратегическото планиране. Тя подготвя лидерите за различни бъдещи възможности и им помага да разработят гъвкави стратегии за управление на потенциални кризи.

Чрез комбинирането на тези елементи, ABVT предоставя на бизнеса **практическа добавена стойност**, като предлага ясна рамка за стратегически дискусии и вземане на решения в ерата на AI.

Чеклист към модела ABVT (Траектория на бизнес стойността на AI)

Чеклиста е разработен, за да помогне оценката на текущото състояние на AI инициативите на компанията и да идентифицира следващите стъпки по траекторията на стойността. За всяко ниво и фаза, следва оценка доколко организацията отговаря на посочените критерии.

Част 1: Оценка на трите нива на зрялост

Технологично ниво

- **Начален етап:**
 - Имате ли поне един пилотен AI проект?
 - Използвате ли готови AI инструменти (напр. чатбот)?
 - Измервате ли базови метрики като процент автоматизация и скорост на отговор?
- **Междинен етап:**

- Разработвате ли собствени AI модели?
- Интегрирани ли са AI решенията в основните ви бизнес системи?
- Следите ли ефективността на моделите с метрики като точност и отклонение?
- **Напреднал етап:**
 - Експериментирате ли с генеративен AI или автономни агенти?
 - Могат ли вашите AI системи да се учат и да се подобряват сами?
 - Имате ли изградена вътрешна инфраструктура за AI, включително MLOps?

Организационно ниво

- **Начален етап:**
 - Имате ли малък екип, който работи по AI проекти?
 - Провеждат ли се обучения за основните ползи от AI?
 - Разбират ли лидерите как AI може да се приложи в техните отдели?
- **Междинен етап:**
 - Включен ли е AI в ключови бизнес процеси като маркетинг или продажби?
 - Имате ли стратегия за обучение на персонала, за да работи с AI инструменти?
 - Има ли дефинирани нови роли, свързани с AI (напр. **Data Scientist**)?
- **Напреднал етап:**
 - Използва ли се AI като стратегически ресурс за вземане на решения за бизнеса?
 - Имате ли **AI одиторив** структурата ви?
 - Създава ли AI нови бизнес модели, а не просто оптимизира съществуващите?

Макро ниво

- **Начален етап:**
 - Запознати ли сте с основните регулации като GDPR?
 - Обсъждате ли рисковете за сигурността и поверителността на данните?
 - Анализирате ли публичното мнение за AI във вашата индустрия?
- **Междинен етап:**
 - Имате ли план за съответствие с предстоящи регулации като **EU AI Act**?
 - Измервате ли общественото доверие в AI решенията си?
 - Инвестирате ли в киберсигурност, насочена към AI модели?
- **Напреднал етап:**
 - Водите ли дискусии с регулаторни органи или други компании за стандартизация?
 - Имате ли протоколи за отговорно внедряване на AGI-кандидати?
 - Подготвяте ли се за социални промени, свързани с AI, като например влиянието върху заетостта?

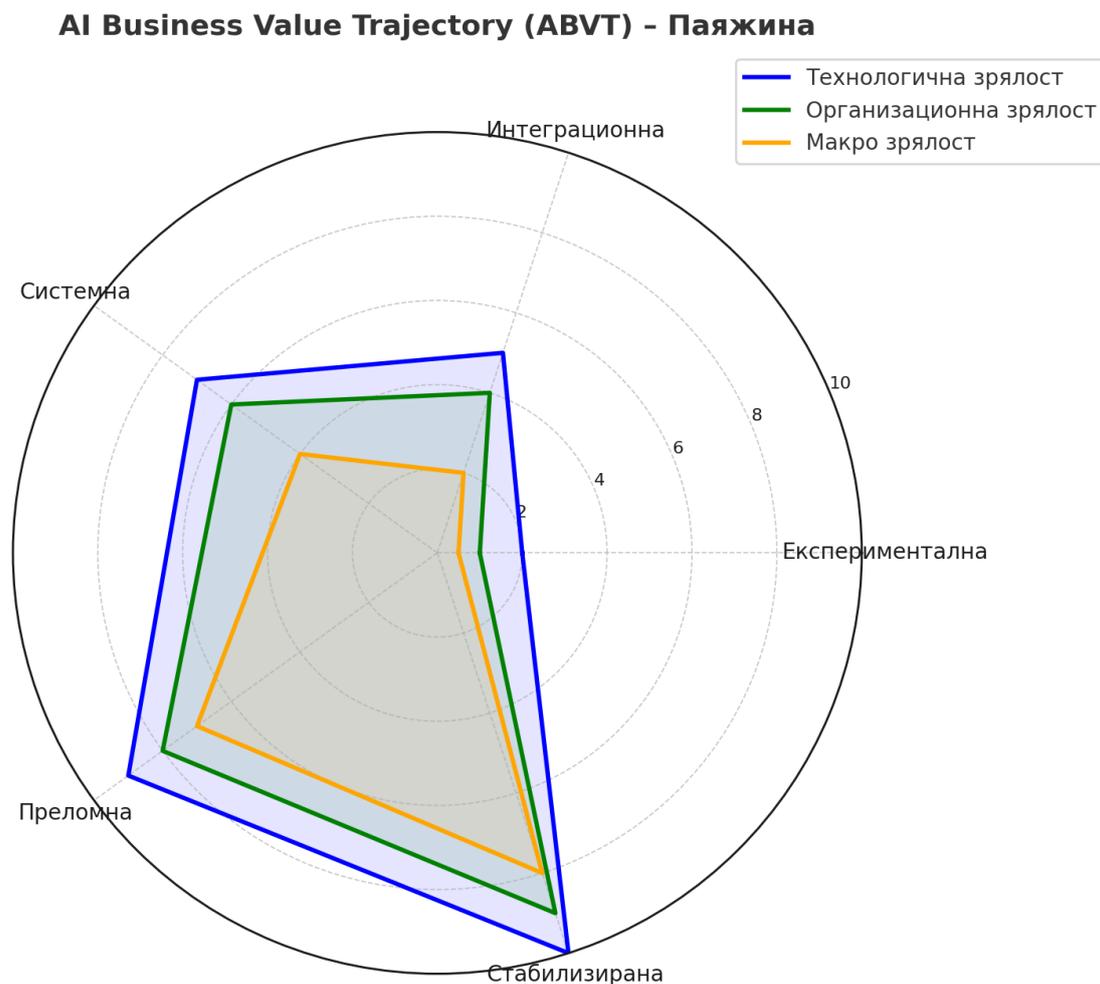
Част 2: Оценка на фазите на имплементация

- **Фаза 1: Експериментална**

- Разработили ли сте поне един пилотен проект с ограничен бюджет и ясна цел?
- **Фаза 2: Интеграционна**
 - Вградили ли сте успешно AI решения в поне един от основните си бизнес процеси?
- **Фаза 3: Системна**
 - Използвате ли AI, за да симулирате сложни сценарии или да оптимизирате цялостни процеси?
- **Фаза 4: Преломна/Определяща**
 - Обмисляте ли внедряването на AI модели, които могат да вземат автономни решения?
- **Фаза 5: Стабилизирана**
 - Разработвате ли вътрешни политики и стандарти за управление на напреднали AI системи?

Използвайте този чеклист като компас, за да определите къде се намира вашата организация по траекторията на бизнес стойността на AI и да планирате следващите си стъпки.

Фигура 13. Визуализация на чеклист за AI ABVT



Теми за дискусия Глава 17

- *Кои метрики са по-надеждни за дългосрочно управление – ROI или TCO (Total Cost of Ownership)?*
- *Как да избегнем капана на „PR пилоти“ и да измерим реална трансформация?*
- *Следва ли да се дефинират **нови AI-специфични KPI**, като скорост на моделна итерация или ниво на организационна адаптация?*
- *Как да се управлява рискът от vendor lock-in при използване на външни AI платформи?*
- *Може ли зависимостта от доставчици да се превърне в конкурентно предимство (напр. чрез ексклузивни партньорства)?*
- *Ще доведе ли AI до **олигополизация** на цели индустрии, където само най-големите играчи могат да извлекат стойност?*

18. Дигиталният разрив. Изоставането на България в ерата на AI и последиците за конкурентоспособността на бизнеса

"В глобална икономика, където технологията определя конкурентоспособността, дигиталното изоставане не е просто статистическа аномалия – то е съществена заплаха за икономическата съдба на нацията."

Марио Драги, бивш председател на ЕЦБ, The Future of European Competitiveness (2024)

В съвременната глобална икономика, дигиталната трансформация вече не е стратегическа опция, а императив за оцеляване. Компаниите, които не успяват да интегрират дигитални технологии и в частност изкуствен интелект (AI) в своите бизнес модели, рискуват да бъдат елиминирани от по-агресивни, технологично напреднали конкуренти. Това важи с особена сила за малките и средни предприятия (МСП), които формират гръбнака на европейската и българската икономика.

България се намира в парадоксална ситуация, страната разполага с отлична дигитална инфраструктура, динамичен ICT сектор и най-развита технологична екосистема в Югоизточна Европа (SeeNext Insights, 2025), но заема последно място в Европейския съюз по навлизането на цифровите технологии в икономиката и обществото според индекса DESI (Digital Economy and Society Index) на Европейската комисия (Европейска комисия, 2024). Този дисбаланс между техническите възможности и практическото приложение създава структурен недостиг в конкурентоспособността, който заплашва дългосрочната жизнеспособност на българския бизнес.

Настоящата глава анализира дълбочината на дигиталния разрив между България и останалите държави-членки на ЕС, изследва причините за това изоставане и оценява неговите последици върху рентабилността, конкурентоспособността и растежа на българските компании. Анализът се основава на официални статистически данни, академични изследвания и актуални доклади на водещи международни организации (OECD, МВФ, McKinsey, PwC).

Дигиталното изоставане на България и в частност ниското внедряване на AI не е просто технологичен проблем, а съществена заплаха за конкурентоспособността на целия бизнес, ограничава достъпа до международни пазари, намалява рентабилността и обрича българските компании на перманентна позиция на догонвачи в глобалната икономика.

Европейският дигитален пейзаж – неравномерно развитие и нарастващи различия

Европейският съюз е възприел дигиталната трансформация като стратегически приоритет още с публикуването на Цифровия компас към 2030 (Digital Compass 2030) през 2021 г. Документът определя четири ключови цели до 2030 г. (Европейска комисия, 2021):

1. Цифрови умения: Минимум 80% от населението на възраст 16-74 години да притежава базови цифрови умения.

