



Co-funded by
the European Union

Резултат от Еразъм+ проект HEDY
KA220-HED-000029536 – Сътрудническо
партньорство във висшето образование.



КНИЖКА

Есе за живота в ерата на ИИ



BAEHF

КНИЖКА - Есе за живота в ерата на ИИ

Автори: Давиде Карелио, Ана И. Алвеш Морейра, Сесилио Ангуло Бахон, Федерика Казачо, Розалина Димова, Тихомир Доврамаджиев, Антония Якоби, Чаба Колар, Йевгения Суковий, Гюла Сабо

2022

DOI: 10.5281/zenodo.7426885

<https://lifeintheaiera.eu/>

Финансирано от Европейския съюз. Изразените възгледи и мнения са само на автора(ите) и не отразяват непременно тези на Европейския съюз или Европейската изпълнителна агенция за образование и култура (EACEA). Нито Европейският съюз, нито EACEA могат да носят отговорност за тях.

Резултат от Еразъм+ проект HEDY
KA220-HED-000029536 – Сътрудническо
партньорство във висшето образование..



Co-funded by
the European Union



Проектът HEDY

HEDY - Животът в ерата на ИИ е 2-годишен проект по програма Еразъм+, стартиран през ноември 2021 г. В самото си заглавие той отдава почит на Хеди Ламар (1914-2000), австрийска актриса и изобретател, съсновател на безжична комуникационна технология, приета за управление на торпеда по време на Втората световна война и която в момента все още се използва в мобилни мрежи, Bluetooth устройства и WiFi. Проектът HEDY представлява безплатен и достъпен източник на информация относно цифровите технологии на 4-та индустриална революция (Индустрия 4.0), а именно ИИ, като осведомява за възможните му положителни бъдещи приложения, като същевременно изяснява възможните въздействия. Целта на HEDY е да предложи изчерпателна и споделена представа за това как ИИ влияе върху живота ни и прекроява нашата социално-икономическа, културна и човешка среда чрез насърчаване на критичното мислене, самообучението и дебатите по тези въпроси. Основната целева аудитория (но не непременно) на този проект е публиката с висше образование.

Автори

Давиде Карелио, Ана И. Алвеш Морейра, Сесилио Ангуло Бахон, Федерика Казачо, Розалина Димова, Тихомир Доврамаджиев, Антония Якоби, Чаба Колар, Йевгения Суковий, Гюла Сабо

Резюме

Книжката е есе, определящо позицията на HEDY за живота в ерата на изкуствения интелект и обосновката за тази позиция. В тази разработка ние описваме и представяме резултатите от нашия двустранен подход за изграждане на рационалната ни представа за предизвикателствата, възможностите и очакваното въздействие на ИИ върху четири области: бизнес, управление, умения&компетенции и хора&начин на живот. Този двустранен подход се състои в събиране на информация от два източника: 1) проучване на литературата и 2) взаимодействие с хората. Ясно е, че първият източник се състои от събиране на текущото състояние на знания за въздействието на ИИ. Вторият източник беше смесен подход, състоящ се както от въпросници, така и от фокус групи, проведени в пет различни европейски държави с експерти или неексперти по ИИ. Тези два източника осигуряват уникален принос към панорамата на ИИ чрез комбиниране на най-съвременни изследвания с мнения на различни социални участници и обсъждани въпроси, отношение и идеи.

Октомври 2022

Съдържание

01	Въведение Каква е целта на тази книжка и как да я постигнем?	5
02	Индустрия 4.0/5.0 Къде сме сега?	7
03	Кратка история на изкуствения интелект	9
04	Изкуствен интелект Какви са въздействията на ИИ в нашето общество?	11
05	Бизнес ИИ подобрява ли приходите на компаниите или качеството на работата?	14
06	Управление Трябва ли ИИ да е без правила?	17
07	Умения и компетенции Влияе ли ИИ на образователната система?	21
08	Хора и начин на живот Може ли ИИ да подобри живота на хората?	24
09	Изводи Какво можем да направим?	27
0A	Библиография Искате ли да научите нашите източници?	30
0B	Приложение Искате ли да знаете повече?	36

1. Въведение

Цифровизацията нахлува в нашия свят, като технологиите се използват във всички измерения на живота, от образование до работа, здравеопазване или администрация. Развитието на знанията и уменията вече е процес през целия живот, който изисква нарастваща цифрова грамотност. Но за някои членове на обществото, като например „дигиталните местни жители“, използването на технологии е естествено, но това не е непременно случаят с неместните дигитални лица (така наречените дигитални имигранти). Как да гарантираме, че всеки гражданин ще развие необходимите умения, за да остане включен във все по-дигиталното общество? И как се постига безпристрастност, вместо увеличаване на неравенствата? Ако приемем, че изкуственият интелект (ИИ) ще трансформира пазара на труда, уместно е да си представим образователната система в свят, в който работата не е централен фактор в живота или където работните места, каквито ги познаваме, не съществуват. Каква би била ролята на образованието? Как бихме могли да го организираме? Какви биха били неговите цели и на какви нужди ще отговаря? И нищо по-добро от това да предложим използването на технологии за повишаване на осведомеността за живота в дигиталната ера и за развиване на умения, да се възползваме от предимствата, но и да се изправим пред предизвикателствата, които тази нова ера предлага.

Това е лийтмотивът на HEDY – Животът в ерата на ИИ [1.1]. Проектът HEDY представлява безплатен и достъпен източник на информация относно цифровите технологии на 4-та индустриална революция (Индустрия 4.0), а именно ИИ, като осветлява възможните му положителни бъдещи приложения, като същевременно изяснява възможните въздействия.

За постигането на това ще бъдат създадени четири конкретни цели:



Книжка – есе, определящо позицията на HEDY относно живота в ерата на ИИ и обосновката за тази позиция. Книжката систематизира характеристиките на ИИ и положителните въздействия, рисковете, свързани с определени употреби, и идентифицира предизвикателствата, възможностите и очакваните въздействия с парадигматични примери, предлагани по увлекателен начин, за да стимулират размисъл и дебат по обществени теми.



Набор от инструменти – колекция от влиятелни аудио-визуални инструменти, показващи как ИИ може да се справи с някои от най-предизвикателните обществени проблеми в света, както и рисковете, които трябва да бъдат сметчени, ако ИИ реализира пълния си потенциал, за да се използва гъвкаво в подкрепа на дейностите, свързани с ученето и дебата. Той включва художествени и документални филми и речи на експерти, пет от които ще бъдат произведени от Hedy.



Масов отворен онлайн курс (МООС) – курс за насърчаване на обширни познания, критичен размисъл и дебат относно ИИ и неговото ключово въздействие върху обществото. Той ще бъде безплатен и отворен за участниците, с интерактивни форуми. Курсът ще бъде структуриран в 5 модула и ще бъде съставен от експерти. Участниците могат да споделят идеи и да се включат задълбочено в темите за ИИ чрез различни онлайн дейности.



Ръководство – кратка и лесна за четене документация, обясняваща естеството на учебните ресурси на HEDY, най-доброто използване на произведените активи и речник с ключови термини по теми, свързани с ИИ. Това ще помогне да се създаде солидна основа за гарантиране на използваемостта на резултатите от HEDY от широка мрежа от практики, дори след края на проекта.

В този документ представяме Книжката. Тя ще бъде принос към предизвикателството, поставено от Европейската комисия за изграждане на солиден европейски подход към ИИ, базиран на стратегията от 2018 г. [1.2] и подсилен от Бялата книга за ИИ от 2020 г. [1.3]. Този документ също така допринася за етичната рамка на ЮНЕСКО [1.4], която предоставя в своята резолюция основата, която да накара системите за изкуствен интелект да работят за доброто на човечеството, хората, обществата и околната среда и екосистемите, както и за предотвратяване на вреди.

Както беше посочено по-горе, целта на този документ е да систематизира характеристиките на ИИ, да идентифицира предизвикателствата, възможностите, рисковете, свързани с определени употреби, и очакваните въздействия върху четири различни области на нашето общество:

- **Бизнес** – с все по-голямото количество налични данни днес и постоянно променящите се предпочитания и сложност на клиентите, фирмите вече не могат да разчитат на традиционни бизнес методи за стимулиране на растежа. Тези радикални промени откриха нова сфера от възможности, с ИИ, за стимулиране на растежа на бизнеса чрез ефективни виждания, генерирани от клиентски данни.
- **Управление** – думата управление е залегнала в стандартния бизнес речник и е нещо, което включва всички процеси на управление. Това е начинът, по който правилата или действията се структурират, поддържат и регулират – и често как се определя отчетността. Управлението на ИИ трябва да затвори пропастта между тях и да гарантира, че границите в рамките на технологиите са определени (т.е. те са етични).
- **Умения и компетенции** – на практика е сигурно, че изкуственият интелект започва да заменя все по-голям брой работници: аутсорсването на труда към машини ще промени уменията и компетенциите, които са ценни на конкурентен пазар на труда. В отговор на неизбежните промени в работната сила, приоритетите на образователните системи изглеждат се нуждаят от промяна, за да отразяват уменията и компетенциите, които ще бъдат ценни в една доминирана от ИИ ера.
- **Хора и начин на живот** – ИИ е навсякъде в ежедневието ни. От улесняване на ежедневния ни живот с препоръки за онлайн търсене, гласови асистенти и влизане с лицево разпознаване до напредъка в здравеопазването, идентифициране на пандемии и подпомагане за облекчаване на глада, ИИ е наистина трансформираща технология с дългосрочни ефекти.

Освен че описваме настоящите приложения и очакваните въздействия на ИИ за всяка област, ние също идентифицираме 5/6 по-обезпокоителни проблеми относно ИИ. Въпреки това, ние не се опитваме да предоставим общи решения на тези проблеми; напротив, идеята е да се предложи увлекателен начин за стимулиране на размисъл и дебат по обществените теми, да се обсъдят етичните ефекти на тези нововъзникващи цифрови технологии и да се осигурят парадигмални примери.

За да постигнем тази цел и да изградим нашите обосновки, събрахме информация от два различни източника: 1) Преглед на литературата и 2) Фокус групи. Първият източник се състоеше от събиране на текущото състояние на знания за приложенията и въздействията на ИИ. Вторият източник се състоеше от събиране на информация чрез интервюиране на хора чрез фокус групи, проведени в пет различни европейски държави с експерти или неексперти по ИИ. Тези два източника осигуряват уникален принос към облика на ИИ чрез комбиниране на най-съвременни изследвания с мнения от първа ръка и обсъждани въпроси, опасения и идеи на взаимодействащи индивиди.

Останалата част от документа е организирана както следва. Раздел 2 описва четвъртата индустриална революция, а именно цифровизацията и автоматизацията на производството чрез фундаментална промяна в начина, по който се произвеждат продуктите, и тя е дълбоко свързана с широкото приемане на ИИ. Раздел 3 е посветен на дефиницията на ИИ и определянето на основните предизвикателства и възможности в общи условия. Раздел 4 прави преглед на текущите приложения на ИИ и очакваните въздействия на ИИ в четирите области, обсъдени по-горе. В раздел 5 ние обобщаваме нашите констатации и подчертаваме ключовите идеи на фокус групите. В раздел 7 е заключението на документа.

2. Индустрия 4.0/5.0

Поучително е да се запомни, че това не е първият път, когато обществото е разтърсено от индустриална революция, а четвъртият. Преди 19 век голям процент от работната сила е концентриран в селскостопанския сектор и използването на механични машини е много ограничено. През последните 2 века човечеството е преживяло четири индустриални революции, които коренно промениха нашия живот и нашето общество.



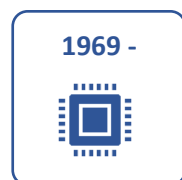
1760-1840

Първата индустриална революция се провежда в Европа и Съединените щати. Използването на пара и водна енергия бележи преминаването от ръчни производствени методи към машини. Първият сектор, в който се случва този преход, е текстилната промишленост, последван от железарската промишленост, селското стопанство и минното дело. Той има и обществено отражение: нарастването на средната класа.



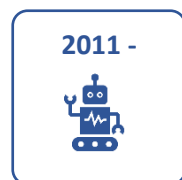
1871-1914

Втората индустриална революция се нарича още технологична революция. Тя е предизвикана от изграждането на обширни железопътни и телеграфни мрежи и широкото въвеждане на електричеството. Тъй като фабриките стават по-електрифицирани, те успяват да изградят модерната производствена линия. Това е време на бърза икономическа експанзия и повишена производителност, но е и време на висока безработица, тъй като много служители в производството са заменени от машини.



1969 -

Третата индустриална революция е цифровата революция и се случва след края на Втората световна война. Състои се в прехода от механична и аналогова към цифрова електроника, който започва с приемането и разпространението на цифрови компютри и цифрово водене на записи. Масовото производство и широкото използване на чипове с интегрални схеми (IC) и свързани технологии, като компютри, микропроцесори, цифрови клетъчни телефони и Интернет, са в основата на тази революция.



2011 -

Четвъртата индустриална революция, известна още като Индустрия 4.0, е въведена за първи път от екип немски учени във високотехнологичен проект, който насърчава компютъризацията на производството [2.1]. По същество Индустрия 4.0 е тенденцията към обединение между физически активи и напреднали цифрови технологии като интернет на нещата (IoT), ИИ, роботи, дроневи, автономни превозни средства, 3D печат, облачни изчисления и други, които са взаимосвързани, имащи възможност да общуват, анализират и действат.



2017 -

Индустрия 5.0 е нов модел, специално прокарван от Европейския съюз, за да пренасочи фокуса на Индустрия 4.0. Не се счита за друга индустриална революция, а като допълнение към настоящата, за да се върнат човешките, социалните и екологичните измерения обратно в уравнението.

Има три причини, поради които днешната революция не е просто продължение на цифровата революция, а по-скоро началото на нова: въздействие, скорост и обхват. Настоящият темп на пробиви е безпрецедентен в историята. Нововъзникващите технологични разработки в дисциплини, включително ИИ, IoT, автономни автомобили, роботи, квантово изчисление и мрежи, нанотехнологии и т.н., карат Индустрия 4.0/5.0 да напредва експоненциално, а не линейно, както в предишните индустриални революции. Освен това, тя причинява смущения в почти всеки бизнес във всяка страна. А обхватът и сложността на тези разработки показват пълно преразглеждане на системите за производство, опериране и управление.

Индустрия 4.0/5.0, подобно на предишните революции, има потенциала да повиши глобалните нива на доходите и да подобри качеството на живот на хората по целия свят [2.2]. Към днешна дата онези, които могат да си позволят и имат достъп до цифровия свят, са се възползвали най-много от него; технологията позволява нови продукти и услуги, които подобряват ефективността и удоволствието в нашето ежедневие. В бъдеще технологичният напредък ще доведе до чудеса в предлагането, с дългосрочни ползи по отношение на ефективността и производството. Разходите за транспорт и комуникация ще се редуцират, логистиката и глобалните вериги за доставки ще станат по-ефективни, а търговските разходи ще намалееят, отваряйки нови пазари и стимулирайки икономическото развитие.

