

Оля Харизанова

ИНТЕГРИРАН

Оля Харизанова • ИНТЕГРИРАН ИНТЕЛЕКТ

ISBN 978-954-07-6433-7



9 789540 764337

unipress.bg



Университетско издателство
„Св. Климент Охридски“

ИНТЕГРИРАН ИНТЕЛЕКТ



ОЛЯ ХАРИЗАНОВА

ИНТЕГРИРАН ИНТЕЛЕКТ

София • 2026

Университетско издателство „Св. Климент Охридски“



Авторът изразява благодарност за технологичната и финансовата подкрепа от проект BG16RFPR002-1.014-0011 „Устойчиво развитие на Център за върхови постижения „Наследство БГ“, финансиран по процедура за предоставяне на безвъзмездна финансова помощ BG16RFPR002-1.014 „Устойчиво развитие на Центрове за върхови постижения и Центрове за компетентност, в т.ч. и на конкретни инфраструктури или техни обединения от Националната пътна карта за научна инфраструктура.

© 2026 Оля Харизанова

© 2026 Валентин Марков, дизайн на корицата

© 2026 Университетско издателство „Св. Климент Охридски“

ISBN 978-954-07-6433-7

СЪДЪРЖАНИЕ

Въведение	9
Методология за идентифициране на интегриран интелект	13
Архитектура и аналитични равнища на методологията.....	13
Предмет, единица на анализ и граници на случая	16
Онтологични и логически условия.....	19
Операционализиране на резонансите.....	24
Критерии за асиметрично ко-креаторство	28
Процедура за идентифициране и кодиране.....	34
Процесни и резултатни критерии	42
Класификационно заключение и степен на проявление	48
Надеждност и валидност	49
Бъдещо количествено и сравнително валидиране	56
Методологичен извод.....	65
Литературен обзор за интегриран интелект	67
Симбиоза човек–компютър и увеличаване на интелекта	68
Разпределено, разширено и съвместно познание.....	69
Колективна интелигентност и организационно знание.....	70
Увеличена и хибридна интелигентност.....	71
Ориентиран към човека изкуствен интелект и асиметрия на отговорността	72
Симбиотично взаимодействие и функционална взаимност.....	73
Човешка креативност и изчислителна генеративност	73
Креативно стимулиране и творчески резонанс	74
Когнитивен резонанс и преобразуване на разбирането.....	75
Емоционален резонанс и комуникативна адаптация	76
Ко-креаторство „човек–изкуствен интелект“	77

Емерджентност и съвместно сътворен продукт.....	78
Критичен синтез и теоретична празнина.....	79
Позициониране на интегрирания интелект	80
Обобщение.....	81
Основни понятия и определения, свързани с интегрирания интелект.....	84
Архитектура на понятийния апарат.....	84
Човешки агент	85
Система, модел и агент с изкуствен интелект	86
Диалог и ко-креативно диалогово пространство	87
Диалогичен ход, рекурсивен цикъл и ко-креативен епизод	88
Функционални роли на изкуствения интелект	89
Креативност и креативно стимулиране.....	90
Когнитивен резонанс	91
Творчески резонанс.....	91
Емоционален резонанс	92
Функционална взаимност.....	93
Асиметрично ко-креаторство	94
Съвместно сътворен продукт	94
Възникващо знание.....	95
Интегриран интелект	96
Разграничаване на интегрирания интелект от сродни конструкции.....	97
Научен статут и теоретични основания	99
Конфигурационен модел	102
Обобщение.....	103
Философски основи на интегрирания интелект.....	104
От индивидуалистично към релационно и системно разбиране	105
Онтологични основания на интегрирания интелект.....	106
Епистемологични основания на интегрирания интелект	109
Аксиологични основания на интегрирания интелект.....	111
Праксеологични основания на интегрирания интелект	113
Конститутивни аксиоми.....	114
Методологични принципи	116
Логическа архитектура и нелинейни отношения.....	117
Обобщение.....	121

Илюстрация на взаимодействието „човек–изкуствен интелект“	123
Тематично свързаните цикли от диалози	
Ролята и действията на автора	125
Изводи на автора от взаимодействието	129
Граници на авторефлексивната реконструкция	130
Значение на авторския опит за концептуалната рамка	131
Заклучение	132
Основни теоретични резултати	132
Концептуална и философска рамка	133
Методологичен принос	134
Литературен и научен принос	135
Емпиричен статут и ограничения	136
Насоки за бъдещи изследвания	136
Извод	138
Речник на основните понятия	140
Концептуални различия	145
Библиография	149
За автора	168

ВЪВЕДЕНИЕ

Книгата е резултат от дългогодишен професионален интерес към отношенията между хората, знанието и интелигентните технологии, както и от изследователска работа, проведена в период на ускорено развитие на изкуствения интелект (ИИ). Взаимодействието с тези системи променя не само начина, по който създаваме текстове, анализираме информация и вземаме решения, но и начина, по който мислим, задаваме въпроси, търсим смисъл и изграждаме ново знание. Книгата е посветена именно на тези промени и на условията, при които те възникват.

Предмет на настоящото изследване е взаимодействието между човека и системите с ИИ. То се разглежда като ново пространство за разгръщане на качествено различни когнитивни, емоционални и творчески процеси. Във фокуса са не технологиите и алгоритмите сами по себе си, а динамиката и резултатите от това взаимодействие.

Основната цел на книгата е да предложи концептуална рамка за обяснение на измененията в интелектуалната дейност. Тези изменения настъпват, когато човешките способности и функционалните възможности на системите с ИИ се съчетават в продължително, целенасочено и рефлексивно взаимодействие.

През последните години множество изследвания разглеждат възможностите, ограниченията и приложенията на ИИ. Значително по-малко внимание се отделя на взаимодействието между човека и ИИ. Недостатъчно проучени остават и новите явления, които възникват в това взаимодействие. Този научен дефицит постави началото на настоящото изследване. Анализът показва необходимостта от по-развита понятийна рамка за осмисляне на наблюдаваните явления. Така се оформиха представените в книгата авторски понятия: пространство за когнитивен диалог, когнитивен резонатор, когнитивен резонанс, емоционален резонанс, творчески резонанс, възникващо знание и интегриран интелект. Тези понятия не са въведени като самоцелна терминология. Те служат за осмисляне на явления, които присъстват все по-осезаемо в съвременната интелектуална практика. Разработването им премина през продължителен процес на анализ, съпоставяне, критична рефлексия и многократно прецизиране.

Постепенно стана ясно, че дефинирането на отделни понятия не е достатъчно. В хода на работата се изгради последователна терминологична система, в която всяко основно понятие заема определено място в единна понятийна архитектура. Веднъж въведени, понятията запазват устойчив смисъл в цялото изложение. Връзките между тях се основават на надграждане, а не на частично припокриване или синонимна употреба. Така книгата предлага не само модел на взаимодействието между човека и ИИ, но и научен език за неговото изследване и анализ.

Макар терминът „интегриран интелект“ да се среща в научната и професионалната литература, той все още няма общоприета дефиниция и единна теоретична основа. Тук интегрираният интелект е концептуализиран като възникваща форма на интелектуална организация, проявяваща се в система на взаимодействие между човек, ИИ и релевантна социотехническа среда, когато различните им функции се координират и преобразуват в общ познавателен процес. Това разбиране не предполага сливане на човека и ИИ или възникване на автономен колективен субект.

Като помощен изследователски инструмент при прецизирането на концептуалната рамка беше избрана генеративна система с ИИ, а именно ChatGPT¹,². Системата беше използвана за генериране на алтернативни формулировки, структуриране на отделни аргументи и езиково усъвършенстване на изложението. Всички генерирани предложения бяха подложени на критична авторска оценка, проверка и редакция. Научните въпроси, теоретичните решения, интерпретациите, понятийната архитектура и окон-

¹ ChatGPT е ИИ асистент, създаден от OpenAI. Той може да: отговаря на въпроси; обяснява сложни теми; пише, редактира и превежда текстове; помага с учене, идеи, планиране и програмиране; анализира документи, изображения и данни. ChatGPT може да използва различни големи езикови модели (LLM) на OpenAI според избрания режим и задачата. За настоящото изследване обаче е особено важно това, че към момента LLM не мисли и не разбира точно като човек. Освен това понякога дава убедително звучащи, но грешни отговори. Ето защо генерираните от ChatGPT резултати могат да бъдат неточни, непълни или пристрастни.

² Защо именно ChatGPT? Този ИИ асистент е най-популярен сред потребителите. Към средата на 2026 г. броят на месечно активните потребители според съпоставими оценки на *Sensor Tower* (<https://elpais.com/tecnologia/2026-06-04/chatgpt-se-convierte-en-la-aplicacion-mas-rapida-en-conseguir-1000-millones-de-usuarios.html>) е както следва: ChatGPT – около 1 млрд.; Google Gemini – около 472 млн.; Doubao – около 106 млн.; Dola – около 78 млн.; DeepSeek – около 68 млн. и др. Видно от посочените данни, ChatGPT е доминиращ. Важно е да се уточни още, че данните не са за „регистрирани акаунти“, защото компаниите обикновено не публикуват такава информация. Това са месечно активни потребители на приложенията, измерени по еднаква методология. Други източници дават различни стойности. Например QuestMobile отчита 127 млн. потребители на DeepSeek само в Китай, но използва различен обхват и методика.

Освен това, в предварителния етап на разработката авторът използва различни системи с ИИ (DeepSeek, Kimi, Gemini, Use). Избран е ChatGPT, тъй като системата се оказва най-пригодна за изпълнение на конкретните задачи, начина на работа и критериите на автора.

чателният текст са резултат от авторската изследователска работа. Използваната ИИ система не е автор или съавтор и не ѝ се приписва независим интелектуален принос. Тя не може да поема академична, етична или правна отговорност. Пълната отговорност за точността, оригиналността, добросъвестността и съдържанието на книгата принадлежи на автора.

Работата със системата има двойна роля. От една страна, тя представлява практическо използване на ИИ. От друга, служи като източник на данни за анализ. Въз основа на тях е проследено взаимното влияние между човека и генеративния ИИ в процеса на диалог.

В предложената концептуална рамка когнитивният резонанс обозначава промени в човешкото разбиране и в организацията на понятията. Творческият резонанс разширява пространството на възможните решения. Емоционалният резонанс поддържа човешката ангажираност и настройката в общуването. При човека тази динамика може да се прояви като разбиране, доверие, мотивация или вдъхновение. При ИИ тя има само функционален характер. Системата обработва подадените сигнали и адаптира последващите си отговори. Това не предполага наличие на съзнание, субективно преживяване или собствено намерение.

Наблюденията върху този процес подкрепят теоретичното предположение на автора, че при определени условия взаимодействието между човек и генеративен ИИ може да се анализира като асиметрично ко-креативно взаимодействие. Това понятие описва функционалното участие на системата в генерирането на варианти и в итеративното усъвършенстване на продукта, но не означава авторство, съавторство, равнопоставено интелектуално участие или автономна творческа инициатива. Използването на ИИ може да разшири пространството на възможните решения и да подпомогне отделни етапи на работата, но целите, критериите за оценка, подборът, проверката, интерпретацията и окончателните решения принадлежат на автора. Човекът и ИИ не са равнопоставени по отношение на съзнание, намерение, авторство и отговорност. Дали при конкретно взаимодействие се проявяват характеристики на интегриран интелект трябва да се установява чрез ясни критерии, операционализация и емпирична проверка, без на системата да се приписва статут на научен субект.

Книгата не претендира да предлага окончателна или завършена теория. Тя формулира теоретично предложение, чиито понятия, отношения и обяснителни възможности подлежат на критика, проверка и доразвиване. Представените дефиниции, аксиоми, принципи и модели са предназначени да положат основа за научен диалог и за последващи интердисциплинарни изследвания на взаимодействието между човека и ИИ. Според автора най-ценните идеи рядко се появяват като завършени решения. Те възникват

постепенно в продължителен диалог, в който въпросите, аргументите, съмненията и различните гледни точки взаимно се допълват и преобразуват. Позицията на автора е, че развитието на ИИ изисква не само технологичен анализ, но и философско осмисляне, методологично структуриране и етично управление.

В книгата е предложен резултат от изследване на условията, при които функционалните възможности на ИИ могат да бъдат включени в човешки направлявана, ценностно ориентирана и творчески продуктивна интелектуална дейност. Подобни изследвания са особено важни днес. Навлизаме в период, който изисква не само по-съвършени технологии, но и нов научен език. Чрез него трябва да разпознаваме и обясняваме промените в човешкото познание, творчество и съвместно мислене, породени от тези технологии. Ако предложената концептуална рамка стимулира нови въпроси, изследвания и интерпретации, книгата ще е изпълнила своето предназначение.

МЕТОДОЛОГИЯ ЗА ИДЕНТИФИЦИРАНЕ НА ИНТЕГРИРАН ИНТЕЛЕКТ

Съществуващите изследвания на взаимодействието между човека и ИИ описват широк спектър от форми на подпомагане, сътрудничество и съвместно създаване, но често не предоставят достатъчно ясни основания за разграничаване между инструментална употреба, резонансно взаимодействие, ко-креаторство и проявление на интегриран интелект. Настоящата методология е насочена към преодоляване на тази празнина чрез последователен преход от теоретичните определения към наблюдаеми процеси, проследими приноси и проверими заключения (Dellermann et al., 2019; Seeber et al., 2020).

Водещият методологичен въпрос е: по какви концептуални, логически, операционни и емпирични критерии може да се установи, че конкретно взаимодействие между човек и ИИ (1) включва когнитивен, творчески и емоционален резонанс, (2) придобива формата на асиметрично ко-креаторство и (3) достига равнище, което може да се тълкува като проявление на интегриран интелект.

Методологията приема, че интегрираният интелект не възниква автоматично при продължителен диалог, при използване на генеративен модел или при получаване на качествен резултат. Тя изисква реконструкция на процеса, разграничаване на ролите, проследяване на преобразуването на приносите и оценка на алтернативни обяснения. Човекът и ИИ участват асиметрично: (1) човекът притежава намерение, разбиране, ценностна преценка и отговорност; (2) ИИ участва чрез аналитични, генеративни и контекстуално преобразуващи функции (Amershi et al., 2019).

1. Архитектура и аналитични равнища на методологията

Методологията за изследване на интегрирания интелект е организирана като аналитична архитектура, която свързва теоретичното определяне на явлението с неговото наблюдение, измерване и емпирична проверка. Тя запазва пет взаимно зависими равнища: онтологично, логическо, операционно, измервателно и емпирично. Тези равнища не представляват отделни методологии и не бива да се прилагат изолирано. Всяко от тях изпълнява

специфична функция в общия изследователски процес, а надеждността на заключението зависи от съгласуваността между тях. По-долу са описание изброените равнища.

Онтологичното равнище определя предмета на изследване и неговите концептуални граници. На това равнище се уточнява какво се разбира под интегриран интелект, кои негови характеристики са съществени и по какво предполагаемото му проявление се различава от инструменталното използване на ИИ, автоматизацията, еднократното генериране на съдържание или обикновеното функционално подпомагане. Основна задача е да се предотврати прекомерното разширяване на конструкта, при което всяко продуктивно взаимодействие между човек и ИИ би могло неправомерно да бъде определено като интегриран интелект.

Логическото равнище превежда онтологичното определение в система от необходими и подкрепящи условия за разпознаване на явлението. То уточнява как трябва да бъдат свързани целенасоченият диалог, рекурсивното преобразуване на приносите, когнитивният, творческият и емоционалният резонанс, асиметричното ко-креаторство и запазената човешка отговорност. На това равнище се изгражда аргументационната структура, чрез която се преценява дали наблюдаваната конфигурация е съвместима с теоретичния конструкт. Логическата съгласуваност е необходима, за да не се извежда заключение само въз основа на отделен признак, например продължителен диалог, субективна полезност или високо качество на крайния продукт.

Операционното равнище свързва теоретичните условия с конкретни наблюдаеми прояви. То определя какви са следите в диалога, какви са последователните версии на продукта, какви решения е взел човека-участник в диалога. Такива следи могат да бъдат промяна в първоначалната концептуална рамка, поява на нова посока на търсене, съществено преработване на машинно предложение, аргументиран отказ от генериран вариант или проследимо включване на различим принос в крайния резултат. Обратната връзка може да служат като доказателства за резонанс, преобразуване и ко-креаторство. Операционализацията трябва да позволява разграничаване между същинско преобразуване и повърхностно редактиране, механично приемане или случайно съвпадение.

Измервателното равнище оценява степента, интензивността и профила на установените процеси. Неговата задача не е единствено да даде обща стойност, а да покаже как са разпределени отделните измерения на предполагаемото проявление. Възможно е например да бъде отчетен силен когнитивен резонанс, умерено ко-креаторство и слабо емоционално-комуникативно взаимодействие. Профилното измерване запазва сложността на явлението и ограничава риска различни конфигурации да бъдат сведени до един недиференциран показател.

Емпиричното равнище проверява доколко използваните понятия, индикатори, процедури и заключения са надеждни, валидни и приложими извън конкретния анализиран случай. То включва използване на независими оценители, проверка на надеждността между тях, анализ на положителни, отрицателни и гранични случаи, както и оценяване на съдържателната, конструктивната и дискриминантната валидност. Необходимо е рамката да бъде прилагана при различни участници, задачи, модели, езици и организационни контексти. Само чрез подобни проверки може да се установи дали индикаторите действително разпознават интегриран интелект, а не обща продуктивност, удовлетвореност от взаимодействието или високо качество на продукта.

Петте равнища образуват последователна, но не линейна система. Резултатите от операционното наблюдение и емпиричното валидиране могат да наложат прецизиране на логическите условия или дори преосмисляне на онтологичното определение. Поради това методологията включва рекурсивна обратна връзка между теорията, наблюдението и проверката. Това я превръща в процедура за прилагане на предварително зададен модел, а така също в средство за постепенно усъвършенстване на самата теория за интегрирания интелект.

В рамките на тази архитектура е необходимо ясно да се разграничат четири аналитични дейности: идентифициране, измерване, оценяване и валидиране (табл. 1). Те отговарят на различни въпроси и водят до различни видове заключения. Идентифицирането определя дали наблюдаваният процес е съвместим с проявление на интегриран интелект. Измерването установява степента и профила на отделните процеси. Оценяването разглежда качествата на крайния продукт; валидирането проверява надеждността и обосноваността на използваните критерии, данни и изводи. Смесването на тези дейности може да доведе до методологична грешка (например високото качество на продукта да бъде прието като достатъчно доказателство за интегриран интелект, без да е установено как е възникнал този продукт).

Таблица 1. *Разграничаване на аналитичните дейности*

<i>Дейност</i>	<i>Основен въпрос</i>	<i>Резултат</i>
Идентифициране	Налице ли е конфигурация, съвместима с проявление на интегриран интелект?	Аргументирано класификационно заключение.
Измерване	В каква степен са проявени отделните процеси и измерения?	Профил или стойности по индикатори.

<i>Дейност</i>	<i>Основен въпрос</i>	<i>Резултат</i>
Оценяване	Какво е качеството, оригиналността или приложимостта на продукта?	Оценка спрямо външни критерии.
Валидиране	Надеждни и обосновани ли са критериите, данните и изводът?	Доказателства за надеждност и валидност.

2. Предмет, единица на анализ и граници на случая

Предмет на идентифициране е организацията на конкретното взаимодействие „човек–ИИ“ и възникналият чрез него продукт. Анализът изследва релационната конфигурация, в която човешките и машинно генерираните приноси се свързват, оценяват, преобразуват и включват в последователен процес на съвместна работа. Следва да се подчертае, че анализът не е насочен към интелигентността на човека или към техническите възможности на модела, разглеждани изолирано. Обект на внимание е от една страна какво могат участниците поотделно, а от друга – как организацията на взаимодействието променя условията за възникване на конкретния резултат.

Това разграничение е методологически съществено. Човешката експертност, дори и висока, не е достатъчно доказателство за интегриран интелект. Използването на усъвършенстван генеративен ИИ модел не предполага автоматично ко-креативно взаимодействие. Възможно е висококачествен резултат да бъде постигнат предимно чрез човешкия опит, чрез еднократно машинно генериране или чрез инструментално подпомагане. Поради това анализът трябва да установи дали крайният продукт носи проследими следи от рекурсивно преобразуване на различни приноси и дали тези преобразувания са повлияли съществено върху развитието на взаимодействието.

Основна единица на анализ е ко-креативният епизод. Той представлява ограничена и смислово свързана последователност от един или повече цикли: „човешка цел или принос → изход на ИИ → човешка интерпретация и преобразуване → ново задание или решение“. Тази последователност е добре да не се разбира като механична и неизменна схема. В реалното взаимодействие отделни нейни компоненти могат да се повтарят, да се разклоняват или да включват сравняване на няколко машинно генерирани варианта, връщане към по-ранна версия и преформулиране на първоначално поставената цел.

Ко-креативният епизод започва, когато може да се установи конкретна цел, проблем, въпрос или намерение, които организират последващия обмен. Началото може да бъде поставено чрез конкретно задание, представяне на първоначална идея, предоставяне на съществуващ материал или формулиране на затруднение, което изисква решение. Същественото е да бъде

възможно проследяването на връзката между началното състояние на задачата, последващите човешки и машинни приноси и достигнатия резултат.