2. Цифрова инфраструктура: Всички домакинства да имат гигабитова свързаност; всички населени места – 5G покритие.
3. Цифровизация на бизнеса: 75% от предприятията да използват облачни услуги, големи данни (big data) и изкуствен интелект; 90% от МСП да достигнат поне базово ниво на цифров интензитет.
4. Цифровизация на публичните услуги: 100% от ключовите обществени услуги да бъдат налични онлайн.

Постигането на тези цели е критично за запазването на конкурентоспособността на Европа спрямо САЩ и Китай. Според доклада на Марио Драги The Future of European Competitiveness (2024), ЕС изостава значително от САЩ по производителност на труда, технологични иновации и икономически растеж. Около 70% от разликата в БВП на глава от населението между ЕС и САЩ се обяснява с по-ниската производителност в Европа (Draghi, 2024). Дигитализацията и AI се разглеждат като основни инструменти за преодоляване на този разрыв.

Индексът DESI: Инструмент за измерване на дигиталния напредък

Индексът за навлизането на цифровите технологии в икономиката и обществото (DESI) е годишен доклад на Европейската комисия, който измерва дигиталното представяне на държавите-членки по четири основни стълба:

- Човешки капитал (Human Capital) – цифрови умения на населението и специалисти в ИТ.
- Свързаност (Connectivity) – наличие и качество на широколентови мрежи.
- Интеграция на дигитални технологии (Integration of Digital Technology) – използване на дигитални решения от бизнеса.
- Цифрови обществени услуги (Digital Public Services) – e-government, цифрово здравеопазване и др.

Резултатите от DESI 2024 разкриват съществени различия между държавите-членки:

Класация DESI 2024	Държава	Резултат (от 100)
1	Финландия	~70
2	Дания	~70
3	Нидерландия	~68
...
Среден за ЕС	–	52.6
26	Гърция	~42
27	Румъния	~38
28	България	~36

Източник: Европейска комисия, 2024; Eurostat, 2024

България заема последно място в ЕС, като резултатът ѝ е 31% под средния за Съюза. Въпреки че страната демонстрира годишен растеж от 9% през последните пет години,

този темп е недостатъчен за догонване на останалите държави-членки, които също се развиват (Европейска комисия, 2024). Анализът на DESI индекса разкрива **три ясно формирани групи** от държави в ЕС:

Група А: Цифрови лидери (>60 точки)

Включва североевропейски държави – **Финландия, Дания, Швеция, Холандия**, които се характеризират с високо ниво на цифрови умения, широко внедряване на технологии в бизнеса и отлични обществени дигитални услуги. Тези страни често се използват като **еталони (benchmarks)** за останалата част на Европа.

Група Б: Средноразвити (45-60 точки)

Обхваща западноевропейски държави като **Германия, Франция, Австрия**, както и някои държави от Централна Европа (**Естония, Чехия**). Тези страни показват **балансирано развитие** по различните стълбове на DESI.

Група В: Дигитално изоставащи (<45 точки)

Включва предимно държави от Южна и Източна Европа – **Гърция, Румъния, България**. Тези страни изпитват **системни предизвикателства** във всички измерения на дигитализацията.

България не просто заема последно място в DESI класацията – разликата с европейските лидери е съществена. Финландия vs. България: Разлика от ~34 точки (почти двойно по-високо представяне). Средно ниво за ЕС vs. България: Разлика от ~16 точки (31% по-ниско представяне), което означава, че българската икономика функционира на нивото на дигитализация, което европейските лидери са постигнали преди повече от десетилетие.

Критични слабости по стълбове на DESI:

Стълб	България	ЕС медиана	Разлика
Използване на интернет	67%	85%	-18 п.п.
Онлайн банкиране	13%	66%	-53 п.п.
Електронна търговия	~25%	~55%	-30 п.п.
Базови цифрови умения	~31%	~54%	-23 п.п.
МСП с базово ниво на цифров интензитет	25%	55%	-30 п.п.

Източник: Eurostat, 2024; Европейска комисия, 2024

Особено тревожен е фактът, че 24% от българите никога не са използвали интернет – най-високият процент в целия ЕС (Eurostat, 2024). Това създава структурен дефицит в човешкия капитал, който ограничава възможностите на бизнеса да внедрява дигитални решения. България демонстрира парадоксален профил: отлична инфраструктура, но слабо практическо приложение.

Силни страни:

- 5G покритие: България има най-доброто 5G покритие в региона на Югоизточна Европа (SeeNext Insights, 2025).
- ICT сектор: Генерира 8.2% от националната брутна добавена стойност (при средно за ЕС: 5.3%) (НСИ, 2024).
- Фиксирана широколентова свързаност: България е сред лидерите в ЕС по скорост на фиксираня интернет.
- Технологична екосистема: Най-развита в Югоизточна Европа, с активно венчърно инвестиране в AI (SeeNext Insights, 2025).

Слаби страни:

- Цифрови умения: Критично ниски сред населението.
- Бизнес трансформация: Само 10/15 точки в стълба "Бизнес трансформация и автоматизация" (SEE Digitalization 4.0).
- AI внедряване: Едва 6.47% от предприятията с 10+ служители използват AI (Eurostat, 2024).
- Културна съпротива: 47% от МСП не виждат смисъл от внедряване на AI (TechNews.bg, 2025).

Този парадокс може да се обясни с дисбаланса между инфраструктурни инвестиции и развитието на човешкия капитал, както и с липсата на стратегическо мислене в българския бизнес относно дигитализацията. Можем да обобщим, че България има най-силната инфраструктура, но най-слабата бизнес трансформация. Анализите по темата предполагат, че техническите възможности сами по себе си не са достатъчни – необходими са организационна култура, лидерство и стратегическо мислене за превръщането им в конкурентно предимство.

Интеграция на AI, новата граница на дигиталното развитие

Изкуственият интелект се превърна в основен двигател на глобалната икономика. Според McKinsey (2024), AI може да добави до 15% към глобалния БВП до 2035 г., или около \$15.7 трилиона (McKinsey Global Institute, 2023). Водещите икономики – САЩ и Китай – инвестират агресивно в AI:

- **САЩ:** \$109.1 млрд. частни инвестиции в AI (2024)
- **Китай:** \$9.3 млрд. частни инвестиции в AI (2024)
- **Великобритания:** \$4.5 млрд. частни инвестиции в AI (2024)

Източник: Stanford HAI AI Index Report, 2025

Европа изостава значително от САЩ и Китай по частни AI инвестиции, но води по регулаторна рамка (AI Act) и етични стандарти. Според доклада на OECD *The Adoption of Artificial Intelligence in Firms* (2025), използването на AI в Европа нараства, но остава фрагментирано по страни и сектори. Според Eurostat (2024), 13.5% от предприятията с 10+ служители в ЕС използват AI технологии. Това представлява 60% увеличение

спрямо 2023 г. (8.5%), което демонстрира ускорена адопция на AI, особено след появата на генеративните AI модели като ChatGPT (OpenAI, 2022).

Използване на AI по страни (2024):

Държава	Процент фирми, използващи AI	Категория
Дания	>25%	Лидер
Швеция	>25%	Лидер
Финландия	>22%	Лидер
Франция	~18%	Напреднал
Германия	~16%	Напреднал
ЕС среда	13.5%	–
Полша	<5%	Изоставащ
Румъния	<5%	Изоставащ
България	6.47%	Изоставащ

Източник: Eurostat, 2024

България изостава с 52% от средното за ЕС и 4 пъти от лидерите.

Глобална тенденция е приемането на AI силно да корелирана с размера на фирмата. Според OECD (2024), големите предприятия са 3.67 пъти по-склонни да използват AI спрямо малките:

Размер на фирма	AI адопция (ЕС 2024)
Големи (250+ служители)	41.17%
Средни (50-249 служители)	20.97%
Малки (10-49 служители)	11.21%

Източник: Eurostat, 2024; OECD, 2024

България следва същата тенденция, но на по-ниско ниво:

Размер на фирма	AI адопция (България 2024)	Разлика спрямо ЕС
Големи	~15-20% (оценка)	-21 до -26 п.п.
Средни	~8-10% (оценка)	-11 до -13 п.п.
Малки	3-6%	-5 до -8 п.п.

Източник: НСИ, 2024; експертни оценки на PwC Bulgaria, 2025

Това дигитално разделение между големи и малки фирми е особено проблематичен за България, тъй като 92% от българските предприятия са микро, малки или средни (НСИ, 2023). Следователно, огромната част от българския бизнес остава извън AI революцията. Използването на AI не е равномерно разпределено и по сектори. Според OECD (2025), най-високи нива на AI внедряване се наблюдават в:

Високо-технологични сектори:

- **ICT услуги:** 48.7% (ЕС среда)
- **Финансови услуги:** ~35%
- **Професионални, научни и технически дейности:** ~25%

Ниско-технологични сектори:

- **Настаняване и хранене:** 5.4% (OECD медиана)
- **Строителство:** ~7%
- **Търговия на дребно:** ~10%

Според анализ на PwC Югоизточна Европа (2025), в България най-активни в AI интеграцията са:

1. **Финтех:** Използване на AI за риск анализ, измами, чатботове
2. **Здравеопазване:** Пилотни проекти за диагностика и анализ на медицински данни
3. **Логистика:** Оптимизация на маршрути и управление на складове
4. **Търговия на дребно:** Динамично ценообразуване и таргетирана реклама

Въпреки това, общото приемане на AI в България остава критично ниска:

- **Използване на големи данни (Big Data):** Само 6% от фирмите (при ЕС цел: 75% до 2030)
- **Облачни услуги (Cloud Computing):** Само 10% от фирмите (при ЕС цел: 75% до 2030)
- **AI:** Само 3% от фирмите (при ЕС цел: 75% до 2030)

Източник: НСИ, 2024; Европейска комисия, 2024

Единственият сектор в България, който демонстрира световно ниво на AI използване, е IT индустрията. Според проучване на Dev.bg (2025) сред българските IT специалисти:

- 88% използват AI инструменти в работата си
- 60% използват AI ежедневно (основно за писане, редактиране и тестване на код)
- Основни инструменти: GitHub Copilot, ChatGPT, Claude, Cursor

Това създава огромен дисбаланс между IT сектора и останалата част от икономиката. IT специалистите са авангардът на дигиталната трансформация, но техните практики не се транслират в традиционните индустрии поради липса на дигитални умения, организационна култура и стратегическо лидерство.