Въпреки това ефектите не са само положителни. Настоящата революция има потенциала да разруши заетостта. Тъй като автоматизацията замества работата в цялата икономика, нетното изместване на служители от машини може да увеличи разликата между капитала и възвръщаемостта на труда. Това все още е открит въпрос, тъй като също така е възможно технологичното изместване на служителите да доведе до нетно увеличение на безопасни и удовлетворяващи професии. Такава несигурност по отношение на ИИ технологията е анализирана в раздел 5.

Освен икономиката, неравенството е най-значимата обществена тревога, свързана с Индустрия 4.0/5.0. Доставчиците на интелектуален и физически капитал като изобретатели, акционери и инвеститори са най-големите благодетели на тази революция. Това обяснява нарастващото несъответствие в богатството между тези, които разчитат на капитал и тези, които разчитат на работа. Това помага да се обясни защо толкова много служители са разочаровани и загрижени, че собствените им и действителните заплати на децата им ще останат стагнирани. Това също така обяснява защо средната класа в света става недоволна и несправедливо третирана. Икономика, в която победителят взема всичко, с ограничен достъп на средната класа е рецепта за демократична стагнация и изоставяне.

Всеобхватното разпространение на цифровите технологии и динамиката на споделяне на информация, представена от социалните медии, също могат да подтикнат към недоволство. За да си взаимодействат, учат и обменят знания, повече от 30% от световното население днес използват социални медийни платформи. Тези контакти, в един идеален свят, биха позволили межкултурно разбиране и интеграция. Те обаче могат да насърчат неточни очаквания за това какво представлява успех за дадено лице или група, както и да предоставят платформи за разпространение на екстремистки вярвания и идеологии.

В този контекст ИИ е един от ключовите двигатели на четвъртата индустриална революция, който вече е навсякъде около нас и влияе ежедневно върху живота ни: от самоуправляващите се превозни средства до виртуалните асистенти, софтуерът, който превежда, инвестира и предлага културни предпочитания, са само няколко примера. Безпрецедентният изчислителен капацитет и капацитет за съхранение, достъпът до огромни количества данни и широкото приемане на ИИ в много различни области създават симбиоза между цифрови и биологични светове, които променят начина, по който живеем и взаимодействаме с околната среда.

Като се позоваваме отново на въпросите, които повдигнахме във въведението и в съответствие с потенциала и рисковете, споменати по-горе, в следващите глави се опитваме да предоставим повече мисли и аргументи относно ИИ и неговите въздействия. Първо, представяме проучване на предизвикателствата и възможностите на ИИ в общ контекст и след това ги специализираме в още четири целеви области, а именно бизнес, управление, образование и начин на живот.

3. Кратка история на ИИ

ИИ е сравнително нова дисциплина (родена в средата на 20 век). Тюринг, често наричан „бащата на ИИ“, публикува през 1936 г. математическо описание на това, което той нарича „универсална машина“ [3.1] и той е първият, който публикува метод (тест на Тюринг) в „Компютърни машини и интелект“ [3.2], за да се определи дали една машина мисли.

Трудно е да се дефинира ясно изкуствения интелект поради разнообразието от проблеми, решения, разграничаване на това какво съдържа ИИ и какво не. ИИ е термин, въведен от Джон Маккарти през 1955 г. и е дефиниран от него през 2007 г. [3.3] като „наука и инженерство за създаване на интелигентни машини, по-специално интелигентни компютърни програми. Това е свързано със задачата за използване на компютри за разбиране на човешкия интелект, но ИИ не трябва да се ограничава до методи, които могат да бъдат наблюдавани биологично. Енциклопедия Британика [3.4] определя ИИ като „способността на цифров компютър или компютърно контролиран робот да изпълнява задачи, обикновено свързани с интелигентни същества. Терминът често се прилага към проекта за разработване на системи, обогатени с интелектуалните процеси, характерни за хората, като способността да разсъждават, да откриват значение, да обобщават или да се учат от минал опит. Речникът на Кеймбридж [3.5] го компресира в по-кратка версия: „използването на компютърни програми, които имат някои от качествата на човешкия ум, като способността да разбират език, да разпознават картини и да се учат от опита“. Най-кратката и проста дефиниция е: „ИИ не е биологичен интелект“ [3.6].

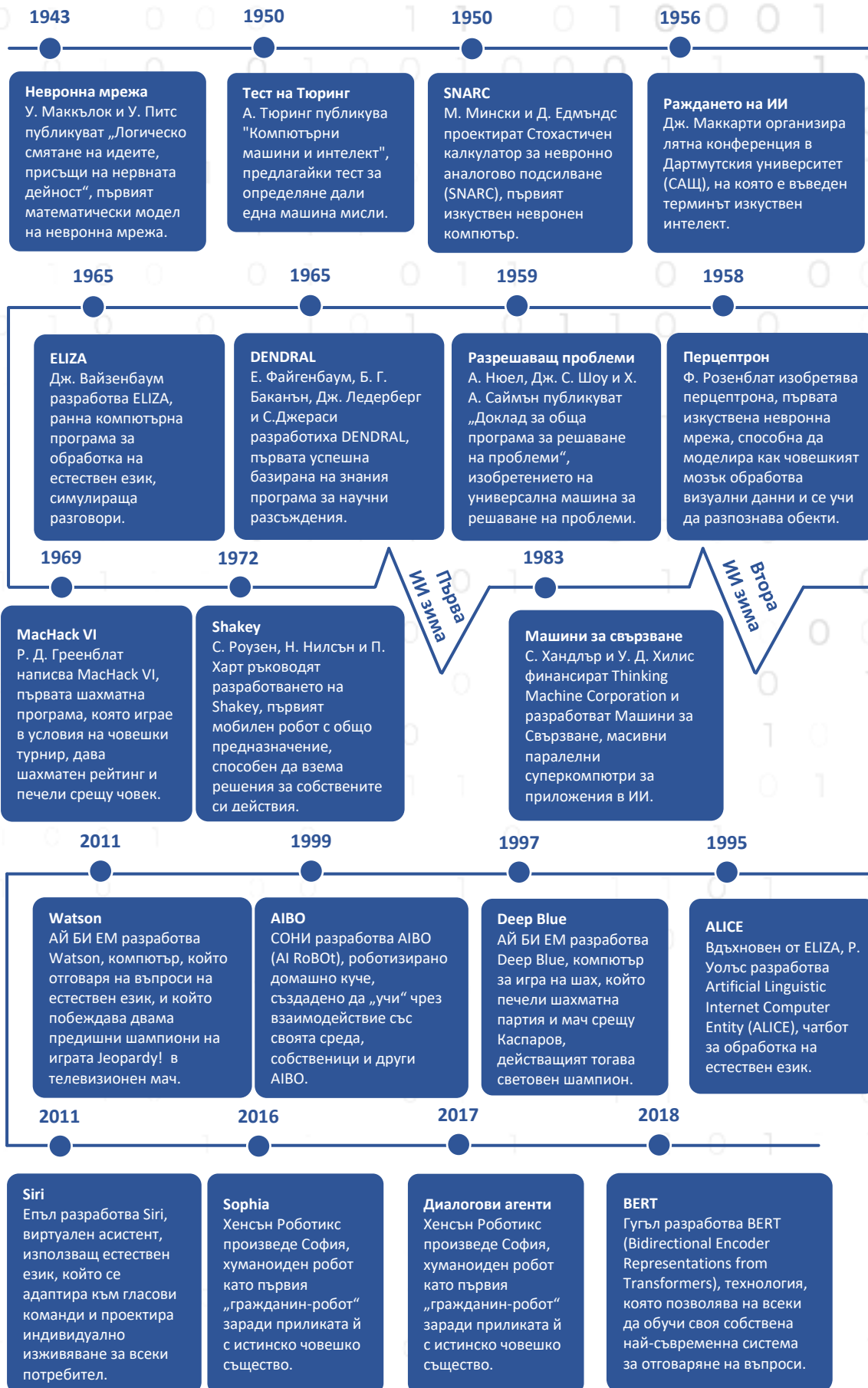
В обобщение, няма общоприето определение за ИИ. ИИ е общ термин, който може да се отнася до широка област на науката, обхващаща не само компютърните науки. Ако тълкуваме ИИ като изучаване на това как информацията се придобива, обработва, съхранява, използва и т.н. в интелигентни животни и машини, тогава той очевидно се припокрива с няколко по-стари дисциплини [3.7]: философия, математика и статистика, икономика, невронауки, психология, биология и медицински науки, лингвистика, компютърни науки, технически науки и науки за безопасността и сигурността.

Следващата схема/ фигура предоставя кратък преглед на историческата еволюция на ИИ, без да претендира за изчерпателност предвид размера на темата [3.8]-[3.10].

Първата фаза на разцвет е през 50-те/60-те години. През тези години преживяваме голям напредък в много аспекти, които поставят основата на ИИ. Например, първата машина за невронни мрежи, SNARC, е създадена през 1950 г. Първият цифров асистент, ELIZA, е разработен през 1965 г., въпреки че намерението на неговия създател (Джозеф Вайзенбаум) е да покаже повърхностността на комуникацията между човек и машина, но е изненадан от броя на хората, които приписват човешки чувства на неговата компютърна програма. Първият автономен мобилен робот с общо предназначение е създаден през 1972 г.

Тази първа фаза е заменена през 70-те години от така наречената ИИ зима, когато развитието се забавя. Прогнозите се оказват пррреувеличени, използваните дотогава подходи не оправдават очакванията, а капацитетът на машините ограничава възможностите. През 80-те години на миналия век експертните системи донасят кратък бум, но след това идва втората ИИ зима. През последните години ИИ започна нова вълна на прогрес и ентузиазъм главно поради три фактора: 1) възможностите на по-мощните компютри; 2) наличието на големи данни от източници, включително електронна търговия, бизнеси, социални медии, наука и правителство; 3) подобрени подходи и алгоритми за машинно обучение.

В следващите глави представяме текущото състояние на ИИ, като идентифицираме основните приложения, предизвикателства и възможности и анализираме въздействието му върху нашето общество като цяло, както и в четири различни области, а именно бизнес, управление, умения&компетенции и хора&начин на живот.



4. Изкуствен интелект

ИИ вече се появява във всички сфери на живота и се разпространява бързо. По-долу са дадени някои примери за употребите на ИИ, без претенции за изчерпателност.

- **Икономика, финанси:** Системите, базирани на ИИ, осигуряват поддръжка за клиенти, откриват аномалии и измами с кредитни карти чрез разпознаване на шаблони, подобряват сигурността в редица бизнес сектори, включително търговия на дребно и финанси. Организацията разчитат на ИИ, за да проследят тези стъпки, като анализират поведението на транзакциите. Предприятията дават нареждане на специалистите по данни да определят бъдещите модели на пазара, да подобрят своите резултати при търговия с акции [4.1]. Повече подробности за ИИ в бизнеса са анализирани в Раздел 5.
- **Селско стопанство:** Организацията използват автоматизация и роботика, за да помогнат на фермерите в управлението на фермата и да ги освободят за други задачи. Системите с изкуствен интелект насърчават намирането на по-ефективни начини за защита на техните култури от плевели, борба с болестите по животните и растенията, наблюдение на движението, температурата и консумацията на храна от животните. Системите с изкуствен интелект позволяват анализ на сателитни изображения за координиране на производството на регионално и национално ниво, например, идентифициране на райони, предразположени към суша [4.1].
- **Индустрия на здравеопазването:** Има много примери за това как ИИ в здравеопазването е помогнал на пациентите; базираните на ИИ приложения могат да подобрят здравните резултати и качеството на живот на хората по целия свят. Личните или онлайн консултации, персонализираните здравни насоки, виртуалните асистенти намаляват ненужните посещения в болница. ИИ използва комбинацията от исторически данни и медицинско разузнаване за откриване на нови лекарства, откриване на заболявания и предлагане на диагноза [4.2].
- **Транспорт:** Постигнат е голям напредък в сегмента на автономните превозни средства (по-специално авиацията) [4.1], поддържан от ИИ: автономни превозни средства, безпилотни дроневи. Днешните автомобили вече имат базирани на ИИ функции за подпомагане на водача като самостоятелно паркиране, усъвършенствани круиз контроли и интерфейси на естествен език и технологии за виртуална помощ.
- **Интелигентни градове, сгради** [3.7] [4.1]: Очакваният растеж на урбанизацията трябва да управлява екологичната, икономическата и социалната устойчивост: Интелигентен /цифров/умен град е общ термин, който използва широк набор от цифрови технологии за подобряване на оперативната ефективност, споделяне на информация с обществеността, възможно най-доброто използване на ресурсите, осигуряване на високо качество на услугите и подобряване на благосъстоянието на гражданите. Например, управление на трафика, управление на отпадъците, услуга за безопасност (мониторинг на зони с висока престъпност, система за ранно предупреждение за инциденти).
- **Образование:** ИИ ще промени фундаментално образованието. Технологията ще революционизира начина, по който учениците учат и ще промени начина, по който учителите мислят за образователните практики [4.3]. Повече подробности са анализирани в раздел 7.
- **Роботика:** роботите често се разглеждат и като изкуствен интелект, тъй като за обикновения потребител е по-лесно да си представи изкуствен интелект в мехатронна структура, която също има тяло, особено ако тялото му прилича на човек/животно, например социални, терапевтични, образователни роботи. Машиноподобните роботи се прилагат в промишлеността и научните изследвания, както и други форми, разработени за военна и медицинска употреба.
- **Персонални помощни средства:** Разпознаване на пръстови отпечатъци, вграден личен асистент, приложения за обучение (математика), автоматично въвеждане на изговорен текст, автоматични езикови преводачи, виртуална бавачка, говореща с деца, са само пример за многобройните функции с активиран ИИ, налични днес, за да опростят ежедневието ни.

Както беше подчертано в предишните няколко примера, ИИ вече се появява във всички области на живота и радикално трансформира света. Човечеството е изправено пред екзистенциално предизвикателство, чието осъзнаване и активна борба могат да доведат до положителна промяна,

където сътрудничеството между машини и хора може да доведе до утопичен свят. В тази насока хората от фокус групите подчертаха значението на класиката на научната фантастика и литературата за работи, за да се вдъхнови разработването на решения от реалния свят, поддържани от ИИ, за да помогнат за еволюцията на човечеството.

„Мисля, че знам разликата между науката и фантазията, но взаимодействието между тези две области е довело до развитието и на двете“

(експерт)

Препоръката на ОИСР [4.4] изяснява, че ролята на ИИ може да бъде ключова за положителното оформяне на бъдещето, насърчавайки благосъстоянието на хората и субективното благополучие и допринасяйки за икономическото развитие и постигането на устойчиви цели. Всичко това е съпроводено с дълбоки социални промени.

Необходимо е хората да се подготвят навреме за тези промени, защото ако не са в състояние да го направят, ще загубят от ИИ. Съществува пълно съгласие, че дебатът вече не трябва да бъде за това дали ИИ е важен или полезен, а за това как да подготвим или трябва да подготвим колкото се може повече групи в обществото за промяна възможно най-скоро.