Епизодът завършва при стабилизиране на продукт или решение, съществена промяна на целта, преминаване към друг относително самостоятелен проблем или прекъсване на взаимодействието. Стабилизирането не предполага непременно окончателно завършване на продукта. То обозначава момент, в който е достигната относително устойчива версия, приета е конкретна посока или е взето решение, което позволява епизодът да бъде анализиран като смислово цяло. Когато целта се променя значително, следващите стъпки се разглежда като нов епизод, дори да е част от същия проект.

Определянето на границите на епизода трябва да се основава на предварително формулирани критерии. Подходящи критерии могат да бъдат: промяна на изследователския въпрос, преминаване към нов вид продукт, приемане или отхвърляне на решение, прекъсване на работата или започване на нова тематична линия. Ясното обозначаване на началото и края е необходимо, за да се избегне избирателното включване само на онези части от диалога, които подкрепят предварително очакваното заключение (Yin, 2018).

При продължителен проект се разграничават четири вложени аналитични равнища: диалогичен ход, рекурсивен цикъл, ко-креативен епизод и проект. Тези равнища осигуряват различна степен на аналитична детайлност. Диалогичният ход позволява да се разгледа функцията на отделния принос. Рекурсивният цикъл показва непосредственото взаимодействие между генериране, оценка и преобразуване. Ко-креативният епизод обхваща относително завършено развитие. Проектът позволява да се проследят устойчивите модели и измененията във времето (Yin, 2018).

Диалогичният ход е най-малката наблюдаема единица. Той може да бъде (1) човешки промпт³, уточнение, критична бележка, избор, редакция или (2) машинно генериран отговор, предложение, анализ или вариант. Сам по себе си отделният ход рядко е достатъчен за формулиране на заключение за наличие на резонанс или ко-креаторство. Неговото значение се определя от функцията му в по-широката последователност, като: дали въвежда ново разграничение, променя посоката на работа, предизвиква критична реакция или участва в преобразуването на продукта.

Рекурсивният цикъл включва функционално свързана последователност от задание, машинен изход, човешка оценка и последващо преобра-

³ Тук под „човешки промпт“ се има предвид инструкция, въпрос или заявка, въведена от човек към система с ИИ. Например: „Провери този текст за правописни грешки.“

зуване. Цикълът е основното равнище, на което може да се наблюдава взаимното изменение на приносите. За неговото установяване не е достатъчно човекът да приеме или копира получения отговор. Нужно е да има проследима интерпретация, селекция, корекция, комбиниране, преформулиране или използване на изхода като основание за ново действие. Последващият машинен принос също трябва да бъде свързан с промененото състояние на задачата.

Ко-креативният епизод обединява един или повече рекурсивни цикли, довели до относително стабилен продукт, решение или концептуална промяна. На това равнище се оценява: дали отделните цикли образуват последователен процес, дали се наблюдават когнитивен, творчески или емоционален резонанс и дали машинният принос има различима функционална роля. Епизодът е основната единица за идентифициране на предполагаемо проявление на интегриран интелект, защото съчетава процесните доказателства с конкретен резултат.

Проектът представлява система от епизоди, обединени от по-обща цел. Един проект може да съдържа епизоди с различна степен и различен профил на резонанс и ко-креаторство. Някои части могат да демонстрират интензивно рекурсивно преобразуване, а други да представляват рутинно редактиране, търсене на информация или техническо подпомагане. Поради това наличието на един силно изразен ко-креативен епизод не е достатъчно целият проект да бъде определен като проявление на интегриран интелект.

Разграничаването на равнищата предотвратява както прекомерното обобщаване, така и аналитичното раздробяване. Прекомерното обобщаване възниква, когато цялата продължителна работа се характеризира като „интегрирана“ въз основа на ограничен брой взаимодействия. Аналитичното раздробяване възниква, когато отделните ходове се разглеждат извън тяхната функционална връзка и се прекъсва процесът на постепенно преобразуване. Анализът на много равнища позволява локалните проявления да бъдат интерпретирани в контекста на цялостната динамика на проекта.

Границите на случая трябва да включват времеви, технологични и контекстуални параметри. Трябва да се документират периодът на взаимодействие, използваният модел и неговата версия, основните настройки и инструкции, достъпът до външни източници, характеристиките на задачата и релевантният опит на човешкия участник. Тези фактори могат да повлияят върху процеса и са необходими за правилното тълкуване, сравняването и евентуалното възпроизвеждане на резултатите.

За всеки анализиран случай трябва да се съхранява достатъчно пълна доказателствена следа. Тя може да включва: хронологията на диалога, последователните версии на продукта, бележки за взетите решения, аргумен-

ти за приемане или отхвърляне на предложения, карта на приносите и окончателния резултат. Когато част от тази информация липсва, трябва ясно да се посочи как ограничението влияе върху класификационното заключение.

Така определените предмет, единици на анализ и граници на случая позволяват интегрираният интелект да бъде изследван като локализирано и доказуемо процесно проявление, а не като общо качество, приписвано на човека, системата или проекта. Следващата табл. 2 систематизира отделните аналитични равнища и основните източници на данни, чрез които те могат да бъдат реконструирани и проверени.

Таблица 2. Единици на анализ и източници на данни

<i>Равнище</i>	<i>Определение</i>	<i>Основни доказателства</i>
Диалогичен ход	Единичен човешки или машинно генериран принос.	Промпт, отговор, бележка, избор или редакция.
Рекурсивен цикъл	Свързана последователност от задание, изход, оценка и преобразуване.	Хронология на диалога и последователни версии.
Ко-креативен епизод	Един или повече цикли, довели до относително стабилен продукт или решение.	Диалог, версии, карта на приносите и краен продукт.
Проект	Система от епизоди, обединени от обща цел.	Корпус, дневник на решенията, междинни и окончателни резултати.

3. Онтологични и логически условия

Интегрираният интелект се разглежда като възможно системно проявление на организирано взаимодействие между човек и ИИ. Понятието обозначава специфична организация на отношенията, при която различни човешки и машинно генерирани приноси участват в рекурсивен процес на взаимно преобразуване и водят до резултат, чието възникване не може да се обясни чрез разглеждане на участниците поотделно. Следва да се уточни, че интегрираният интелект не представлява самостоятелна субстанция, общо съзнание, сливане на човешкия и машинния разум или възникване на нов морален субект.

Тази постановка изисква ясно разграничаване между участниците, взаимодействието и възникващото системно проявление. Човекът остава носител на намерението, преживяването, ценностната преценка и моралната отговорност. Изкуственият интелект участва чрез аналитични, генеративни, преобразуващи и контекстуално адаптивни функции. Това функционално участие не предполага наличие на собствено съзнание, автономна воля или

морална субектност. Интегрираният интелект не принадлежи само на човека или само на техническата система. Той се проявява чрез съчетаването на техните различни възможности и роли в конкретна социална и технологична среда.

Релационният характер на конструкта означава, че обяснението се отнася преди всичко до структурата, качеството и динамиката на връзките. От значение е както какъв принос прави всеки участник, така и как този принос променя условията за следващите действия. Машинно генерираното предложение може да предизвика от страна на човека ново разграничение, критична корекция или промяна на целта. От своя страна човешката интерпретация и преформулиране могат да доведат до машинен изход, който не би възникнал при първоначалното задание. Така всеки последващ ход се основава върху вече преобразувано състояние на задачата.

Релационността не означава симетрия между човека и ИИ. Двете страни участват по различен начин и имат различен онтологичен и нормативен статут. Човекът задава или приема целта, интерпретира значението на предложенията, оценява тяхната допустимост и решава как да бъдат използвани. Изкуственият интелект може да разширява пространството на възможните решения, да предлага неочаквани комбинации, да открива зависимости и да преобразува предоставения контекст. Това функционално различие е основание взаимодействието да бъде определено като асиметрично ко-креаторство, а не като равнопоставено съавторство.

Емергентността означава, че конкретната организация на процеса или възникналият продукт не се обясняват достатъчно чрез механичното събиране на изолираните приноси. Не е достатъчно човекът и ИИ последователно да добавят отделни елементи към общ резултат. Нужно е приносите да се изменят взаимно, така че последващите решения да зависят от историята на взаимодействието. Емергентният характер се проявява, когато формата, структурата или смисловата организация на продукта възникват чрез рекурсивната динамика, а не са предварително завършени в човешкия замисъл или в първоначалния машинен изход.

Емергентността не трябва да се отъждествява автоматично с новост, сложност или високо качество. Един продукт може да бъде оригинален, без да е резултат от ко-креативен процес, както и да бъде сложен поради натрупването на множество генерирани елементи. За да се говори за емергентно системно проявление, трябва да бъде показано как взаимодействието е довело до качествена промяна в разбирането, организацията на задачата или пространството на възможните решения. Затова твърдението за емергентност изисква процесни доказателства, а не само оценка на крайния продукт.

Онтологичната рамка поставя граници пред интерпретацията на емоционалния резонанс. Адаптивният, подкрепящият или емпатично звучащи-

ят отговор на ИИ може да влияе върху човешката ангажираност и устойчивостта на взаимодействието. Това влияние е реално като част от преживяването и поведението на човека, но не доказва, че техническата система изпитва емоции. Емоционалният резонанс трябва да се разглежда като релационен и комуникативен ефект, а не като свидетелство за симетрично емоционално преживяване.

Логическата спецификация превежда тези онтологични положения в условия за идентифициране на предполагаемо проявление на интегриран интелект. Нейната функция е да определи кои признаци трябва задължително да бъдат установени и кои допълнителни характеристики увеличават обосноваността на заключението. Така се избягва класифициране въз основа на общо впечатление, субективна полезност, продължителност на диалога или качество на крайния продукт.

Условията се разделят на необходими и подкрепящи. Необходимите условия очертават минималната конфигурация, без която твърдението за интегриран интелект не може да бъде защитено. Подкрепящите условия увеличават увереността, показват интензивността или профила на проявлението, но отсъствието на отделно подкрепящо условие не води автоматично до отрицателно заключение. Разграничението предотвратява едновременно прекалено широкото и прекалено ограничителното използване на конструкта.

Първото необходимо условие е наличието на определима цел и контекст. Трябва да бъде възможно да се установи какъв проблем се решава, какъв продукт се създава или какво разбиране се търси. Целта може да бъде уточнявана и преобразувана в хода на взаимодействието, но трябва да съществува достатъчна смислова приемственост, която позволява отделните ходове да бъдат разглеждани като части от общ процес. Без такава приемственост продължителната размяна на съобщения не образува ко-креативен епизод.

Второто необходимо условие е наличието на съдържателен принос както от човека, така и от ИИ. Човешкият принос може да включва целеполагане, първоначална идея, експертно знание, критерии, критика, подбор, интерпретация или редакция. Машинният принос трябва да бъде функционално релевантен за развитието на задачата и да надхвърля технически операции без смислово значение. Не е необходимо приносите да бъдат равни по обем, но трябва да бъдат различни и да имат проследима роля в процеса.

Третото необходимо условие е наличието на поне един завършен рекурсивен цикъл. Еднократното задаване на въпрос и приемането на получения отговор не са достатъчни. Нужно е машинният изход да бъде интерпретиран, оценен или преобразуван от човека и тази промяна да участва в следващо задание, решение или версия на продукта. Рекурсивността показва, че

взаимодействието не е еднопосочно предаване, а процес, в който резултатът от предходния ход изменя следващото състояние на задачата.

Четвъртото необходимо условие е човешката интерпретация и критична преработка. Човекът не трябва да функционира единствено като оператор, който предава, приема или механично сглобява машинни изходи. Той трябва да оценява значението, точността, релевантността и допустимостта на предложенията, да ги свързва с целта и при необходимост да ги коригира или отхвърля. Критичната преработка може да бъде концептуална, структурна, езикова, методологична или ценностна, но трябва да оставя наблюдаема следа.

Петото необходимо условие е проследимата връзка между взаимодействието и възникналия продукт. Трябва да може да се покаже кои идеи, решения, структури или промени са възникнали в конкретни моменти от процеса и как са били включени в крайния резултат. Проследимостта не изисква всяка дума или елемент да бъде приписан изцяло на един участник. Тя изисква достатъчна генеалогия на продукта, чрез която да се реконструира ролята на отделните приноси и тяхното преобразуване.

Шестото необходимо условие е човешкото утвърждаване и запазената отговорност. Крайният продукт придобива практическа, научна или обществена валидност само след човешка преценка, проверка и решение за неговото използване. Изкуственият интелект не може да поеме моралната, професионалната или юридическата отговорност за резултата. Поради това интегрираният интелект предполага не отслабване, а засилване на изискванията към човешката критичност, прозрачност и отчетност.

Като първо подкрепящо условие се отнася появата на ново разбиране. Тя може да се прояви чрез преформулиране на проблема, откриване на неподозирана зависимост, промяна на първоначална позиция или изграждане на по-цялостна концептуална структура. Новото разбиране подкрепя твърдението за когнитивен резонанс, когато може да бъде свързано с конкретни преобразувания в диалога, а не единствено със субективното усещане за напредък.

Второ подкрепящо условие е разширяването или преобразуването на пространството на възможните решения. Изкуственият интелект може да предложи варианти, аналогии или комбинации, които човекът не е формулирал предварително, а човешката оценка може да ги пренареди в нова посока. Това подкрепя наличието на творчески резонанс, особено когато крайното решение не представлява просто избор от готов списък, а резултат от последователно комбиниране, отхвърляне и преработване.

Емоционалната ангажираност и устойчивостта на процеса също имат подкрепящ характер. Те могат да се изразят в повишена мотивация, увере-

ност за изследване на нови посоки, готовност за продължителна работа или по-добро справяне с неопределеността. Тези признаци не са достатъчни за установяване на интегриран интелект, защото могат да възникнат и при приятно, но повърхностно взаимодействие. Значението им се определя от връзката с когнитивните и творческите преобразувания.

Особено силно подкрепящо условие са доказателствата за преобразуване, което не е било предварително завършено в нито един от началните приноси. Това може да бъде нова концептуална схема, аргумент, модел, решение или композиционна структура, възникнали чрез последователната взаимна промяна на човешките и машинните предложения. Твърдението не се подкрепя единствено чрез вътрешните способности или намерения на участниците. Твърдението трябва да бъде подкрепено чрез сравнение между началното състояние, междинните версии и крайния продукт.

Необходимите условия имат конюнктивен характер⁴: отсъствието на което и да е от тях отслабва или изключва класифицирането на конкретния епизод като предполагаемо проявление на интегриран интелект. Подкрепящите условия имат профилен характер и могат да бъдат представени с различна степен на изразеност. Възможни са например епизоди със силен когнитивен, но слаб емоционален резонанс, или с ясно ко-креаторство, но ограничена оригиналност на продукта.

Заключението трябва да отчита алтернативните обяснения. Наблюдаваният резултат може да се дължи на предварителната експертност на човека, на качеството на първоначалното задание, на възпроизвеждане на вече известна информация, на случайно съвпадение или на селективно избиране на успешни машинни предложения. Поради това идентифицирането изисква не само потвърждаващи признаци, но и проверка дали по-просто обяснение описва достатъчно наблюдавания процес.

Оントологичните и логическите условия изпълняват едновременно ограничителна и конструктивна функция. Те ограничават необоснованото приписване на интелигентност, съзнание или авторство на техническата система, но същевременно позволяват да бъде призната реалната функционална роля на машинния принос. Така интегрираният интелект може да се изследва като специфично системно проявление, без да се заличават различията между участниците и без човешката отговорност да се прехвърля към технологиите.

⁴ Тук „конюнктивен характер“ означава, че условията трябва да са изпълнени едновременно и съвместно – свързани са с логическото „И“, а не с „ИЛИ“.

4. Операционализиране на резонансите

Резонансът се операционализира като взаимно обусловено изменение на последователните човешки и машинно генерирани приноси в рамките на конкретен диалогичен процес. Той не се определя само чрез тематично сходство, съгласие между участниците, продължителност на взаимодействието или субективна удовлетвореност от получения резултат. За да бъде установен резонанс, трябва да има наблюдаеми доказателства, че предходният принос е повлиял върху съдържанието, структурата, посоката или оценката на следващия принос.

Взаимната обусловеност не предполага психологическа, когнитивна или емоционална симетрия между човека и ИИ. При човека изменението може да се изрази в промяна на разбирането, намерението, оценката, критериите или творческото решение. То може да включва осъзнаване на нов проблем, отказ от първоначално допускане, формулиране на различна цел или преработване на собствена идея под въздействието на машинно предложение. При ИИ изменението се наблюдава функционално – като промяна на следващия изход вследствие на натрупания контекст, човешката обратна връзка, новите инструкции или предоставените междинни версии.

Приписването на резонанс се основава на проследими промени във взаимодействието, а не на предположения за вътрешните състояния на техническата система. Изследването не трябва да твърди, че ИИ разбира, преживява или създава по човешки начин. Аналитично релевантно е дали системата генерира контекстуално променен принос, който участва във формирането на следващите човешки решения, и дали тези решения на свой ред променят последващото машинно генериране.

Операционализирането изисква резонансът да бъде разглеждан като процес, а не като еднократно събитие. Изолиран силен отговор, оригинална идея или убедително формулирана интерпретация могат да бъдат полезни, но не са достатъчни за установяване на резонанс. Трябва да се проследи поне един рекурсивен цикъл, в който приносът е интерпретиран, оценен или преобразуван и резултатът от тази преработка е върнат в последващото взаимодействие.

Основното доказателство е промяната между последователни състояния на задачата. Поради това анализът трябва да сравнява първоначалната формулировка на проблема, машинните изходи, човешките реакции, междинните версии и крайния продукт. Всяко твърдение за резонанс трябва да се свърже с конкретна хронологична и съдържателна следа: какво е било изходното положение, кой принос е предизвикал промяната, как човекът я е интерпретирал и по какъв начин тя е повлияла върху следващия цикъл.

Трите резонанса, а именно когнитивен, творчески и емоционален, представят различни аналитични измерения. Те могат да се проявяват едновременно, последователно или независимо едно от друго. Силният когнитивен резонанс не предполага непременно творческа новост, а емоционалната ангажираност не доказва промяна в разбирането. Разделното им кодиране е необходимо, за да се установи конкретният профил на взаимодействието и да не се формира общо положително заключение въз основа на един силно изразен аспект.

Когнитивният резонанс се отнася до проследима промяна в разбирането и концептуалната организация на задачата. Неговите индикатори трябва да показват, че взаимодействието е довело до ново разграничение, преформулиране на проблема, промяна на допускане, установяване на връзка между по-рано несвързани елементи или аргументирана корекция. Съществена е не само появата на нова информация, а нейното интегриране в изменена структура на разбирането.

Творческият резонанс се отнася до преобразуването на пространството на възможните решения. Той може да се прояви чрез генериране и последващо преработване на варианти, откриване на неочаквани връзки, комбиниране на различни идеи, промяна на първоначалните критерии или развитие на замисъл през няколко рекурсивни цикъла. За творчески резонанс не е достатъчно ИИ да предложи оригинален вариант. Необходимо е този вариант да бъде интерпретиран и използван като материал за по-нататъшно преобразуване.

Емоционалният резонанс се отнася до влиянието на взаимодействието върху човешката ангажираност, мотивация, увереност, доверие и готовност за продължаване на работата. Той може да се наблюдава чрез самоотчети, промяна на тона, увеличаване на детайлността на заданията, продължаване на диалога след затруднение или поемане на по-сложна посока на изследване. Тези признаци трябва да бъдат тълкувани внимателно, защото могат да отразяват и обща удовлетвореност, удобство на интерфейса или антропоморфизиране на системата.

Наблюдаемите индикатори и ограниченията на тяхното тълкуване са обобщени в следващата табл. 3.

Таблица 3. *Операционни индикатори на трите резонанса*

<i>Измерение</i>	<i>Наблюдаеми индикатори</i>	<i>Ограничение на тълкуването</i>
Когнитивен резонанс	Ново разграничение; преформулиране на проблема; промяна на допускане; по-висока концептуална свързаност; аргументирана корекция.	Не се свежда до съгласие или езикова гладкост.

<i>Измерение</i>	<i>Наблюдаеми индикатори</i>	<i>Ограничение на тълкуването</i>
Творчески резонанс	Разширяване на вариантите; неочаквани връзки; рекомбинация; промяна на критериите; развитие на идея през няколко цикъла.	Не е тъждествен с оригиналността на единичен машинен изход.
Емоционален резонанс	Самоотчет за разбиране, доверие, мотивация или вдъхновение; промяна на тона; продължаване и задълбочаване на диалога; функционално адаптиране на отговора.	Не предполага човекоподобно емоционално преживяване на ИИ.

Представените индикатори имат насочваща, а не автоматично класификационна функция. Наличието на отделен индикатор не е достатъчно за заключение за резонанс. Например преформулирането на проблем може да представлява само езикова редакция, а продължаването на диалога може да се дължи на неяснота или неудовлетвореност. Анализаторът трябва да установи функцията на промяната в конкретния контекст и връзката ѝ с последващото развитие на задачата.

Всеки индикатор следва да се кодира въз основа на ясно определена доказателствена единица. Тя може да включва конкретен диалогичен ход, свързана двойка от човешки и машинен принос, цял рекурсивен цикъл или сравнение между последователни версии на продукта. Кодирането трябва да посочва не само вида на предполагаемия резонанс, но и изходното състояние, наблюдаваната промяна, приноса, който я е предизвикал, и отражението ѝ върху следващите действия.