Икономически последици от дигиталното изоставане

Връзката между дигитализация и икономически растеж е добре документирана в академичната литература. Brynjolfsson & McAfee (2014) твърдят, че информационните технологии са general-purpose technology (GPT) – т.е. технология, която трансформира широк спектър от индустрии и създава значителни производителни печалби

(Brynjolfsson & McAfee, 2014). Емпиричните изследвания потвърждават тази теза, Acemoglu & Restrepo (2020) твърдят, че Автоматизацията и AI водят до увеличение на производителността на труда, но и до дислокация на работна сила в краткосрочен план (Acemoglu & Restrepo, 2020). OECD (2019) - страните с по-висока цифрова зрялост демонстрират по-висок растеж на общата факторна производителност (TFP) (OECD, 2019). European Commission (2020) - Пълната дигитализация на ЕС може да добави €2.2 трилиона към БВП до 2030 г. (Европейска комисия, 2020). Дигиталните технологии намаляват транзакционните разходи, подобряват координацията, ускоряват иновациите и позволяват мащабиране на бизнес модели (Evans & Gawer, 2016).

Множество изследвания оценяват макроикономическото въздействие на AI:

McKinsey Global Institute (2023):

- AI може да добави **\$13 до \$15.7 трилиона към глобалния БВП до 2030 г.**
- Това представлява **увеличение от 1.2% годишно в глобалния растеж**
- **Лидерите в AI** ще се ползват от **20-25% по-висок растеж** спрямо изоставащите

Goldman Sachs (2023):

- AI може да **увеличи глобалния БВП с 7% за десетгодишен период**
- **Производителността на труда** може да нарасне с **1.5% годишно**

IMF (2025) – *Artificial Intelligence and Productivity in Europe:*

- Оценява потенциален годишен TFP растеж от **0.4 до 1.3 процентни пункта** в страни с висока AI експозиция (като САЩ)
- За **Европа:** Очакван TFP растеж от **0.07 до 1.1 процентни пункта годишно**
- **България:** При активно внедряване на AI, потенциалният годишен растеж може да достигне до **1.3%** (IMF, 2025)

Критичен извод, е че страните, които активно внедряват AI, ще реализират значителни производствени печалби, докато изоставащите ще изпитват нарастващо изоставане в конкурентоспособността.

Цената на изоставането, сценарии за внедряване на AI в бизнеса и икономически ефекти за България

Сценарий 1: Статуквото (Business-as-Usual)

При сегашния темп на AI внедряване (6.47% от фирмите), България няма да постигне целите на ЕС до 2030 г. (75% от фирмите да използват AI). Разликата в производителността на труда спрямо европейските лидери ще продължи да се увеличава. Количествена оценка на загубата: ако приемем, че пълната дигитализация може да добави 8% към БВП на ЕС до 2035 г. (Implement Consulting, 2024), а България

постига само 25% от този потенциал поради ниското внедряване, загубата за българската икономика е:

- БВП на България (2024): ~€90 млрд.
- Потенциална печалба при пълна дигитализация: €90 млрд. × 8% = €7.2 млрд.
- Реализирана печалба при 25% внедряване: €7.2 млрд. × 25% = €1.8 млрд.
- Загубена възможност: €7.2 млрд. – €1.8 млрд. = €5.4 млрд. годишно

Това е 6% от БВП, което е еквивалентно на годишните преки инвестиции в България.

Сценарий 2: Активна дигитализация (Catch-Up)

При агресивна стратегия за AI внедряване, при която България достига европейската средна (13.5%) до 2027 г. и 50% AI адопция до 2030 г., потенциалните печалби са:

- Годишен TFP растеж: +0.6 до +0.8 процентни пункта
- Кумулативна печалба до 2035 г.: €8-10 млрд.

На корпоративно ниво, AI внедряването води до измерими подобрения в рентабилността:

PwC (2025) – доклад "Value in Motion":

- Компании, които са внедрили AI, отчитат 15-20% по-ниски разходи
- Спестяване на разходи: 20-30% при правилно внедряване (PwC, 2025)

McKinsey (2024):

- Lighthouse factories (фабрики с напълно интегриран AI) постигат:
 - До 40% увеличение в производителността на труда
 - До 50% намаляване на lead times
 - 20-30% намаляване на дефектите

Дигитализация и конкурентоспособност

Майкъл Портьър (Porter, 1985) дефинира конкурентоспособността като способността на фирмата да създава и задържа конкурентно предимство чрез по-ниски разходи или диференциация (Porter, 1985). В дигиталната ера, тези две стратегии са тясно свързани с използването на технологии:

- Разходно лидерство (Cost Leadership), постига се чрез автоматизация, оптимизация на веригите на доставки и предиктивна аналитика (AI-driven cost reduction).
- Диференциация, постига се чрез персонализация, по-добро клиентско обслужване и иновативни продукти (AI-enabled differentiation).

Brynjolfsson, Rock & Syverson (2021) твърдят, че AI е "intangible capital" – нематериален актив, който мултиплицира ефекта на други инвестиции и създава трайно конкурентно предимство (Brynjolfsson, Rock & Syverson, 2021). Глобалният индекс за конкурентоспособност (Global Competitiveness Index) на Световния икономически форум (WEF) включва цифрова готовност като ключов стълб за оценка на националната конкурентоспособност (WEF, 2019). Страните с по-високо ниво на дигитализация демонстрират:

- По-висока производителност на труда
- По-висок БВП на глава от населението
- По-големи преки чуждестранни инвестиции (FDI)

Емпирични доказателства на теоретичните твърдения на авторите са скандинавските страни Дания, Швеция, Финландия, които заемат водещи позиции в индекса DESI и имат БВП на глава от населението >€50 000, което е двойно повече от България (~€25 000). Друг пример е Естония - малка държава с агресивна дигитална стратегия (e-Residency, e-Government) привлича високотехнологични инвестиции и постига годишен растеж над 3%.

Фирми от дигитално изоставащи страни изпитват структурни недостатъци при конкуренция на международните пазари:

1. **По-високи разходи:** Липсата на автоматизация води до по-високи оперативни разходи.
2. **По-бавни процеси:** Ръчни процеси забавят времето на доставка на продукти (time-to-market).
3. **По-ниско качество:** Липсата на управляван от AI контрол води до по-висок процент дефекти.
4. **По-слабо обслужване:** Липсата на AI чатботове и предиктивна аналитика води до по-лошо клиентско изживяване.

Глобалните вериги на стойност (Global Value Chains – GVCs) изискват високо ниво на дигитална интеграция. Водещите multinational corporations (MNCs) като Amazon, Walmart, Siemens изискват от доставчиците си:

- Проследяване на стокити/инвентара в реално време Real-time inventory tracking (IoT и облачни системи)
- AI предиктивни анализи за прогнозиране на търсенето (predictive analytics)
- Автоматизиран контрол на качеството (Automated quality control, computer vision AI)
- Дигитално фактуриране и използване на блокчейн смарт договори (Digital invoicing and blockchain) за проследимост

Български фирми, които не отговарят на тези изисквания, са изключени от GVCs. Според OECD Trade in Value Added (TiVA) database (2023), българското участие в GVCs

е ниско – само 45% от експорта е част от глобални вериги, спрямо 60-70% за Чехия и Полша. Основната причина: ниска дигитална зрялост на българските доставчици.

Дигиталното изоставане на България може да се обясни с три взаимно усилващи се дефицита:

1. Дефицит на човешки капитал (Human Capital Deficit)

- 31% от българите имат базови цифрови умения (при ЕС: 54%)
- Специалисти в ИКТ: Само 3.9% от работната сила (при ЕС: 4.8%)
- AI експерти: Критичен недостиг (според LinkedIn Talent Insights, България има 10 пъти по-малко AI специалисти на глава от населението спрямо Великобритания)

2. Дефицити в организационна култура (Organizational Culture Deficit)

- 47% от МСП не виждат стойност в AI (TechNews.bg, 2025)
- Липса на стратегическо мислене: Само 15% от българските фирми имат AI стратегия (експертна оценка, PwC Bulgaria)
- Съпротива срещу промяна: Особено в традиционни индустрии (земеделие, строителство)

3. Дефицит на финансиране (Financing Deficit)

- Венчърни инвестиции в AI: България привлича <€50 млн. годишно (Dealroom.co, 2024), спрямо €4.5 млрд. във Великобритания
- R&D разходи: България инвестира 0.84% от БВП в R&D (НСИ, 2023), което е 3 пъти по-малко от средното за ЕС (2.3%)

Тези три дефицита се засилват взаимно: липсата на умения води до ниска иновативност, която води до ниски инвестиции, които водят до липса на обучение – порочен кръг на изоставане. Въпреки, че България е приела "Концепция за развитие на изкуствения интелект в България до 2030 г." (МС, 2021), но имплементацията е бавна. Липсват:

- Конкретни цели и KPIs
- Бюджет и ресурси
- Координационни механизми между министерства

България е силно зависима и от чуждестранни технологични доставчици:

- Облачни услуги: Почти изцяло от Amazon AWS, Microsoft Azure, Google Cloud
- AI модели: OpenAI, Anthropic, Google
- ERP системи: SAP, Oracle – основно чуждестранни решения

Това създава стратегически риск: при геополитически напрежения или промени в ценообразуването, българските фирми нямат алтернативи.