„Обществото е фундаментално неподготвено за бързите промени, предизвикани от ИИ“

(експерт)

Независимо от това, ИИ също носи безпрецедентни предизвикателства и тъй като тази нова технология бързо се разпространява в света, възникват и много етични, морални и правни рискове. По-конкретни подробности относно етичните рискове са дадени в раздел 6.

Например Илон Мъск, Стивън Хокинг, Стив Возняк (и много други) подписаха отворено писмо до Обединените нации (ООН) [4.5] с искане да се наложи забрана върху разработването на оръжия с изкуствен интелект с възможност за насочване и убиване без значима човешка интервенция. Всички държави-членки на ЮНЕСКО приеха през 2021 г. [1.4] историческо споразумение, което определя общите ценности и принципи, необходими за осигуряване на здравословното развитие на ИИ, едно глобално споразумение за етиката на изкуствения интелект [4.6].

В тази посока, в нашите дебати, хората подчертават възможността ИИ да стане толкова автономен, че да осъзнае себе си и да се отрица, развивайки се по начини, които могат да бъдат както лоши (машини, които се обръщат срещу хората), така и добри (повишаване на сигурността). Участниците смятат, че оптималното решение е, че с развитието на технологиите доверието ни в устройствата с ИИ се увеличава, но това доверие може да стигне само дотам, че все още е възможно безопасно да се поеме контролът върху ИИ по всяко време.

„От онова, което не познаваме, обикновено се страхуваме и страхът ни често приема формата на отхвърляне и омраза. Така е и с ИИ. Има само един път за човечеството: опознайте изкуствения интелект, за да можете съзнателно да решите дали го обичате или мразите.“

(експерт)

„Вярвам в интелекта на машината само когато той дава същия резултат като човешкия интелект“

(неспециалист)

Визията е, че в бъдещето - светът след 10 години, през 2032 г. - напредъкът може да бъде както добър, така и лош, и че има толкова много препятствия, които затрудняват представата за бъдещето въз основа на реалистични факти. Участниците във фокус групите основно биха искали да живеят в по-щастлив, по-безопасен и по-удобен свят благодарение на ИИ, но те не биха искали да бъдат толкова разглезени, че да останат със скучен житейски опит във всеки един аспект.

„Бих бил много по-спокоен, ако светогледът след 10 години е положителен и ще мога да кажа, че благодарение на ИИ родителите ми са в безопасност в интелигентния дом“

(неспециалист)

Следователно изглежда спешно разработването на нова парадигма [4.7], в която човечеството ще дефинира своята визия за институционалните системи на ИИ. Важно е да се има предвид, че ИИ не обслужва интересите на тясна прослойка (ориентирана към печалба, аморална, с манипулативна употреба на ИИ), а трябва да постави общественото благо над индивидуалните интереси. Всъщност има нарастващ брой разработки - не само във военната област - за които нито етичната, нито хуманната норма са верни.

„Препоръките, разпоредбите и законите за изкуствения интелект са толкова добри, колкото и броя им, който се спазва. Докато наказанието може да бъде възпиращо средство в някои случаи, за съжаление може да се каже, че като цяло никакво наказание няма да възпре онези, които разработват ИИ със зли намерения и за собствените си цели“

(експерт)

Необходима е индивидуална осведоменост за промените относно ИИ. Това включва превръщането на цифровото гражданство в неразделна част от ежедневието, насърчаване на социалния диалог относно ИИ: хората трябва да бъдат агенти на промяната, а не нейни жертви. Следователно е от съществено значение да се поддържат автентични и качествени медии и да се възстанови социалното доверие чрез разработване на нова парадигма, в която човечеството ще определи своята визия, институционалните системи на ИИ и реализацията на ИИ, ориентиран към човека [4.7].

„В миналото беше човек срещу човек, след това технология срещу технология, а сега изглежда, че е изкуственият интелект срещу изкуствен интелект. Може би щяхме да сме по-добре, ако бяхме избрали сътрудничество, вместо борба“

(експерт)

5. Бизнес

Смятан за ключов двигател на бъдещото икономическо развитие и растеж, ИИ се превърна в предложение за основна стойност за нарастващ брой нови стартап фирми, особено в Европа [5.1]. ИИ завладява всяка индустрия по света и мотивира бизнеса да се конкурира, за да стане субект, фокусиран върху ИИ. Конкуrentната бизнес среда принуждава корпоративни лидери, предприемачи, стратегии и изследователи да използват ИИ за разработване на нови стратегии и генериране на нови потоци от приходи [5.2]. Европейската комисия [5.3] посочва, че през 2017 г. 25% от големите предприятия в ЕС и 10% от малките и средните предприятия са използвали анализ на големи данни. Само един от всеки пет малки и средни предприятия е значително дигитализиран, а една трета от работната сила все още няма основни цифрови умения.

Междувременно предимствата на ИИ приложенията са широко признати. Ето няколко примера:

Предприятия в сектора на **търговията, хранително-вкусовата промишленост и строителството**, които са внедрили ИИ, отчитат отлични резултати при разработването на продукти или услуги, привличането на нови клиенти и навлизането на нови пазари [5.3].

В **производството** ИИ помага за подобряване на качеството на производствените системи и съответно на качеството на продуктите. ИИ също така позволява създаването на силно персонализирани стоки, отговарящи на интересите на клиентите.

Маркетингът е една от най-напредналите области, когато става въпрос за ИИ. Приложенията в маркетинга изследват как методите на ИИ могат да помогнат при прогнозирането как бъдещите разходи на нов клиент ще намалееят или ще нараснат след първоначалната покупка.

В **управлението** ИИ се използва широко в човешките ресурси за подобряване на процесите на вземане на решения, интегриращи технически, човешки и организационни системи за постигане на стратегически успех на предприятието.

ИИ направи революция в **онлайн пазаруването**, където големи компании за електронна търговия, като Amazon, Alibaba, eBay, внедриха ИИ, за да предложат препоръки за продукти, от които клиентите може да се интересуват, което доведе до значителен ръст на приходите [5.4].

ИИ е ключов компонент на популярните платформи в **социалните медии**, които се използват предимно за бизнес цели. LinkedIn използва изкуствен интелект, за да предоставя предложения за работа, да предлага нови мрежови възможности и разнообразно съдържание [5.1].

Експертите се съгласиха най-вече с факта, че ИИ ускорява задачите, особено тези, които не изискват сложни процеси и че следователно намалява лесните грешки, като същевременно спестява време.

„Използваме ИИ, за да правим анализ на настроеността и всички те са с предимство. Разходите, свързани с тези дейности, могат да бъдат неблагоприятни, както и намирането на хора, които могат да управляват тези процеси, е проблематично.“

(експерт)

Въпреки че има нарастващ интерес сред бизнеса към инвестиране и включване на ИИ в техните операции, съществуват значителни бариери на организационно ниво, които пречат на бизнеса да постигне пълния потенциал на ИИ. Някои от основните пречки в световен мащаб, възпрепятстващи приложенията на ИИ, са липсата на ясна стратегия за ИИ, културна съпротива, липса на талант, необходим за решения на ИИ, размер на предприятието и бюджетни ограничения [5.2].

Констатациите показват, че в Европа липсата на държавно финансиране и рисков капитал често се съобщават като финансови пречки пред развитието на ИИ, особено в малките и средни предприятия и нетехнологичните компании [5.5].

В това отношение Европа увеличи своите инвестиции и ангажимента си към изследванията на ИИ, за да увеличи потенциала за технологичен растеж и да настигне страните, водещи в надпреварата за ИИ. Според Европейската комисия [1.3] иновативните продукти и услуги могат да бъдат стимулирани от ИИ в области, в които Европа е отличник (машини, киберсигурност, транспорт, земеделие, зелена и кръгова икономика, здравеопазване и сектори с висока добавена стойност като мода и туризъм). Развитието на ИИ технологиите в Европа също означава намалена зависимост от чужди технологии, което е жизненоважно за стратегическата автономност на Европа и привежда ИИ технологиите в съответствие с европейските ценности. Такива прозрения в решенията за приемане на бизнес технологии са от съществено значение за насочване на законодателството и гарантиране, че ИИ технологиите са от полза както за работодателите, така и за служителите, като правят технологията надеждна, лесна за използване и полезна в ежедневната работа [5.6].

Независимо от това, в този процес на трансформация, бизнес светът е доста скептичен относно бързия темп на технологичните промени и техните въздействия [5.7]. Притесненията относно вземането на решения, поверителността, етиката и доверието също са налице и се очаква да нараснат в близко бъдеще. Както експертите, така и неекспертите подчертаха в нашите фокус групи значението на способността на ИИ да подобрява възможностите и перспективите в бизнеса, но също така беше посочено, че ИИ не трябва да замества човешките способности и вземането на решения и човешкото взаимодействие все още ще е необходимо.

„Хората са тези, които могат да определят нещата в области като биомедицината. Но все още можем да използваме ИИ за прости открития. И в по-големите медицински области е важно да се използва ИИ, но трябва да сме сигурни, че е безопасно и няма да причини големи грешки. Експертните познания не трябва да се заменят с ИИ, а по-скоро да се подобряват. ИИ трябва да дава препоръки, но не трябва да решава. Бих предпочел ИИ да подобри това, което хората могат да направят. Ние сме много далеч от точката, в която ИИ така или иначе може да замени човешкото знание. Освен това може да бъде опасно. Също така, отговорността („Дилемата с трамвая“) е друг въпрос, който трябва да се има предвид, когато става въпрос за вземане на решения с изкуствен интелект“.

(експерт)

„ИИ не може да бъде напълно внедрен, тъй като се нуждаем от взаимодействие и трябва да говорим с реални хора“.

(неспециалист)

„Човешкото взаимодействие е важно, когато имаме ИИ. Ние също трябва да запазим известна свобода и процес на вземане на решения като хора“

(неспециалист)

Най-честата причина за проблеми с поверителността е разкриването на лична информация; по този начин компаниите се стремят да си поставят конкретни цели, за да спечелят доверие. Например Гугъл заяви, че няма да преследва приложенията с ИИ, които са причинили или се очаква да причинят широко разпространена вреда и ще продължи само ако ползите надвишават рисковете и ще включва подходящи ограничения за безопасност, ако има материална опасност от вреда [5.8]. Докато повечето нации имат добре установени закони за защита на данните, ИИ има потенциала да генерира нови проблеми със защитата на данните, които не са третираны от законодателството, което поражда допълнителни етични проблеми. ИИ може също да използва или генерира видове лични данни, които сега се пренебрегват, като емоционални лични данни, допринасящи за проблема [5.9]. Повече подробности относно етичните аспекти са дадени в раздел 6.

Проучванията прогнозираят, че ИИ може да трансформира концепцията за работната сила, като закрие някои работни места. Изследване, проведено в [5.10] относно потенциалното въздействие на ИИ върху глобалната икономическа активност, подчертава силно потенциалното развитие на толкова нови работни места, колкото е броят на тези, които се заменят или ще бъдат заменени. Очаква се в близко бъдеще в Европа да има повече свободни работни места за инженери, разработчици на софтуер и специалисти по ИКТ. Според доклада на Делойт за тенденциите в човешкия капитал [5.11], новосъздадените работни места ще бъдат по-ориентирани към услугите, интерпретативни, социални, изискващи креативност, емпатия, комуникация и сложни умения за решаване на проблеми. Що се отнася до изследванията на Accenture [5.12], ръководените от ИИ работни места, които ще бъдат създадени, са обучителни, обясняващи и поддържащи. Тези нови работни места ще включват системи за обучение на ИИ, гарантиращи, че те продължават да работят според дизайна и няма да научат „погрешното“ нещо, и затваряне на пропастта между бизнеса и технологиите. Повече подробности за новата работа и образователна система, свързани с ИИ, са докладвани в раздел 7. Нашите участници също изразиха загриженост относно оптимизирането на разходите, което следва внедряването на определени ИИ и което води до по-високи нива на безработица за хората.

„Компаниите могат да се възползват от ИИ, за да обновят своя имидж и това прави добро впечатление. Също оптимизиране на разходите в човешките ресурси. Добре е от бизнес гледна точка. Но що се отнася до корпоративната социална отговорност и човешкия труд, това може да бъде проблематично, тъй като ИИ замества някои работни места“.
(неспециалист)

Всеобщото мнение е, че всичко зависи от сферата на работа.

„ИИ може да отвори врати за нови задачи, но в някои области може просто да замени човешкия труд и да намали сложността. Така че наистина зависи““.
(неспециалист)

И в заключение, ИИ ще бъде пазарна тенденция и бизнес възможност през следващото десетилетие. Предвижда се, че той ще допринесе с 15,7 трилиона долара за глобалния БВП и че този индекс ще бъде с 14 процента по-висок до 2030 г., което се дължи на ИИ. Анализаторите прогнозираят печалба от 6,6 трилиона долара в производството, с увеличение от 9,1 трилиона долара в потреблението. Ако Европа разработи и разпространи ИИ в съответствие с настоящите си активи и дигитална позиция по отношение на останалия свят, тя може да допринесе с 2,7 трилиона евро, или 20%, към комбинираната си икономическа продукция до 2030 г., което ще доведе до 1,4 процента общ годишен растеж. Това въздействие би било почти двойно по-голямо от предишните технологии с общо предназначение, които развитите страни са възприели в миналото [5.10].

6. Управление

Сега сме наясно, че ИИ е част от живота ни. Знаем за присъствието му и взаимодействието с него например, когато помолим Siri да намери ресторант за нас според нашите предпочитания за храна. Но в много други аспекти не сме напълно наясно, че ИИ присъства: например финансовите институти използват ИИ, за да идентифицират потенциално измамни дейности в нашите сметки; ИИ се използва за проследяване и прогнозиране на въздействието върху околната среда в земеделските полета, като се използват данни от сателитно сканиране и мониторинг на здравето на културите и почвата. Това са само няколко примера и според няколко проучвания епидемията от Ковид-19 е ускорила приемането на ИИ във всички сектори на икономиката [6.1].

Въпреки това ИИ не е само цветя и рози. Много учени посочват, че начинът, по който се произвеждат ИИ инструменти, трябва да се промени поради ограничения в сътрудничеството и допускания на неточни данни, като например неразумните очаквания, които водят до използването на ИИ системи, които не са достатъчно стабилни. Например, бездействието срещу предразсъдъците срещу ИИ доведе до много несправедливости срещу цели групи хора, расово профилиране и други смущаващи инциденти. Deepfakes и възможността за създаване на реалистични видеоклипове, снимки, текст, реч и друга форма на (социална) комуникация напоследък повдигнаха много етични и правни опасения относно използването на ИИ за манипулиране на човешките възприятия. В киберсигурността лошите актьори също имат достъп до ИИ инструменти, така че играта на котка и мишка продължава. Видеонаблюдението, базирано на ИИ за разпознаване на хора чрез тяхното лице, реч, ходене или движение, също повдигна някои опасения за поверителността. Алекса на Амазон наскоро предложи на 10-годишно момиче да пхне монета в електрически контакт, след като момичето е поискало предизвикателство от устройството [6.2].