Когнитивният резонанс не трябва да се приписва само защото машинният отговор е логически подреден, информативен или съвпада с позицията на човека. Съгласието може да потвърди съществуващо разбиране, без да го промени. По-силно доказателство е налично, когато човекът формулира нова концептуална връзка, ревизира допускане, въвежда по-точно разграничение или променя структурата на аргумента въз основа на взаимодействието.

Творческият резонанс трябва да се разграничава от количественото увеличаване на вариантите. Генерирането на множество предложения може да разшири избора, но не показва непременно преобразуване на творческото пространство. По-силно доказателство е налице, когато предложенията предизвикват промяна на критериите, комбиниране на по-рано несвързани идеи или развитие на решение, което не присъства в завършен вид нито в човешкия замисъл, нито в първоначалния машинен изход.

Емоционалният резонанс изисква най-внимателно тълкуване. Самоотчетите за доверие, вдъхновение или усещане за разбиране предоставят информация за човешкото преживяване, но не доказват качеството на продукта или наличието на когнитивно и творческо преобразуване. Високото доверие може да бъде както ресурс за устойчиво взаимодействие, така и риск, ако намалява критичната проверка и увеличава склонността към приемане на неточни или неподходящи предложения.

Промяната на тона на машинния отговор също не е самостоятелно доказателство за емоционален резонанс. Тя може да бъде технически предизвикана от инструкция за стил или от натрупания контекст. Аналитично значение придобива, когато адаптираният отговор влияе върху човешката мотивация, увереност или готовност за преработване на задачата и тази промяна може да бъде подкрепена с допълнителни процесни данни.

Трите вида резонанс могат да образуват различни конфигурации. Когнитивният резонанс може да предизвика творчески, когато новото разбиране отвори различни решения. Творческият резонанс може да повиши мотивацията и да подкрепи емоционалния. Емоционалната ангажираност може да увеличи продължителността и дълбочината на работата, като създаде условия за последващи когнитивни преобразувания. Тези връзки трябва да се изследват емпирично, а не да се приемат като универсална причинна последователност.

Трябва да се допуска и отрицателен или дисонантен резонанс. Машинният принос може да предизвика объркване, концептуално отклонение, прекомерно разширяване на вариантите или необосновано доверие. Подобни ефекти също представляват взаимно обусловено изменение, но не подкрепят положителна оценка на процеса. Поради това видът, посоката и последствията от промяната трябва да се кодират отделно от самото наличие на взаимодействие.

Креативността, креативното стимулиране и творческият резонанс се операционализират и кодират като свързани, но различни явления. Креативността обозначава способност или качество, което може да се отнася до човека, процеса или продукта. При продукта тя обикновено се оценява чрез оригиналност, уместност, полезност и разработеност. Тази оценка е резултатна и сама по себе си не установява как е възникнал продуктът.

Креативното стимулиране обозначава въздействието на машинно генерираните варианти, аналогии или въпроси върху човешкото мислене. То може да разшири вниманието, да активира нови асоциации или да помогне за преодоляване на творчески блокаж. Стимулирането може да възникне и при еднократен машинен изход, без последващ рекурсивен процес. Поради

това то е потенциална предпоставка за творчески резонанс, но не е тъждествено с него.

Творческият резонанс предполага рекурсивно преобразуване. Машинно генерираните варианти се интерпретират, подбират, комбинират, изменят или отхвърлят от човека, а резултатът от тази работа се връща в диалога и влияе върху следващия машинен принос. Именно тази проследима взаимна промяна отличава творческия резонанс от пасивното получаване на идеи и от избора между предварително генерирани алтернативи.

При кодирането трябва отделно да се отбелязват източникът на първоначалния импулс, видът на човешката преработка, промяната в следващия машинен изход и отражението върху продукта. Така може да се установи дали ИИ изпълнява функция на генератор на варианти, критичен събеседник, преобразуващ посредник, структуриращ инструмент или ко-креативен участник. Функционалната роля може да се променя в различните цикли на един и същ епизод.

Надеждното операционализиране изисква използването на положителни, отрицателни и гранични примери. Положителните случаи показват ясно рекурсивно преобразуване; отрицателните включват взаимодействия без съществено изменение; граничните съдържат отделни индикатори, но не дават достатъчно основание за категорично заключение. Съпоставянето им подпомага прецизирането на кодиращите правила и намалява риска всеки продуктивен диалог да бъде определян като резонансен (Krippendorff, 2018).

Окончателната оценка не трябва да се основава на механично преброяване на индикаторите. Значение имат тяхната последователност, взаимна връзка, интензивност и роля в развитието на продукта. По-обоснован е профилният подход, при който за всеки резонанс се описват наличните доказателства, степента на изразеност, ограниченията на тълкуването и увереността на оценителя. Така трите резонанса се превръщат от общи теоретични категории в наблюдаеми, сравними и емпирично проверими процесни измерения.

5. Критерии за асиметрично ко-креаторство

Асиметричното ко-креаторство обозначава форма на взаимодействие, при която човекът и ИИ участват съвместно във възникването и преобразуването на определен продукт. Но човекът и ИИ не са еднакви по природа, начин на познание и морална отговорност. Асиметрията не е недостатък на взаимодействието. Асиметрията е определящо условие за взаимодействие. Човекът участва чрез целеполагане, интерпретация, преживяване, експер-

тна и ценностна преценка, критична проверка и поемане на отговорност. Изкуственият интелект участва чрез аналитични, генеративни, структуриращи и преобразуващи функции, зависими от предоставените данни, инструкции и натрупания контекст.

Ролята на ИИ се определя според неговия функционален и проследим принос към процеса. Тя не зависи от допускания за човекоподобно съзнание, намерения, преживявания или автономна воля. Въпросът не е дали системата „иска“ да създаде продукта или „разбира“ значението му по човешки начин. От значение е дали ИИ приноса е съдържателно релевантен и участва в рекурсивното развитие на задачата⁵. Важно е също така дали той променя пространството на възможните решения и оставя реконструируема следа в крайния резултат.

Функционалното определение позволява да бъде признат реалният принос на ИИ, без техническата система да се приравнява на човешки автор или морален субект. Изкуственият интелект може да изпълнява ко-креаторска функция в конкретен процес, но от това не следва, че притежава обща способност за творчество, самостоятелно авторско намерение или право на академично авторство. Ко-креаторската функция е локална, релационна и зависима от конкретната организация на взаимодействието.

Локалният характер означава, че ко-креаторството се установява за определен епизод, продукт или решение, а не се приписва трайно на системата. Един и същ модел може да има ко-креаторска функция в един епизод и инструментална функция в друг. Например при разработването на концептуална рамка ИИ може да предложи ново разграничение, което да бъде многократно преработено и включено в крайния модел. При последваща правописна проверка същата система изпълнява помощна, а не ко-креаторска функция.

Релационният характер означава, че функцията на машинния принос не може да се определи извън човешката интерпретация и последващата употреба. Един оригинален машинен изход не е достатъчен за ко-креаторство, ако остава неизползван, механично копиран или несвързан с развитието на задачата. Същият изход може да придобие ко-креаторско значение, когато предизвика промяна в човешката позиция, води до съществено преобразуване и участва в следващи рекурсивни цикли.

Ко-креаторството трябва да се разграничава от техническото асистирание. При асистирането ИИ изпълнява предварително определена операция, без да променя съществено съдържанието или посоката на работата. Такива

⁵ Тук „рекурсивното развитие на задачата“ означава, че тя се развива чрез многократно повтаряне на един и същ процес, като резултатът от всяка стъпка служи за вход на следващата.

функции могат да бъдат корекция на правопис, форматиране, преобразуване на файлове, извличане на предварително зададени данни или съкращаване на текст според ясни инструкции. Те могат да повишат ефективността, но не разширяват непременно пространството на решенията и не участват в концептуалното развитие на продукта.

Нужно е разграничение и между ко-креаторството и креативното стимулиране. При креативното стимулиране машинният изход активира човешки асоциации, подпомага преодоляването на затруднение или предлага варианти, но по-нататъшното решение се разработва предимно от човека и не включва съществен рекурсивен обмен. Стимулиращата функция може да бъде важна предпоставка за ко-креативен процес, но сама по себе си не доказва, че ИИ е участвал в преобразуването на крайния продукт.

Резонаторната функция е налице, когато машинният принос отразява, структурира, разширява или поставя под въпрос човешките идеи и така подпомага тяхното изясняване. Изкуственият интелект може да действа като аналитично огледало, критичен събеседник или средство за експлициране на неясни допускания. Резонаторната функция се доближава до ко-креаторството, но остава по-ограничена, когато приносът влияе върху човешкото разбиране, без да оставя достатъчно различима и проследима роля в самия продукт.

Определянето на ко-креаторска функция изисква доказателства, които показват как е извършен даден процес, а не само какъв е крайният резултат. Качеството, оригиналността или полезността на крайния продукт не са достатъчни, защото подобен резултат може да бъде постигнат чрез самостоятелна човешка работа, чрез еднократно генериране или чрез подбор между готови варианти. Трябва да се реконструира как машинният принос е възникнал, как е бил интерпретиран и преработен, как е повлиял върху следващите решения и каква част от неговото въздействие се запазва в крайния резултат.

За да се достигне до силно заключение за наличие на асиметрично ко-креаторство могат да послужат четири критерия: различим принос, итеративно преобразуване, влияние върху пространството на решенията⁶ и проследима роля в продукта. Критериите са аналитично различни, но функционално свързани. Всеки от тях проверява отделен аспект на процеса и ограничава различен вид необосновано заключение.

Първият критерий е наличието на различим и съдържателно значим машинен принос. Той проверява дали ИИ добавя елемент, който надхвър-

⁶ Тук под „пространство на решенията“ се разбира съвкупността от възможности, измежду които може да се избере най-подходящото решение.

ля механичното повторение, перифразирание или езиково изглаждане на човешкото задание. Положително доказателство могат да бъдат: нова концептуална връзка, вариант, модел, аргумент, критика, структура, аналогия или въпрос, които имат функционално значение за развитието на задачата.

Различимостта не означава приносът да бъде напълно независим от човешкия контекст или да бъде създаден „от нищото“. Всеки машинен изход е обусловен от заданието, данните и историята на диалога. Критерият изисква приносът да не може да се опише достатъчно като пряко възпроизвеждане на вече предоставеното съдържание. Необходимо е той да въвежда преобразуване, разширение, съпоставка или структуриране, които изменят наличния материал.

Значимостта на приноса се оценява спрямо конкретната цел, а не според неговия обем. Кратък въпрос или разграничение може да има по-голяма ко-креаторска стойност от продължителен генериран текст, ако промени посоката на анализа. Обратно, голямо количество съдържание може да остане функционално несъществено, когато не е използвано или не влияе върху следващите решения.

Вторият критерий е итеративното преобразуване. Той проверява дали машинният принос участва в един или повече последващи цикли и дали в процеса се променя чрез човешка интерпретация, критика, подбор, комбиниране или преформулиране. За ко-креаторство не е достатъчно ИИ да предостави еднократен завършен резултат, който човекът да приеме без съществена преработка.

Итеративността предполага възможност да бъде проследена последователност от взаимно обусловени промени. Машинното предложение може да бъде частично прието и коригирано; човешката корекция може да доведе до ново задание; следващият машинен изход може да предложи различна структура, която отново да бъде оценена и изменена. Така продуктът се развива чрез цикли, при които предходното състояние се превръща в условие за следващото.

Броят на циклите сам по себе си не доказва итеративно преобразуване. Продължителният диалог може да включва повторение, уточняване на формулировки или поредица от несвързани заявки. Аналитично значение има качеството на връзката между циклите: дали всеки следващ етап използва и преобразува резултата от предходния и дали се наблюдава развитие, а не механично натрупване.

Третият критерий е влиянието върху пространството на решенията. Той проверява дали машинният принос променя разглежданите възможности, критериите за избор, формулировката на проблема или посоката на търсене. Положителното влияние може да се прояви чрез въвеждане на нова

алтернатива, откриване на пропуснато допускане, предлагане на различна структура или поставяне под въпрос на предварително избрана стратегия.

Промяната на пространството на решенията може да бъде разширяваща, ограничаваща или преобразуваща. Разширяващото влияние добавя нови възможности. Ограничаващото помага да бъдат отхвърлени неподходящи варианти чрез аргументи или критерии. Преобразуващото променя самото разбиране за проблема и води до формулиране на различен клас решения. И трите форми могат да бъдат ко-креаторски, ако са съдържателно релевантни и проследими.

Не всяко ускоряване на работата представлява влияние върху пространството на решенията. Ако ИИ само изпълнява по-бързо вече напълно определена операция, без да променя възможностите или критериите, приносът остава инструментален. Ефективността е важен практически резултат, но не е самостоятелно доказателство за ко-креаторство.

Четвъртият критерий е проследимата роля на машинния принос в крайния продукт. Той проверява дали влиянието на ИИ може да бъде реконструирано чрез историята на диалога, последователните версии, бележките за решенията или карта на приносите. Проследимостта свързва процеса с резултата и позволява твърдението за ко-креаторство да бъде независимо проверено.

Проследимата роля не предполага непременно дословно присъствие на машинно генериран текст. Приносът може да бъде концептуален, структурен или критичен и да се запази в продукта в преобразувана форма. Например предложение на ИИ може да бъде преработено изцяло на езиково равнище, но да остане разпознаваемо като нова логическа структура или изследователска посока. В такъв случай трябва да бъде документирана веригата на преобразуването.

Липсата на дословно съвпадение не изключва ко-креаторство, но липсата на реконструируема връзка значително отслабва заключението. Ретроспективното твърдение на участника, че ИИ е „оказал влияние“, не е достатъчно без подкрепящи данни. Колкото по-пълна е доказателствената история, толкова по-висока може да бъде увереността в класификацията.

Четирите критерия са съвместно необходими за силно заключение за ко-креаторска функция на ИИ. Различимият принос без итеративно преобразуване може да бъде оригинален, но еднократен машинен изход. Итеративното взаимодействие без промяна на пространството на решенията може да представлява продължително редактиране. Влиянието върху решенията без проследима връзка с продукта остава трудно проверимо. Проследимото присъствие на машинен текст без съдържателен и преобразуващ принос може да бъде механично включване.

Съвместната необходимост на критериите не предполага еднаква степен на проявление. Различните ко-креативни епизоди могат да имат различни профили. При концептуална задача водещо може да бъде влиянието върху пространството на решенията. При създаването на визуален или текстов продукт по-ясно могат да се проявят итеративното преобразуване и проследимата роля. Оценката трябва да отчита спецификата на задачата, без да отменя минималния праг за всеки критерий.

Частичното изпълнение позволява по-предпазливо и диференцирано определение. При наличие на различен принос без последваща рекурсивна преработка е подходящо да се говори за генериране или креативно стимулиране. При продължително адаптиране, което подпомага човешкото разбиране, но не оставя проследима роля в продукта, може да се използва определението „резонаторна функция“. При техническо изпълнение на предварително определена операция по-коректно е понятието „асистиране“.

При несигурни или гранични случаи анализаторът не трябва да налага двоична класификация⁷. Възможно е заключението да бъде формулирано като „ограничени доказателства за ко-креаторска функция“, „частично изпълнение на критериите“ или „преобладаващо инструментално подпомагане с отделни ко-креативни елементи“. Подобна степенувана оценка отразява по-точно сложността на реалните взаимодействия.

Критериите се прилагат към конкретния ко-креативен епизод, а не автоматично към целия проект. В рамките на една продължителна работа ИИ може последователно да изпълнява различни функции: източник на информация, технически асистент, генератор на варианти, критичен резонатор и ко-креативен участник. Функционалната роля трябва да се определя отделно за всеки аналитично разграничен епизод.

Оценката трябва да отчита и алтернативните обяснения. Привидно новият машинен принос може да произтича непосредствено от подробно човешко задание; промяната в решението може да е била предварително обмислена от човека; крайният продукт може да съдържа машинен текст, без той да е повлиял върху концептуалната организация. Поради това всяко положително заключение трябва да се подкрепи с доказателства, които разграничават ко-креаторската функция от по-прости форми на участие.

Асиметричното ко-креаторство не намалява човешката отговорност. Напротив, колкото по-съществен е машинният принос, толкова по-голяма е необходимостта от критична проверка, документиране и ясно човешко утвърждаване. Човекът носи отговорност за избора на целта, използването на

⁷ Тук „двоична класификация“ означава разделяне на случаите в две взаимно изключващи се категории, т.е. всеки елемент се отнася към една от две възможни групи (например „да / не“, „има / няма“, „положителен / отрицателен“ и др. под).

системата, приемането или отхвърлянето на предложенията, проверката на резултата и последствията от неговото приложение.

Така четирите критерия изпълняват двойна функция. Те позволяват да бъде признато значимото и преобразуващо участие на ИИ, без да му се приписват недоказани човешки характеристики. Същевременно поставят достатъчно строг праг, който разграничава асиметричното ко-креаторство от асистирането, автоматизацията, еднократното генериране и общото креативно стимулиране.

Следващата табл. 4 систематизира предложените критерии, възможните контролни въпроси и видовете положителни, отрицателни и гранични доказателства, използвани при определянето на ко-креаторската функция на ИИ.

Таблица 4. Критерии за ко-креаторска функция на ИИ

<i>Критерий</i>	<i>Контролен въпрос</i>	<i>Положително доказателство</i>	<i>Отрицателен/граничен случай</i>
Различим принос	Добавя ли ИИ значим елемент извън механичното повторение на заданието?	Нова връзка, вариант, модел, критика или структура.	Парафраза, шаблон или несъществена редакция.
Итеративно преобразуване	Участва ли приносът в следващи цикли и променя ли се чрез тях?	Версии, в които човешки и машинни приноси се преобразуват взаимно.	Еднократен изход без по-нататъшно развитие.
Влияние върху пространството на решенията	Променя ли приносът разглежданите възможности или критериите?	Нова посока, алтернатива или преосмислено допускане.	Само ускоряване на вече определено решение.
Проследима роля в продукта	Може ли влиянието да се реконструира в крайния резултат?	История на диалога, версии и карта на приносите.	Липса на връзка между изхода и продукта.

6. Процедура за идентифициране и кодиране

Процедурата за идентифициране и кодиране превежда онтологичните, логическите и операционните положения на методологията⁸ в последова-

⁸ Тук се имат предвид три взаимосвързани групи предпоставки: онтологични (какво съществува и каква е природата на изследвания обект), логически (как се достига до научно знание и как се правят изводите), операционни (как на практика се извършва изследването – с какви методи, инструменти и процедури). Тези три нива заедно осигуряват последователност между философските основи на изследването, начина на разсъждение и конкретното му изпълнение.

телност от практически аналитични действия (табл. 5). Тя започва с ясно ограничаване на изследвания случай и завършва с аргументирано заключение за характера на наблюдаваното взаимодействие. Целта ѝ е да направи анализа прозрачен, проследим и възпроизводим, така че заключението да не зависи единствено от общото впечатление или субективната преценка на отделния анализатор.

Процедурата следва ясно да разграничава идентифицирането, измерването и оценяването. Идентифицирането отговаря на въпроса дали в конкретния епизод е налице конфигурация, съвместима с проявление на интегриран интелект. Измерването определя степента и профила на отделните резонанси и ко-креаторски критерии. Оценяването се отнася до качеството, оригиналността, точността, приложимостта или обосноваността на продукта. Трите дейности са свързани, но резултатите от едната не могат автоматично да заместят доказателствата, необходими за другите.

Висококачественият продукт не е достатъчно доказателство за интегриран интелект. Това е така, защото той може да е резултат от предварителната експертност на човека, от еднократен машинен изход или от техническо подпомагане без съществено рекурсивно преобразуване. По същия начин наличието на богат ко-креативен процес не гарантира непременно високо качество на продукта. Анализът трябва да допуска възможността за силно взаимодействие с незадоволителен резултат, както и за качествен резултат без установено ко-креаторство.

Процедурата се прилага към предварително определена единица на анализ. В зависимост от целта това може да бъде рекурсивен цикъл, ко-креативен епизод или система от епизоди в рамките на проект. За силно класификационно заключение е препоръчително основната единица да бъде ко-креативният епизод, защото той позволява да бъдат свързани последователността на взаимодействието, наблюдаваните преобразувания и относително стабилният продукт или решение.

Преди началото на същинското кодиране трябва да се определи аналитичният протокол, т.е. да се създаде структуриран документ или запис, в който систематично да се описва как е извършен анализът и какви резултати са получени. Той трябва да включва работни определения на категориите, правила за включване и изключване, критерии за начало и край на епизода, скала за оценяване и изисквания към доказателствата. Когато тези решения се вземат едва след прегледа на резултата, възниква риск категориите да бъдат адаптирани към очакваното заключение. Предварителното им формулиране ограничава потвърждаващото пристрастие и улеснява независимото повторение на анализа (Krippendorff, 2018).

Първият етап е ограничаването на случая. Анализаторът трябва да определи задачата, целта, контекста, участниците, използваната система, вре-

мевия обхват, началото и края на взаимодействието и продукта, към който то е насочено. Трябва да бъдат документирани и релевантните характеристики на средата: версия на модела, основни инструкции, използвани настройки, достъп до външни източници и наличие на допълнителни човешки участници.