Последици за българския бизнес и сценарии за развитие

Сценарий А: "Статукво" (Business-as-Usual)

Предпоставки:

- AI приемането расте с 2-3 процентни пункта годишно (текущ темп)
- Държавата не предприема значими действия за стимулиране на дигитализацията
- МСП запазват резистентност към дигитални технологии

Резултати до 2030 г.:

- AI приемане: ~15% от фирмите (все още под ЕС средното от 2024 г.)
- DESI класация: България остава на последни позиции
- Производителност на труда: Расте с 1-1.5% годишно (под ЕС средното от 2-2.5%)
- Експортна конкурентоспособност: Намалява, български фирми губят пазарни дялове

Икономически ефект:

- Загубени възможности: €5-6 млрд. годишно
- Отлив на таланти: 10,000+ AI специалисти емигрират годишно
- Затваряне на фирми: Неконкурентни компании обявяват несъстоятелност

Социални последици:

- Увеличаване на безработицата в традиционни индустрии
- Нарастване на неравенството между IT и non-IT сектори
- Демографска криза: Млади, образовани хора напускат страната

Сценарий Б: "Частично догонване" (Partial Catch-Up)

Предпоставки:

- Държавата активира програми за дигитализация на МСП (ваучери, обучения)
- София Tech Park AI Factory привлича €90 млн. и създава 500 AI работни места
- Големите компании инвестират в AI, МСП остават изоставащи

Резултати до 2030 г.:

- AI приемане: ~25% от фирмите (достигане на ЕС средното от 2024 г.)
- DESI класация: България се изкачва на 22-24 място
- Производителност на труда: Расте с 2-2.5% годишно
- Експортна конкурентоспособност: Стабилизира се, но не догонва лидерите

Икономически ефект:

- Реализирани възможности: €2-3 млрд. годишно
- Задържане на таланти: 50% от AI специалистите остават в България
- Възникване на AI champions: 5-10 български компании стават регионални лидери

Предизвикателства:

- Все още съществен гар спрямо лидерите
- МСП остават изоставащи – 80% от фирмите не използват AI

Сценарий В: "Пълно догонване" (Full Catch-Up)

Предпоставки:

- Национална AI стратегия с бюджет от €500 млн. (2025-2030)
- Задължително AI обучение за всички средни училища
- Данъчни стимули за AI инвестиции (20% данъчен кредит за AI R&D)
- Публично-частно партньорство (PPP) за AI инфраструктура

Резултати до 2030 г.:

- AI приемане: ~50% от фирмите (над ЕС средното)
- DESI класация: България влиза в топ 15
- Производителност на труда: Расте с 3-4% годишно
- Експортна конкурентоспособност: България става регионален AI хъб

Икономически ефект:

- Реализирани възможности: €5-7 млрд. годишно
- Приток на таланти: Български емигранти се връщат (brain gain)
- Unicorns: Появяват се 2-3 български AI компании с оценка над \$1 млрд.

Изисквания:

- Политическа воля и непрекъснатост на политиката
- Машабни инвестиции в образование и инфраструктура
- Културна промяна в бизнес мисленето

Дигиталното изоставане на България не е статистическа аномалия, а структурен проблем, който заплашва дългосрочната конкурентоспособност на българския бизнес. Въпреки отличната инфраструктура и динамичния IT сектор, страната заема последно място в ЕС по навлизането на цифровите технологии в икономиката (DESI, 2024) и изостава четворно от европейските лидери по AI внедряване (6.47% vs. 25%+).

Последиците от това изоставане са сериозни и количествени:

1. Загубени възможности: До €5-6 млрд. годишно (6% от БВП) при запазване на статуквото.
2. Намаляваща конкурентоспособност: Български фирми губят пазарни дялове на международните пазари поради по-високи разходи и по-бавни процеси.
3. Ограничен достъп до GVCs: Липсата на дигитална зрялост изключва българските компании от глобалните вериги на стойност.
4. Отлив на таланти: 10,000+ AI специалисти емигрират годишно към по-развити икономики.
5. Социално неравенство: Нарастващ гар между IT и non-IT сектори създава "две скорости" в икономиката.

Но България има и силни страни, които могат да бъдат използвани като трамплин за дигитален скок:

- Най-доброто 5G покритие в региона
- Силен ICT сектор (8.2% от БВС)
- Най-развита технологична екосистема в Югоизточна Европа
- София Tech Park AI Factory с €90 млн. инвестиция

България може да избере статуквото и да се примири с перманентно изоставане, или може да избере активна дигитализация и да се превърне в регионален AI хъб. Разликата между двата сценария е €5-7 млрд. годишно – или 10% от БВП. Всяка година на изоставане увеличава разликата с лидерите. Както заяви Марио Драги в доклада си за европейската конкурентоспособност: „Европа се изправя пред избор: да се трансформира, или да изчезне бавно.“ (Draghi, 2024). Това важи с още по-голяма сила за България.

Теми за дискусия Глава 18

- *Дали дигиталното изоставане на България е неизбежно, или може да бъде преодоляно чрез правилни политики? Защо да, или защо не?*
- *Кои сектори в българската икономика имат най-голям потенциал за AI трансформация и защо? (Разгледайте финтех, земеделие, туризъм, производство)*
- *Парадоксът на България: Отлична инфраструктура, слабо приложение. Какви са основните причини за този дисбаланс? (Човешки капитал? Култура? Регулация?)*
- *Дали малките и средните предприятия могат да си позволят AI, или това е привилегия на големите корпорации? Обсъдете модели за достъпност (SaaS, консорциуми, публични облаци).*
- *Сценариите за развитие: Кой сценарий е най-вероятен за България до 2030 г. – Статукво, Частично догонване, или Пълно догонване? Аргументирайте.*

Заклучение. Бъдещето е сега

Изминахме пътешествие през най-фундаменталните аспекти на изкуствения интелект – от неговата история и революционен потенциал, през икономическите му ефекти, до етичните предизвикателства и практическите му приложения. Тази книга дава карта на терен, който се променя с всеки изминал ден. Тя не е просто наръчник, а призив за действие, който обобщава няколко ключови истини:

Уроците на миналото са нашето най-силно оръжие

Видяхме, че провалите на AI проекти като IBM Watson Health и Zillow Offers са толкова ценни, колкото и успехите. Те ни научиха, че технологията сама по себе си не е достатъчна. Успехът се крие в **правилното дефиниране на проблема**, наличието на **качествени данни** и силна **етична рамка**. Познаването на симптомите на провала е първата стъпка към изграждането на устойчиви и надеждни AI системи.

AI е най-мощният двигател на промяната

Разбрахме как изкуственият интелект не просто автоматизира процеси, а пренаписва основните правила на икономиката. От **персонализация** на мащаб до създаване на **нови източници на приходи** и оптимизация на разходите, AI превръща бизнеса в динамичен и интелигентен организъм. Той променя самата същност на конкуренцията, като поставя акцент върху данни, скорост и иновации.

Препоръки за успеха в ерата на AI

За бизнеса

- **Инвестирайте в хората, а не само в технологиите.** Успешното внедряване на AI изисква **културна трансформация**, която приема промяната. Обучавайте екипите си, насърчавайте експериментите и създавайте среда, в която грешките се разглеждат като възможност за учене.
- **Започнете с малки, но важни проекти.** Не се опитвайте да решите всички проблеми наведнъж. Идентифицирайте конкретни бизнес казуси, където AI може да добави

реална стойност (например, автоматизация на рутинни задачи или подобряване на клиентското преживяване) и след това постепенно мащабирайте.

- **Етиката и регулацията са ваше конкурентно предимство.** Прозрачността и отговорното използване на данни ще изградят доверие у клиентите и ще ви предпазят от правни рискове.

За студентите

- **Приемете ученето за начин на живот.** Най-ценното умение в ерата на AI не е знанието на един конкретен език за програмиране, а способността да се адаптирате и да учите нови неща постоянно.
- **Изграждайте интердисциплинарни умения.** Бъдещите лидери ще бъдат тези, които разбират както технологиите, така и бизнеса, етиката и социалните аспекти на AI. Не се ограничавайте в една хипотеза или специалност.
- **Не се страхувайте да експериментирате.** AI е област, която се развива бързо. Използвайте всеки шанс да създадете прототип, да работите по реален проект и да се учите от своите грешки.

За хората

- **Бъдете адаптивни и отворени към промяната.** Не гледайте на AI като на заплаха, а като на инструмент, който може да ви помогне да бъдете по-ефективни и креативни. Роботите няма да заменят хората, но хората, които използват AI, ще заменят тези, които не го правят.
- **Развивайте уникалните си човешки умения.** AI ще автоматизира рутинните задачи, което ще ни позволи да се фокусираме върху неща, в които машините не са добри – **критично мислене, емпатия, креативност и емоционална интелигентност.**
- **Подгответе се за бъдещето.** Не чакайте промяната да дойде, а бъдете част от нея. Включете AI в своето ежедневие и работа, за да сте успешен участник в новата индустриална ера.

Бъдещето на индустрията е в ръцете на тези, които не се страхуват от промяната, а я приемат с отворено съзнание. Нека тази книга е само началото на вашето вълнуващо приключение в света на изкуствения интелект. Успех!

БИБЛИОГРАФИЯ

Монографии и книги

1. Agrawal, A., J. Gans, A. Goldfarb. Prediction Machines: The Simple Economics of Artificial Intelligence. Cambridge, MA: Harvard Business Review Press, 2018.
2. Anderson, C. The Long Tail: Why the Future of Business is Selling Less of More. New York: Hyperion, 2006.
3. Barto, A. G., R. S. Sutton. Reinforcement Learning: An Introduction. 2nd ed. Cambridge, MA: The MIT Press, 2018.
4. Berger, J. Contagious: Why Things Catch On. New York: Simon & Schuster, 2013.
5. Breazeal, C. Designing Sociable Robots. Cambridge, MA: MIT Press, 2002.
6. Brynjolfsson, E., A. McAfee. The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. New York: W. W. Norton & Company, 2014.
7. Christensen, C. M. The Innovator's Dilemma: When New Technologies Cause Great Firms to Fail. Boston: Harvard Business School Press, 1997.
8. Drucker, P. F. Innovation and Entrepreneurship: Practice and Principles. New York: Harper & Row, 1985.
9. Drucker, P. F. Management Challenges for the 21st Century. New York: HarperBusiness, 1999.
10. Hastie, T., R. Tibshirani, J. Friedman. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. New York: Springer, 2009.
11. Iansiti, M., K. R. Lakhani. Competing in the Age of AI. Boston: Harvard Business Review Press, 2020.
12. Knight, F. H. Risk, Uncertainty and Profit. Boston: Houghton Mifflin, 1921.
13. Kotler, P., H. Kartajaya, I. Setiawan. Marketing 5.0: Technology for Humanity. Hoboken: Wiley, 2021.
14. Marcus, G. Rebooting AI: Building Artificial Intelligence We Can Trust. New York: Pantheon, 2022.
15. Mitchell, T. M. Machine Learning. New York: McGraw Hill, 1997.
16. Moore, G. A. Crossing the Chasm: Marketing and Selling Technology Products to Mainstream Customers. New York: HarperBusiness, 1991.
17. O'Neil, C. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown, 2016.
18. Pariser, E. The Filter Bubble: What the Internet Is Hiding from You. London: Penguin UK, 2011.
19. Picard, R. W. Affective Computing. Cambridge, MA: MIT Press, 1997.