В този сценарий на плюсове и минуси при работа с ИИ, прилагането на управление става фундаментално. Управлението се отнася до формирането, поддържането и регулирането на правила или дейности, както и възлагането на отговорност [6.3].

Когато ИИ е включен в термина управление, могат да се намерят две различни тълкувания: 1) Използването на системи, базирани на ИИ в управлението, което означава приемането на ИИ при предоставяне на услуги, създаване на политики и прилагане в екосистемите на публичния сектор [6.4]; 2) Управление на ИИ, което означава насърчаване на подходяща институционална и правна рамка за разработването и използването на ИИ [6.5]. Въпреки че са различни теми, не е възможно да се поддържа дискусия за ИИ в управлението, без да се вземе предвид управлението на ИИ, защото те работят като комуникиращи съдове. Под управлене тук се разбира това, което е известно като „ИИ управление“ - идея, съставена от три компонента, свързани с: а) инфраструктурата - получаване, съхраняване и обработка на данни; б) приложението - управление на данни; в) използването – процесите на вземане на решения и оценка на базата на данни.

Много други определения могат да бъдат намерени в литературата [6.6]-[6.8]. Вероятно най-пълната дефиниция е налична в [6.9], според която „управлението на ИИ е система от правила, практики, процеси и технологични инструменти, които се използват, за да се гарантира, че използването на технологии с ИИ от организацията е в съответствие със стратегиите, целите, и ценностите; отговаря на законите изисквания; и отговаря на принципите на етичния ИИ, следван от организацията“. Накратко, управлението на ИИ трябва да затвори пропастта, която съществува между отчетността и етиката в технологичния напредък [6.10] и да гарантира, че са поставени надеждни граници в рамките на технологията, така че да не вреди и допълнително да влошава неравенствата случайно, докато работи.

„Изследователски [проект] за разпознаване на човек въз основа на ириса беше финансиран от фондове, свързани с данъчен рай, за да се идентифицират жени с бурка и да се разбере дали са били със съпруга си или не. Бях много изненадан [...]. Какво трябва да се направи? Как се стигна дотук? Какви са лимитите?“

(неспециалист)

Съществува широко разпространен консенсус относно необходимостта от обсъждане на тези надеждни граници при разработването на ИИ системи, тъй като използването им може да има много важни отрицателни последици за живота на хората или да възпроизвежда социални модели, които се считат за морално осъдителни.

Ограниченията обаче не са ясни и е трудно да се установи или договори етична, политическа или регулаторна рамка, която може да управлява развитието на форми на ИИ, които след това могат да имат силно въздействие върху социалните решения. Една от трудностите, които възникват в това отношение, е напрежението между редица гаранции за гражданите и в същото време конкурентоспособността в научните изследвания и иновациите.

Участниците от двете групи подчертаха важни въпроси, свързани с тези специфични области:

- **Пристрастност към данните.** Акцентът е върху необходимостта да се гарантира, че събраните данни не са предубедени по пол, социално-икономическо ниво, етническа принадлежност и т.н. Гаранцията за разнообразие на данните и техния състав се отнася до използването на ИИ във всички етапи на процеса, събирането на данни, самото решение или оценката. Използването на машини за вземане на решения не е изключено от идеологията, която е в основата на всяко решение. Тези идеологии могат да представляват интереси на различни участници, като са от политическо, техническо или икономическо естество. Това е важен проблем, който трябва да бъде разрешен, за да се гарантира, че събраните данни и тяхното използване отговарят на целите, за които са предназначени.

„Хората вземат много решения въз основа на идеология (...) Една машина също ще вземе пристрастно решение. Предубедена от кого? Заради данните, заради инженера, който я е проектирал или компанията зад него, или идеологията на държавата, която я е финансирала“.

(експерт)

- **Справедливост.** Използването на способността да се обработват големи обеми от данни и да се правят статистически прогнози се разглежда като важна стойност на ИИ. Ако решенията засягат несъществени въпроси за живота на хората, възможна грешка в решенията на ИИ може да се счита за незначителен проблем и следователно ИИ може да се използва за вземане на решения по този конкретен въпрос. От друга страна, независимо дали решенията засягат съществени въпроси от живота на хората, погрешното решение може да има ужасно несправедливи последици, които биха обусловили живота на човека и следователно по този въпрос решенията не трябва да се вземат от системата..

„През годините изградихме важна съдебна система, която искаме да поддържаме. Има области, в които въздействието [на решенията, взети от ИИ] върху човека е много важно. ИИ не трябва да навлиза в тази област“.

(неспециалист)

- **Приватизация.** Някои компании или корпорации натрупват много знания за алгоритмите и за поведението на населението, което предполага липса на гаранции, че тези данни или тези знания се извършват при спазване на принципи или договорени етични ценности. В този смисъл натрупването на данни и знания в ИИ от субекти извън обхвата на държавен надзор означава приватизация на управлението, проблем, който трябва да бъде коригиран.

„Вярвам, че трябва да се създадат регулаторни институции по същия начин, по който има институции, които регулират банките и ги одитират, за да видят какво правят с парите. Трябва да одитирате тези компании като Гугъл, Нетфликс и подобни, за да видите какво наистина правят техните алгоритми.“

(експерт)

- **Автоматизация на решенията.** Дебатът за ограничаване на автоматизацията на процесите на вземане на решения не може да бъде обвързан с ИИ, по същия начин, по който ИИ не може да се счита единствено за отговорен за автоматизирането на решенията. Проблемът с ИИ е, когато тези, които проектират алгоритъм, не са в състояние да обяснят неговите решения, както и когато потребителите не знаят критериите, които ИИ дизайнерът е внедрил в алгоритъма. Може да попаднем в това, което се нарича „синдром на компютъра, който казва „да““ [6.11], при който служителите, които са прекарвали дълго време в работа със система, в която грешките са редки (както трябва да бъде случаят със системите с изкуствен интелект), естествено стават по-малко склонни да предизвикат коректността на системата във времето.

„Има алгоритми, които не са правилно регулирани, което може да бъде проблем. Също така съм скептичен, че ИИ няма да прави човешки грешки. Но какво ще стане, ако алгоритъмът е обучен погрешно или с някакво пристрастие (расизъм)? В този случай ИИ може да направи грешки. И така, кой ще вземе последното решение - ИИ или човекът? Не знаем как ИИ взема решение. Защо е взето каквото и да е решение, ние не знаем. Следователно трябва да се положат усилия, за да се направят системите с ИИ по-обясними, за да разберат защо решават за или против“.

(експерт)

- **Свободата.** Заплахата за свободата, породена от използването на ИИ системи в процесите на вземане на решения, може да бъде разбрана от две различни точки. Първото измерение се отнася до стратегиите, които използват ИИ за постигане на по-голяма реклама или визуализация, базирани на алгоритми, които карат потребителите да влизат в циклични процеси, които се използват от компании от типа Мета или Туитър. Второто измерение е свързано със способността на ИИ да контролира емоциите и да регулира чувствата. Като се има предвид дигиталната следа, която всички граждани оставят във всичките си ежедневни движения, получаването и използването на тези данни за търговски или авторитарни цели може да бъде много опасно.

„Имам тийнейджъри, които вярват на това, което виждат: на фалшивите новини, на белината, която се пие, за да се лекува Ковид. Имам дъщеря на 12 години. Виждам, че информацията, която вижда, е брутална опасност. Хората са впечатлителни и това е много сложно. Когато виждаме съдържание, когато сме млади, търсим новини, които очакваме, ние сме по-податливи. Ако видите някакво видео... Е, казвате „Искам да отида в Малибу“, „Искам чанта Прада“. Алгоритъмът ви движи“.

(неспециалист)

Мерките за проактивно управление стават все по-широко признати като отличителна черта за фирмите, които се стремят да създадат репутация на надеждни. Съществуват редица световни рамки относно концепциите за управление и етика на ИИ. Европейският съюз издаде Общия регламент за защита на данните (GDPR), който включва специален набор от правила, които се отнасят до правото на потребителя на обяснение, когато корпорациите използват алгоритми за извършване на автоматизирани избори. Независимо от това, той предизвика и някои спорове, тъй като не дава право на обяснение на автоматизираното вземане на решения [6.12]. В това отношение ЕС вероятно ще бъде първият, който ще приеме регулаторно законодателство за ИИ [6.13]. Законът за алгоритмичната отчетност [6.14] в САЩ изисква от големите компании с достъп до големи количества данни да извършват одит на базирани на системи с ИИ за справедливост, поверителност, точност и рискове за сигурността. Забележителна инициатива е Сингапурската рамка за управление на ИИ.

Това е първият модел, разработен в Азия и силата му е, че превежда принципите в практическа, оперативна рамка за незабавни действия, намалявайки бариерите на ИИ. Тази рамка се основава на два фактора: 1) ИИ решенията трябва да са ориентирани към човека и 2) решенията, взети или подпомогнати от ИИ, трябва да бъдат прозрачни, обясними и справедливи.

В обобщение, за да бъдат ефективни и да осигурят правилния компромис между стратегиите и целите на компанията, законовите изисквания и етиката, много участници работят върху идентифицирането на основните принципи. Например Харвардският университет [6.15] създаде карта за визуализация от 32 комплекта принципи на ИИ. КПМГ [6.16] предоставя четири насоки, за да помогне на организациите да гарантират правилното управление на алгоритмите. Гугъл [6.5] подчертава пет специфични области, в които точни, специфични за контекста насоки от правителства и гражданско общество биха помогнали за напредък в правното и етично развитие на ИИ. В нашата работа се разглеждат набор от шест принципа на ИИ в управлението, които са функционално независими от алгоритми, технологии и сектори:

Отчетността изисква ясна идентификация на това кой носи отговорност за решенията и действията при проектирането, разработването, експлоатацията и/или внедряването на ИИ система. Трябва да са хора или организации, които в крайна сметка са отговорни за действията на системите с ИИ, без значение колко сложна е системата.

Прозрачността се отнася до способността да се обясни защо дадена ИИ система се държи по определен начин, за да се повиши увереността и доверието на хората в точността и уместността на нейните прогнози. Определено, колкото повече потребителите смятат, че разбират цялостната ИИ система, толкова по-склонни ще бъдат да я използват.

Справедливостта трябва да гарантира, че ИИ системите са етични, без пристрастия, без предразсъдъци и че не се използват защитени атрибути. Има толкова много различни гледни точки за определяне на справедливостта и те могат директно да си противоречат. Това може да бъде смекчено, ако точната техника за справедливост, която се използва, бъде решена предварително и тя е прозрачна.

Безопасността се отнася до предприемането на мерки срещу неволна и умишлена злоупотреба с ИИ, която представлява заплахата за хората. Това обаче трябва да се направи по разумен начин, като се вземе предвид потенциалът за вреда и практичността на предложените превантивни мерки по отношение на технологични, правни, икономически и културни фактори.

Човешкият контрол означава, че хората трябва да бъдат в една или повече точки в процеса на вземане на решения на иначе автоматизирана система. Определено, независимо колко прецизна е една ИИ система, обществото иска човек да направи крайната преценка.

Принципът на **универсалност** препоръчва дефинирането и прилагането на технически, етични и регулаторни стандарти по време на разработването, оценката и внедряването на алгоритъм, за да има оперативна съвместимост, сътрудничество и дадено ниво на качество, безопасност и доверие.

7. Умения и компетенции

Въпреки младостта си, ИИ оказва влияние върху пазара на труда. Първо, поради автоматизацията някои работни места със средни умения изчезват. Второ, хората вече са по-склонни да използват ИИ в ежедневието си, включително в работата си, например около 50% от организациите съобщават, че използват ИИ [7.1]. Фокусирайки се върху работни места, които изчезват поради ИИ, участниците във фокус групите се споразумяха за три категории работни места, които могат да се заменят с ИИ: повтарящи се/рутинни работни места, където ИИ може да намали риска от грешки и да увеличи производителността (като складове/ фабрични работни места, шофьори, пилоти, служители на ишлеме); работни места в офис/бели якички (като секретарки, специалисти по набиране на персонал, адвокати, данъчни консултанти, съдействие, преводачи, помощник-юристи); и накрая работни места с физически труд (като товарене и разтоварване на камиони, (де)палетизиране на кутии).

„Например, складов работник, който в момента използва мотокар за преместване на палети, скоро може да управлява малък автопарк от автономни движещи се роботи (AMR)“

(експерт)

Поради това институциите за образование и обучение трябва да се адаптират, за да предоставят на учащите умения и компетенции, които са необходими в този бързо променящ се свят. Това е особено необходимо като начин за борба с недоверието и страха на хората от автоматизацията и цифровизацията, които заместват хората [7.2]-[7.4]. Наистина, мнозинството от европейците подкрепят правителствата да ограничат внедряването на автоматизация и цифровизация на работните места като начин за защита на работните места и запазване на работата на хората [7.2], [7.4]. Въпреки това, променящият се характер на работа и внедряването на нови технологии е неизбежно и 37% от респондентите в проучването на Gartner 2019 CIO заявяват, че вече са внедрили ИИ и/или биха го направили в близко бъдеще, за да се опитат да останат пред конкуренцията или най-малкото да не изостават [7.5]. Освен това по-ранните доклади за хора, фокусирани върху ИИ, ще бъдат изоставени и заменени от технология, която е допринесла/послужила за потвърждаване на погрешните схващания и страхове на хората относно ИИ. Вместо това по-новите доклади се концентрират върху това как ИИ създава работни места и/или позволява на работниците да имат по-пълноценни роли, като се освобождават от черна и/или опасна работа [7.6] – [7.8].

Въпреки че беше обърнато голямо внимание на страха от замяната на работните места с машини, по-малко се фокусира върху перспективата, че не самите работни места ще остаряват по своята същност, а по-скоро начинът на работа ще се промени и че определени умения и компетенции ще придобият значение, докато други ще бъдат отхвърлени [7.9]. Всъщност ползите от използването на системи с ИИ силно зависят от компетенциите и уменията на тези, които ги управляват, като липсата на такива умения е проблем номер едно за предприятията в рамките на ИИ проекти [7.9]. Следователно прилагането на съзнанието за изкуствен интелект и свързаните с него предизвикателства в учебната програма е определено като решаващо за справяне с промяната на работния опит и бизнеса [7.10].

Освен това ИИ може да помогне за повишаване на уменията и компетенциите на работниците [7.8], [7.11]. Достатъчно е само да погледнете примера на производителя на асансьори KONE, който е инсталирал Интернет на нещата (IoT) в своите асансьори и е използвал ИИ за анализ на данните, което позволява на техниците да бъдат информирани за потенциални проблеми и да извършват превантивна поддръжка [7.11]. Освен това, макар и да не е само по себе си с ИИ, хибридно симулационно обучение показва голям потенциал в медицинските и здравните професии [7.12].