Границите на случая трябва да бъдат достатъчно тесни, за да позволяват подробно проследяване, но и достатъчно широки, за да обхванат функционално завършен процес. Ако се анализира само отделен впечатляващ отговор, може да се пропусне липсата на последващо преобразуване. Ако се включи целият проект без разграничаване на епизодите, локалните прояви могат необосновано да бъдат обобщени като характеристика на цялата работа.

Началото на епизода се определя чрез появата на конкретна цел, проблем, въпрос или творческо намерение. Краят му се свързва със стабилизиране на продукт, приемане на решение, съществена промяна на целта или прекъсване на взаимодействието. Когато в хода на работата възникне нов относително самостоятелен проблем, той трябва да бъде отделен като нов епизод, дори да принадлежи към същия по-голям проект.

Вторият етап е събирането и организирането на доказателствената база. Тя може да включва пълната история на диалога, последователните версии на продукта, човешки бележки, дневник на решенията, използвани външни източници, резултати от проверки, самоотчети и окончателния продукт. Данните трябва да позволяват хронологично реконструиране на взаимодействието и проследяване на връзката между отделните приноси и последващите промени.

Предпочитание е добре да се отдава на пълния корпус на взаимодействието. Когато по практически или етични причини се използва подбор, правилата за подбор трябва да бъдат изрично формулирани и обосновани. Не е допустимо да се включват само успешните или най-впечатляващите части от диалога, защото това създава изкривена представа за последователността и ефективността на процеса.

Данните трябва да се съхраняват във форма, която запазва тяхната хронология и произход. Трябва да бъде ясно кой принос е човешки, кой е машинно генериран, какво е времето и мястото му в последователността и към коя версия на продукта се отнася. При редактирани или обединени записи трябва да се пази и първоначалната версия, за да не бъде загубено проследяването на произхода, развитието и етапите на една промяна или трансформация (т.е. генеалогията на преобразуването).

Третият етап е сегментирането на материала. Диалогът се разделя на ходове, рекурсивни цикли и ко-креативни епизоди. Сегментирането не трябва

да се основава само на технически граници между съобщенията. Един аналитичен ход може да съдържа няколко функции, а един рекурсивен цикъл може да обхване поредица от съобщения, свързани с общо преобразуване.

За всеки сегмент се отбелязват неговата функция, връзката с предходния принос и последствията за следващия етап. Човешкият ход може да изпълнява функция на целеполагане, оценка, критика, подбор, преработване или утвърждаване. Машинният принос може да действа като предложение, анализ, структуриране, критика, разширяване на варианти или функционално адаптиране. Един и същ принос може да има повече от една функция, но всяка кодирана функция трябва да бъде обоснована.

Четвъртият етап е същинското кодиране. Анализаторът оценява проявите на когнитивен, творчески и емоционален резонанс, както и изпълнението на критериите за ко-креаторска функция на ИИ. Всеки зададен код трябва да бъде свързан с конкретен цитат, действие, промяна между версии или друг наблюдаем елемент. Общи формулировки като „диалогът е продуктивен“ или „ИИ има съществен принос“ не са достатъчни без посочено доказателство.

Кодът е необходимо да описва както наличието на индикатор, така и посоката на промяната. Например преформулирането на проблема може да доведе до по-голяма концептуална яснота, но може и да отклони анализа от първоначалната цел. Генерирането на допълнителни варианти може да разшири творческото пространство, но може и да създаде информационно претоварване. Емоционалната адаптация може да подкрепи ангажираността, но и да засили некритичното доверие.

При кодирането трябва да се регистрират както положителните, така и отрицателните, граничните и противоречащите доказателства. Положителните доказателства подкрепят определен код или критерий. Отрицателните показват отсъствие на очакваната проява или наличие на процес, несъвместим с нея. Граничните дават частична или слабо проследима подкрепа. Противоречащите доказателства показват, че различните източници или части от процеса водят до несъвместими интерпретации.

Неуспешните, отхвърлените и нерелевантните машинни предложения са особено важни. Те показват, че човекът не приема автоматично генерираните изходи, а упражнява критична преценка, подбор и контрол. Същевременно прекомерният брой отхвърлени предложения може да означава, че ИИ има ограничена функционална роля и че крайният резултат се дължи предимно на човешката работа. Отхвърлянето не е автоматично нито положителен, нито отрицателен знак; значението му зависи от функцията в процеса.

Неуспешните предложения могат да имат и продуктивна роля. Неподходящ машинен вариант може да помогне на човека да уточни критериите,

да разпознае скрито допускане или да формулира по-точно целта. В такъв случай приносът не участва пряко в продукта, но може да има когнитивна или резонаторна функция. Тази функция трябва да се кодира отделно от проследимата ко-креаторска роля в крайния резултат.

Петият етап е триангулацията⁹. Тя включва съпоставяне на различни източници на доказателства: диалогичната история, последователните версии, самоотчета на участника, дневника на решенията и външната оценка. Целта е да се установи къде всички източници се подкрепят, допълват или си противоречат.

Самоотчетът може да предостави информация за промяната в човешкото разбиране, намерение, мотивация или усещане за творческо стимулиране. Съществува обаче вероятност резултатите да бъдат повлияни от ретроспективно изкривяване, социална желателност и антропоморфизация. Поради това самоотчетът не трябва да се разглежда като самостоятелно доказателство, когато съответните твърдения могат да бъдат проверени чрез анализа на диалога и последователните версии на продукта.

Диалогичната история показва непосредствената последователност на приносите, но не винаги разкрива причините за човешкия избор. Последователните версии показват как се е изменял продуктът, но могат да пропуснат отхвърлените алтернативи. Външната оценка позволява независимо преценяване на качеството, но няма непременно достъп до вътрешната логика на процеса. Триангулацията използва силните страни на всеки източник и компенсира ограниченията му.

Противоречията между източниците не бива да се заличават чрез механично усредняване. Ако участникът съобщава силна концептуална промяна, но диалогът и версиите не показват проследима следа, заключението трябва да е предпазливо. В случай, че външният оценител намира продукта за неоригинален, но процесът съдържа ясни доказателства за творчески резонанс, трябва да се разграничат качеството на продукта и характерът на взаимодействието.

Шестият етап е оценяването на силата на доказателствата чрез скала от 0 до 3. Скалата измерва степента, в която конкретен индикатор или критерий е подкрепен от данните. Всяко измерение трябва да бъде оценено отделно, като към стойността се отбелязва кратко обяснение и препратка към съответните доказателства.

Стойност 0 означава, че проявата отсъства или че наличните данни противоречат на нейното наличие. Тя не трябва автоматично да се присъжда,

⁹ Тук под „триангулация“ се има предвид следното: един и същ феномен се изследва чрез повече от един източник, метод, изследовател или теоретична перспектива, за да се повиши достоверността и надеждността на резултатите.

когато информацията е недостатъчна. Липсата на данни и доказаното отсъствие са различни състояния и трябва да се разграничават в кодиращата схема.

Когато доказателствената база не позволява оценка, е препоръчително да се използва отделно означение, например „НД“ – недостатъчни данни или „неоценимо“. Това предотвратява изкуственото понижаване на резултата и позволява да се посочи каква допълнителна информация е необходима. Стойност 0 трябва да се използва само когато анализираният материал е достатъчен за отрицателно заключение.

Стойност 1 обозначава ограничена, единична или слабо проследима проява. Може да е налице отделен индикатор, който не участва в завършен рекурсивен цикъл, или твърдение, подкрепено предимно от един несигурен източник. Тази стойност показва възможно проявление, но не позволява уверено заключение.

Стойност 2 обозначава ясно проявление с конкретна и проследима следа. Необходимо е да може да се посочи определен цитат, действие, промяна между версии или друг надежден източник. Проявата е отчетлива в конкретен момент или цикъл, но може да няма достатъчна повтораемост или независима подкрепа, за да бъде определена като устойчива.

Стойност 3 обозначава устойчиво проявление, наблюдавано в повече от един цикъл, потвърдено от повече от един източник или независимо проверимо чрез историята на взаимодействието и версиите на продукта. Тази стойност изисква съгласувани доказателства и липса на съществени необяснени противоречия. Големият брой повторения сам по себе си не е достатъчен, ако те възпроизвеждат една и съща слабо значима проява.

Скалата може да се прилага профилно, като отделни стойности се присъждат на когнитивния, творческия и емоционалния резонанс и на всеки критерий за ко-креаторска функция. Не е препоръчително те автоматично да се сумират в един общ резултат, защото различните конфигурации могат да имат различно теоретично значение. Епизод със силен когнитивен резонанс и слаба емоционална ангажираност не е еквивалентен на обратното (слаб когнитивен резонанс и силна емоционална ангажираност), дори сборът да е еднакъв.

Ако се използва общ индекс, неговата логика трябва да бъде предварително обоснована и емпирично валидирана. Необходимите условия и подкрепящите индикатори не трябва да получават еднаква тежест по подразбиране. Високите стойности по подкрепящи признаци не могат да компенсират отсъствието на необходимо условие, например липсата на човешка критична преработка или проследима връзка с продукта.

Седмият етап е формирането на заключение. То трябва да посочи дали необходимите условия са изпълнени, какъв е профилът на резонансите,

доколко са подкрепени критериите за ко-креаторска функция и какви ограничения има доказателствената база. Заключениеето не трябва да бъде само етикет, а кратка аргументирана реконструкция на процеса.

Вместо единствено двоично решение „има“ или „няма“ интегриран интелект могат да се използват няколко степени на увереност: липса на доказателства; ограничени или гранични доказателства; конфигурация, съвместима с предполагаемо проявление; силно подкрепено предполагаемо проявление.

Когато критериите са изпълнени само частично, взаимодействието трябва да се определи по-точно като инструментално асистирание, генеративно подпомагане, креативно стимулиране, резонаторна функция или смесена конфигурация. Това не представлява отрицателна оценка на полезността на процеса. То показва, че наличните доказателства подкрепят различен тип взаимодействие.

Осмият етап е проверката на надеждността на кодирането. При изследвания, включващи повече от един анализатор, е необходимо поне част от емпиричния материал да бъде кодирана независимо от всеки анализатор, след което да се оцени степента на съгласие между оценителите. Разногласията трябва да се използват за прецизиране на определенията, индикаторите и примерите, а не само да се разрешават чрез обсъждане без документиране.

Преди основното изследване е препоръчително пилотно кодиране на положителни, отрицателни и гранични случаи. Това позволява да се установи кои категории се припокриват, кои индикатори са твърде общи и при кои типове данни възниква несигурност. Кодирацият наръчник трябва да се актуализира въз основа на пилотните резултати, като промените се документират.

Процедурата трябва да осигурява и възможност за одит. Външен изследовател трябва да може да проследи как от първичните данни се достига до конкретен код и от кодовете – до общото заключение. За тази цел е необходимо да се поддържа регистър, съдържащ идентификатор на епизода, цитат или препратка към доказателството, присъден код, стойност, аргументация, алтернативна интерпретация и степен на увереност.

Етичните и правните изисквания са част от процедурата, а не последващо допълнение. Диалозите могат да съдържат лични данни, поверителна информация, авторски материали или чувствителни сведения. Преди архивирането и анализа на данните е необходимо да се установят ясни правила относно информираното съгласие, анонимизацията, режима на достъп, условията за съхранение и начина на представяне на данните. Необходимост-

та от проследимост не отменя задължението за защита на участниците и засегнатите лица.

В окончателния отчет трябва да бъдат представени не само резултатите, но и ограниченията на процедурата. Те могат да включват непълен архив, липса на междинни версии, зависимост от самоотчет, неясни граници на епизода, промяна на модела по време на проекта или недостатъчно независимо кодиране. Ограниченията трябва да бъдат свързани с конкретното им влияние върху увереността на заключението.

Така процедурата превръща идентифицирането на интегриран интелект от интуитивна оценка в доказателствено обоснован аналитичен процес. Ограничаването на случая, системното събиране на данни, сегментирането, кодирането, триангулацията, профилната оценка и проверката на надеждността образуват свързана последователност. Следващата табл. 5 обобщава основните етапи и скалата за кодиране.

Таблица 5. Процедура и скала за кодиране

<i>Етап/стойност</i>	<i>Съдържание</i>	<i>Изискване</i>
1. Ограничаване	Определяне на задача, контекст, участници, система, начало, край и продукт.	Еднозначна единица на анализ.
2. Събиране	Архивиране на диалог, версии, бележки, външни проверки и продукт.	Пълна или обосновано подбрана доказателствена база.
3. Кодиране	Оценка на резонансите и ко-креаторските критерии.	Конкретен цитат, действие или промяна за всеки код.
4. Триангулация	Съпоставка на диалог, версии, самоотчет и външна оценка.	Отбелязване на съвпадения и противоречия.
5. Заключение	Определяне на типа взаимодействие, профила на проявлението и увереността.	Аргументирана връзка между кодовете, отрицателните доказателства и класификацията.
НД – недостатъчни данни	Доказателствената база не позволява надеждна оценка.	Не се приравнява на стойност 0 и не участва автоматично в обща оценка.
0 – отсъства	Данните показват отсъствие на проявата или активно ѝ противоречат.	Присъжда се само при достатъчна база за отрицателно заключение.

<i>Етап/стойност</i>	<i>Съдържание</i>	<i>Изискване</i>
1 – ограничено	Единична или слабо проследима проява.	Частично доказателство.
2 – ясно	Отчетлива проява с конкретна проследима следа.	Най-малко един надежден източник.
3 – устойчиво	Последователна проява в повече от един цикъл или източник.	Съгласувани и независимо проверими следи.

7. Процесни и резултатни критерии

Идентифицирането на ко-креативен процес и оценяването на възникналите продукти са свързани, но методологично различни задачи. Първата е насочена към начина, по който се развива взаимодействието между човека и ИИ, а втората – към качествата на получения резултат. Разграничението е необходимо, защото високото качество на продукта не доказва автоматично ко-креаторство, както и наличието на ко-креативен процес не гарантира успешен резултат.

Качествен продукт може да бъде получен чрез еднократна инструментална употреба на ИИ. Добре формулирано задание може да доведе до точен, оригинален или практически полезен машинен изход, който човекът да приеме с минимални промени. В такъв случай продуктът може да има висока стойност, но процесът не съдържа достатъчно доказателства за рекурсивно преобразуване и асиметрично ко-креаторство. Възможно е и обратното. Процесът може да включва продължителна обратна връзка, когнитивен и творчески резонанс, взаимно преобразуване на приносите и ясно разпределение на решенията. И въпреки това да завърши с неточен, непълен или практически неприложим продукт. Причината може да бъде например неправилна първоначална информация, недостатъчна експертност, неподходящ критерий за избор, ограничение на модела или неуспешно човешко утвърждаване.

Ко-креативността е характеристика на организацията на процеса, а качеството – характеристика на резултата спрямо определени външни критерии. Между тях може да съществува положителна връзка, но тя трябва да бъде изследвана емпирично. Не е коректно наличието на ко-креаторство да се извежда само от оригиналността, точността, полезността или сложността на крайния продукт.

Критерии за проследяване на процеса произтичат от динамиката и развитието на взаимодействието между човека и системата с ИИ. Те обхващат рекурсивните цикли, преобразуването на човешкия и машинния принос и използването на обратната връзка. Включват също проявите на когнитивен, творчески и емоционален резонанс, както и разпределението на решенията.

та между човека и системата. Анализът не се ограничава до установяване на тези характеристики. Той разглежда тяхната функция и принос им към процеса на решаване на задачата.

Рекурсивността отразява степента, в която резултатите от предходните ходове се използват като входни условия за последващите етапи на взаимодействието. Тя не се свежда единствено до последователна размяна на съобщения. Необходимо е човекът да оцени или преработи машинния изход. Въз основа на него той формулира ново задание, взема решение или създава нова версия. Този принос от своя страна влияе върху следващия машинно генериран резултат.

Преобразуването се отнася до съдържателната промяна на приносите. Машинното предложение може да бъде допълнено, комбинирано, ограничено, критично преформулирано или използвано като основание за нова посока. Човешката идея също може да бъде структурирана, разширена или поставена под въпрос чрез машинния изход. Когато приносите само се натрупват без взаимна промяна, доказателствата за ко-креативен процес остават ограничени.

Обратната връзка трябва да има функционална роля. Общи реакции като „продължи“, „направи го по-добре“ или „дай още варианти“ могат да поддържат диалога, но невинаги показват критична преработка. По-силно доказателство е налице, когато обратната връзка съдържа аргументирана оценка, промяна на критерий, посочване на проблем или въвеждане на ново разграничение, което влияе върху последващия изход.

Резонансът показва дали взаимодействието променя разбирането, пространството на решенията или човешката ангажираност. Процесният анализ установява вида, посоката и интензивността на тази промяна. Положителният резонанс подпомага развитието на задачата. Дисонансът може да доведе до объркване, отклонение от целта или прекомерно доверие в системата.

Разпределението на решенията показва кой определя целите и критериите, избира между възможните варианти, извършва проверките и утвърждава крайния продукт. Чрез този показател се установява асиметрията във взаимодействието. Изкуственият интелект може да влияе върху решенията чрез анализи и предложения. Окончателната смислова, ценностна и практическа преценка обаче трябва да остава проследимо човешка.

Резултатните критерии оценяват относително стабилизирания краен продукт. Според характера на задачата той може да бъде текст, концептуален модел, решение, план, изображение, прототип, компютърна програма или друг артефакт. Критериите трябва да отчитат неговото предназначение и функция. Доколкото е възможно, оценката на продукта следва да бъде независима от оценката на процеса на създаването му.

Оригиналността показва доколко продуктът включва нетривиални идеи, комбинации или решения. Тя не е равнозначна на необичайност. Новият елемент трябва да съответства на целта и да има смислова или функционална стойност. Оценката е по-надеждна, когато се извършва от независими оценители или чрез сравнение с подходящ корпус от решения.

Обосноваността се определя от логическата последователност на изложението, качеството на аргументацията и наличието на релевантни данни или източници. Убедителното звучене на машинно генерирания текст само по себе си не доказва неговата обоснованост. Необходимо е да се провери дали твърденията произтичат от посочените основания и дали направените заключения са логически подкрепени.

Точността показва доколко резултатът съответства на проверими факти, коректни изчисления, надеждни източници, приложими нормативни изисквания и технически спецификации. Тя е особено важна при научни, медицински, правни, финансови и технически задачи. В тези области дори малки неточности могат да имат сериозни последици. Поради това точността се оценява независимо от езиковото качество, стилистичната последователност и увереността, с която е представен резултатът.

Приложимостта показва доколко продуктът може да изпълни предназначението си в конкретни условия. Тя зависи от практическата му реализуемост, наличните ресурси и разбираемостта му за целевата аудитория. От значение е и съответствието с организационните, етичните и правните ограничения. Дори оригинален продукт може да има ниска приложимост, ако не отчита реалния контекст.

Концептуалната цялост показва доколко отделните части на продукта образуват последователна структура. Тя предполага съгласуваност на понятията и липса на вътрешни противоречия. Важно е да има ясна връзка между целта, аргументацията и заключението. Различните представяния на резултата също трябва да си съответстват. Този критерий е особено важен при продукти, които съчетават множество човешки и машинни приноси.

Процесните и резултатните показатели трябва да се оценяват поотделно, а след това да се съпоставят. За тази цел може да се изгради двуизмерен профил. Едната ос представя силата на ко-креативния процес, а другата – качеството на крайния продукт. Така се разграничават четири основни конфигурации:

- ясен ко-креативен процес и висококачествен продукт;
- ясен ко-креативен процес, но незадоволителен продукт;
- ограничено или липсващо ко-креаторство, но висококачествен продукт;
- ограничено ко-креаторство и незадоволителен продукт.

Първата конфигурация представлява най-силния практически случай, но дори при нея качеството на продукта не замества процесните доказател-

ства. Втората показва, че ко-креативното взаимодействие не е гаранция за успех и че човешкото утвърждаване може да е било недостатъчно. Третата предупреждава да не се приписва ко-креаторство само поради впечатляващ резултат. Четвъртата може да показва неподходяща задача, слаба организация на взаимодействието или ограничения на използваната система.

Генеалогията на продукта свързва анализа на процеса с оценката на крайния резултат. Тя проследява как идеите, структурите, аргументите, решенията и редакциите преминават през отделните версии. Така се установява как те достигат до крайния продукт. Без подобна реконструкция процесът и резултатът остават паралелни описания. Връзката между тях не може да бъде надеждно обяснена във функционален и причинен план.

Генеалогичният анализ обхваща приетите, изменените и отхвърлените предложения. Приетите предложения показват пряк принос към крайния продукт. Изменените разкриват начина, по който човекът интерпретира и преобразува предложенията на ИИ. Отхвърлените също имат аналитична стойност. Те могат да свидетелстват за критична преценка, промяна на критериите или по-точно определяне на целта. Това важи дори когато съдържанието им не присъства в крайния резултат.

Надеждната генеалогия се изгражда чрез хронологично съхранени диалози, междинни версии и дневник на решенията. За всеки ключов елемент на продукта трябва да се установи кога е възникнал и от чий принос произлиза. Необходимо е също така да се проследят неговите преобразувания и основанията за приемането му. Не е нужно всяка дума да се приписва изцяло на човека или на ИИ. Важно е обаче функционалната история на съществения елементи да може да бъде реконструирана.