20. Poole, D., A. Mackworth. *AI: Foundations of Computational Agents*. 3rd ed. Cambridge: Cambridge University Press, 2023.
21. Porter, M. E. *Competitive Advantage: Creating and Sustaining Superior Performance*. New York: Free Press, 1985.
22. Reeves, B., C. Nass. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. Cambridge: Cambridge University Press, 1996.
23. Rogers, E. M. *Diffusion of Innovations*. New York: Free Press, 1962.
24. Russell, S., P. Norvig. *Artificial Intelligence: A Modern Approach*. 3rd ed. Upper Saddle River: Prentice Hall, 2010.
25. Schumpeter, J. A. *Capitalism, Socialism and Democracy*. New York: Harper & Brothers, 1942.
26. Schwartz, B. *The Paradox of Choice: Why More Is Less*. New York: Harper Perennial, 2004.
27. Tapscott, D., A. Tapscott. *Blockchain Revolution: How the Technology Behind Bitcoin Is Changing Money, Business, and the World*. New York: Portfolio, 2016.
28. Tauter, K., P. DiMaggio. *The Digital Transformation of Banking: Exploring Technological and Organizational Change*. London: Routledge, 2020.
29. Turkle, S. *Alone Together: Why We Expect More from Technology and Less from Each Other*. New York: Basic Books, 2011.
30. Zuboff, S. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs, 2019.

Статии в научни списания

31. Acemoglu, D. *Is AI About to Revolutionize Productivity?* Project Syndicate, 2023.
32. Ayers, J. W. et al. *Comparing Physician and Artificial Intelligence Chatbot Responses to Patient Questions Posted to a Public Social Media Forum*. *JAMA Internal Medicine*, 2023.
33. Barney, J. *Firm Resources and Sustained Competitive Advantage*. *Journal of Management*, 1991, vol. 17, № 1, pp. 99–120.
34. Bender, E. M., B. Friedman. *Data statements for natural language processing*. *Transactions of the Association for Computational Linguistics (TACL)*, 2018, vol. 6, pp. 123–135.
35. Bengio, Y. *Learning deep architectures for AI*. *Foundations and Trends in Machine Learning*, 2009, vol. 2, № 1, pp. 1–127.
36. Bresnahan, T. F., M. Trajtenberg. *General Purpose Technologies 'Engines of Growth'?* *Journal of Econometrics*, 1995, vol. 65, № 1, pp. 83–108.
37. Chen, H. et al. *Large Language Models in Mental Health Support: A Systematic Review*. *JMIR Mental Health*, 2024, vol. 11, № 1.

38. Cooper, R. G. AI and New Product Development. *Research-Technology Management*, 2024, vol. 67, № 1, pp. 22–34.
39. Darling, K. Law, Robotics, and the Morality of Artificial Empathy. SSRN, 2016.
40. Davis, F. D. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 1989, vol. 13, № 3, pp. 319–340.
41. Duncan, R. B. The Ambidextrous Organization: Designing Dual Structures for Innovation. *Organizational Dynamics*, 1976, vol. 5, № 4, pp. 2–17.
42. Gomez-Uribe, C. A., N. Hunt. The Netflix Recommender System: Algorithms, Business Value, and Innovation. *ACM Transactions on Management Information Systems (TMIS)*, 2015, vol. 6, № 4, pp. 1–19.
43. Gray, H. M., K. Gray, D. M. Wegner. Dimensions of Mind Perception. *Science*, 2007, vol. 315, № 5819, p. 1618.
44. Hayes, R. H., S. C. Wheelwright. Link Manufacturing Process and Product Life Cycles. *Harvard Business Review*, 1979, vol. 57, № 1, pp. 133–140.
45. Hayes, R. H., S. C. Wheelwright. The Dynamics of Process-Product Life Cycles. *Harvard Business Review*, 1979, vol. 57, № 2, pp. 127–136.
46. Iansiti, M., K. R. Lakhani. The Truth About Blockchain. *Harvard Business Review*, 2017, vol. 95, № 1, pp. 118–127.
47. Kahneman, L. M. et al. The Effects of Generative AI on Worker Productivity (NBER Working Paper). *The Quarterly Journal of Economics*, 2025.
48. LeCun, Y. et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 1998, vol. 86, № 11, pp. 2278–2324.
49. Linden, G., B. Smith, J. York. Amazon.com Recommendations: Item-to-Item Collaborative Filtering. *IEEE Internet Computing*, 2003.
50. Meyer, J. G. et al. ChatGPT and large language models in academia. *BioData Mining*, 2023, vol. 16, № 1.
51. Mitchell, M. et al. The debate over understanding in AI's large language models. *PNAS*, 2023, vol. 120, № 24.
52. Rochet, J.-C., J. Tirole. Platform Competition in Two-Sided Markets. *Journal of the European Economic Association*, 2003, vol. 1, № 4, pp. 990–1029.
53. Rochet, J.-C., J. Tirole. Two-Sided Markets: A Progress Report. *The RAND Journal of Economics*, 2006, vol. 37, № 3, pp. 645–667.
54. Schröder, T. et al. Monitoring ML models: A categorization of challenges. *Patterns*, 2022, vol. 3, № 11, p. 100587.
55. Seitz, H. Artificial empathy in healthcare chatbots: A critical discussion of authenticity and manipulation. *Ethics and Information Technology*, 2024.

56. Sharkey, N., A. Sharkey. Social implications of robot deception and manipulation. *Ethics and Information Technology*, 2020, vol. 22, № 4, pp. 313–324.
57. Sharon, T. Techno-solutionism and the overlooked risks of 'empathic' medical chatbots. SpringerLink, 2025.
58. Silver, D. et al. Mastering the game of Go with deep neural networks and tree search. *Nature*, 2016, vol. 529, № 7587, pp. 484–489.
59. Sohail, A. et al. Using large language models to facilitate academic work. PMC, 2025.
60. Sull, D. N. Why Good Companies Go Bad. *Harvard Business Review*, 1999, vol. 77, № 4, pp. 42–52.
61. Turing, A. M. Computing Machinery and Intelligence. *Mind*, 1950, vol. 59, № 236, pp. 433–460.
62. Wieczorek, O. The Bot Delusion: Large language models and anticipated scientific effects. ScienceDirect, 2025.
63. Zsifkovits, M., M. Taisch. The Product Customization Process in Relation to Industry 4.0 and Digitalization. *Applied Sciences*, 2021, vol. 11, № 3, p. 539.

Доклади и изследвания

64. Bank for International Settlements. Governance of AI adoption in central banks. Basel: BIS, 2025.
65. Chui, M. et al. The economic potential of generative AI: The next productivity frontier. McKinsey & Company, 2023.
66. Davenport, T. H., R. Ronanki. Artificial intelligence for the real world. *Harvard Business Review*, Jan–Feb 2018.
67. EU High-Level Expert Group on AI. Trustworthy AI. Brussels: European Commission, 2019.
68. Farayola, O. A. et al. Innovative Business Models Driven by AI Technologies: A Review. 2023.
69. Jacobides, M. G., M. D. Ma. Assessing the expected impact of Generative AI on the UK competitive landscape. Institute of Directors (IoD) & London Business School Policy Paper, May 2024.
70. Jordan, M. I. Artificial Intelligence — The Revolution Hasn't Happened Yet. EECS at Berkeley (Essays and Commentaries), 2019.
71. McKinsey & Company. The Economic Potential of Generative AI: The Next Productivity Frontier. McKinsey Global Institute, 2023.
72. McKinsey & Company. Technology Trends Outlook 2025. McKinsey & Company, 2025.
73. McKinsey & Company. Beyond the Hype: Capturing the Potential of AI & GenAI in TMT. McKinsey & Company, 2024.

74. McKinsey & Company. The state of AI in 2024. McKinsey & Company, 2024.
75. McKinsey & Company. The state of AI: Global survey. McKinsey & Company, 2025.
76. MIT Sloan Management Review & BCG. Artificial Intelligence and Business Strategy reports. 2022–2024.
77. Reim, W. et al. Implementation of Artificial Intelligence (AI): A Roadmap for Business Model Innovation. 2020.
78. Singla, A. et al. Beyond the hype: Capturing the potential of AI & gen AI in TMT. McKinsey & Company, February 2024.
79. Stanford HAI. AI Index Report 2025. Stanford: Stanford Human-Centered AI Institute, 2025.

Конференции и сборники

80. Bender, E. M. et al. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 2021.
81. Breck, E. et al. The ML Test Score: A rubric for ML production readiness. In: IEEE Big Data, 2017.
82. Gebru, T. et al. Datasheets for datasets. arXiv:1803.09010, 2018.
83. Holland, S. et al. The dataset nutrition label. arXiv:1805.03911, 2018.
84. Mitchell, M. et al. Model cards for model reporting. In: Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT), 2019, pp. 220–229.
85. Sculley, D. et al. Hidden technical debt in machine learning systems. In: Proceedings of the 28th International Conference on Neural Information Processing Systems (NeurIPS), 2015, pp. 2503–2511.
86. Brown, T. B. et al. Language Models are Few-Shot Learners. arXiv preprint arXiv:2005.14165, 2020.
87. Colombo, L. et al. SaulLM-7B: legal domain LLM. arXiv, 2024.
88. Huang, L. et al. A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions. arXiv, 2024.
89. Kalyan, S. S. et al. A Survey of GPT-3 Family Large Language Models Including ChatGPT and GPT-4. arXiv, 2023.
90. Naveed, H. et al. A Comprehensive Overview of Large Language Models. arXiv, 2023.
91. Sakib, M. N. et al. Risks, Causes, and Mitigations of Widespread Deployments of Large Language Models (LLMs): A Survey. arXiv, 1 август 2024.
92. Welivita, A., P. Pu. Are Large Language Models more Empathetic than Humans? A Comparison of Empathetic Responses. arXiv preprint arXiv:2401.03748, 2024.

93. Xie, Y. et al. A Methodology for Scaling and Explainability in LLM Empathy Evaluation. arXiv preprint, 2024.

94. Yuan, Z. et al. Code Llama: Open Foundation Models for Code. arXiv preprint arXiv:2308.12950, 2023.

95. Zhao, W. X. et al. A Survey of Large Language Models. arXiv, 2023.

Регулативни документи

96. European Commission. EU Artificial Intelligence Act – Questions & Answers. Brussels: European Commission, 2024.

97. European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data (GDPR). Official Journal of the European Union, L 119, 2016.