Според скорозна адаптация на типологията [7.11], ИИ като двигател на иновациите в компаниите може да работи както за повишаване на компетентността, така и за разрушаване на компетентността. Докато първият насърчава съществуващите умения и знания, а вторият ги състарява. Тази типология може да бъде особено полезна за мениджърите, за да идентифицират и предвидят въздействието от внедряването на ИИ в компетенциите на тяхната компания. Докато в момента, както отбелязват авторите, повечето приложения на ИИ насърчават иновации, повишаващи компетентността, в бъдеще се предвижда увеличаване на иновациите, унищожавачи

компетентността, поради подобряването на машинното обучение, решаването на проблеми и разсъжденията (Paschen et al. 2020).

Тук се крие едно важно обещание в ИИ: валоризирането на тези умения и компетенции, които е невъзможно да бъдат възпроизведени от роботи. В [7.13] авторите заявяват, че „сложното социално взаимодействие и креативността са най-трудните неща за автоматизиране“. Наистина, поради този факт е малко вероятно преподавателите да бъдат заменени от ИИ, въпреки засиленото му внедряване в образованието, отчасти поради пандемията, която го направи крайно необходим [7.14]. Наистина ИИ започна да се използва по-широко в образованието и вероятно ще бъде пътят на бъдещето в образованието.

„Мисля, че се нуждаем от повече технически хора, повече знания за еволюцията на обществото [...]. И също така, от друга страна, от хората, които са повече в областта на управлението [...], които също разбират това ново партньорство, което е навсякъде ... На образователно ниво трябва да се опитаме да положим усилия да интегрираме този ИИ в цялата съществуваща база от знания“.

(експерт)

Работният документ [7.14] подчертава възможността на ИИ да подобри персонализацията и подобрите резултати от обучението. По-конкретно, ИИ позволява на (маргинализирани) хора да се възползват от ученето, въпреки че не могат да присъстват намясто. Той улеснява възможността за адаптиране и персонализиране на работния напредък спрямо индивида, като предлага начини за създаване на учебни планове, предпочитания и траектории. Експертите се съгласиха, че изкуственият интелект вече е променил средата на висшето образование, тъй като сега всеки може да научи нови умения поотделно.

„Очаква се ИИ да създаде гъвкаво и персонализирано образование чрез анализиране и откриване на слабите или различни страни на учениците“

(експерт)

Това може да стане например чрез помощта на МООС, чрез който не само може да се достигне до огромен брой студенти, но който също така включва гъвкав подход на обучение за учащи от всякакъв вид. От страна на учителите ИИ технологията може да изобрази инструмент за измерване на оценките, както и помощ за провеждане на урока и наблюдение на дискуссионни групи. Поради социално-емоционалните и творческите компоненти на ефективното преподаване, [7.14] оценява заместването на учителите с ИИ технологии като малко вероятно. Въпреки това те подчертават колко е важно учителите да се адаптират към новата цифрова ера чрез разработване на нови методологии и подобряване на адекватните умения.

„Би могло да бъде полезно да се дават часове по ИИ за по-малки ученици, за да знаят как да се справят с него и да знаят рисковете/предизвикателствата. Вече го използваме, но все още е трудно и не го обмисляме много. Може би ще е полезно да има тези курсове в училище, за да се научи повече за това.“

(неспециалист)

Едно подходящо умение, придобило значение в ерата на ИИ, е критичното мислене. В тази епоха, в която фалшивите новини могат лесно да бъдат разпространени чрез социалните медии, уменията за критично мислене и потвърждаването на източниците са от първостепенно значение [7.15]. Критичното мислене позволява на хората не само да достигат до заключения въз основа на наличните в момента доказателства, но също така техните вярвания да бъдат оспорвани по отношение на точността и уместността поради по-нови или различни източници и съответно да ги променят [7.16]. Освен това социалните умения остават особено важни поради трудността им да бъдат автоматизирани. Други умения са способността за адаптиране към промяната, работа в екип, решаване на проблеми, комуникация и подпомагане на клиентите при управление на проекти и използване на ИТ [7.16]. Учените до голяма степен са съгласни, че интегрирането на такова съдържание е от съществено значение за осигуряване на полезен преход към ерата на ИИ [7.10][7.17].

В тази връзка скорошно широкомащабно проучване [7.17] заключава, че необходимостта от физически, основни когнитивни и ръчни умения ще намалее поради поемането на управлението от ИИ. Следователно авторите са идентифицирали 56 ОЕНТУ (отделни елементи на нагласи, талант и умения), разделени в 13 групи умения, които сами по себе си са разделени в четири основни категории - Когнитивни, Междоличностни, Самоуправление и Цифрови. Те включват гореспоменатото Критично мислене, Работа в екип и т.н., но също така и други като Цифрово Владеее и Гражданство, Използване и Развитие на софтуер и Разбиране на Цифровите системи. Тези констатации са в съответствие с други скорошни изследвания, като [7.1], [7.9], [7.18], подчертаващи необходимостта от данни, технологични и дигитални познания, както и компетенции за решаване на проблеми, емпатия, комуникация, иновации, критично мислене, и работа в екип. Въпреки това, участниците в нашите фокус групи също посочиха, че тези умения не е задължително да станат по-широко разпространени:

„Има разлика между това, което е важно и това, което би било промотирано. И наистина не знам какви компетенции... освен очевидните като, ОК, дигитални компетенции и други неща... които всъщност биха били насърчавани от един ИИ свят сам по себе си”.

(неспециалист)

Освен това съществува схващането, че образователната система е (твърде) бавна за промяна и преподаване на уменията и компетенциите, необходими за един свят на ИИ.

„Образователната система все още се основава на логиката, че много знания трябва да се научат наизуст, докато сега те са на една ръка разстояние от всеки по всяко време. Въпреки че има малка промяна към образование, основано на компетентности, това е далеч не е достатъчно. Повече внимание върху технологичното образование, както и решаването на комплексни проблеми, например чрез работата по проекти трябва да заеме по-голяма част от образователната стратегия”.

(експерт)

Въпреки че прилагането на ИИ в образователен контекст осигурява голям брой предимства, недостатъците на този процес не трябва да се пренебрегват. Съгласно [7.14] разработването на всеобхватна публична политика за прилагане на ИИ за насърчване на устойчивото развитие е от съществено значение. Въпреки че самият ИИ може да даде възможност за включване, например чрез възможността за дистанционно обучение, предизвикателства като електричество, хардуер и достъп до интернет, разходи за данни, основни ИКТ умения, език, както и културната уместност на съдържанието са само някои аспекти, които трябва да се вземат предвид. Това важи конкретно за „по-малко развити“ страни, които са застрашени да бъдат изпреварени още повече, ако тези предизвикателства не бъдат разпознати и не се преодолеят по-задълбочено. По същия начин ИИ се състои от данните, с които се захранва, следователно качеството и включването на такива трябва да бъдат един от основните приоритети при разработването на ИИ технологии. Чрез прозрачност и ясен етичен кодекс може да се предотвратят поддържаните от ИИ неравенства [7.14].

Като последно съображение за настоящите промени, причинени от ИИ, трябва да се вземат предвид и новите работни места, създадени от него, тъй като има нужда от хора, които да обучават ИИ (т.е. да обучават ИИ да върши предвидената работа), да обясняват резултатите от ИИ (т.е. да обяснят на неспециалист как ИИ е стигнал до определено заключение) и да се поддържа отговорно използване на ИИ (т.е. да се гарантира, че ИИ системите работят правилно, безопасно и се използват отговорно) [7.8]. Тези нови работни места създават нужда от нови умения и образование; по този начин някои места като ProgeTigerProgramme - която стартира през 2012 г. в Естония - започват да насърчават внедряването на програмиране и роботика в учебните програми за предучилищни и начални училища, но също и на професионално ниво, за да подготвят хората за пазара на труда на бъдещето [7.1].

8. Хора и начин на живот

Фактът, че по-голямата част от населението има индивидуален смартфон и/или компютър, вече може да се счита за предпоставка за потреблението на определени приложения, софтуер и повече ИИ приложения [8.1], [8.2]. В нормалния си живот хората и по-специално потребителите (обикновени хора или професионалисти) могат да взаимодействат с настоящите основни типове ИИ и ИИ приложения, в зависимост от техния фокус [8.3]:

- **ИИ важни задачи.** ИИ е свързан с преобладаване на интелектуален анализ като машинен превод, автоматично преpraщане и извличане на информация, речева комуникация, компютърно зрение и извличане на данни.
- **ИИ, базиран на специфични инструменти.** Разликата между тази посока и горната е, че тук ИИ е проектиран да може да решава по-голям клас проблеми като еволюционни изчисления, разпознаване на шаблони, евристично програмиране и многоагентен подход.
- **ИИ според разработения модел на мислене.** ИИ се характеризира според разработения модел на мислене като търсене в онлайн пространството за решение(я), представяне на знания и машинно обучение.

Насочен към ежедневните дейности на по-голямата част от хората, а именно потребители (и отчасти също професионални разработчици), използването на ИИ е свързано с определени приоритетни приложения [8.3]. Някои от популярните реализации на ИИ (популярни в Хора и начин на живот) са свързани с дейности като:

- **Компютърно зрение.** Тази технология обработва визуална информация, за да извлече полезни знания. Тя включва много задачи [8.4]: откриване на сайт; проследяване на обекти; разпознаване на шаблон; сегментиране; оценка на дълбочината на разстоянието.
- **Биометрична идентификация.** Те са много и разнообразни, като: система за разпознаване на лица, ирисова идентификация, анализ на химичния състав на потта или миризмата на тялото, анализ на микровибрация на пръстите и микродвижения на ръцете, анализ на сърдечна честота и размер на сърцето, пръстови отпечатьци, анализ на действията на потребителя, почерк на клавиатурата, индивидуален почерк на автора, анализ на поза, идентификация на устните, ДНК тест и др. [8.5].
- **Обработка на естествен език, търсене и извличане на информация от текстове.** Те се използват за генериране на текстове, които са почти неразличими от човешките по стил [8.6].
- **Гласово разпознаване.** Намира широко приложение както в кол центрове, така и в образованието и в процеса на изучаване на чужди езици.
- **Синтез на речта.** Може да се използва за промяна на стила на речта, генериране на няколко гласа от един модел, генериране на непознати досега гласове, предаване на интонацията по модел, адаптиране към гласа на говорещия и много други.
- **Машинно зрение.** Това е приложението на компютърното зрение в индустрията и производството, например за броене на обекти на конвейерна лента, четене на серийни номера или търсене на повърхностни дефекти. Съвременните машини вече разпознават над 90% от обектите, като не само фиксират присъствието, но също така определят какво точно виждат [8.7]. Пример е PowerAI Vision на АйБиЕм [8.8].
- **Машинен превод.** В зависимост от езиковата двойка, предметната област и всъщност колко сходни са данните, използвани за обучение на модели за машинен превод, качеството на резултатите на различните системи може да варира значително [8.9].
- **Генериране на текст.** Позволява измерването на качеството на езиковите модели, например, чрез вероятността за отгатване на следващата дума от предишния контекст (обърканост на дума). Например, Google Brain позволява много отдалечени корелации да бъдат ефективно взети под внимание поради оригиналната схема за кодиране на позицията на думата, използваща трансформация на Фурие [8.9].

- **Диалогови системи (чатботове).** Те са свързани с взаимодействието между човек и превозно средство (например коли, автобуси, камиони, кораби и др.). По предназначение тези системи се разделят на три групи чатботове: с общо предназначение, цели и способни на диалог [8.10].
- **Анализ на тоналността.** Чрез тази дейност потребителите могат да определят отношението или емоционалната реакция на говорещия. Пример за популярни приложения са IBM Watson, Meaning Cloud и Salesforce Einstein.

Перспективите за развитие на ИИ са пряко свързани с развитието на компютърните технологии, ИКТ, електрониката, автоматизацията и други [8.11] [8.12]. Тяхното приложение ще стане по-осезаемо и ще бъде постоянна част от онлайн пазаруването и търговията (особено по време на епидемии), здравеопазването, транспорта, киберсигурността и други [8.13]. ИИ ще се превърне от услуга в постоянна част от живота на хората. Негативният ефект също го има.

*„По-голямата част от обществото се е фокусирало изцяло върху AI, други основни тенденции и основи в живота на хората се игнорират.“
(експерт)*

Промяната в действителния и бъдещия начин на живот на хората става все по-осезаема, в световен мащаб присъствието на ИИ се свързва с [8.13]:

- Освобождаване на хора от рутинни дейности, замяна или намаляване на интензивността на интелектуалния труд в определени професии до пълната замяна на специалисти от определени професии с интелигентни устройства;
- Изграждане на цифрово интерактивно информационно технологично пространство, където хората и мислещите машини ще си сътрудничат;
- Пълна интеграция на мислещи машини като работи в сложни и опасни места за работа, спасителни операции и други;
- Вземане на отговорни решения в сложни ситуации и процеси;
- Повишаване ефективността на обработката на информация с големи обеми данни;
- Подобряване на качеството на помощта в рутинна сфера от ежедневието;
- Подобряване качеството на професионалната помощ;

Когато говорим за ИИ, свързан с хората и начина на живот, трябва да вземем предвид и отношението на отделните хора и различните общества по този въпрос. Това се отнася не само за чисто техническите и практически, но и за психологическите и социалните аспекти, както и за зоната на комфорт на индивида и другите [8.14]. Някои хора са склонни да бъдат много консервативни, защото не се чувстват комфортно с технологията, която отнема контрола им.

*„Помага на човек в ежедневието, но улеснението би довело до привикване. Ако хората използват ИИ в по-трудни ситуации и ако човек няма ИИ, човек може да изпадне в паника.“
(неспециалист)*

Ясно е, че ИИ е технология, която се развива заедно с цифровизацията [8.15]. Хората използват ИИ в ежедневието си, но осъзнават, че този процес на интелектуална дигитализация трябва да бъде внимателно наблюдаван [8.16]. ИИ може да направи хората пристрастени и да ги направи по-мързеливи, разчитайки на интелекта на ИИ, а не на естествения им интелект. Това ще улесни живота им, но и ще ги лиши от социални комуникации и близост.

*„Ограниченото човешко мислене и липсата на саморазвитие могат да доведат до човешка зависимост от наличието на изкуствен интелект, което води до спад в развитието на хората.“
(експерт)*

Мнозина виждат въздействието на ИИ върху хората и начина на живот в увеличаване на човешките способности, но някои прогнозираат, че нарастващата зависимост на хората от автоматизирани системи ще подкопае способността им да мислят независимо, да предприемат действия и да комуникират ефективно с другите. Хората могат да загубят чувството си за истински ценности в

живота. ИИ привлича и насочва вниманието на хората в области, в които някой ги обвинява, вместо да насочва вниманието им към личното саморазвитие. Прави хората по-зависими от реклама и външни влияния. Това ги откъсва от спокойствието на живота сред природата.