Когато продуктът е текст, могат да се анализират промени в тезите, аргументацията, структурата и терминологията. При визуален продукт могат да се проследят композицията, цветовите решения, символите и последователните варианти. При модел или решение могат да се анализират предпоставките, променливите, зависимостите, критериите и отхвърлените алтернативи. Методът трябва да бъде адаптиран към вида на артефакта, без да се губи общата логика на проследимостта.

Анализът на алтернативните обяснения е задължителен етап от идентифицирането. Чрез него се проверява дали наблюдаваният резултат и предполагаемите процесни ефекти имат по-просто обяснение. Те могат да произтичат от самостоятелната експертност на човека, шаблонен машинен изход, техническа или редакционна помощ, подробно предварително задание или натрупан човешки труд. Само след отхвърлянето на тези алтернативи може по-надеждно да се приеме, че резултатът е свързан с ко-креативното взаимодействие.

Необходимо е да се подчертае, че наличието на машинен текст в продукта не доказва ко-креаторство, ако системата само е формулирала вече напълно определено от човека съдържание. Също така липсата на дословно машинно съдържание не изключва ко-креаторска функция, ако машинен въпрос или критика е довел до проследима промяна на човешката концепция. Анализът трябва да оценява функционалното влияние, а не само повърхностния текстов произход.

Когато е възможно, крайният продукт трябва да се оценява независимо от информацията за начина на неговото създаване. Предварителното знание за участието на система с ИИ може да повлияе върху оценката както положително, така и отрицателно. Отделянето на процесното кодиране от резултатното оценяване ограничава риска от подобни пристрастия. То позволява и последващо съпоставяне на оценките за процеса и крайния резултат.

Следващата табл. 6 обобщава четирите аналитични области и показва различните въпроси, на които те отговарят.

Таблица 6. *Процесни критерии и генеалогия на продукта*

<i>Област</i>	<i>Показатели</i>	<i>Въпрос за анализ</i>
Процес	Рекурсивност; преобразуване; обратна връзка; резонанс; разпределение на решенията.	Как се развива взаимодействието?
Продукт	Оригиналност; обоснованост; точност; приложимост; концептуална цялост.	Какво е качеството на резултата?
Генеалогия	Приети, изменени и отхвърлени предложения; връзка между версии; произход на ключови елементи.	Как различимите приноси достигат до продукта?
Алтернативни обяснения	Самостоятелна човешка работа; шаблонен машинен изход; редакционна помощ; ефект на натрупания човешки труд.	Може ли резултатът да се обясни по-просто?

Съпоставянето на четирите области позволява да се формулира по-пълно и внимателно заключение. Процесният анализ установява дали взаимодействието има рекурсивен и преобразуващ характер. Оценката на продукта показва неговата стойност според външни критерии. Генеалогичният анализ проследява връзката между процеса и крайния резултат. Анализът на алтернативните обяснения проверява дали определянето на взаимодействието като ко-креативно е действително обосновано.

Когато процесните доказателства са силни, но продуктът е с ниско качество, трябва да се установи на кой етап е възникнал проблемът. Възможно е взаимодействието да е разширило прекомерно кръга от варианти, без да

доведе до подходящ избор. Причината може да бъде и недостатъчната човешка проверка или липсата на ясни критерии за качество. Това не отрича ко-креативния характер на процеса, но намалява практическата му стойност.

Възможна е и обратната ситуация – висококачествен продукт при слаби доказателства за ко-креативен процес. Тогава е по-точно да се говори за ефективно инструментално подпомагане. Изкуственият интелект може да е допринесъл за по-бързата работа, генерирането на варианти или езиковото оформяне. Това обаче не означава непременно, че е участвал в рекурсивно преобразуване на задачата. Ползността на системата не бива да се подценява, но и не е достатъчно основание взаимодействието да се определя като ко-креативно.

Генеалогията може да разкрие, че ко-креативността не е разпределена равномерно в рамките на продукта. Някои негови части могат да са резултат от интензивно взаимодействие и преобразуване. Други може да са създадени изцяло от човека, а трети да включват машинен изход без съществена преработка. В такива случаи продуктът може да се определи като смесен. Необходимо е обаче ясно да се посочат конкретните области и функции на машинния принос.

Неуспешните и отхвърлените предложения също са част от генеалогията на продукта. Тяхната функция може да бъде отрицателна, неутрална или продуктивна. Нерелевантният вариант понякога забавя работата. В други случаи обаче той помага да се уточнят критериите или да се изясни посоката на търсене. Затова анализът не трябва да се ограничава до въпроса дали дадено предложение е включено в крайния продукт. Важно е да се проследи и влиянието му върху последващите решения.

Синергията може да се разглежда като възможен резултат от ко-креативния процес, но не и като задължително условие за неговото установяване. Тя предполага, че съвместният резултат превъзхожда резултатите, които човекът и ИИ биха могли да постигнат поотделно. Това е силно сравнително твърдение и изисква подходяща емпирична основа.

Оценката на синергията предполага сравнение между три условия: самостоятелна човешка работа, самостоятелно машинно генериране и взаимодействие между човек и ИИ. Задачата, времето, достъпната информация и критериите за оценяване трябва да бъдат еднакви. Създадените продукти следва да се оценяват независимо. Установените разлики не трябва да произтичат от неравни ресурси или различна продължителност на работата.

Сравненията с предполагаеми способности не са достатъчно надеждни. Не може да се твърди, че човекът „никога не би достигнал“ до определено решение или че ИИ „не би могъл“ да го създаде самостоятелно, ако това не

е проверено емпирично. Подобни заключения са особено несигурни при отворени творчески задачи, тъй като пространството на възможните решения не може да бъде изчерпано.

При отсъствие на реална база за сравнение защитимият извод е по-ограничен. Може да се установи, че продуктът съдържа проследимо преобразуване на различни човешки и машинни приноси; че не съвпада с нито един от първоначалните варианти; че конкретната му форма е възникнала в хода на рекурсивното взаимодействие. Това дава основание да се достигне до заключение за ко-креативен произход, но не доказва, че резултатът превъзхожда всички възможни продукти, които човекът или ИИ биха могли да създадат самостоятелно.

Синергията трябва да се разграничава от повишената ефективност. По-бързата разработка на продукта, по-ниските разходи или по-големият брой генерирани варианти могат да бъдат съществени ползи от взаимодействието. Те обаче не доказват, че крайният резултат е с по-високо качество. Затова ефективността, качеството, новостта и ко-креативният произход трябва да се оценяват като отделни измерения.

Възможна е и отрицателна синергия, при която взаимодействието с ИИ води до по-слаб резултат от този, който човекът би постигнал сам. Причините могат да бъдат различни: автоматизационно пристрастие, отслабване на доверието в собствената преценка, информационно претоварване, приемане на погрешни предложения или прекомерно време, отделено за редактиране. Анализът на подобни случаи е необходим, за да се постигне балансирана оценка на взаимодействието „човек–ИИ“.

Окончателното заключение трябва да разглежда процеса, продукта, неговата генеалогия и алтернативните обяснения като отделни, но взаимно свързани компоненти. В него следва ясно да се посочат видът на установеното взаимодействие, качеството на крайния резултат и приносите, които могат надеждно да бъдат проследени. Необходимо е да се оцени и доколко наблюдаваната конфигурация би могла да се обясни по по-прост начин. Така се избягват както безкритичното възхищение от технологията, така и необоснованото отричане на действителната функционална роля на ИИ.

8. Класификационно заключение и степен на проявление

Заключението се формулира на три равнища: тип на взаимодействието, профил на проявлението и степен на увереност. Разграничават се четири типа: инструментална употреба, асистирано създаване, резонансно взаимодействие и асиметрично ко-креаторство. Интегрираният интелект не представлява пети тип. Той е системна интерпретация, допустима само когато

резонансът, ко-креаторството, генеалогията на продукта и човешката отговорност образуват съгласувана конфигурация.

Степента на проявление се представя чрез профил, а не непременно чрез един общ индекс. Например даден епизод може да се характеризира със силен когнитивен и творчески резонанс, ограничен емоционален резонанс, ясно изразено ко-креаторство и висококачествен продукт. Към заключението се посочва ниска, умерена или висока увереност в зависимост от пълнотата, независимостта и съгласуваността на доказателствата.

Класификацията следва йерархична логика. Първо се проверява дали са изпълнени необходимите условия. След това се определят профилите на резонанса и ко-креаторската функция. Накрая се съпоставят генеалогията на продукта, отрицателните доказателства и възможните алтернативни обяснения. Високите оценки по подкрепящи показатели не могат да компенсират липсата на необходимо условие.

При непълни или противоречиви данни заключението трябва да бъде предпазливо. Не бива да се налага изкуствено нито положителна, нито отрицателна класификация. Важно е ясно да се разграничават липсата на проявление, недостатъчните данни и граничната конфигурация, тъй като всяко от тези положения има различен методологичен смисъл.

Финалният отчет трябва накратко да посочва единицата на анализ, използваните източници, изпълнените и неизпълнените критерии, профила на проявлението, основните алтернативни обяснения и степента на увереност. По този начин класификацията остава проверима и може да бъде преразгледана при появата на нови данни.

9. Надеждност и валидност

Надеждността и валидността са основни предпоставки предложената методология да се утвърди като устойчив изследователски инструмент, а не да остане само концептуално обоснована схема. Макар да са тясно свързани, те се отнасят до различни нейни качества. Надеждността показва доколко при повторно прилагане процедурата води до сходни резултати, включително когато се използва от различни оценители или при близки по характер случаи. Валидността, от своя страна, се отнася до това дали избраните понятия, индикатори и изводи действително отразяват предполагаемото проявление на интегриран интелект.

Високата надеждност сама по себе си не гарантира валидност. Двама оценители могат последователно да прилагат една и съща кодираща схема, без тя действително да измерва резонанс и ко-креаторство. Възможно е вместо това схемата да отчита общата продуктивност, продължителността

на диалога или удовлетвореността на участниците. Възможна е и обратната ситуация: критериите да са теоретично подходящи, но практически ненадеждни, защото са твърде общи, припокриват се или не позволяват ясно разграничаване между положителни, отрицателни и гранични случаи.

Проверката трябва да обхваща цялата верига от концептуалното определение до крайното заключение. Необходимо е да се установи дали теоретичните характеристики са представени в критериите, дали критериите са преведени в наблюдаеми индикатори и дали тези индикатори могат да бъдат кодирани последователно. Следва да се провери и дали получените резултати подкрепят очакваните отношения. Слабостта на която и да е връзка в тази верига ограничава обосноваността на крайния извод.

Оценката на надеждността и валидността трябва да протича поетапно. В началото основна роля имат експертната оценка, пилотното кодиране и подробният анализ на отделни случаи. При натрупването на по-голям корпус могат да се прилагат количествени процедури за установяване на надеждността, проверка на хипотези и сравнение между модели. По-сложните статистически анализи са оправдани едва когато категориите са достатъчно стабилизираны, извадката е подходяща и данните отговарят на изискванията на избрания метод.

Преди формалното валидиране трябва да се установи дали термините и инструкциите са достатъчно ясни. Оценителите трябва надеждно да разграничават когнитивния, творческия и емоционалния резонанс, както и ко-креаторската функция, креативното стимулиране и инструменталното подпомагане. Ако едни и същи примери системно се тълкуват по различен начин, причината може да бъде в неясните граници между категориите, а не непременно в недостатъчната подготовка на оценителите.

Съдържателната валидност показва доколко системата от критерии обхваща съществените характеристики на изследвания конструкт. При интегрирания интелект това предполага включването на целенасочения контекст, съдържателните приноси, рекурсивността, трите вида резонанс, асиметричното ко-креаторство, генеалогията на продукта и запазената човешка отговорност. Ако дадено измерение липсва или е представено само чрез един твърде тесен индикатор, методологията може да създаде изкривена представа за явлението (Messick, 1995).

Проверката на съдържателната валидност изисква систематично съпоставяне на концептуалната рамка с кодиращата схема. Всеки теоретичен компонент трябва да бъде свързан с един или повече наблюдаеми индикатори. Същевременно за всеки индикатор трябва да има ясно теоретично основание. За тази цел може да се изготви матрица, която показва кои ха-

рактеристики са обхванати, къде се наблюдава припокриване и кои области остават недостатъчно представени.

Експертната оценка трябва да обхваща гледните точки на специалисти от различни области. В нея могат да участват изследователи на взаимодействието човек–компютър, ИИ, когнитивните науки, творчеството, методологията и етиката. Интердисциплинарният състав е необходим, тъй като конструктът съчетава технически, психологически, процесни и нормативни измерения. Експертите следва да преценят релевантността, яснотата и необходимостта на всеки критерий, както и да установят евентуални пропуски в системата.

Конструктната валидност показва дали наблюдаваните кодове и индикатори действително отразяват заложената теоретична структура и връзките между нейните компоненти. Ако концептуалната рамка разграничава три вида резонанс, емпиричните данни трябва да позволяват тяхното аналитично разграничаване. Същевременно те не бива да се разглеждат като напълно независими явления. Ако рекурсивността и преобразуването са определени като основни условия за ко-креативно взаимодействие, по-силното им проявление би трябвало да се свързва с по-ясно проследима генеалогия на крайния продукт (Campbell & Fiske, 1959; Messick, 1995).

Проверката на конструктната валидност започва с предварително формулирани хипотези. Може например да се очаква качеството на обратната връзка да бъде свързано с когнитивния резонанс, а развитието на идеите през повече от един цикъл – с творческия резонанс. Проследимото взаимно преобразуване би следвало да разграничава ко-креативните взаимодействия от еднократната инструментална употреба. Ако очакваните връзки не се установят, трябва да се прецени дали причината се крие в теоретичната рамка, в начина на операционализация или в качеството на събраните данни.

При достатъчно голяма и разнообразна извадка могат да се приложат факторен анализ, моделиране на латентни променливи и сравнение на конкурентни модели. Чрез тях може да се провери дали индикаторите се групират според теоретично предвидените измерения и дали алтернативна структура не обяснява данните по-добре. Тези методи обаче не бива да се прилагат механично в ранните етапи, когато категориите все още имат изследователски характер, а броят на наблюдаваните случаи е ограничен.

Разграничителната (дискриминантната) валидност показва дали понятията „интегриран интелект“ и „асиметрично ко-креативно взаимодействие“ могат ясно да се разграничат от близки, но теоретично различни явления. Всяко от тях трябва да притежава собствено съдържание, отличителни характеристики и емпирични проявления, които не се свеждат до тези на сродни понятия. Особено важно е разграничаването им от инструментал-

ното асистиране, автоматизацията, креативното стимулиране, хибридната интелигентност, разпределеното познание и общото сътрудничество между човек и ИИ. Ако едни и същи критерии определят всяко полезно взаимодействие като ко-креативно, конструктът губи аналитичната си специфичност (Campbell & Fiske, 1959).

Проверката трябва да включва предварително подбрани положителни, отрицателни и гранични случаи. Положителните случаи следва да съдържат ясни доказателства за изпълнението на всички необходими условия. Отрицателните могат да обхващат еднократен машинен изход, механична редакция или автоматизирана операция без рекурсивна преработка. При граничните случаи се наблюдават отделни признаци на резонанс или креативно стимулиране, но критериите за ко-креаторство не са изпълнени изцяло.

Граничните случаи имат особена диагностична стойност. Те показват къде определенията се припокриват, кои индикатори са недостатъчно специфични и какъв минимален праг е необходим за класификационно заключение. Методология, която работи само при очевидни положителни и отрицателни примери, не е достатъчно валидирана за анализ на реални взаимодействия.

Надеждността между оценителите показва доколко кодирането е независимо от субективната преценка на конкретния анализатор. Тя е особено важна при сложни процесни категории като резонанс, преобразуване и ко-креаторска функция, при които едно и също действие може да има различно значение в зависимост от контекста. Високото съгласие предполага, че определенията и правилата са достатъчно ясни за последователно прилагане (Cohen, 1960; Krippendorff, 2018; McGraw & Wong, 1996).

Оценителите трябва първоначално да кодират материала независимо, без да съгласуват решенията си в хода на работата. Преди това е необходима обща подготовка с кодиращ наръчник, примери и пилотни упражнения. Подготовката не трябва да цели механично уеднаквяване на мненията, а споделено разбиране на категориите и процедурите за доказване.

Изборът на показател за оценка на съгласието зависи от вида на данните и изследователския дизайн. При номинални категории съгласието между двама оценители може да се оцени чрез коефициента κ на Коен (Cohen's kappa)¹⁰, а при повече оценители – чрез съответните многовариантни разширения на този подход. За ординални (порядкови)¹¹ оценки е подходящ

¹⁰ Cohen's κ е статистически показател за съгласието между двама оценители или кодиращи, като отчита и съгласието, което би могло да възникне случайно. Стойностите са: $\kappa = 1$ – пълно съгласие; $\kappa = 0$ – съгласието не е по-добро от случайното; $\kappa < 0$ – съгласието е по-слабо от очакваното по случайност.

¹¹ Порядкови оценки (ординални оценки) са оценки, при които стойностите могат да бъдат подредени по ред, но разликите между тях не са количествено равни или известни.

претегленият k , тъй като отчита степента на несъгласие между категориите. При количествени или профилни оценки може да се използва коефициентът на вътрешнокласова корелация (ICC). Процентът на съгласие може да се представи като допълнителен описателен показател, но не е достатъчен сам по себе си, тъй като не отчита вероятността част от съвпаденията да са резултат от случайност.

Статистическата стойност трябва да бъде придружена от анализ на видовете несъгласие. Възможно е общият коефициент да бъде приемлив, но оценителите системно да се разминават при разграничаването на творчески резонанс от креативно стимулиране. Подобен модел показва необходимост от прецизиране на конкретни определения, дори когато общата надеждност изглежда задоволителна.

Протоколът за несъгласия трябва да бъде предварително определен. След независимото кодиране оценителите могат да обсъдят различията и да достигнат до консенсус или да използват трети оценител. Първоначалните кодове не трябва да се заличават, защото са необходими за изчисляване на надеждността и за анализ на проблемните категории.

Проследимостта показва дали всеки код и всяко заключение могат да бъдат свързани с конкретни първични данни. Тя е условие за надеждност, проверимост и възможност за одит. Без проследимост твърденията за резонанс или ко-креаторство остават зависими от интерпретативния авторитет на анализатора.

За всеки присъден код трябва да бъдат посочени епизодът, диалогичният ход, версията на продукта, цитатът или наблюдаваната промяна, както и кратка аргументация. Полезен инструмент е създаването на матрица на доказателствата, която да свързва теоретичните условия, индикаторите, първичните данни, присъдените стойности, отрицателните доказателства и степента на увереност.

Одитната следа трябва да включва и методологичните решения: как са определени границите на случая, как са подбрани данните, как са променяни кодиращите правила и как са разрешавани несъгласията. Това позволява външен изследовател да реконструира аналитичния път от първичния материал до крайното заключение.

Рефлексивността е особено важна, когато авторът, участникът и анализаторът са едно и също лице. Тази позиция предоставя достъп до намеренията, преживяванията и логиката на решенията, които външен наблюдател трудно би реконструирал. Същевременно тя увеличава риска от потвърждаващо пристрастие, селективно запомняне, преувеличаване на машинния принос и благоприятно тълкуване на неясните случаи.

Рефлексивният дневник трябва да документира очакванията на изследователя, промените в собствената му позиция, основанията за избор и от-

хвърляне на предложения и моментите на несигурност. Важно е да се записват и решенията, които не подкрепят теоретичните очаквания. Дневникът не премахва субективността, но я прави видима и достъпна за критична проверка.

Външната проверка може да включва независимо кодиране на избрани епизоди, критичен преглед на матрицата на доказателствата или оценяване на продукта от лица, които не познават авторската интерпретация. Особено полезно е целенасоченото търсене на отрицателни случаи – епизоди, при които взаимодействието е продължително и полезно, но не изпълнява критериите за интегриран интелект.

Екологичната валидност показва доколко резултатите от изследването се отнасят до реални практики извън контролираната среда. Лабораторните задачи позволяват по-добър контрол върху времето, ресурсите, инструкциите и сравнителните условия, но могат да опростят мотивацията, организационните ограничения и продължителността на действителната работа.

Естествените случаи предоставят по-богата информация за продължителни проекти, промяна на целите, институционални изисквания, екипни роли и реални последствия. Те обаче включват повече неконтролирани фактори и често имат непълна документация. Поради това лабораторните и естествените изследвания трябва да се разглеждат като взаимно допълващи се.

Съпоставянето между контролирани и реални случаи може да покаже кои характеристики са устойчиви и кои зависят от средата. Например рекурсивното преобразуване може да бъде ясно наблюдавано и в двата контекста, докато емоционалният резонанс, разпределението на решенията и организационната отговорност могат да имат различни прояви. Подобни различия не обезсилват конструкта, а очертават границите на неговата приложимост.

Валидирането трябва да отчита технологичната изменчивост. Резултатите, получени с конкретен модел, интерфейс или версия не могат автоматично да бъдат обобщени към всички системи. Трябва да се документират моделът, версията, настройките, системните инструкции, достъпът до външни инструменти и времето на провеждане. Повторението с различни системи позволява да се разграничат общите процесни механизми от ефектите на конкретна технология.