98. European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (AI Act). Official Journal of the European Union, L 2024/1689, 2024.

99. NIST. Artificial Intelligence Risk Management Framework (AI RMF 1.0). Gaithersburg: National Institute of Standards and Technology, 2023.

100. NIST. Generative AI Profile (AI RMF Companion). Gaithersburg: National Institute of Standards and Technology, 2024.

101. OECD. Recommendation of the Council on Artificial Intelligence; OECD AI Principles. Paris: OECD, 2019.

Корпоративни източници и изследвания

102. Accenture. AI and the Future of Work: данни за намаляване на разходите, цитирани в текста.

103. Amazon. Reuters Report: AI Hiring Tool Found to Be Biased Against Women.

104. DBS Bank. Annual Report: AI/ML Use Cases. Singapore: DBS Bank, 2024.

105. Gartner. AI TRiSM overview and guidance. Stamford: Gartner, 2023–2024.

106. Gartner. AI in logistics and inventory optimization. Stamford: Gartner, 2024.

107. IBM. AI FactSheets: Increasing trust in AI services. IBM Journal of Research and Development, 2018–2019.

108. Pfizer. Virtsant Case Study: AI Drug Discovery Cuts Years Off Development Time.

109. Shell Global. Predictive Maintenance with AI – Case Study.

110. Stripe. Radar Fraud Detection – Expanding to ACH and SEPA.

111. UPS. AI-Driven Logistics Optimization.

112. Visa. Advanced Authorization Fraud Detection System.

Блог публикации и онлайн ресурси

113. Credo AI. The difference between MLOps & AI Governance and why you need both to deliver on Responsible AI. Credo AI Company Blog, 1 декември 2022.
114. Darnell, D. MLOps and AI Governance, A Difference of Perspective. Dataiku blog, 21 юни 2023.
115. Morris, B. Synthesia: Can a Synthetic Media Company Prevent the Use of Their Tech for Deepfaking? Should It? 2025.
116. Rieben, R. LLM & Gen-AI Risk Management: Guide to Navigating Security. Linford & Company LLP, 10 април 2024.
117. Trigyn Technologies. Strategies to Manage Risk in Enterprise LLM Implementations. Trigyn Technologies, 20 ноември 2024.
118. Rowland, S., D. Weik. Generative AI Governance Emerging. MLOps Podcast #247 [Аудио подкаст], 17 юли 2024.

Академични институции и изследователски центрове

119. American Bar Association. Navigating the AI Employment Bias Maze. Chicago: ABA, 2024.
120. Brookings Institution. Algorithmic bias detection and mitigation. Washington: Brookings, 2019.
121. Digital Adoption. 5 Real-life Examples of AI Bias. 2025.
122. UC Law SF Scholarship Repository. Algorithmic Bias: AI and the Challenge of Modern Employment Practices. San Francisco: UCSF, 2022.
123. University of Minnesota. Algorithmic Bias in Job Hiring. Minneapolis: University of Minnesota, 2024.
124. Yale Insights. AI and the Law: Responsibility, Transparency, and Regulation. New Haven: Yale University.
126. Yale Insights. The Ethical Dilemmas of AI and How to Avoid Them. New Haven: Yale University.
127. Metcalfe, R. Metcalfe's Law. (Теоретичен принцип за мрежови ефекти)
128. Acemoglu, D., & Restrepo, P. Robots and jobs: Evidence from US labor markets. Journal of Political Economy, 2020, 128(6), 2188–2244.
129. Brynjolfsson, E., Rock, D., & Syverson, C. The productivity J-curve: How intangibles complement general purpose technologies. American Economic Journal: Macroeconomics, 2021, 13(1), 333–372.
130. Draghi, M. The future of European competitiveness. Европейска комисия, 2024.
131. Европейска комисия (2020). Shaping Europe's Digital Future. <https://ec.europa.eu/digital-strategy>

132. Европейска комисия (2021). Digital Compass 2030: The European way for the Digital Decade. COM(2021) 118 final.
133. Европейска комисия (2024). Digital Economy and Society Index (DESI) 2024. <https://digital-strategy.ec.europa.eu/en/policies/desi>
134. Eurostat (2024). Enterprises using AI technologies. <https://ec.europa.eu/eurostat>
135. Evans, P. C., & Gawer, A. (2016). The Rise of the Platform Enterprise: A Global Survey. The Center for Global Enterprise.
136. Filippucci, F., et al. (2025). Artificial Intelligence and Productivity in Europe. IMF Working Paper WP/2025/067.
137. International Monetary Fund (2025). AI Preparedness Index. <https://www.imf.org/external/datamapper/datasets/AIPI>
138. McKinsey (2024). Accelerating Europe's AI adoption: The role of sovereign AI capabilities. <https://www.mckinsey.com>
139. Национален статистически институт (2023). Иновационна дейност на предприятията. <https://www.nsi.bg>
140. Национален статистически институт (2024). Информационно общество. <https://www.nsi.bg>
141. OECD (2019). Going digital: Shaping policies, improving lives. OECD Publishing.
142. OECD (2024). The impact of AI adoption on market dynamics. <https://www.oecd.org>
143. OECD (2025). AI adoption by small and medium-sized enterprises. OECD Discussion Paper.
144. OECD/BCG/INSEAD (2025). The Adoption of Artificial Intelligence in Firms. OECD Publishing.
145. Oxford Insights (2024). Government AI Readiness Index 2024. <https://oxfordinsights.com>
146. World Economic Forum (2019). The Global Competitiveness Report. <https://www.weforum.org>
147. Goldman Sachs (2023). The potentially large effects of artificial intelligence on economic growth. Global Economics Analyst.
148. Implement Consulting (2024). Europe's AI Innovation Opportunity. <https://publicpolicy.google>
149. PwC (2025). Value in Motion: Global AI Study. <https://www.pwc.com>
150. SeeNext Insights (2025). SEE Digitalization 4.0 Index. <https://seenext.org>

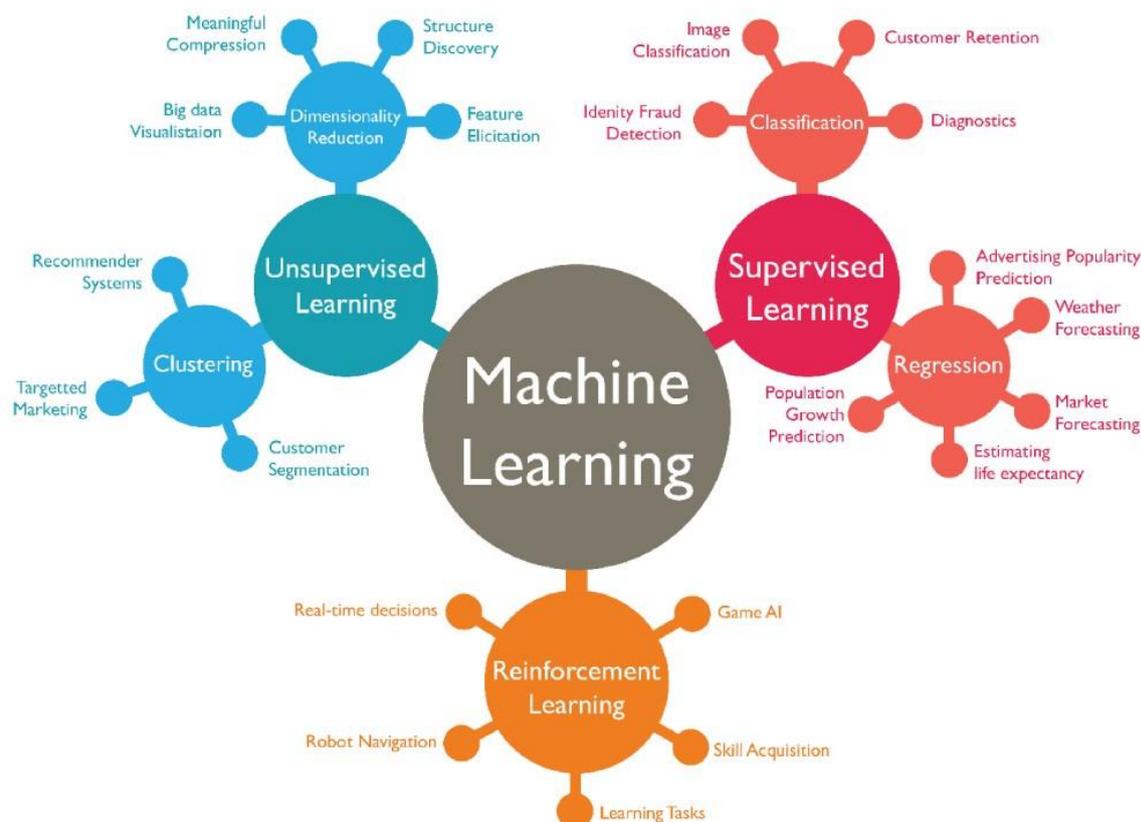
ПРИЛОЖЕНИЕ

ОСНОВНИ ТЕРМИНИ СВЪРЗАНИ С AI

Generative AI (генеративен изкуствен интелект) е клон на изкуствения интелект, който се фокусира върху създаването на ново съдържание - текст, изображения, музика, код, видео и други форми на данни. За разлика от традиционния AI, който анализира и класифицира съществуваща информация, генеративният AI може да произвежда оригинално съдържание въз основа на научени модели.