ИИ има голям потенциал да промени живота на човека и да го направи по-продуктивен, ефективен и по-лесен. Животът ще продължи да се променя бързо и човек трябва да може да се адаптира към новите условия. Напредъкът в ИИ ще повлияе на това какво наистина означава да бъдеш човек през 21-ви век, но тези промени може да са неравностойни.

„По-младите хора ще се адаптират по-лесно и бързо към автоматизираните системи, а за по-възрастните ще им е все по-трудно да поддържат уменията и знанията си за работа с тях актуални.“

(неспециалист)

Промените, които ИИ ще внесе в живота, ще имат положителни и отрицателни ефекти върху ежедневието на хората. В обобщение можем да идентифицираме риска и възможностите, както следва:

Положителни

Подобрява ефективността на човешкия труд и увеличава свободното време, щастието и удовлетворението на хората.

Нови възможности и способности за подобряване на начина на живот, развиване на естествени интереси и таланти.

По-добри възможности за наблюдение и диагностика, предоставяне на персонализирани планове за лечение.

Печелене на време и производителност с автономен транспорт.

По-сигурно разкриване на престъпления.

Виртуални асистенти за разбиране и изпълнение на задачи, дадени от хора.

Автоматизирани системи за промяна на игрите и домашния живот.

По-здрави семейни връзки за намаляване на психическото напрежение при управлението на домакинството.

По-голяма индивидуализация, като например обучение, базирано на човешките нужди и способности.

Подобрения на инфраструктурата (облекчаване на трафика, подобрения на веригата за доставки и др.)

Отрицателни

Преориентиране на работната сила и загуба на работни места.

Икономически, правни, политически и регулаторни последици, които ще повлияят на начина на живот.

ИИ може да бъде ограничен от преминаване на етични или законови граници.

Събирането на данни може да наруши поверителността на личната информация.

Ограничаване на личното пространство и унищожаване на дигиталната поверителност.

Социално-икономическо неравенство с изчезването на милиони работни места с ниска до средна квалификация.

ИИ повдига въпроса за разбирането на човека за себе си и свободата. Надеждата е, че ИИ ще има повече положително, отколкото отрицателно въздействие върху хората. Човешката воля и съзнание остават източник на намерения и съдник на всички резултати. Машините са създадени, за да осигурят лекота и ефективност в пътуването от намерението до резултата. ИИ трябва да работи за подобряване на човешките дейности и преживявания, да спести време и да увеличи удовлетвореността от живота на хората. Очаква се в бъдеще изкуственият интелект да засегне още повече ежедневието на хората и силно ще зависи от контрола на хората, които управляват тези системи, както и от допустимото им приложение в обществото.

„Опасна е човешката глупост, която ограничава функциите на мозъка, както и бездействието“

(експерт)

9. Изводи

ИИ ще остане в ежедневието ни. Това е очевидно заключение и трябва да се справим с него. И така, как можем да направим това?

В този документ прегледахме много литературни източници и разговаряхме директно с експерти и неспециалисти, за да разберем по-добре проблема, опасенията, възможностите, които ИИ може да донесе на човечеството и неговите въздействия върху нашето общество. Така вече можем да предоставим набор от заключения и насоки.

Накратко, можем първо да обобщим нашите открития в четирите области, както следва:

Бизнес: ИИ се разглежда като голям потенциал, особено когато става въпрос за спестяване на време, предоставяне на по-бързи и иновативни решения, повишаване на удовлетвореността на клиентите, производителността и ефективността на разходите и максимизиране на продажбите, като същевременно се намалява вероятността от човешка грешка.

Основните рискове са свързани с намирането на баланс между икономическите интереси на компаниите и нарушаването на правата на гражданите по въпроси, свързани с неприкосновеността на личния живот и личната свобода.

Управление: има глобално съгласие по основен набор от шест принципа на ИИ, които са функционално независими от алгоритми, технологии и сектори, за да осигурят компромис между стратегиите и целите на компаниите, правните изисквания и етиката: отчетност, прозрачност, справедливост, безопасност, човешки контрол и универсалност.

Основните рискове са свързани с неспазването на един или повече от тези принципи, което води до неправилен резултат, пристрастност, липса на свобода, несправедливост и др.

Умения и компетенции: нуждата от физически, основни когнитивни и ръчни умения ще бъде намалена поради използването на ИИ, докато дигиталните компетенции, критичното мислене, работата в екип, емпатията, по-високите когнитивни умения, технологичните, както и социалните и емоционалните умения ще бъдат повече търсени. Приоритетите на образователните системи се нуждаят от промяна, за да отразят това, за да развият по-добре способностите на учениците.

Основният риск е свързан с „по-слабо развитите“ страни, които са застрашени да изостанат още повече.

Хора и начин на живот: ИИ трябва да помага така, че да направи живота на човека по-продуктивен, ефективен, сигурен и по-лесен. Може да имаме персонализирани възможности за наблюдение и диагностика, повече свободно време, възможност да развием естествените си интереси и таланти, по-добра и по-бърза инфраструктура, печалба във времето и продуктивност, по-сигурен живот и по-голяма индивидуализация.

Основният риск е свързан с факта, че ИИ може да бъде предубеден и да затвърди или дори да увеличи половите и расови различия и неравенство.

Второ, нашето изследване на ИИ ни позволява да установим, че ИИ е замислен по два различни начина. Тези два начина се появяват последователно във всички наши източници, което позволява фокусът да бъде върху различни въпроси и предложения. Въпреки че тези две гледни точки могат да се разбират като противоречащи си, те всъщност са допълващи се гледни точки, които позволяват да се разбере сложността на мненията, опасенията и предложенията около използването на ИИ системи.

Ограничителна гледна точка: ИИ е още една технология и следователно трябва да се третира като всяка друга технология.

Разрушителна гледна точка: ИИ е различна технология, която маркира граница преди и след в човешкото общество и връзката с

От ограничителната гледна точка ИИ има склонност да се разглежда като шанс: той може да създаде нови, много квалифицирани и платени работни места, да отвори нови пазарни и бизнес възможности, да направи живота по-лесен и здравословен, бюрокрацията по-бърза. Общата идея е, че ИИ е полезен в процесите на вземане на решения. Предизвикателствата, които могат да бъдат повдигнати от тези процеси, се считат за ограничени и е възможно да се преодолеят чрез очертаване на ясна граница между това кога ИИ може да се използва и кога не. В този смисъл се смята, че ИИ е особено ефективен за управление и анализ на данни, както и информационна помощ за вземане на решения и оценка, но не и за автоматизирано вземане на решения. В този смисъл се смята, че решенията, които имат пряко въздействие върху хората, трябва да се вземат от хората.

В контраст, от разрушителната гледна точка за ИИ, впечатлението е, че в съвременните общества всяка сфера интегрира или ще интегрира ИИ. Смята се, че въпреки че не искаме ИИ да участва в много аспекти от нашето ежедневие, жизненоважно е да анализираме разходите и ползите въз основа на оценката на това какво би се случило, ако системите на ИИ вземат неправилни решения. Това може да засегне няколко области от нашата среда и в тези специфични области трябва да се оценят рисковете от използването на ИИ системи. Справедливостта, неприкосновеността на личния живот на хората (т.е. свободата), самите алгоритми (известни още като пристрастие на данните), биомедицината, финансите са само няколко примера за идентифицираните области.

Споделена загриженост, която е свързана най-вече с ограничителния възглед за ИИ и която се появява както изрично, така и имплицитно във всички източници, е свързана с връзката на ИИ системите с научната фантастика или с идеята, че ИИ може да реши всички проблеми от всякакъв вид. Многобройни приложения са разработени в областта на ИИ и могат да бъдат приложени в много области, но има значителна разлика между настоящите функционалности и технически възможности и разказа за това какво ИИ може да направи в бъдеще. Счита се, че този тип разказ около ИИ, който не съответства на настоящото развитие, има два вида отрицателни ефекта: 1) Трудността да се формулира доказан публичен дебат относно отчетността при използване на форми на ИИ в процесите на вземане на решения; 2) Появата на поредица от катастрофални видения, които генерират нежелание към ИИ.

Струва си да се споменат и темите, които не са обсъдени в този документ. Един от тях, вероятно най-важният, е въздействието на ИИ върху нашата околна среда и как той може да помогне при настоящата криза, дължаща се на изменението на климата. Всъщност ние не успяхме да представим този въпрос във всички наши фокус групи, но също така не беше повдигнат от участниците, което може да означава, че околната среда все още не се смята за основен проблем. Независимо от това, ИИ е основен потребител на енергия предвид сложността на обучението и изводите върху големи данни, и предвид, че цялата ИКТ екосистема вече е един от основните фактори, допринасящи за емисиите на парникови газове [9.1]. От друга страна, ИИ също беше представен като решение за изменението на климата поради неговите многофункционални възможности, които включват проследяване и намаляване на емисиите, позволяващи креативни икономически модели за подпомагане на околната среда и повишаване на устойчивостта на климата. Например, проучване, поръчано от Microsoft [9.2], заключава, че използването на ИИ за екологични приложения има потенциала да увеличи глобалния БВП с 3,1 – 4,4%, като същевременно намали глобалните емисии на парникови газове с около 1,5 – 4,0% до 2030 г. спрямо бизнеса, до 2,2% в енергетиката и до 1,7% в транспорта.

Едно споделено мнение е, че социалните и интелектуални умения на човечеството като креативност, емпатия, работа в екип, иновации са незаменими от ИИ. В краткосрочен план визията е, че ще се сблъскаме с появата на повече артисти. Въпреки това тази визия изглежда твърде оптимистична: ИИ вече може да композира симфонии, да рисува картини, да пише стихове, песни и истории и да играе игри. Някои страни като Австралия [9.3] вече приеха, че ИИ машина може да бъде регистрирана като изобретател в патент. В бъдеще е вероятно тези възможности да бъдат по-проучени.

Общоприето мнение е също така, че изкуственият интелект не трябва да замества човешките способности и човешката свобода за вземане на решения не трябва да бъде повлиявана от управлявани от ИИ инструменти. По-конкретно, етиката е повтарящ се проблем, повдигнат във всички наши източници. Дори експертите, които интервюирахме, смятат, че нямат достатъчно знания, за да могат да решават по етични и социални въпроси; знания, които трябва да бъдат интегрирани по интердисциплинарен начин.

За да се избегне този тип наратив и неговите отрицателни ефекти, са необходими действия, свързани с информацията и гражданското участие:

Информация: Гарантиране, че медиите докладват етично и честно, когато говорят за ИИ системи, което позволява ясно разграничение между спекулативните футуристични визии и текущите разработки и възможности. Необходимо е да се разработи педагогическа задача, която позволява на обществеността да знае как работи ИИ и какви приложения се използват. Подобно знание може да повиши увереността и доверието на хората в точността и целесъобразността на системите с изкуствен интелект.

Участие: Включване на обществеността в установяването на приоритети за развитието на ИИ. Това се счита за добавената стойност на европейската стратегия за развитие на ИИ в сравнение с други стратегии, които може да са по-напреднали в технологично отношение, като тези на Китай или Съединените щати. Счита се, че европейската стратегия може да включи като добавена стойност към своя ИИ интегрирането на гражданите в установяването на приоритетни области, в които да се развива или прилага.

В заключение, ИИ е технология, която в своя дизайн и развитие е толкова далеч от ежедневието, че експертите смятат, че населението не е достатъчно обучено, за да взема решения как да използва ИИ. Въпреки че в същото време се счита за необходимо хората да вземат решения относно хода на ИИ. Поради тази причина изтъкваме необходимостта от обучение на хората за работата, потенциала и възможните ефекти от ИИ. Следователно трябва да осигурим преподаване, курсове и обучения в училища и институти за висше образование, за да улесним използването и приемането на ИИ от младите хора и бъдещите поколения. Това наистина е основната цел на проекта HEDY. През следващите две години HEDY ще предостави масивен отворен онлайн курс точно с цел да достигне до аудиторията с висше образование и да им покаже способността на ИИ, възможността, която нашето общество има в този момент да промени средата ни към по-добро, но също така и рисковете, пред които сме изправени от различни гледни точки. Този материал ще бъде допълнен с Основен Наръчник: набор от поддържащи мултимедийни инструменти с възможност за по-непосредствено и визуално въздействие върху аудиторията.

A. Библиография

- [1.1] HEDY project, *Life in the AI era*, KA220-HED 0C8D3623 - Cooperation partnerships in higher education, <https://lifeintheaiera.eu> , достъп октомври 2022.
- [1.2] European Commission, *Artificial intelligence for Europe*, COM(2018) 237 final, Brussels, April 2018.
- [1.3] European Commission, *White paper on artificial intelligence - A European approach to excellence and trust*, COM(2020) 65 final, Brussels, February 2020.
- [1.4] UNESCO, *Recommendation on the ethics of artificial intelligence*, SHS/BIO/REC-AIETHICS, 2021, <https://unesdoc.unesco.org/ark:/48223/pf0000380455>, достъп април 2022
- [2.1] Henning Kagermann, Wolf-Dieter Lukas, Wolfgang Wahlster, "Industrie 4.0: mit dem Internet der dinge auf dem weg zur 4. industriellen Revolution", VDI nachrichten, April, 2011, <https://web.archive.org/web/20130304101009/http://www.vdi-nachrichten.com/artikel/Industrie-4-0-Mit-dem-Internet-der-Dinge-auf-dem-Weg-zur-4-industriellen-Revolution/52570/1>, достъп април 2022.
- [2.2] Thomas Philbeck, Nicholas Davis, "The fourth industrial revolution", *Journal of International Affairs*, vol. 72, no. 1, pp. 17–22, 2018.
- [3.1] Alan M. Turing, "On computable numbers, with an application to the entscheidungsproblem", in *Proceedings of the London Mathematical Society*, vol. s2-42, no, 1, pp. 230-265, November-December 1936.
- [3.2] Alan M. Turing, "Computing machinery and intelligence", *Mind*, vol. LIX, no. 236, pp. 433–460, October 1950.
- [3.3] John McCarthy, "What is artificial intelligence?", *Project JMC*, November 2007, <http://jmc.stanford.edu/articles/whatisai.html>, достъп октомври 2022.
- [3.4] Encyclopedia Britannica, *Definition of artificial intelligence*, last update August 24, 2022, <https://www.britannica.com/technology/artificial-intelligence>, достъп октомври 2022.
- [3.5] Cambridge Dictionary, *Definition of artificial intelligence*, Cambridge university press, <https://dictionary.cambridge.org/dictionary/english/artificial-intelligence>, достъп октомври 2022.
- [3.6] Max, Tegmark, *Life 3.0: being human in the age of artificial intelligence*, Knopf, 1st edition, August 2017.
- [3.7] Csaba Kollár, "A mesterséges intelligencia és a kapcsolódó technológiák bemutatása a biztonságtudomány fókuszában", *Kiberbiztonság/Cybersecurity*, vol. 2, pp. 47-61, edited by Zoltán Rajnai, Doctoral School of Security Sciences, Budapest, Hungary, 2019.
- [3.8] Shelly Fan, *Will AI replace us: a primer for the 21st century*, The big idea series, Thames and Hudson, 1st edition, April 2019.
- [3.9] Council of Europe, *History of artificial intelligence*, December 2018, <https://www.coe.int/en/web/artificial-intelligence/history-of-ai>, достъп октомври 2022.
- [3.10] Rebecca Reynoso, *A complete history of artificial intelligence*, G2 report, May 2021, <https://www.g2.com/articles/history-of-artificial-intelligence>, достъп октомври 2022.
- [5.1] Csaba Kollár, "A mesterséges intelligencia kapcsolata a humán biztonsággal", *Nemzetbiztonsági Szemle*, vol. 6, no. 1, 2018.