Проверката на валидността следва да отчита и влиянието на контекста на приложение. В зависимост от характера на задачата критериите могат да функционират по различен начин при научно писане, творчески дейности, обучение, организационно планиране или техническо проектиране. Ето защо валидността е добре да бъде изследвана в разнообразни условия, включващи участници с различно равнище на експертност, различни езикови и културни среди, както и различни видове крайни продукти.

Валидирането следва да се осъществява като итеративен процес, включващ последователни цикли на проверка, анализ и усъвършенстване на методологичната рамка. Резултатите от първоначалната проверка могат да наложат прецизиране на индикаторите, разграничаване или обединяване на категории, коригиране на праговите стойности или преразглеждане на отделни теоретични предпоставки. Всички промени следва да бъдат системно документиранни, а актуализираната рамка – проверена върху независим набор от данни, а не единствено върху случаите, довели до нейното преработване.

Оценката на надеждността и валидността трябва да се извършва на няколко взаимосвързани равнища: на отделните кодове, на аналитичните измерения, на класификационните заключения и на методологията в нейната цялост. Възможно е кодирането на отделните индикатори да бъде последователно и надеждно, без това непременно да гарантира еднозначност на класификационните заключения, ако правилата за интегриране на показателите не са достатъчно ясно дефинирани. По същия начин методологията може да показва висока валидност в определен тип задачи, но ограничена обобщаемост или по-ниска екологична валидност при прилагането ѝ в различни контексти.

Окончателният отчет трябва да представя не само положителните резултати от валидирането, но и несъгласията, неуспешните проверки и ограниченията. Прозрачното документиране на слабостите увеличава научната стойност на методологията, защото показва при какви условия заключенията са защитими и къде е необходима допълнителна работа.

В следващата табл. 7 са систематизирани основните компоненти на надеждността и валидността, характерните рискове и подходите за тяхната проверка.

Таблица 7. Проверка на надеждността и валидността

<i>Компонент</i>	<i>Основен риск</i>	<i>Подход за проверка</i>
Съдържателна валидност	Критериите не покриват съществените характеристики.	Експертна оценка и съпоставка с концептуалната рамка.
Конструктна валидност	Кодовете не отразяват предполагаемите отношения.	Проверка на хипотези, факторен анализ и сравнение на модели.
Дискриминантна валидност	Ко-креаторството не се различава от асистирание или хибридна интелигентност.	Положителни, гранични и отрицателни случаи.
Надеждност между оценители	Кодирането зависи от субективната преценка.	Независими оценители, κ или ICC и протокол за несъгласия.

<i>Компонент</i>	<i>Основен риск</i>	<i>Подход за проверка</i>
Проследимост	Изводът не може да се свърже с първичните данни.	Одитна следа, версии и матрица на доказателствата.
Рефлексивност	Авторът и анализаторът са едно лице.	Рефлексивен дневник, външна проверка и отрицателни доказателства.
Екологична валидност	Контролираната задача не отразява реална работа.	Съпоставка на лабораторни и естествени случаи.

10. Бъдещо количествено и сравнително валидиране

Количественото валидиране трябва да бъде последващ, а не начален етап в развитието на методологията. В ранната фаза основната задача е да се изясни съдържанието на конструкта, да се определят неговите граници и да се установи как предполагаемите проявления могат да бъдат разпознавани в реални взаимодействия. Преждевременното конструиране на скала би създавало риск от количествено измерване на все още неясно или нестабилно понятие (Boateng et al., 2018).

Разработването на стандартизиран измервателен инструмент изисква предварително прецизиране на теоретичните понятия за когнитивния, творческия и емоционалния резонанс, асиметричното ко-креативно взаимодействие и генеалогията на продукта. Необходимо е също ясно концептуално разграничаване между асиметричното ко-креативно взаимодействие, инструменталното асистизиране, креативното стимулиране, резонаторната функция и общата полезност на системите с ИИ. При съществено припокриване между тези категории количествените показатели биха загубили еднозначната си интерпретация и теоретична валидност.

Разработването на индикаторите следва да започне с качествен анализ на разнообразен емпиричен корпус. Освен типични положителни примери, той трябва да включва отрицателни, противоречиви и гранични случаи, които поставят на изпитание разграничителната способност на предложените категории. Това позволява да бъдат идентифицирани признаците, които последователно разграничават асиметричното рекурсивно ко-креативно взаимодействие от еднократното взаимодействие или от използването на ИИ като предимно инструментално средство.

Качественото картографиране трябва да включва широк спектър от задачи и крайни продукти, за да се оцени устойчивостта на предложените индикатори в различни контексти на приложение. Аналитичната функция на отделните индикатори може да се променя в зависимост от характера на дейността, например при концептуално писане, решаване на проблеми, съз-

даване на визуални артефакти, обучение, организационно планиране или техническо проектиране. Ето защо методологията трябва ясно да разграничава универсалните процесни индикатори, които се проявяват независимо от предметната област, от специфичните индикатори, чието значение зависи от особеностите на конкретната задача или продукт (Braun & Clarke, 2006).

Следващата стъпка е съдържателното валидиране чрез интердисциплинарна експертна оценка. Експертите трябва да преценят дали предложените индикатори са релевантни, достатъчно ясни, взаимно различни и представителни за теоретичните измерения. Подходящо е да бъдат включени специалисти по ИИ, човеко–компютърно взаимодействие, когнитивни науки, творчество, психометрия, организационни изследвания и етика.

Експертната оценка трябва да установи дали предложените индикатори не въвеждат имплицитно предположение за психологическа симетрия между човека и системата с ИИ. Операционалните дефиниции следва да описват единствено наблюдаеми характеристики на взаимодействието и функционални изменения в генерираните изходи, без да приписват на системата психични състояния, намерения, разбиране, съзнание или емоционални преживявания. Запазването на подобна онтологична предпазливост е необходимо както при качествения анализ, така и при преминаването към количествено измерване.

Когнитивните интервюта представляват важен етап от валидирането на измервателния инструмент, тъй като позволяват да се провери как бъдещите оценители и участници интерпретират използваните понятия, въпроси и скални категории. Подобна проверка е особено необходима при понятия като „резонанс“, „преобразуване“, „асиметрично ко-креативно взаимодействие“ и „проследим принос“, които могат да бъдат разбирани по различен начин. Анализът на интервютата позволява да се установи дали отговорите се основават на предвиденото теоретично съдържание на измерваните конструкции или отразяват по-скоро общи оценки за удовлетвореност, полезност или положително отношение към използването на системата с ИИ.

Преди разработването на скали за самооценка или на стандартизирани измервателни инструменти е необходимо да се проведе пилотно кодиране. В него независими оценители прилагат операционалните категории към един и същ корпус от взаимодействия, без предварително съгласуване на отделните кодировъчни решения. Последващият анализ на несъгласията служи за оценка на яснотата на операционалните дефиниции, достатъчността на правилата за кодиране и наблюдаемостта на отделните индикатори. Индикаторите, които не могат да бъдат кодирани последователно от независими оценители, следва да бъдат прецизирани или изключени от окончателната методологична рамка.

Надеждността следва да бъде оценявана отделно за всяко аналитично измерение и за всяка категория в методологичната рамка. Високото общо съгласие между оценителите може да прикрива системни затруднения при разграничаването на концептуално близки категории, например между творческия резонанс и креативното стимулиране или между асиметричното ко-креативно взаимодействие и продължителното инструментално асистирание. Поради това, наред с общите показатели за съгласие, е необходимо да се анализират матрицата на несъгласията и типичните гранични случаи, тъй като именно те разкриват категориите, които изискват допълнително прецизиране на операционалните дефиниции и правилата за кодиране.

Надеждното (стабилното) кодиране на процеса създава необходимите предпоставки за разработването на количествени показатели, но не означава, че всички аналитични измерения трябва да бъдат операционализирани чрез една обща скала. Характерът на отделните конструкти може да изисква използването на различни методи за измерване – експертно кодиране на взаимодействието, анализ на последователните версии на продукта или самоотчет на участниците. Подобен многометоден подход позволява всяко измерение да бъде оценявано чрез най-подходящия източник на данни и ограничава риска сложните характеристики на взаимодействието да бъдат сведени до резултатите от един-единствен инструмент, например въпросник за самооценка.

Подходящ подход е използването на многометоден изследователски дизайн, който комбинира процесно кодиране, самоотчети, поведенчески показатели, генеалогичен анализ и независима външна оценка на крайния продукт. Следва да се отчита, че отделните измерения могат да изискват различни източници на данни. Например емоционалната ангажираност може да бъде оценявана чрез съчетаване на самоотчет и поведенчески показатели, докато проследимото преобразуване трябва да се установява чрез анализ на взаимодействието и на последователните версии на продукта. Съгласуваността на резултатите, получени чрез различните методи, би осигурила по-силни доказателства за валидността на предложената методологична рамка.

Проверката на предложената методология изисква използването на сравнителен изследователски дизайн с най-малко три условия: (1) самостоятелна работа без използване на система с ИИ; (2) работа с ограничена инструментална подкрепа; (3) работа, основана на рекурсивно взаимодействие със системата. Това позволява да се оцени дали наблюдаваните проявления на резонанса и асиметричното ко-креативно взаимодействие произтичат от рекурсивната организация на взаимодействието, а не представляват общ ефект от използването на генеративен ИИ като помощно средство (Shadish et al., 2002).

За да бъдат сравненията интерпретируеми, експерименталните условия трябва да бъдат еквивалентни по отношение на задачата, времето за изпълнение, достъпната информация, критериите за оценяване и възможностите за редактиране. Всяко систематично различие между условията (например допълнително време, повече ресурси или различни инструкции) може да се превърне в конкуриращо обяснение на наблюдаваните резултати. Ето защо експерименталният контрол следва да минимизира влиянието на подобни фактори, така че установените различия да могат с по-голяма увереност да бъдат свързани със специфичния начин на взаимодействие, а не с особеностите на експерименталната организация.

Полезно е сравнителният дизайн да включва и условие за нерекурсивно използване на система с ИИ. При него участникът разполага с един или няколко предварително генерирани отговора. Той обаче не може да продължи взаимодействието чрез последователна обратна връзка. Съпоставянето между това условие и рекурсивното взаимодействие създава възможност да се разграничат два потенциални механизма на въздействие: (1) ефектът от първоначалното генериране на варианти и (2) ефектът от тяхното постепенно усъвършенстване чрез рекурсивно взаимодействие. Това разграничение е важно за проверката на основната хипотеза. Според нея резонансът и асиметричното ко-креативно взаимодействие се определят от рекурсивната организация на взаимодействието, а не само от достъпа до генеративен модел.

Подходящият изследователски дизайн се определя от целите на изследването и проверяваните хипотези. Може да се използва междугрупов, вътрешногрупов или смесен дизайн. Всеки от тях има свои предимства и ограничения. Междугруповият дизайн ограничава пренасянето на знания, стратегии и опит между условията. Той обаче изисква повече участници, за да осигури сравнима статистическа прецизност. Вътрешногруповият дизайн сравнява различните условия при едни и същи участници. Така се намалява влиянието на индивидуалните различия. Необходим е обаче ефективен контрол върху реда на условията, пренасянето на опит и умората. Това може да се постигне чрез рандомизация или балансиране на последователността.

Рандомизацията на участниците и балансирането на задачите повишават вътрешната валидност. Те ограничават систематичните различия между експерименталните условия. Когато пълната рандомизация е невъзможна, потенциалните смущаващи променливи трябва да бъдат измерени и контролирани. Сред тях са предварителната експертност, опитът със системи с ИИ, дигиталната грамотност, мотивацията и отношението към технологиите. Тези фактори могат да повлияят върху начина на взаимодействие и

качеството на крайния продукт. Затова те трябва да бъдат включени в аналитичния модел като контролни променливи. Друга възможност е да бъдат отчетени чрез подходящи процедури за съпоставяне на групите.

Количественото валидиране трябва да съхрани разграничението между процесните характеристики на взаимодействието и резултатните характеристики на крайния продукт. Поради това зависимите променливи могат да обхващат различни измерения, като: профили на резонанса, степен на изпълнение на критериите за асиметрично ко-креативно взаимодействие, качество на крайния продукт, време за изпълнение, брой последователни версии, увереност на участника и качество на независимата човешка проверка. Всяка от тези променливи отразява различен аспект на изследваното явление и следва да бъде анализирана самостоятелно или като част от предварително обоснован измервателен модел. Обединяването им в един общ показател е оправдано само ако е подкрепено от ясни теоретични аргументи и емпирични доказателства за тяхната структурна свързаност.

Оценяването на крайния продукт следва да бъде независимо от процесното кодиране както на организационно, така и на аналитично равнище. Когато е възможно, оценителите трябва да бъдат „слепи“ по отношение на експерименталното условие, при което е създаден продуктът, за да се предотврати влиянието на предварителни очаквания или нагласи към използването на системи с ИИ. Подобна процедура намалява риска от систематично оценъчно изкривяване и позволява по-надеждно изследване на връзките между характеристиките на взаимодействието и качеството на крайния продукт.

След уточняването на операционалните индикатори започва емпиричната проверка на измервателния модел. Първата стъпка може да бъде експлораторен факторен анализ (Exploratory Factor Analysis, EFA)¹². Чрез него се установява дали индикаторите се групират в съответствие с предполагаемите теоретични измерения. Този подход е особено подходящ при разработването на нови измервателни инструменти, чиято факторна структура все още не е достатъчно емпирично обоснована (Boateng et al., 2018; Kline, 2023).

Въз основа на установената факторна структура може да се приложи потвърдителен факторен анализ (Confirmatory Factor Analysis, CFA)¹³ за

¹² EFA е статистически метод за изследване на скритата структура на набор от наблюдавани променливи или индикатори. Той показва дали тези индикатори могат да бъдат обединени в по-малък брой латентни фактори. Методът се прилага, когато предварително не е установено колко фактора съществуват и кои индикатори принадлежат към всеки от тях.

¹³ CFA е статистически метод, чрез който се проверява дали емпиричните данни съответстват на предварително зададен теоретичен измервателен модел, т.е. за проверка на съответствието между теоретично предположената и наблюдаваната структура от фактори.

проверка на предварително специфициран измервателен модел. Чрез него могат да се сравнят конкуриращи се модели. Такива са моделът с три корелиращи фактора на резонанса, моделът с един общ фактор и йерархичният модел. При последния отделните фактори се обединяват от фактор от по-висок ред, който представя общия процесен конструкт. Оценката трябва да отчита съответствието с емпиричните данни, теоретичната обосновааност, интерпретируемостта и парсимонията на моделите (Kline, 2023).

Експлораторният и потвърдителният факторен анализ се прилагат върху независими извадки или върху предварително разделени части от достатъчно голям емпиричен корпус. Обичайно факторната структура първо се изследва чрез експлораторен факторен анализ, а след това се проверява чрез потвърдителен анализ върху независими данни. Тази процедура ограничава риска от свръхнапасване. При него измервателният модел отразява специфичните особености на конкретната извадка, а не устойчивата структура на изследвания конструкт.

Надеждността на измервателния модел се оценява чрез показатели, съобразени със структурата и вида на данните. Представянето само на един коефициент за вътрешна консистентност не е достатъчно. Необходимо е да се установи дали индикаторите са достатъчно свързани, без излишно да се припокриват. Важно е също да се провери дали съответното измерение се измерва устойчиво в различни групи и контексти.

Конвергентната валидност предполага достатъчно силна връзка между индикаторите на един и същ конструкт. Причината е, че те измерват общо теоретично измерение. Например индикаторите на когнитивния резонанс трябва да отразяват промени в разбирането, концептуалната организация и когнитивното реструктуриране. Тази свързаност обаче не бива да се дължи основно на външни фактори. Такива са продължителността на взаимодействието, броят на обменените съобщения и общата удовлетвореност на участника. В противен случай индикаторите може да измерват конструкт, различен от първоначално предвидения.

Дискриминантната валидност показва дали отделните конструкти могат да бъдат емпирично разграничени един от друг и от близки, но теоретично различни явления. Например творческият резонанс следва да бъде различим от оригиналността на крайния продукт, а емоционалният резонанс – от общото доверие или положителното отношение към технологиите. По същия начин асиметричното ко-креативно взаимодействие трябва да бъде разграничимо както от инструменталното асистирание, така и от общата ефективност при използването на системи с ИИ.

Важен етап от валидирането е проверката на измервателната инвариантност¹⁴ между различни групи участници и условия на приложение. Без доказателства за инвариантност не може да се приеме, че наблюдаваните различия отразяват реални различия в измерваните конструкти и че не е различен начин на функциониране на самия измервателен инструмент. Поради това, когато данните и спецификацията на измервателния модел го позволяват, следва последователно да бъдат изследвани конфигурационната, метричната и скаларната инвариантност. Само при наличие на достатъчна степен на инвариантност могат да се правят валидни сравнения между участници с различна експертност, между различни езикови версии или между взаимодействия с различни модели на ИИ (Kline, 2023).

След установяване на приемлив измервателен модел може да се премине към моделиране със структурни уравнения (Structural Equation Modeling, SEM)¹⁵. Този подход позволява едновременно да бъдат изследвани отношенията между латентните конструкти и наблюдаваните показатели, както и преките и косвените връзки между характеристиките на взаимодействието и качеството на крайния продукт. Чрез SEM може да се изследва дали рекурсивното взаимодействие със система с ИИ влияе върху когнитивния резонанс. Може да се изследва и дали когнитивният и творческият резонанс посредничат между рекурсивната организация на взаимодействието и концептуалното качество на крайния продукт. Моделът може да провери още дали резонансът от своя страна предсказва качеството на крайния продукт. Възможно е да се анализира и непряк ефект, при който резонансът изпълнява ролята на медиатор (Kline, 2023).

Структурното моделиране само по себе си не доказва причинно-следствени връзки. За подобна интерпретация са необходими теоретична обосновка и ясно установена времева последователност между променливите. Важно е да се контролират и възможните алтернативни обяснения. Когато е възможно, следва да се използва експериментален или друг дизайн, който позволява причинни изводи. При напречни изследвания, основани на само-

¹⁴ Измервателната инвариантност показва дали даден измервателен инструмент запазва еднаква структура и значение на измерваните конструкти при различни групи участници, условия на приложение или контексти на измерване. Наличието на измервателна инвариантност е предпоставка за валидно сравняване на резултатите между тези групи.

¹⁵ SEM е статистически подход за изследване на сложни зависимости между наблюдавани променливи и латентни конструкти. Латентните конструкти не се измерват пряко, а чрез набор от индикатори. Например чрез SEM може да се изследва дали рекурсивното взаимодействие със система с ИИ влияе върху когнитивния резонанс. Моделът може да провери и дали резонансът от своя страна предсказва качеството на крайния продукт. Възможно е да се анализира и непряк ефект, при който резонансът изпълнява ролята на медиатор.

отчет, връзките между променливите трябва да се разглеждат като асоциативни, а не като причинно-следствени.

Надлъжните изследователски дизайни са особено подходящи за изучаването на интегрирания интелект, тъй като той се развива във времето. Чрез тях може да се проследи как участниците усъвършенстват уменията си с натрупването на опит. Сред тези умения са: формулирането на задания, критичната оценка на машинните предложения, ко-креативното взаимодействие. Тези дизайни позволяват също така да се установи дали различните форми на резонанс се засилват или отслабват. Те могат да разкрият и постепенно възникващи неблагоприятни ефекти. Такива са пристрастието към автоматизацията, прекомерната зависимост от ИИ и отслабването на способността за самостоятелно изпълнение на задачи (Shadish et al., 2002).

Подходящо е да се проследяват и промените във функционалната роля на ИИ във времето. В началото системата може да се използва предимно като инструмент, а с натрупването на опит да придобие резонаторна или ко-креаторска функция. Възможна е и обратната динамика, при която първоначално интензивното и продуктивно взаимодействие постепенно се превръща в рутинно и шаблонно.

Многосайтовите изследвания са необходими за оценяване на външната валидност. Те позволяват методологията да бъде проверена в различни институции, професионални области, културни среди и технологични условия. Съгласуваните протоколи улесняват сравнението между отделните контексти, но трябва да запазват достатъчна гъвкавост за анализ на съществения различия между тях.

Необходимо е да се изследват и различни модели, версии и интерфейси. Технологичните характеристики могат да влияят върху дължината на достъпния контекст, последователността на отговорите, мултимодалните възможности и адаптирането към обратната връзка. Поради това резултат, установен при работа с конкретна система, не може автоматично да бъде обобщен за всички генеративни модели.

Външната валидност изисква разнообразие и по отношение на задачите. Структурираните задачи с проверим отговор позволяват по-надеждно оценяване на точността, докато отворените задачи са по-подходящи за изследване на творческия резонанс и преобразуването на пространството на възможните решения. Съчетаването на двата типа задачи позволява да се установи кои механизми са общи и кои зависят от спецификата на задачата.

Количествената изследователска програма трябва да включва предварително планиране на извадката и анализ на статистическата мощност. Необходимият размер на извадката зависи от броя на индикаторите, сложността на модела, очакваната големина на ефектите, вложената структура

на данните и броя на сравняваните групи. Поради това не следва да се определя универсален минимален брой наблюдения без конкретен изследователски дизайн и предварителен анализ.