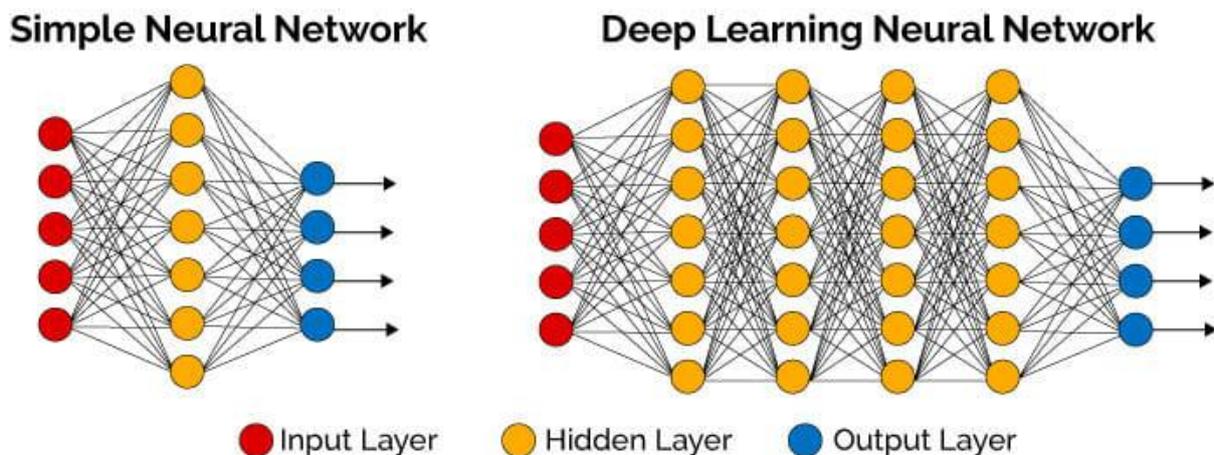
Машинно обучение (Machine Learning). Фундаментът на Generative AI. Машинното обучение позволява на системите да се обучават от данни без да бъдат експлицитно програмирани за всяка задача. Моделите "учат" чрез разпознаване на модели в големи обеми данни.

Тигура 14. Типове машинно обучение



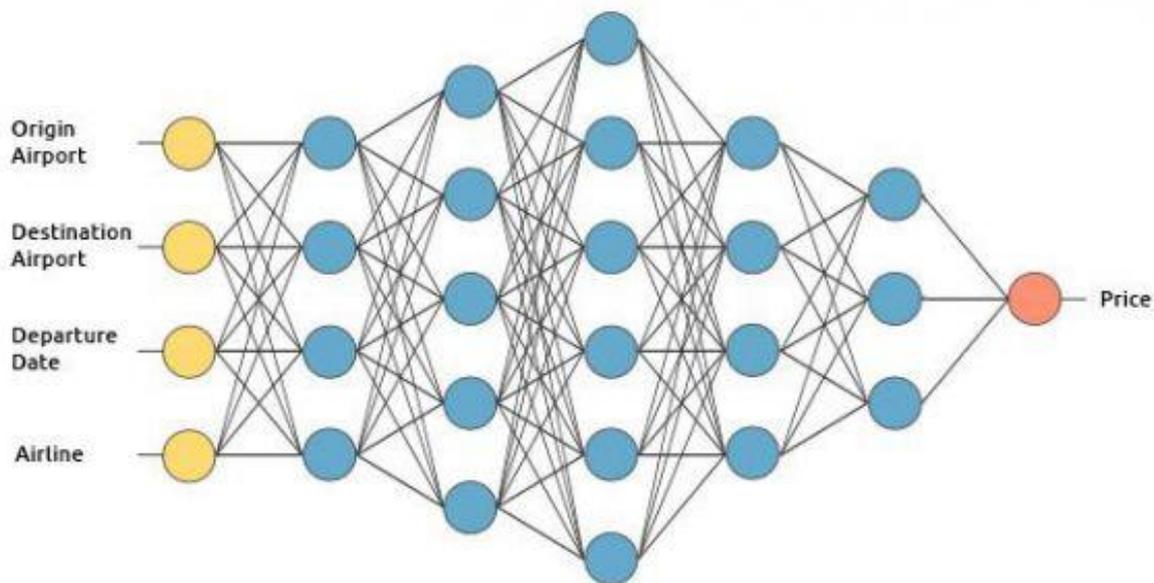
Дълбоко обучение (Deep Learning). Подмножество на машинното обучение, използващо изкуствени невронни мрежи с множество слоеве. Тези многослойни мрежи могат да обработват сложни данни и да извличат абстрактни представяния.

Фигура 15. Дълбоко обучение, принципи и основни слоеве



Невронни мрежи (Neural Networks). Математически модели, вдъхновени от структурата на човешкия мозък. Състоят се от взаимосвързани възли (неврони), които обработват информация чрез слоеве. Мрежата се състои от входен слой (input layer), скрити слоеве (hidden layers) и изходен слой (output layer). Всеки неврон изпраща резултат към следващите, а чрез обучение моделът минимизира грешките с помощта на алгоритъм като backpropagation. Пример за разпознаване на изображение (Image Recognition), малка невронна мрежа може да различава хоризонтални, вертикални и диагонални линии в 2x2 пиксела. Този модел показва как локални шаблони се комбинират, за да се формират по-сложни представяния.

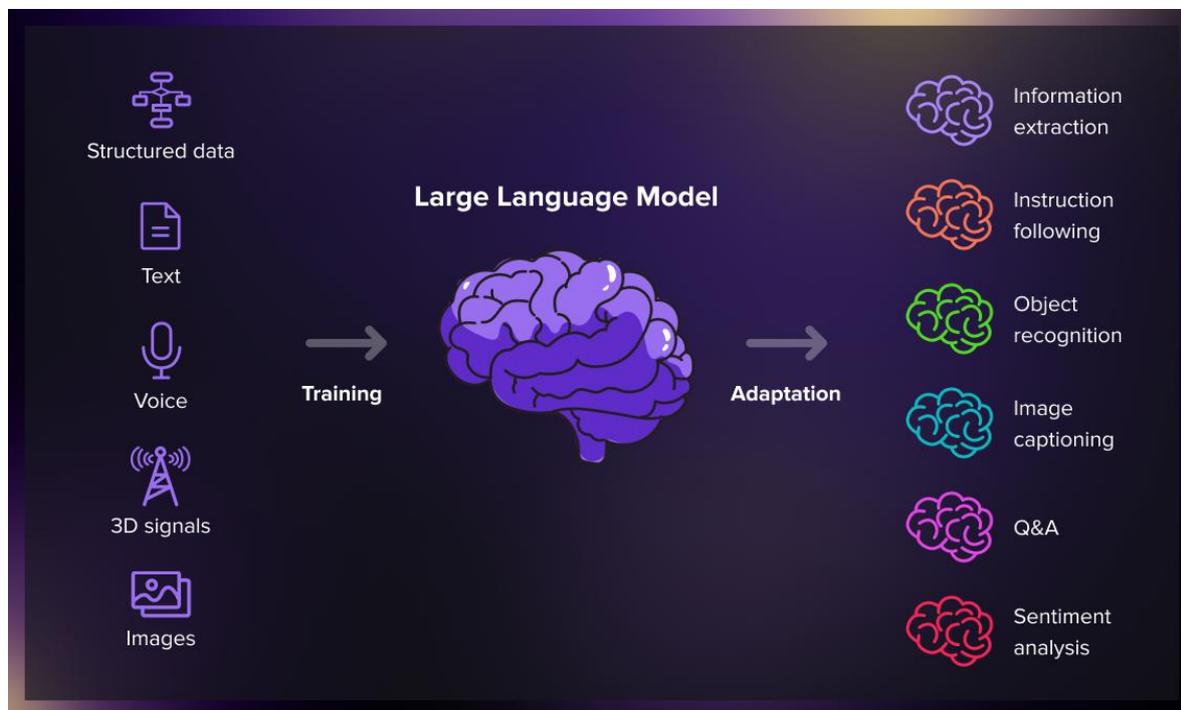
Фигура 16. Принципи на работа на невронни мрежи



Трансформери (Transformers). Модерна архитектура на невронни мрежи, въведена през 2017 г. Използва механизъм за "внимание" (attention mechanism), който позволява на модела да се фокусира върху най-релевантните части от входните данни. Основа за съвременни езикови модели като GPT и BERT.

Large Language Models (LLM). Големи езикови модели, обучени на огромни текстови корпуси. Те могат да разбират контекст, генерират кохерентен текст и изпълняват разнообразни езикови задачи.

Фигура 17. Големи езикови модели



Езикови модели (Language Models). Езиковите модели предсказват следващата дума въз основа на предишните. Те работят с вероятностни разпределения $P(w_t | w_{\{t-n+1\}, \dots, w_{\{t-1\}}})$.

N-gram модели. Изразът “of May” може да бъде последван от “and”, “next”, “doth”. Моделът избира най-вероятната дума. Ограничение: къс контекст и ниска семантична дълбочина.

Контекстуални представяния (Word Embeddings). Embeddings представляват вектори, които улавят смисъла на думите. Пример: “king – man + woman \approx queen”. Това позволява контекстно разбиране.

Autocomplete on steroids. Generative AI може да се разглежда като разширено автодовършване – “autocomplete on steroids”. Моделите предсказват всяка следваща дума динамично, като постоянно обновяват контекста.

Prediction is Knowledge. Знанието в Generative AI е вероятно. Моделът не „помни“ факти, а избира най-вероятната дума. Следователно знанието = високоправдоподобно предсказание.

Настройка на LLM

Prompt/промпт. Входната инструкция или въпрос, който подавате на генеративен AI модел. Качеството на промпта значително влияе на генерирания отговор. Най-актуалните модели на структуриране на промпт са:

- **TIC (Task, Instructions, Context)**, модела задава базови изисквания за промпт да се постави ясна задача, очакване, да се дадат точни инструкции по задачата, да се даде контекст роля на LLM.
- **Chain-of-Thought (CoT) Prompting** популярна техника, при която, моделът се насърчава да мисли стъпка по стъпка, Показва междинните стъпки на разсъждение, Подобрява точността при сложни задачи
- **Tree-of-Thoughts (ToT) Prompting** е разширена версия на CoT, при която моделът изследва множество алтернативи, създава "дърво" от възможни решения, Оценява всеки клон преди да избере най-добрия
- **ICL (In-Context Learning)** е обучение в контекста чрез примери. Как работи - предоставяте примери в самия prompt, Моделът учи от тях без допълнително обучение известно още като "few-shot learning"
- 4D модел. **D1, Define (Определи целта/ролята):** Каква е задачата и кой „си“ ти? **D2 Data (задайте данните/контекста)** Факти, ограничения, аудитория, тон. **D3 Directions (Инструкции)** Стъпки, какво да се включи/изключи, критерии за качество. **D4 Deliverable (Изход)** Формат (списък/таблица/JSON), дължина, език, check за готовност.