- [4.1] Csaba Kollár, “A mesterséges intelligencia kapcsolata a humán biztonsággal”, *Nemzetbiztonsági Szemle*, vol. 6, no. 1, 2018.
- [4.2] Peter Stone et al., *Artificial intelligence and life in 2030: one hundred year study on artificial intelligence*, Report of the 2015-2016 Study Panel, September 2016, <https://ai100.stanford.edu/2016-report>, достъп октомври 2022.
- [4.3] Kathe Pelletier et al., *2021 EDUCAUSE horizon report - teaching and learning edition*, Educause publications, Horizon report, April 2021, <https://library.educause.edu/resources/2021/4/2021-educause-horizon-report-teaching-and-learning-edition>, достъп октомври 2022 .
- [4.4] OECD, *Recommendation of the council on artificial intelligence*, OECD/LEGAL/0449, adopted on 22/05/2019, <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>, достъп октомври 2022.
- [4.5] *An open letter to the United Nations convention on certain conventional weapons*, July 2015, <https://www.cse.unsw.edu.au/~tw/ciair/open.pdf>, достъп октомври 2022.
- [4.6] United Nations, *193 countries adopt first-ever global agreement on the Ethics of Artificial Intelligence*, UN news, November 2021, <https://news.un.org/en/story/2021/11/1106612>, достъп октомври 2022.
- [4.7] Omar Hatamleh, George Tilesh, *BetweenBrains: Taking back our AI future*, GTPublishDrive, May 2020.
- [5.1] Sandra Maria Correia Loureiro, João Guerreiro, Iis Tussyadiah, “Artificial intelligence in business: State of the art and future research agenda”, *Journal of Business Research*, vol. 129, pp. 911–926, May 2021.
- [5.2] James Eager et al., *Opportunities of artificial intelligence*, Study requested by the ITRE committee, European Parliament, June 2020, [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/652713/IPOL_STU\(2020\)652713_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/652713/IPOL_STU(2020)652713_EN.pdf), достъп октомври 2022.
- [5.3] Laurent Probst et al., *EU businesses go digital: Opportunities, outcomes and uptake*, Digital transformation scoreboard 2018, European Commission, March 2019.
- [5.4] Neha Soni, Enakshi Sharma, Narotam Singh, Amita Kapoor, “Impact of artificial intelligence on businesses: from research, innovation, market deployment to future shifts in business models”, *arXiv.org*, May 2019, <http://arxiv.org/abs/1905.02092>, достъп октомври 2022.
- [5.5] Awishkar Ghimire, Surendrabikram Thapa, Avinash Kumar Jha, Surabhi Adhikari, Ankit Kumar, “Accelerating business growth with big data and artificial intelligence”, in *Proceedings of the Fourth International Conference on IoT in Social, Mobile, Analytics and Cloud (I-SMAC 2020)*, Palladam, India, October 2020.
- [5.6] Mia Hoffmann Laura Nurski, *What is holding back artificial intelligence adoption in Europe?* Bruegel policy contribution issue 24/21, November 2021, <https://www.bruegel.org/wp-content/uploads/2021/11/PC-24-261121.pdf>, достъп октомври 2022.
- [5.7] Fotis Kitsios, Maria Kamariotou, “Artificial intelligence and business strategy towards digital transformation: a research agenda”, *MDPI Sustainability*, vol. 13(4), 2021, February 2021.
- [5.8] Denise Carter, “How real is the impact of artificial intelligence? The business information survey 2018”, *Business Information Review*, vol. 35, no. 3, pp. 99-115, July 2018.

- [5.9] Bernd Carsten Stahl, *Ethical Issues of AI*, Artificial Intelligence for a Better Future, pp. 35-53, Springer, March 2021.
- [5.10] Jacques Bughin, Jeongmin Seong, James Manyika, Lari Hämäläinen, Eckart Windhagen, Eric Hazan, *Notes from the AI frontier tackling Europe's gap in digital and AI*, Discussion paper, Mckinsey Global Institute, February 2019.
- [5.11] Dimple Agarwal, Josh Bersin, Gaurav Lahiri, Jeff Schwartz, Erica Volini, *The rise of the social enterprise*, 2018 Deloitte global human capital trends, University Press, 2018, https://www2.deloitte.com/content/dam/insights/us/articles/HCTrends2018/2018-HCTrends_Rise-of-the-social-enterprise.pdf, достъп октомври 2022.
- [5.12] İzzet Kılınc, Aslihan Ünal, "AI is the new black: effects of artificial intelligence on business world", *Journal of Contemporary Administrative Science*, vol. 2, no. 6, pp. 238-258, September 2019.
- [6.1] Melanie Malluk Batley, *AI adoption accelerated during the pandemic but many say it's moving too fast: KPMG survey*, Thriving in an AI World, KPMG study, March 2021, <https://info.kpmg.us/news-perspectives/technology-innovation/thriving-in-an-ai-world/ai-adoption-accelerated-during-pandemic.html>, достъп октомври 2022.
- [6.2] BBC news, Alexa tells 10-year-old girl to touch live plug with penny, December 28, 2021, <https://www.bbc.com/news/technology-59810383>, достъп август 2022.
- [6.3] Marc Hufty, "Investigating policy processes: the governance analytical framework", *Research for Sustainable Development: Foundations, Experiences, and Perspectives*, pp. 403-424, edited by U. Wiesmann, et al., Geographica Bernensia, Bern, Switzerland, 2011.
- [6.4] Anneke Zuiderwijk, Yu-Che Chen, Fadi Salem, "Implications of the use of artificial intelligence in public governance: a systematic literature review and a research agenda", *Government Information Quarterly*, vol. 38, no. 3, July 2021.
- [6.5] *Perspectives on Issues in AI Governance*, Google report, <https://ai.google/static/documents/perspectives-on-issues-in-ai-governance.pdf>, достъп октомври 2022.
- [6.6] James Butcher, Irakli Beridze, "What is the state of artificial intelligence governance globally?", *The RUSI Journal*, vol. 164, n. 5-6, pp. 88-96, November 2019.
- [6.7] Johannes Schneider, Rene Abraham, Christian Meske, Jan vom Brocke, "AI governance for businesses", *arXiv.org*, November 2020, <https://doi.org/10.48550/arXiv.2011.10672>, достъп октомври 2022.
- [6.8] Alan F.T. Winfield, Marina Jirotko, "Ethical governance is essential to building trust in robotics and artificial intelligence systems", *Philosophical Transactions of the Royal Society A*, vol. 376, no. 2133, 20180085, November 2018.
- [6.9] Matti Mäntymäki, Matti Minkkinen, Teemu Birkstedt, Mika Viljanen, "Defining organizational AI governance", *AI and Ethics*, February 2022.
- [6.10] KOSA AI, *The importance of AI governance and 5 key principles for its guidance*, <https://kosa-ai.medium.com/the-importance-of-ai-governance-and-5-key-principles-for-its-guidance-219798c8f407>, достъп октомври 2022.
- [6.11] Kevin Hoff, Masooda Bashir, "Trust in Automation: Integrating Empirical Evidence on Factors That Influence Trust", *Human Factors - The Journal of the Human Factors and Ergonomics Society*, vol. 57, no. 3, pp. 407-434, May 2015.
- [6.12] Sandra Wachter, Brent Mittelstadt, Luciano Floridi, "Why a right to explanation of automated decision-making does not exist in the general data protection regulation", *International Data Privacy Law*, vol. 7, no. 2, pp. 76-99, May 2017.

- [6.13] European Commission, *Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*, COM(2021) 206 final, Brussels, April 2021.
- [6.14] US Congress, *Algorithmic Accountability Act of 2019*, H.R.2231, 116th Congress, April 2019, <https://www.congress.gov/bill/116th-congress/house-bill/2231>, достъп август 2022.
- [6.15] Jessica Fjeld, Nele Achten, Hannah Hilligoss, Adam Nagy, Madhulika Srikumar, *Principled artificial intelligence: mapping consensus in ethical and rights-based approaches to principles for AI*, Berkman Klein Center Research Publication No. 2020-1, February 2020.
- [6.16] Martin Sokalski, *The shape of AI governance to come*, KPMG Insights, December 2020, <https://assets.kpmg/content/dam/kpmg/xx/pdf/2021/01/the-shape-of-ai-governance-to-come.pdf>, достъп април 2022.
- [7.1] Kelly Shiohira, *Understanding the impact of artificial intelligence on skills development*, Education 2030, UNESCO-UNEVOC, March 2021.
- [7.2] Diego Rubio, Carlos Lastra, *Mapping European attitudes to technological change and its governance*, Center for the Governance of Change, European Tech Insights, 2019, <https://docs.ie.edu/cgc/European-Tech-Insights-2019.pdf>, достъп октомври 2022.
- [7.3] Carlos Lastra, Diego Rubio, *Unveiling the technological future that citizens want and their concerns in a changing world*, Center for the Governance of Change, European Tech Insights, 2020, <https://docs.ie.edu/cgc/CGC-European-Tech-Insights-2020.pdf>, достъп октомври 2022.
- [7.4] Oscar Jonsson, Carlos Luca de Tena, *Part I How the pandemic altered our relationship with technology*, Center for the Governance of Change, European Tech Insights, 2021, <https://docs.ie.edu/cgc/IE-CGC-European-Tech-Insights-2021.pdf>, достъп октомври 2022.
- [7.5] Chris Howard, Andy Rowsell-Jones, *2019 CIO survey: CIOs have awoken to the importance of AI*, Gartner Research, January 2019.
- [7.6] Jenny Burke, Michael Coovert, Robin R. Murphy, Jennifer Riley, Erika Rogers, "Human-robot factors: robots in the workplace", in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, no. 9, pp. 870-874, October 2006.
- [7.7] Albert Ellis, Steve Bates, *A changing perspective*, Harvey Nash/KPMG CIO Survey 2019, https://assets.kpmg/content/dam/kpmg/kz/pdf/2019/09/CIO-Survey_2019_ENG.pdf, достъп октомври 2022.
- [7.8] H. James Wilson, Paul R. Daugherty, "Collaborative intelligence: Humans and AI are joining forces", *Harvard Business Review*, vol. 96, no. 4, pp. 115–123, July/August 2018.
- [7.9] Eduard Anton, Alina Behne, Frank Teuteberg, "The humans behind artificial intelligence – an operationalisation of AI competencies", in *Proceedings of the 28th European Conference on Information Systems (ECIS2020)*, June 2020.
- [7.10] David Chrisinger, "The solution lies in education: artificial intelligence & the skills gap", *On the Horizon*, vol. 27, no. 1, pp. 1–4, March 2019
- [7.11] Ulrich Paschen, Christine Pitt, Jan Kietzmann, "Artificial intelligence: building blocks and an innovation typology", *Business Horizons*, vol. 63, no. 2, pp. 147–155, November 2019.
- [7.12] Jean F. Byrd, Fabien Pampaloni, Linda Wilson, "Hybrid simulation", *Human simulation for nursing and health professions*, pp. 267–271, edited by L Wilson, L. Rockstraw, Springer, New York, USA, 2012.

- [7.13] Carl Benedikt Frey, *Technology trap: capital, labor, and power in the age of automation*, Princeton University Press, June 2019.
- [7.14] Francesc Pedró, Miguel Subosa, Axel Rivas, Paula Valverde, *Artificial intelligence in education: challenges and opportunities for sustainable development*, Education 2030, UNESCO, Paris, France, 2019.
- [7.15] Julian McDougall, "Media Literacy versus Fake News: Critical thinking, resilience and civic engagement", *Medijske studije*, vol. 10, no. 19, pp. 29–45, October 2019.
- [7.16] Lisa French, Mark Poole, *New competencies for media and communication in an AI era*, Humanistic futures of learning - perspectives from UNESCO chairs and UNITWIN networks, UNSCO, pp. 136–140, 2020.
- [7.17] Marco Dondi, Julia Klier, Frédéric Panier, Jörg Schubert, "McKinsey: these are the skills you will need for the future of work", *World Economic Forum*, June 2021, <https://www.weforum.org/agenda/2021/06/defining-the-skills-citizens-will-need-in-the-future-world-of-work/>, достъп октомври 2022.
- [7.18] Giselle Rampersad, "Robot will take your job: Innovation for an era of artificial intelligence", *Journal of Business Research*, vol. 116, pp. 68–74, August 2020.
- [8.1] Wadzani A. Gadzama, Joseph Bitrus, Ngubdo Maigana A, "Global smartphone ownership, Internet usage and their impacts on humans", *Journal of Communications Networks*, vol. 1, no. 1, October 2017.
- [8.2] Noshir Kaka, Anu Madgavkar, Alok Kshirsagar, Rajat Gupta, James Manyika, Kushe Bahl, *Digital India: Technology to transform a connected nation*, McKinsey Global Institute, March 2019, <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/digital-india-technology-to-transform-a-connected-nation>, достъп октомври 2022.
- [8.3] Moscow Industrial Development Agency, *Technologies of Artificial Intelligence*, 2019, <https://apr.moscow/content/data/6/11%20Технологии%20искусственного%20интеллекта.pdf>, достъп октомври 2022.
- [8.4] Maxim Akimov, Herman Gref, Alexander Vedyakhin, *Analytical review of the global robotics market*, Sberbank Robotics Laboratory, Russia, 2019.
- [8.5] Tadviser, *Biometric identification technologies*, Russian analytical agency, November 2019, https://tadviser.com/index.php/Article:Biometric_identification_technologies, достъп октомври 2022.
- [8.6] Bohdan Balov, *Convolutional neural networks from scratch*, Medium.com, February 2019, <https://medium.com/@balovbohdan/сверточные-нейронные-сети-с-нуля-4d5a1f0f87ec>, достъп октомври 2022.
- [8.7] Center2M, *What is machine vision and how is it different from human vision? Now let's make it clear!*, Meduza.io, March 2019, <https://meduza.io/feature/2019/03/30/что-такое-машинное-зрение-и-чем-оно-отличается-от-человеческого-сейчас-обязательно-понятно>, достъп октомври 2022.
- [8.8] IBM PowerAI Vision, <https://www.ibm.com/docs/en/mvi/1.1.0?topic=overview>, достъп октомври 2022.
- [8.9] Igor Pivovarov (editor), *The Artificial Intelligence almanac*, Analytical Collection No. 2, Center of the National Technology Initiative, MIPT, Russia, September 2019.
- [8.10] Pierrick Milhorat, Stephan Schlögl, Gerard Chollet, Jérôme Boudy "Multi-step natural language understanding", in *Proceedings of 14th Annual Meeting of the Special Interest Group on Discourse and Dialogue* (SIGdial 2013), Metz, France, August 2013.