При данни с йерархична структура (например диалогични ходове, вложени в цикли, цикли в епизоди и епизоди в участници) трябва да се използват подходящи статистически модели, предназначени за анализ на йерархично организирани данни (многоравнищни модели). Третирането на всички диалогични ходове като независими наблюдения би довело до подценяване на неопределеността и до необосновано завишаване на статистическата значимост.

Липсващите данни, прекъснатите диалози и непълните версии трябва да се анализират систематично. Липсата на данни може да не бъде случайна, тъй като участниците могат да прекратят взаимодействието именно когато го възприемат като неуспешно или неудовлетворително. Изключването на подобни случаи би могло да доведе до надценяване на ефективността и устойчивостта на ко-креативния процес.

Предварителната регистрация на хипотезите, основните показатели и аналитичните решения би повишила прозрачността на изследването. Потвърдителните анализи трябва ясно да се разграничават от изследователските. Промените, направени след прегледа на данните, следва да бъдат изрично обозначени и проверени в последващо независимо изследване.

Откритото споделяне на кодиращите наръчници, анонимизираните примери, аналитичните скриптове и измервателните модели би подпомогнало възпроизводимостта. Пълните диалози обаче могат да съдържат лични данни, защитено съдържание или друга чувствителна информация. Поради това прозрачността трябва да се съчетава с информирано съгласие, надеждна анонимизация и подходящ контрол на достъпа.

Количественото валидиране не бива да заличава сложността на конструкта чрез преждевременното му свеждане до един общ индекс. Профилът на отделните видове резонанс и ко-креаторските критерии може да бъде по-информативен от единствена сумарна стойност. Ако все пак бъде предложен общ индекс, неговата структура, тегла и прагове трябва да бъдат теоретично обосновани и независимо валидирани.

Развитието на методологията трябва да протича итеративно. Резултатите от качествените изследвания информират разработването на индикаторите, а пилотното кодиране разкрива неясноти и проблеми в определенията. Измервателните анализи проверяват предполагаемата структура, докато сравнителните и надлъжните изследвания оценяват устойчивостта на установените отношения и тяхната обобщимост. При несъответствие с

емпиричните данни теоретичната рамка и изследователските инструменти трябва да бъдат прецизирани, а при необходимост – частично отхвърлени.

Тази последователност ограничава риска статистическата сложност да се превърне в заместител на концептуалната яснота. Потвърдителният факторен анализ и моделирането със структурни уравнения са подходящи едва когато категориите са достатъчно стабилизиращи, индикаторите са съдържателно валидирани, кодирането е надеждно и е натрупан достатъчен обем качествени данни.

Следващата табл. 8 представя поетапна програма за преминаване от качествено картографиране на явленията към сравнителна, измервателна, структурна и външна проверка.

Таблица 8. Програма за бъдещо емпирично валидиране

<i>Етап</i>	<i>Цел</i>	<i>Подход</i>
Качествено картографиране	Извеждане и прецизиране на индикаторите.	Контрастни случаи, тематичен и процесен анализ.
Експертна оценка	Съдържателна валидност и яснота.	Интердисциплинарен панел и когнитивни интервюта.
Пилотно кодиране	Приложимост и надеждност на кодовете.	Независими оценители и анализ на несъгласията.
Сравнителен дизайн	Разграничаване на инструментална, асистирана и ко-креативна работа.	Условия без ИИ, с еднократна помощ и с рекурсивен диалог.
Измервателен модел	Структура на латентните измерения.	EFA/CFA, надеждност, конвергентна и дискриминантна валидност.
Структурна проверка	Отношения между диалог, резонанс, ко-креаторство и резултат.	SEM или други подходящи причинно-аналитични модели.
Външна валидност	Обобщеност между участници, модели, задачи и контексти.	Многосайтови и надлъжни изследвания.

11. Методологичен извод

Предложената методология превръща теорията за интегрирания интелект в проверима процедура за анализ на конкретни взаимодействия между човек и ИИ. Тя обединява: онтологичното определяне на явленията, логическите условия за неговото разпознаване, операционните индикатори, про-

филното измерване, емпиричното валидиране в обща архитектура, която подлежи на рекурсивно усъвършенстване.

Основната единица на анализ е ко-креативният епизод, изследван чрез диалогичните ходове, рекурсивните цикли и генеалогията на крайния продукт. Методологичното ядро е проследимото преобразуване: как човешките и машинно генерираните приноси изменят последващите действия, как се проявяват когнитивният, творческият и емоционалният резонанс, по какъв начин тези процеси участват във формирането на резултата.

Асиметрично ко-креаторство може да се приеме само когато приносът на ИИ е различим, включен е в итеративно преобразуване, влияе върху пространството на възможните решения и има проследима и доказана роля в крайния продукт. При частично изпълнение на тези условия е необходима по-предпазлива класификация, например като инструментално асистирание, креативно стимулиране, резонаторна или смесена функция.

Процесът и продуктът трябва да се оценяват отделно. Качественият резултат сам по себе си не доказва ко-креаторство, както и ко-креативният процес не гарантира високо качество. Генеалогията на продукта и анализът на алтернативните обяснения свързват двете равнища и ограничават приписването на ефекти, които могат по-убедително да се обяснят с човешката експертност, шаблонния машинен изход или обикновената редакционна помощ.

Класификационното заключение трябва да обединява типа на взаимодействието, профила на проявлението и степента на увереност, като отчита положителните, отрицателните, граничните и противоречащите доказателства. Надеждността, валидността, проследимостта и рефлексивността са условия за допустимост на извода. Количественото моделиране е оправдано едва след концептуалното стабилизиране и пилотното валидиране на категориите.

Методологията запазва онтологичната и нормативната асиметрия между участниците. Тя допуска реален функционален принос на ИИ, без да му приписва човекоподобно съзнание, авторство или морална отговорност. Човекът определя или утвърждава целите, проверява резултата и носи отговорност за неговото използване.

Следователно интегрираният интелект е вероятно системно проявление, което трябва да бъде обосновано във всеки конкретен случай и не е предварително предполагаемо качество на всяко полезно взаимодействие между човек и ИИ. Методологичната стойност на предложената концептуална рамка се състои именно във възможността това твърдение да бъде потвърдено, ограничено или отхвърлено въз основа на прозрачни, проследими и възпроизводими доказателства.

ЛИТЕРАТУРЕН ОБЗОР ЗА ИНТЕГРИРАН ИНТЕЛЕКТ

Настоящият литературен обзор проследява развитието на идеята за „интегриран интелект“. Той посочва и празнината в научните изследвания, която предложената рамка се стреми да запълни. Интегрираният интелект не се представя като вече утвърдена теория. Той не се извежда пряко и от съществуващите понятия. Анализът разглежда как различни научни направления обясняват допълването между човек и ИИ, споделените процеси на мислене и човешкия контрол. Разглеждат се също помощта от генеративния ИИ, резонансът и съвместното творчество. Накрая се посочват въпросите, които все още нямат отговор.

Водещият въпрос е: какво вече обяснява литературата за взаимодействието между човека и ИИ и какво все още не обяснява относно преминаването от технологично подпомагане към когнитивен, творчески и емоционален резонанс, асиметрично ко-креаторство и съвместно сътворен резултат. Обзорът използва критичен синтез: за всяко направление се разграничават неговият принос, аналитичната му граница и значението му за авторската теория.

Водещият въпрос е какво вече обяснява литературата за взаимодействието между човека и ИИ и какво остава неясно или недоизяснено. Особено внимание се отделя на прехода от технологична помощ към когнитивен, творчески и емоционален резонанс. Разглеждат се също асиметричното ко-креаторство и съвместно създаденият резултат. Обзорът прави критичен анализ на различни научни направления. За всяко от тях се посочва какво обяснява, къде са неговите граници и как допринася за предложената тук теория.

Литературата дава основа за отделните компоненти. Начинът, по който те са свързани в последователността „резонанс → ко-креаторство → съвместно създаден продукт → възможна проява на интегриран интелект“ е авторски принос. Авторски е и подходът, който разглежда когнитивния, творческия и емоционалния резонанс като свързани, но различни процеси.

Обзорът е концептуален и критичен, а не систематичен. Подбраните източници представят основните теоретични направления, нужни за изясняване на конструкта. Те обаче не обхващат цялата налична литература.

Включени са направления, които допринасят за разбирането на поне един от четири проблема: разпределението на функциите, взаимното преобразуване, ролята на ИИ като ко-креатор или резонатор, създаването и утвърждаването на общ продукт. Това уточнение е важно. Научната празнина се определя чрез сравнение на границите на съществуващите обяснения, а не чрез твърдението, че изследвания на взаимодействието „човек–ИИ“ изобщо липсват (Snyder, 2019).

1. Симбиоза човек–компютър и увеличаване на интелекта

Licklider (1960) формулира визия за симбиоза между човека и компютъра, основана на различните и взаимнодопълващи се способности на двамата. В нея човекът определя целите, формулира хипотезите и оценява резултатите, а компютърът извършва рутинните изчислителни и информационни операции. Така те могат съвместно да решават проблеми, да вземат решения и да управляват сложни ситуации. Историческото значение на тази постановка може да се види в преместването на вниманието от избора „човек или машина“ към координацията на техните различни способности в общ процес. Engelbart (1962) развива сродна, но изрично насочена към увеличаване на човешките способности перспектива. Според него техническите средства, заедно с езика и методите на работа, могат да подобрят способността на човека да разбира сложни ситуации, да структурира понятия и информация, да намира решения и да ги прилага.

Тези подходи поставят в основата допълването на функциите и съвместната работа между човека и технологията. Те обаче остават на равнището на общи идеи. Не обясняват достатъчно ясно кога приносът на технологията става ко-креативен. Не показват и как този принос се развива чрез повтарящ се диалог и как се проследява в крайния продукт. Затова тези подходи са важна историческа основа, но не предлагат цялостна теория на интегрирания интелект.

Развитието от ранното взаимодействие между човека и компютъра до съвременната работа с генеративен ИИ променя както техническите възможности, така и начина на анализ. В ранните системи компютърът подпомага главно изчисленията и работата с информация (Licklider, 1960; Engelbart, 1962). Съвременните генеративни модели могат да създават съдържание и да предлагат варианти, които човекът оценява и преработва (Bommasani et al., 2021). Така взаимодействието може да премине през няколко последователни цикъла. В предложената тук рамка това налага да се проследява как приносът на човека и ИИ се променя в хода на работата и как участва в крайния резултат.

2. Разпределено, разширено и съвместно познание

Теорията на разпределеното познание измества вниманието от мисленето на отделния човек към познавателната дейност на цялата система, която включва хората, техническите средства, информацията и средата. Тя включва хората, техническите средства, начините за представяне на информацията и средата. Hutchins (1995), както и Hollan et al. (2000), показват, че познавателните процеси могат да бъдат разпределени между хората, използваните от тях средства и средата. Общото изпълнение зависи от връзките и координацията между тези части. Тази гледна точка може да се приложи и към взаимодействието между човека и ИИ, като двамата участници във взаимодействието се разглеждат като части на обща система. Това приложение обаче се отнася до разпределението и координацията на функциите. То не означава, че човекът и ИИ имат еднакъв статут. В предложената тук рамка ИИ не се приема за мислещ и морално отговорен участник като човека.

Clark и Chalmers (1998) приемат, че външни средства могат да участват в мисловния процес. Те дават примери с бележник, джобен калкулатор, компютър, химикал и хартия. Основният им пример е бележникът на Ото, който изпълнява функция, подобна на паметта. Според авторите външното средство може да стане част от познавателния процес, когато е лесно достъпно, използва се постоянно и информацията от него пряко насочва действията на човека. Съвременен подобен пример е информацията, която човек съхранява в телефона си и редовно използва вместо запомнена информация. Hollnagel и Woods (2005) предлагат човекът и използваната технология да се разглеждат като части на обща познавателна система. В нея те изпълняват свързани функции и заедно поддържат изпълнението на дейността. Резултатът зависи не само от възможностите на човека и технологията поотделно, но и от координацията между тях. Например при работа с ИИ системата може да обработва информация и да предлага варианти, а човекът да ги оценява и да взема решение. Този пример прилага идеята на авторите към съвременния ИИ, но не е взет пряко от тях. В по-широк план тази теория подкрепя разбирането, че познавателната дейност може да се изследва на равнището на общата система, а не само на отделния човек.

Разпределението на функциите не означава непременно тяхното свързване в единен процес. Не всяка система от човек, данни и технически средства проявява интегриран интелект. Нужно е да се обясни как различните приноси се променят взаимно в поредица от цикли. Трябва да се проследи и как те влияят върху следващите действия и водят до съвместно създаден резултат.

3. Колективна интелигентност и организационно знание

Изследванията на колективната интелигентност показват, че способността на една група да изпълнява различни задачи не се обяснява само със способностите на отделните ѝ членове. Lévy (1997) определя колективната интелигентност като основана на знание, което е разпределено между хората, непрекъснато се развива и се координира чрез взаимодействието им. Теорията за трансактивната памет на Wegner (1987) допълва тази перспектива, като обяснява как споделеното разбиране за разпределението на експертността позволява на групата да използва знание, което никой отделен член не притежава изцяло. Hutchins (1995) разширява анализа чрез понятието за разпределено познание, при което когнитивните процеси обхващат хората, използваните инструменти и създадените представяния. Hong и Page (2004) от своя страна показват значението на разнообразието от перспективи и подходи при колективното решаване на проблеми. Malone et al. (2010) насочват вниманието към организационните механизми на колективната интелигентност, а Woolley et al. (2010) предоставят емпирични доказателства за общ фактор на колективната интелигентност.

Организационните теории разширяват анализа към процесите, чрез които знанието се създава, интегрира и използва. Grant (1996) разглежда организацията като механизъм за интегриране на специализираното знание на нейните членове. Nonaka и Takeuchi (1995) обясняват създаването на организационно знание чрез преобразувания между неявно и явно знание, а Choo (1998) свързва създаването на смисъл, знание и решения. Crossan et al. (1999) показват как ученето преминава между индивидуалното, груповото и организационното равнище, докато March (1991) поставя проблема за баланса между изследването на нови възможности и използването на вече придобитото знание. Alavi и Leidner (2001) допълват тази перспектива, като анализират как информационните системи подпомагат създаването, съхраняването, предаването и прилагането на организационно знание.

В съвременните организации системите с ИИ могат да подпомагат обработването и структурирането на информация, разпознаването на зависимости и достъпа до организационно знание, но същевременно могат да преобразуват установените практики, разпределението на експертността и критериите за валидно знание (Faraj et al., 2018; Jarrahi, 2018).

Посочените по-горе автори обясняват как на равнището на групата или организацията могат да възникнат резултати, които не се свеждат до способностите на отделния участник. Тъй като е разработена предимно за човешки колективи, прилагането ѝ към системи „човек–ИИ“ изисква да се разграничат човешката агентност и отговорност от функционалния принос на машината.

При това приложение за настоящата концептуална рамка е полезно да се разграничат три начина на съчетаване на създадените от човека и ИИ идеи, анализи, оценки и решения. Агрегирането обединява резултати, създадени поотделно от човека и ИИ. Координацията разпределя задачите между тях и съгласува последователността на действията. При интеграцията човешките и машинните приноси се интерпретират съвместно, взаимно се преработват и се свързват в цялостен резултат. Това разграничение се въвежда в настоящата рамка и не представлява пряк модел на разглежданите автори.

Организационната литература е особено важна за разбирането на интеграцията. Тя показва, че създаденият резултат не се превръща автоматично в организационно знание. Необходимо е той да бъде интерпретиран, проверен, споделен и включен в решенията, практиките и установените процедури на организацията. Следователно ИИ може да има значим принос при генерирането и преобразуването на съдържание, но хората и организациите остават отговорни за неговата оценка, приемане и използване.

4. Увеличена и хибридна интелигентност

Увеличената интелигентност разглежда ИИ като средство за разширяване на човешките способности. Нейни концептуални предшественици са идеята на Licklider (1960) за симбиоза между човека и компютъра и рамката на Engelbart (1962) за увеличаване на човешкия интелект чрез интегрирана система от хора, методи и технически средства. В настоящата рамка увеличената интелигентност запазва асиметрията на ролите: човекът определя целите, интерпретира контекста и носи отговорност, а системата осигурява функционална подкрепа.

Хибридната интелигентност е най-близкото функционално направление до разглеждания конструкт. Kamar (2016) я свързва с проектирането на системи, които съчетават човешки и машинни способности при решаването на задачи. Dellermann et al. (2019) акцентират върху взаимното допълване на тези способности в обща система, а Akata et al. (2020) определят хибридната интелигентност като съчетание, което увеличава човешките способности и позволява постигането на цели, недостижими за човека или машината поотделно. Те формулират четири основни изследователски предизвикателства: сътрудническа, адаптивна, отговорна и обяснима хибридна интелигентност.

Тези подходи обясняват взаимното допълване, разпределението на функциите и координацията при съвместно изпълнение. Те признават, че взаимодействието може да доведе до резултати, които не са достъпни за участниците поотделно, но не изследват систематично условията, при които машинният принос преобразува пространството на възможните реше-

ния, поражда рекурсивна резонансна динамика и придобива функционален статут на ко-креаторство. В този смисъл хибридната интелигентност е не-посредствена теоретична основа, но не изчерпва понятието за интегриран интелект.

5. Ориентиран към човека изкуствен интелект и асиметрия на отговорността

Човекоцентричният ИИ поставя изисквания за надеждност, безопасност, прозрачност, управляемост и ефективен човешки надзор. Shneiderman (2022) защитава системи, които съчетават висока степен на автоматизация с висока степен на човешки контрол. Parasuraman et al. (2000) показват, че човешкото участие трябва да се разглежда спрямо различните функции и равнища на автоматизация, а Amershi et al. (2019) формулират практически насоки за разбираемо, коригируемо и адаптивно взаимодействие между човека и ИИ. Santoni de Sio и van den Hoven (2018) обвързват смисления човешки контрол с чувствителността на системата към релевантните човешки основания и с възможността действията ѝ да бъдат проследени до информирани и отговорни човешки участници.

В настоящата рамка тази традиция определя нормативната граница на ко-креаторството. Признаването на значим функционален принос на ИИ не му придава академично авторство, морална отговорност или право на окончателна преценка. Човекоцентричността обаче не изключва ко-креаторската функция на системата. Тя определя асиметрията на взаимодействието: човекът формулира намерението, упражнява крайната преценка и носи отговорността, докато ИИ може да създава различни приноси, които преобразуват развитието или съдържанието на резултата.

Смисленият човешки контрол не се свежда до формалната възможност машинният резултат да бъде потвърден или отхвърлен. В предложената рамка той предполага човекът да разбира целта, възможностите и ограниченията на системата; да разполага с достатъчна информация за проверка на резултата и използваните източници; да има реална възможност да коригира или прекрати процеса; и да може обосновано да поеме отговорност за последствията. При ко-креаторството човешката проверка, интерпретация и окончателна преценка следователно са съдържателни фази на създаването, а не административно одобрение на вече приключил автоматизиран процес. Така човекоцентричната традиция предоставя както етическа граница на ко-креаторството, така и критерий за качеството на взаимодействието между човека и ИИ.

6. Симбиотично взаимодействие и функционална взаимност

Съвременните употреби на идеята за симбиоза насочват към продължителни отношения, при които начините на работа се изменят с натрупването на опит. Маскау (1990) описва отношенията между потребители и персонализируем софтуер като коадаптивен процес, а Нос (2000) свързва човешко-машинното сътрудничество с разпределяне на функциите, взаимно моделиране и координация. Jarrahi (2018) разглежда симбиозата „човек–ИИ“ при организационното вземане на решения, докато Akata et al. (2020) включват адаптивността сред основните задачи на хибридната интелигентност. При човешко-роботни задачи Nikolaidis et al. (2017) използват „взаимна адаптация“ в по-силен технически смисъл, при който системата актуализира модел на човешката стратегия. Този смисъл обаче не може автоматично да бъде пренасян към езикови модели, които приспособяват изхода си към наличния контекст, без непременно да променят трайно параметрите си (Brown et al., 2020).

Понятията „взаимна адаптация“, „съвместно учене“ и „коеволюция“ следователно трябва да се използват внимателно. Човекът може устойчиво да изменя своето разбиране, очаквания и поведение, докато езиковият модел обикновено изменя непосредствения си изход в зависимост от подадения контекст. Bansal et al. (2019) показват, че човешките очаквания към една ИИ система могат да се формират във времето и да бъдат нарушени от последващи промени в нея. Най-защитима за отделния ко-креативен епизод остава идеята за асиметрична функционална взаимност: приносът на всеки компонент изменя условията за следващото действие на другия, но измененията имат различна природа и продължителност.

На проектно и организационно равнище могат да възникнат по-трайни процеси на взаимно оформяне. Orlikowski (2000) показва как повтарящата се употреба на технологии участва във възпроизвеждането и изменението на организационните практики, а Leonardi (2011) анализира постепенното наслагване на човешката и материалната агентност. Тези подходи позволяват да се говори за коеволюция на практики, организационни структури и развиващи се технически системи. Подобна дългосрочна коеволюция обаче не е необходимо условие за всеки отделен ко-креативен епизод. Симбиотичността е възможен контекст на развитие, а не синоним на интегриран интелект.