Златните правила за създаване на промптове, които работят най-успешно

- Започнете с определяне на „ролята“ на AI. Посочете кой е целевия модел. „Ти си финансов анализатор...“, „Ти си UX дизайнер...“. Това фокусира отговора.
- Започнете с контекст преди задачата, няколко изречения за изходното ниво, защо се дава задачата и каква е целта вт.ч. цел, канал, ограничения (напр. регулаторни/бренд тон).
- Формулирайте една ясна задача. Една основна инструкция е много повече от много дребни. Ако има подзадачи, номерирай ги.
- Посочи желан формат на резултата с конкретна структура - списък, таблица др. Добавете пример на очаквания изход за референция.
- Поставете критерии за качество на задачата - „Използвай кратки изречения, максимум 200 думи, без жаргон, на български, тон: професионално-приятелски.“
- Дайте примери (few-shot). Поосочете няколко примера за желан резултат.
- Дефинирайте входните данни. Поставете данните между ... или <data> ... </data>, за да избегнете смесване с инструкциите.
- Включи проверки и ограничения. „Ако липсва информация, кажи какво липсва.“ „Не измисляй факти; ако не си сигурен – посочи несигурността.“

- Изисквайте източници на информация и обосновка. „Цитирай източници/връзки или отбележи ‘без достъп до източник’. Обясни накратко защо.“
- Управлявайте обема и езика на заданието. Посочете език, дължина, ниво (начално/експертно), тон (неутрален/убедителен).
- Съобразявайте се с ограничението на контекста. За по-обширни материали дайте инструкции да се работи на части. „Обобщи секция по секция, после синтезирай. Работи на партиди“
- В пропта включете критерии за „готовност“. „Считай задачата за завършена, ако (а) форматът е валиден, (б) има проверени мин. три различни източника на информация, (в) всички теми имат кратка обосновка.“
- Винаги задавай въпроси за критичен и скептичен анализ на резултатата и аргументи и презложения за промени
- При всеки резултат за да считате за финален попитайте това ли е най-добрия резултат, на който модела е способен и изискайте още по-добра и задълбочена работа и резултат.
- Винаги питайте има ли модела въпроси за да се справи добре и има ли 100% цялата необходима информация.
- Пишете пропърт подобно на „код“ максимално подробно и с ясни очаквания.

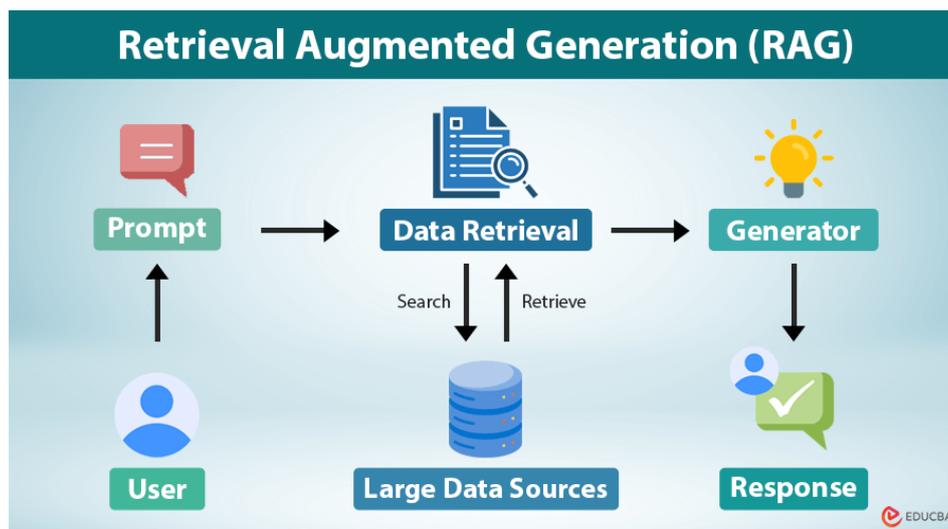
RAG (Retrieval Augmented Generation) е техника в областта на генеративния изкуствен интелект, която комбинира **извличане на информация** (retrieval) с **генериране на текст** (generation). Тази хибридна архитектура позволява на езиковите модели да достъпват външни източници на знание в реално време, вместо да разчитат единствено на знанието, вградено в техните параметри по време на обучението. RAG системата се състои от три основни компонента:

1. **Retrieval (Извличане)** Когато потребител зададе въпрос, въпросът се преобразува в **вектор** (embedding - числово представяне на текста). Системата търси в **база от документи** най-релевантните фрагменти, използвайки семантично сходство, извличат се най-подходящите пасаж/документи, които съдържат релевантна информация.

2. **Augmentation (Обогатяване)**, извлечените документи се комбинират с оригиналния въпрос и се създава **разширен контекст** (augmented context), който включва както въпроса, така и релевантната информация от документите.

3. **Generation (Генериране)**, езиковият модел получава обогатения контекст, генерира отговор, базиран както на своите вградени знания, така и на извлечената информация, отговорът е информиран от конкретни, актуални източници

Фигура 18. RAG модел на настройка на LLM

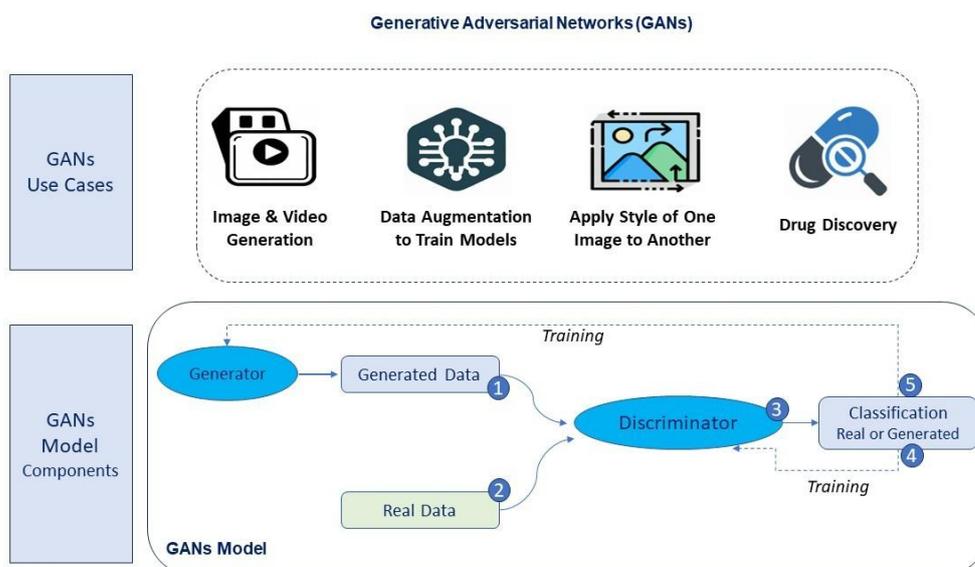


Fine-tuning (Финастройване). Процес на адаптиране на предварително обучен модел към специфична задача или домейн чрез допълнително обучение върху специализирани данни.

Token (Токен). Основна единица за обработка на текст в езиковите модели. Един токен може да бъде дума, част от дума или символ. Моделите имат ограничение за брой токени, които могат да обработят наведнъж.

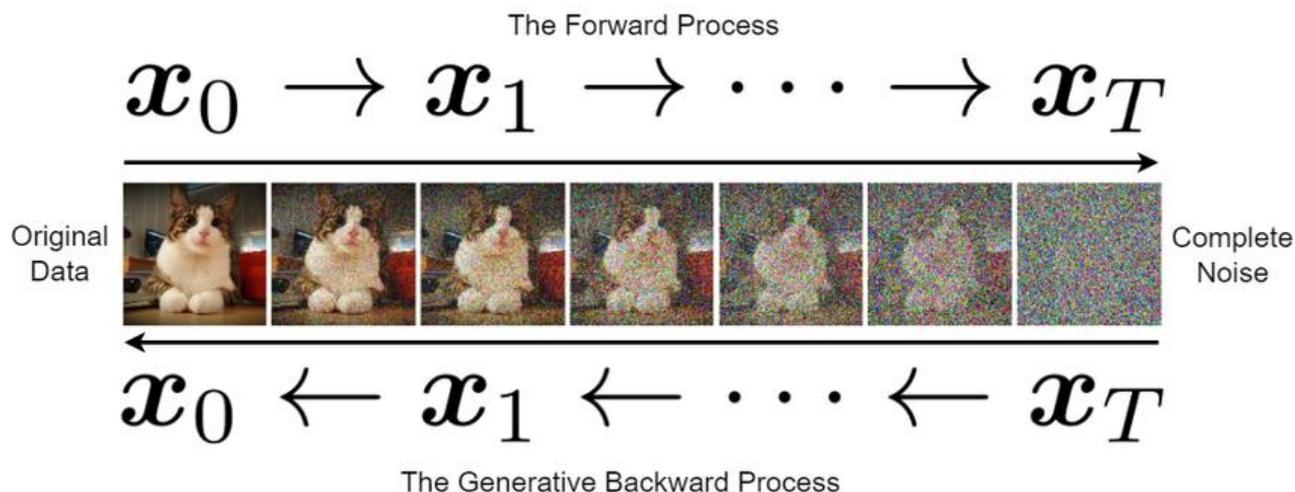
Генеративни състезателни мрежи (GANs - Generative Adversarial Networks). Архитектура, състояща се от два модела: генератор (създава ново съдържание) и дискриминатор (оценява дали съдържанието е реално или генерирано). Двата модела "се състезават", усъвършенствайки се взаимно.

Фигура 19. GANs - Generative Adversarial Networks



Diffusion Models (Дифузионни модели). Нов тип генеративни модели, които създават изображения чрез постепенно "обезшумяване". Използват се в системи като DALL-E, Midjourney и Stable Diffusion.

Фигура 20. Дифузионни модели за генериране на изображения



Ключови архитектури на Gen AI модели

- GPT (Generative Pre-trained Transformer). Семейство модели от OpenAI, базирани на трансформерна архитектура. GPT-3, GPT-4,5 GTP-5 и по-нови версии демонстрират изключителни възможности за генериране на текст.
- BERT (Bidirectional Encoder Representations from Transformers). Модел на Google, който обработва текст в двете посоки едновременно, което подобрява разбирането на контекст.
- VAE (Variational Autoencoders). Модели, които кодират данни в компресирано представяне и след това ги декодират обратно, позволявайки генериране на нови вариации.

Николай Лазаров

Изкуствен интелект в бизнеса

Българска, първо издание

© Автор	Николай Лазаров, 2026
Дизайн на корица	<i>Вирджиния Анастасова</i>
Предпечат	<i>Издателство на УЗФ</i>
Формат	21/29.7
Печат	Издателство на УЗФ „Св. Григорий Боголсов“, 2026

© Издателство на УЗФ „Св. Григорий Боголсов“, 2026

ISBN (pdf): 978-619-7622-88-1