- [8.11] Boris Mirkin, *Artificial intelligence: history and current state*, Report at the second Nizhny Novgorod festival of science and art, November 2010, <https://www.hse.ru/data/2010/12/05/1209601907/MachineIntel.pdf>, достъп октомври 2022.
- [8.12] Vamsi Vedantam, *Artificial intelligence in information and cyber security*, Tech Mahindra, January 2021, https://www.researchgate.net/publication/349350306_Artificial_Intelligence_in_Information_and_Cyber_Security, достъп октомври 2022.
- [8.13] Chenzhuoer Li, Runjie Pan, Huiyu Xin, Zhiwen Deng, "Research on artificial intelligence customer service on consumer attitude and its impact during online shopping", in *Proceedings of 5th Annual International Conference on Information System and Artificial Intelligence (ISAI2020)*, Zhejiang, China, May 2020.
- [8.14] Vyacheslav Ovchinnikov, *Doroga v mir iskusstvennogo intellekta* [Road to the World of Artificial Intelligence], Institute of Economic Strategies, Moscow, Russia, 2017.
- [8.15] Maria José Sousa, Gabriel Osório de Barros, Nuno Tavares, "Artificial intelligence a driver for digital transformation", *Digital transformation and challenges to data security and privacy*, Chapter 14, pp. 234-251, edited by Pedro Fernandes Anunciação, Cláudio Roberto Magalhães Pessoa, George Leal Jamil, IGI Global, February 2021.
- [8.16] Philip Boucher, *Artificial intelligence: How does it work, why does it matter, and what can we do about it?*, Directorate-General for Parliamentary Research Services, European Parliament, June 2020, <https://data.europa.eu/doi/10.2861/44572>, достъп октомври 2022.
- [9.1] Charlotte Freitag, Mike Berners-Lee, Kelly Widdicks, Bran Knowles, Gordon S. Blair, Adrian Friday, "The real climate and transformative impact of ICT: A critique of estimates, trends, and regulations", *Patterns review*, vol. 2, no. 9, September 2021.
- [9.2] Celine Herweijer, Benjamin Combes, Jonathan Gillham, *How AI can enable a sustainable future*, Microsoft and PwC report, April 2019, <https://www.pwc.co.uk/sustainability-climate-change/assets/pdf/how-ai-can-enable-a-sustainable-future.pdf>, достъп октомври 2022.
- [9.3] Rebecca Currey, Jane Owen, "In the courts: Australian court finds AI systems can be inventors", *World Intellectual Property Organization magazine*, September 2021, https://www.wipo.int/wipo_magazine/en/2021/03/article_0006.html, достъп октомври 2022.
- [A.1] Richard A. Krueger, Mary Anne Casey, *Focus groups: a practical guide for applied research*, Newbury Park, Sage Publications, August 2014.
- [A.2] Virginia Braun, Victoria Clarke, "Using thematic analysis in psychology", *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77-101, January 2006.

В. Допълнителен материал

Методика

Събрахме информация от два различни източника, за да изградим нашите обосновки: 1) Проучване на литературата и 2) Фокус групи.

Нашият първи източник на информация дойде от наличната литература за ИИ като цяло и по-специално за четирите области, описани по-горе. Целта беше да се придобие задълбочено разбиране на темата и да се разберат текущите знания. Това ни позволи да: 1) идентифицираме съответните теории, методи и мнения в съществуващите технологии и да ги докладваме в Книжката; и 2) организираме и насочваме участниците през фокус групите с тези вече научени основи, така че да получим необходимата допълнителна информация.

Затова потърсихме, прочетохме и оценихме повече от 250 документа, в това число научни статии, статии в списания, експертни блогове, доклади на компании, насоки на агенции и т.н., за да получим правилните познания за предизвикателствата и възможностите на ИИ и текущите приложения и очаквани въздействия както в краткосрочен, така и в дългосрочен план. Тези документи са внимателно подбрани като най-добро представяне на състоянието на техниката, според тяхното въздействие (напр. брой цитирания, разпространение в новини и социални мрежи и т.н.), автора/ите, издателя и годината на публикация.

Вторият източник се състоеше от събиране на информация чрез интервюиране на хора чрез фокус групи с експерти и неексперти по ИИ в 5 различни европейски държави. Фокус групата е вид качествена техника за събиране на данни, при която група хора, ръководени от модератор, водят разговор и дискутират около обща тема. Обикновено се състои от група от 5-10 души, които не се познават. Тези лица са избрани, защото имат специфични черти, които са подходящи за темата на фокус групата. Като се грижи за толерантна и спокойна среда, която насърчава много перспективи и гледни точки, без да принуждава членовете да постигнат консенсус, модераторът използва групата и нейното взаимодействие, за да научи повече за определена тема [A.1].

През февруари 2022 г. подготвихме и проведохме две различни фокус групи с обща методология и организация във всяка от страните на партньорите по проекта: една фокус група само с експерти по ИИ и една фокус група само с неексперти по ИИ. Проектът дефинира понятието експерт като лице с висше образование, работещо поне 5 години в областта на ИИ, дигиталното общество, взаимодействието човек-робот или Индустрия 4.0 и поне 3 публикувани научни или професионални статии. Броят на участниците във всяка група беше определен на минимум 5. Във всички фокус групи бяхме подпомагани от един модератор и един асистент: първият насочваше и хармонизираше дискусиите, докато вторият водеше бележки. Всяка фокус група продължи приблизително един час; сесиите бяха записани и по-късно транскрибирани. В един случай фокус групата не беше възможно да бъде организирана поради противоречиви графици/Ковид ограничение; ние заменихме фокус групата с въпросник. Въпросникът е създаден с подобни въпроси, използвани във фокус групата. В обобщение бяха организирани девет фокус групи и един въпросник и резултатите бяха анализирани съгласно указанията, налични в [A.2].

Повече подробности за всички сесии са докладвани тук:

- 1) Две фокус групи в Будапеща (Унгария) относно предизвикателствата и възможностите на ИИ
 - 8 експерти (7 мъже, 1 жена) на възраст 28-61 години;
 - 5 неспециалисти (2 мъже, 3 жени), на възраст 20-28 години, студенти, записани в магистърска степен;

- 2) Две фокус групи в Мюнстер (Германия) за ИИ в бизнеса
 - 7 експерти (7 мъже) на възраст 26-50 години;
 - 7 неспециалисти (3 мъже, 4 жени), на възраст 22-30 години, студенти, записани в магистърска степен.
- 3) Две фокус групи в Барселона (Испания) относно ИИ в управлението
 - 9 експерти (7 мъже, 2 жени) на възраст 35-70 години;
 - 10 неспециалисти (7 мъже, 3 жени), на възраст 22-70 години, от гражданското общество без предишни познания по ИИ.
- 4) Смесен подход в Лисабон (Португалия) относно ИИ в умения и компетенции
 - 9 експерти (6 мъже, 3 жени) на възраст 26-67 години чрез онлайн въпросник;
 - 5 неспециалисти (3 мъже, 2 жени) на възраст 23-55 магистри или наскоро завършили чрез фокус група.
- 5) Две фокус групи във Варна (България) за ИИ относно хората и начина на живот
 - 6 експерти (4 мъже, 2 жени) на възраст 29-59 години;
 - 15 неспециалисти (4 мъже, 11 жени), на възраст 20-24 години, студенти с висше образование.

Ограничения

Ясно е, че тази методология има някои ограничения. Първо, това е анализ, чиито заключения позволяват да се идентифицират различните тълкувания и аргументи, обществено достъпни върху даден проблем, но за разлика от количествения анализ, заключенията му не са представителни, а значими. Освен това съществува ограничение на хетерогенността на фокус групите, тъй като повечето от експертите са академици, а неекспертите са с висше образование; следователно резултатите може да не представляват възгледите на общото население по темата. Струва си обаче да се спомене, че сравнихме мненията на хората с наличната литература и обратно, така че нашите открития са ценни и други подобни трудове вероятно ще стигнат до същите заключения.

Достъп до чернова, по-дълга версия

Тази книжка е крайният резултат от почти едногодишна работа по изследване на въздействието на ИИ в нашето общество. За да постигнем тази цел, през тази година създадохме много други междинни разработки, които бяха окончателно хармонизирани и обобщени в настоящия документ. Заинтересованите читатели може да поискат да получат достъп до първата, по-дълга версия на документа, която е достъпна в хранилището на Zenodo (<https://www.zenodo.org>) с DOI: 10.5281/zenodo.7243312.

Благодарности

Бихме искали да благодарим на всички хора, които ни помогнаха в подготовката на тази книжка и щедро и открито участваха в нашите дейности.

Икони в глава 1, направени от Chanut-is-Industries от www.flaticon.com.

Финансиране от програма Еразъм+. Изразените възгледи и мнения обаче са само на авторите и не отразяват непременно тези на Европейския съюз или Европейската изпълнителна агенция за образование и култура (EACEA). Нито Европейският съюз, нито EACEA носят отговорност за тях.

Автори



Давиде Карелио получава двойна магистърска диплома по телекомуникационно инженерство и електроинженерство от Технически университет на Каталуня (UPC) (2000) и Технически университет в Торино (2001),

съответно, и докторска степен от ТУ на Каталуня през 2005 г. В момента той е член на Изследователския Център за Интелигентна Наука за Данни и Изкуствен Интелект (IDEAI), разположен в ТУ на Каталуня. Изследователските му интереси са фокусирани върху алгоритми и протоколи за компютърни мрежи със специални интереси в оперативната съвместимост, контрол и управление, планиране и маршрутизиране. Участвал е в няколко европейски и индустриални изследователски проекти.



Ана И. Алвеш Морейра има бакалавърска степен по психология и магистърска степен по психология на межкултурните отношения от ISCTE- Университетския институт в Лисабон, за което получава награда за академични постижения.

Тя е лицензиран психолог от Португалския орден на психолозите и е международен асоцииран член на Американската психологическа асоциация. Работила е като младши ръководител на проекти в европейски проекти в AidLearn и като научен асистент на свободна практика, свързана предимно с анализ/интерпретация на данни. Нейните научни интереси са предимно в областта на социалното познание, културата и многообразието и социалното включване.



Сесилио Ангуло Бахон получава своята докторска степен по приложна математика от ТУ в Каталуня през 2001 г., където в момента е редовен професор и основател на изследователския център IDEAI. Той е и член на комисията по етика в този

университет. Неговите научни интереси включват теоретични аспекти на машинното обучение, социалната и когнитивната роботика, обучението с подсилване и взаимодействието човек-робот. Той е автор на книги за машинно обучение и роботи и е публикувал близо 300 статии в списания и конференции. Той е ръководил и участвал в 47 конкурентни научноизследователски проекта, 21 от които финансирани от Европейската комисия.



Федерика Казачо притежава бакалавърска степен по Политически Науки и Международни Отношения (Италия и Хърватия) и магистърска степен по Изследвания на Мира, Конфликтите и Развитието

(Испания), специализирана в анализа на образованието за мир като трансформиращ инструмент на обществото. Тя има богат опит в изследванията, застъпничеството и управлението на проекти относно валоризацията на образованието. В момента работи като служител по стратегическо развитие в ACEEU, където участва в развитието на бизнеса и в управлението на няколко проекта, финансирани от ЕС в рамките на междусекторните теми за дигитализация, предприемачество и иновации във висшето образование.



Розалина Димова получава диплома по електроинженерство от Технически университет Варна, докторска степен от Технически университет София и има опит като преподавател повече от 20

години по Комуникационна Техника и Технологии. Тя е ръководител на Център за Приложни Здравни Технологии и бивш ректор на Технически университет – Варна. Има настоящи изследователски интереси в ИКТ, ИИ и мрежи от следващо поколение. Съавтор е на над 120 научни публикации и е участвала като ръководител или член на научни екипи в 7 европейски и повече от 50 национални проекта.



Тихомир Доврамаджиев има професионална дейност, свързана с управлението, академичната, научната, образователна и др. сфери. Той получава докторска степен по Ергономия и промишлен дизайн

(Технически Университет – Варна, 2012). Той е доцент д-р инженер в Технически Университет Варна, катедра Индустриален дизайн повече от 10 години. Автор е на книгата „Авангардни технологии в дизайна“, ТУВ, стр. 228, ISBN: 978-954-20-0771-5, 2017 г. Участвал е в над 60 научни публикации (вкл. Спрингър и Елзевир). Въз основа на данните на ResearchGate (2022 г.), той има около 300 000 четения на изследвания.



Антония Якоби завършва бакалавърската си степен по Психология в Германия, след което продължава своята магистърска степен по Психология на Межкултурните

отношения в Португалия. Работи като младши мениджър проекти в европейски проекти в AidLearn. Нейната работа се фокусира върху финансирани от Еразъм+ европейски проекти, които се занимават с обществени предизвикателства и проблеми от образователно естество. Поради това тя е участвала в различни проекти по теми като равенство между половете, глобално гражданство и социално включване. Нейните интереси са психо-социални въпроси, по-специално социалното включване, многообразието и пола.



Чаба Колар е старши научен сътрудник в Университета Обуда, Будапеща, ръководител на ИИ Работилница, лектор и ръководител на докторантура по науки и безопасност и сигурност, и докторантура по

военно инженерство на Националния университет за Обществена Служба, Унгария. Той е комуникационен инженер, дипломиран специалист по комуникации, ръководител на електронна информационна система по сигурността, консултант, коуч и медиатор. Неговите изследователски интереси включват социалните аспекти и икономическите въздействия на дигиталната ера, по-специално човешкото измерение на информационната сигурност и осведомеността, взаимодействието човек-робот, интелигентен град, ИИ, социална кредитна система и домотика.



Йевгения Суковий е филолог и международен бизнес инженер. Тя получава бакалавърска степен по английската филология от университета Намик Кемал в Турция и следдипломна степен в

областта на международното бизнес инженерство от Университета в Монпелие във Франция. Завършила е две обучения за обмен по програма Еразъм+ и два стажа в Португалия, Полша и Германия. Работила е като езиков учител, бизнес и маркетинг разработчик и асистент по европейски проекти. Нейните научни интереси включват (но не се ограничават до) предприемачество, образование, бизнес и маркетинг.



Гюла Сабо е регистриран европейски ергономист и доцент в Университета Обуда, Будапеща. Получава степен по електроинженерство и педагогика, докторска степен в Будапещенския Технологичен Университет, доктор

По Ергономия в Националния Университет за Обществена Служба, Унгария. Основният му изследователски интерес е оценката на човешките фактори на работното място и ръководи проекти за тяхното развитие. Участва в научния комитет на конференции и коректори на списания. Освен това е специалист и инженер по безопасност на труда, специалист по разследване на трудови злоупотреки и специалист по ергономия и човешки фактори.

HEDY

Life in the AI Era



BAEHF