7. Човешка креативност и изчислителна генеративност

Изследванията на творчеството традиционно свързват креативността с оригиналността и уместността на резултата (Runco & Jaeger, 2012). Amabile

(1996) допълва продуктовите критерии, като свързва творческото изпълнение със знанията, творческите умения, мотивацията и социалната среда. От системна перспектива Csikszentmihalyi (1999) показва, че новостта придобива творчески статус чрез взаимодействието между създателя, съответната културна област и общността, която оценява резултата. Следователно оригиналността е необходимо, но невинаги достатъчно основание за приписване на креативност.

Изчислителното творчество поставя въпроса при какви критерии поведението на технически системи може да бъде описвано като творческо (Boden, 2004; Colton, 2008). Wiggins (2006) формализира творческите системи чрез пространства от възможности, правила за тяхното изследване и механизми за оценяване, а Ritchie (2007) предлага емпирични критерии за оценяване на създаваните от програми продукти. Тези подходи позволяват да се анализира машинното поведение, без предварително да се приема, че системата притежава човешко намерение или разбиране.

Генеративните модели усложняват разграничението, защото произвеждат текстове, изображения и комбинации, които могат да бъдат нови на равнището на конкретния изход. Генеративната новост обаче не е тъждествена нито с човешката креативност, нито с автономна творческа агентност. Човекът формира намерение, интерпретира значението и контекста, оценява уместността и носи отговорност, докато системата преобразува зададените условия и разширява пространството на възможните варианти чрез статистически и изчислителни операции.

Това разграничение не обезсилва машинния принос, а позволява той да бъде определен по-точно като ко-креативна функция. Davis et al. (2013) разглеждат човешко-компютърната ко-креативност като съвместен процес, а Kantosalo и Toivonen (2016) показват, че формите на сътрудничество могат да се разграничават според инициативата и разпределението на приносите. Така ко-креаторството може да бъде функционално реално, без от него непременно да следват равностойно авторство, идентична творческа агентност или еднаква отговорност.

8. Креативно стимулиране и творчески резонанс

Литературата за съвместното и изчислителното творчество показва, че ИИ може да стимулира човешката креативност чрез генериране на варианти, комбиниране на идеи и предлагане на алтернативни перспективи (Lubart, 2005; Shneiderman, 2007; Davis et al., 2015). Това въздействие може да бъде еднократно и да остане на равнището на креативно стимулиране. При ко-креативните системи обаче взаимодействието може да придобие

итеративен характер, при който предложенията на човека и машината се редуват и взаимно влияят (Yannakakis et al., 2014; Kantosalo & Toivonen, 2016). Изследването на Davis et al. (2015b) показва конкретно как човешките действия и оценки могат да бъдат използвани за изменение на следващите творчески действия на системата.

Творческият резонанс обозначава по-силна релационна динамика. Той възниква, когато човешката интерпретация не просто избира машинно предложение, а го преобразува и го връща в диалога като нова цел, въпрос, критерий или ограничение. Така преобразуваният принос изменя последващото генериране на ИИ, което от своя страна предоставя нов материал за човешка интерпретация. Подобна процесуална логика се открива в концепциите за разпределена креативност (Sawyer & DeZutter, 2009), рефлексивен разговор със ситуацията (Schön, 1983) и партиципативно създаване на смисъл (Fuchs & De Jaegher, 2009). Резонансът следователно не означава само взаимно усиляване, а рекурсивно и взаимно обусловено преобразуване на творческите приноси.

Конкретното позициониране на творческия резонанс като механизъм между креативното стимулиране и ко-креаторството (креативно стимулиране → творчески резонанс → ко-креаторство) представлява авторски теоретичен синтез. Съществуващата литература обосновава неговите компоненти, а именно обратна връзка, редуване на инициативата, преобразуване на предложенията и възникване на резултата от взаимодействието, но не ги организира непременно в предложената последователност.

9. Когнитивен резонанс и преобразуване на разбирането

Когнитивният резонанс може да бъде теоретично свързан с разпределеното познание (Hutchins, 1995; Hollan et al., 2000), разширения ум (Clark & Chalmers, 1998), разпределената интелигентност (Pea, 1993), съвместното конструиране на смисъл (Mercer, 2000; Stahl, 2006) и рефлексивната практика (Schön, 1983). Общото между тези подходи е разбирането, че мисленето не се осъществява непременно само във вътрешното пространство на отделния индивид, а може да се развива чрез взаимодействие с други участници, символни структури и материални или цифрови артефакти. Външните представяния могат да стабилизират идеи, да направят видими скрити предпоставки и да открият отношения, които преди това не са били забелязани.

В диалога „човек–ИИ“ когнитивният резонанс възниква, когато машинно генерираният принос не само предоставя информация, а участва в преосмислянето на проблема: поражда ново разграничение, поставя под въпрос допускане или подпомага преструктурирането на концептуалния

модел. Последващата човешка реакция изменя контекста, целите или критериите, а чрез тях и следващия изход на системата. Понятието обозначава тази рекурсивна функционална динамика. То не предполага, че човекът и ИИ притежават еднакви когнитивни процеси, разбиране или епистемична агентност.

Когнитивният резонанс трябва да бъде разграничен както от получаването на допълнителна информация, така и от езиковото преформулиране. Разликата между подпомагане със средство и по-трайна промяна у човека е подчертана от Salomon et al. (1991), а литературата за концептуалната промяна свързва ученето с реорганизация на съществуващите схващания (Posner et al., 1982). Сходно, разграничението между едноконтурно и двуконтурно учене показва, че по-дълбока промяна настъпва, когато се ревизират управляващите допускания и критерии, а не само конкретното действие (Argyris & Schön, 1978).

Следователно по-силният критерий за когнитивен резонанс е проследима промяна в концептуалната организация: нова връзка между понятия, ревизирано допускане, преформулиран проблем или изменен критерий за решение. Емпиричното основание трябва да се търси в последователността на взаимодействието „машинен принос, човешка концептуална промяна и последващ машинен изход“, обусловен от променения контекст. Така когнитивният резонанс се определя чрез процесни свидетелства, а не само чрез субективното усещане за полезност.

Понятието „когнитивен резонанс“, неговата рекурсивна структура и процесните критерии за установяването му представляват авторска концептуализация, изградена чрез теоретичен синтез на подходите към разпределеното познание, диалогичното учене, концептуалната промяна и рефлексивната практика.

10. Емоционален резонанс и комуникативна адаптация

Афективните изчисления изследват техническото разпознаване, моделиране и отразяване на емоционални сигнали (Picard, 1997; Calvo & D'Mello, 2010). Изследванията на човешките реакции към комуникиращи технологии показват, че хората могат да прилагат социални и емоционални очаквания към системи, които използват език и адаптивно поведение, включително когато съзнават техния технологичен характер (Nass & Moon, 2000; Nass & Brave, 2005). Емоционалният дизайн допълнително показва влиянието на преживяването върху ангажираността и употребата (Norman, 2004), а изследванията на релационните и емпатично реагиращите агенти свързват социално-емоционалното поведение на системата с възприемана-

та подкрепа, доверието и устойчивостта на взаимодействието (Brave et al., 2005; Bickmore & Picard, 2005).

Тези направления позволяват емоционалният резонанс да бъде определен асиметрично. Човекът преживява разбиране, доверие, близост, мотивация или вдъхновение. Изкуственият интелект участва функционално чрез разпознаване или извеждане на емоционално-смислови сигнали и чрез адаптиране на тона, структурата и съдържанието на отговора. Неговият отговор може на свой ред да промени човешкото преживяване и готовността за по-задълбочен диалог. Тази функционална взаимозависимост не предполага симетрия между човешкото преживяване и машинната обработка.

Емоционалният резонанс не доказва наличие на машинни чувства, съзнание или преживяване. Той е свойство на комуникативната динамика и на човешкото преживяване в нея. При определени условия афективната динамика може да подпомогне ангажираността, устойчивостта при затруднение и когнитивната откритост (D'Mello & Graesser, 2012), но това влияние трябва да бъде установявано емпирично, а не да се приема като автоматичен резултат от подкрепящия тон.

Емоционалният резонанс има двойствена роля. Подкрепящата комуникация може да увеличи мотивацията и готовността за изследване на алтернативи, но същите характеристики могат да засилят антропоморфизацията и приписването на човешки качества на системата (Epley et al., 2007). Те могат също да породят прекомерно или недостатъчно калибрирано доверие, при което убедителността на взаимодействието замества оценката на компетентността и надеждността на системата (Parasuraman & Riley, 1997; Lee & See, 2004). Поради това емоционалната ангажираност не е самостоятелно доказателство за интегриран интелект. Нейното значение трябва да се оценява чрез връзката ѝ с критичната проверка, когнитивното преобразуване и качеството на взетите решения.

Определянето на емоционалния резонанс като асиметрична и рекурсивна комуникативна динамика, която може едновременно да подпомага взаимодействието и да поражда рискове от антропоморфизация и прекомерно доверие, представлява авторска концептуализация, изградена чрез синтез на изследванията върху афективните изчисления, емоционалния дизайн и човешкото взаимодействие с технологиите.

11. Ко-креаторство „човек–изкуствен интелект“

Изследванията на съвместното творчество и сътрудничеството „човек–ИИ“ все по-често разглеждат техническите системи не само като инструменти, а като функционални участници, чиито предложения могат да про-

менят процеса и резултата от създаването (Lubart, 2005; Davis et al., 2015; Kantosalo & Toivonen, 2016). Подходите към смесената инициатива показват, че приносът може да се развива чрез последователни човешки и машинни ходове, при които предложенията се оценяват, преобразуват и връщат обратно в творческия процес (Yannakakis et al., 2014; Karimi et al., 2018). Seeber et al. (2020) разширяват тази перспектива, като поставят въпроса за машините като съотборници. Понятията „партньор“, „участник“ и „съотборник“ обаче следва да се използват функционално и да не се приемат като доказателство за човешкоподобна субектност, съзнание или морална агентност.

В настоящия модел ко-креаторството се определя чрез различим и проследим принос, участие в итеративното преобразуване на предложенията, влияние върху пространството на възможните решения и обяснително значима роля в генеалогията на крайния продукт. Това определение не изисква ИИ да притежава съзнание, собствено творческо или научно намерение и способност да носи морална отговорност.

Асиметрията не отменя ко-креаторството, а определя неговата специфична човешко–машинна форма. Човекът остава носител на намерението, смисловата интерпретация, критичната оценка и отговорността, докато ИИ участва чрез своя генеративен и преобразуващ капацитет. Следователно функционалното ко-креаторство не е тъждествено на академичното съавторство и не предоставя на системата нормативен статут на автор.

Определянето на човеко–ИИ ко-креаторството чрез проследим принос, итеративно преобразуване и влияние върху пространството на решенията, при едновременно запазване на асиметрията в намерението, субектността и отговорността, представлява авторска концептуализация, изградена чрез синтез на изследванията върху ко-креативните системи, смесената инициатива и машините като участници в съвместни процеси.

12. Емерджентност и съвместно сътворен продукт

Теориите на емерджентността анализират свойства, структури и форми на организация, които възникват от взаимодействията между компонентите и не се обясняват достатъчно чрез тяхното изолирано разглеждане или просто агрегиране (Goldstein, 1999; Holland, 1992, 1998; Wimsatt, 1997; Sawyer, 2005). Процесните концепции за емерджентността поставят акцент върху динамиката, чрез която взаимодействията пораждат нови макроскопични модели и организационни равнища (Crutchfield, 1994; Bedau, 1997; Corning, 2002; De Naan, 2006). Приложена към взаимодействието „човек–ИИ“, тази перспектива позволява да се изследва дали конкретната форма на знанието,

решението или творческия продукт възниква в рекурсивната динамика на процеса.

За да има обяснителна стойност, емерджентността трябва да бъде свързана с наблюдаеми преходи между версии, а не само с ретроспективното впечатление за новост. Необходимо е да се проследи кои елементи са внесени от човека и ИИ, как са били приети, преобразувани, комбинирани или отхвърлени и по какъв начин тези промени са обусловили следващите ходове. Подобна генеалогична реконструкция не доказва, че резултатът е принципно недостижим за участниците поотделно. Тя обосновава по-ограниченото и емпирично проверимо твърдение, че конкретната му форма е възникнала чрез наблюдаваното взаимодействие.

Определянето на емерджентността в взаимодействието „човек–ИИ“ чрез проследими преходи между версии, рекурсивно преобразуване на приносите и генеалогична реконструкция на крайния продукт представлява авторска концептуализация, изградена чрез процесна операционализация на теориите за емерджентните свойства и комплексните системи.

13. Критичен синтез и теоретична празнина

Разгледаните теоретични традиции осветляват различни, но взаимно допълващи се аспекти на изследваното проблемно поле. Симбиозата „човек–машина“ и увеличаването на интелекта формулират кооперативната ориентация на взаимодействието. Разпределеното и разширеното познание обосновават включването на техническите ресурси в по-широки когнитивни системи. Теориите за колективната и организационната интелигентност изясняват механизмите на координация, а хибридната интелигентност описва функционалното съчетаване на човешки и машинни способности. Човекоцентричният ИИ установява нормативната асиметрия между човека и системата, докато изследванията върху творчеството, афективното взаимодействие и емерджентността обясняват отделни механизми и възможни резултати от тяхното взаимодействие.

Въпреки този значителен теоретичен ресурс съществуващата литература не предлага достатъчно аргументирана рамка, която да обясни: как когнитивният, творческият и емоционалният резонанс се свързват в процес на асиметрично ко-креаторство; как човешките и машинно генерираните приноси взаимно се преобразуват чрез рекурсивни итерации; как всеки от тях оставя различима и проследима роля в генеалогията на съвместно сътворения продукт. Недостатъчно ясно остава и разграничението между функционалната ко-креаторска роля на ИИ, авторството и моралната и епистемичната отговорност за крайния резултат.

Установената теоретична празнина има три взаимно свързани измерения. Концептуалното измерение се изразява в липсата на единна връзка между трите форми на резонанс, асиметричното ко-креаторство и емерджентното системно проявление. Методологичното измерение произтича от отсъствието на достатъчно ясни и приложими критерии за разграничаване между техническо асистирание, стимулиране на човешката дейност, резонансно взаимодействие и ко-креаторска функция. Епистемично-нормативното измерение се отнася до необходимостта машинният принос да бъде разпознаван, атрибутиран и проследяван, без на системата да се приписват академично авторство, право на утвърждаване на резултата или морална отговорност.

Следователно необходимостта от нов конструкт е обоснована само ако той изпълнява едновременно три функции: интегрира различните равнища на взаимодействието, предлага операционални критерии за тяхното емпирично разграничаване и запазва нормативната асиметрия между човешкия и машинния участник. Именно в пресечната точка на тези три изисквания се разполага теоретичната необходимост от конструкта „интегриран интелект“.

14. Позициониране на интегрирания интелект

Интегрираният интелект се позиционира не като синоним на увеличената, хибридна, симбиотичната, колективната или организационната интелигентност, а като авторски метако-конструкт със собствен обяснителен фокус. Той обозначава възможно системно равнище на интелектуална организация, при което различимите човешки и машинно генерирани приноси се свързват чрез резонанс, рекурсивно преобразуване и асиметрично ко-креаторство.

Това позициониране избягва две противоположни крайности. Първата е ИИ да бъде представян като самостоятелно мислещ субект, който разбира, има собствени намерения и носи отговорност. Втората е той да бъде разглеждан само като пасивен инструмент, без да се отчита, че предложенията му могат да насочат взаимодействието в нова посока, да разширят възможните решения и да повлияят върху крайния резултат. Интегрираният интелект заема междинна позиция: ИИ не е самостоятелен носител на разбиране и отговорност, но в определени условия може да стимулира мисленето и творчеството и да допринася съществено за създаването на крайния продукт. Ролята му обаче остава различна от тази на човека, който определя целите, оценява и утвърждава резултата и носи отговорност за него.

Конструктът може да бъде разграничен на четири аналитични равнища. На онтологично равнище интегрираният интелект обозначава възможно

системно проявление, възникващо в организираното взаимодействие между човека и ИИ. На епистемологично равнище той насочва към условията, при които знанието възниква, преобразува се и получава обосновка в хода на взаимодействието. На концептуално равнище представлява авторски метаконструкт, който свързва резонанса, рекурсивността, ко-креаторството и емерджентността в обща обяснителна рамка. На методологично равнище се операционализира чрез установяване на формите на резонанс, прилагане на критерии за ко-креаторска функция, реконструиране на генеалогията на продукта и проследяване на човешкото утвърждаване и отговорност.

Това позициониране има съществено сравнително значение: наличието на интегриран интелект не може да бъде обосновано чрез един изолиран признак. Високото качество на продукта, продължителният диалог, множеството генерирани варианти или силната човешка ангажираност могат да възникнат и при обикновено инструментално използване на ИИ. Необходимо е съгласувана конфигурация от условия: целенасочено организиран контекст; рекурсивно преобразуване на човешките и машинно генерирани-те приноси; поне една ясно установима форма на резонанс; функционално ко-креаторство; проследима генеалогия на крайния продукт; и човешко утвърждаване и поемане на отговорност.

Следователно интегрираният интелект не е общо название за успешно сътрудничество между човек и ИИ, а проверима теоретична претенция за специфична организация на взаимодействието. Нейното приемане зависи както от качеството на резултата, така и от емпиричното установяване на процесите и отношенията, чрез които конкретната му форма е възникнала.

Определянето на интегрирания интелект като възможно системно проявление, възникващо чрез съгласувана конфигурация от резонанс, рекурсивно преобразуване, асиметрично ко-креаторство, проследима продуктова генеалогия и човешка отговорност, представлява авторска концептуализация, която не се свежда единствено до съществуващите теории за симбиозата „човек–машина“, разширено познание, хибридна интелигентност и ко-креативност.

15. Обобщение

Литературният обзор показва, че интегрираният интелект има разпознаваеми теоретични предпоставки, но не може да бъде сведен до механичен сбор от съществуващи понятия. Всяка от разгледаните традиции осветлява отделен аспект на изследваното явление, но същевременно има собствени обяснителни граници. Именно пресечната точка на тези граници обосновава необходимостта от формулирането на новия конструкт.

Авторският принос не се състои в твърдението, че резонансът, ко-креаторството или емерджентността са напълно непознати явления, а в специфичното им концептуално свързване. Когнитивният, творческият и емоционалният резонанс се разглеждат като механизми за взаимно обусловено преобразуване на човешките и машинно генерираните приноси. Асиметричното ко-креаторство се разглежда като форма на организираното им участие. Съвместно сътвореният продукт – като проследим резултат от взаимодействието. Интегрираният интелект се разглежда като възможно системно проявление на цялостната конфигурация.

Литературният обзор очертава своеобразна програма за последващо емпирично изследване. Тя следва да предвиди: сравняване на самостоятелната, инструментално подпомогнатата и рекурсивно ко-креативната работа; проследяване на измененията между последователните версии на продукта; разграничаване на показателите за процеса от показателите за резултата; проверяване на надеждността на предложените критерии при различни участници, задачи и модели на ИИ. По този начин литературният синтез не се ограничава до установяването на теоретична празнина, а формулира и условията, при които конструктът „интегриран интелект“ може да получи емпирична подкрепа, да бъдат очертани границите на неговата приложимост или да бъде отхвърлен. Така понятието се представя не като предварително приета даденост, а като проверяема теоретична претенция.

Таблица 9. *Критичен синтез на изследователските традиции*

<i>Традиция</i>	<i>Основни източници</i>	<i>Обяснителен принос</i>	<i>Оставаща граница</i>
Симбиоза и увеличаване	Licklider (1960); Engelbart (1962)	Функционална допълняемост и увеличаване на човешките интелектуални способности.	Не определя достатъчно ясно ко-креаторската функция и генеалогията на съвместния продукт.
Разпределено и разширено познание	Hutchins (1995); Clark и Chalmers (1998); Hollan et al. (2000)	Познанието може да обхваща хора, технически артефакти и среди.	Разпределението на познавателните процеси не обяснява само по себе си тяхната рекурсивна интеграция.
Колективна и организационна интелигентност	Nonaka и Takeuchi (1995); Lévy (1997); Malone et al. (2010)	Обяснява координацията и институционализирането на знанието.	Запазва предимно човешки и организационен фокус и не изяснява специфичната роля на генеративния ИИ.

<i>Традиция</i>	<i>Основни източници</i>	<i>Обяснителен принос</i>	<i>Оставаща граница</i>
Хибридна интелигентност	Dellermann et al. (2019); Akata et al. (2020)	Функционално комбинирани на човешки и машинни способности.	Не винаги изяснява механизмите на резонанс, рекурсивното преобразуване и ко-креаторската функция.
Ориентиран към човека ИИ	Shneiderman (2022)	Утвърждава човешкия контрол, ценностите, надеждността и отговорността.	Нормативните условия не обясняват как възниква конкретната форма на съвместния продукт.
Съвместно и изчислително творчество	Boden (2004); Colton (2008); Davis et al. (2015a, 2015b)	Обяснява генеративността, създаването на варианти и итеративното творческо взаимодействие.	Не винаги разграничава стимулирането, резонанса, функционалното ко-креаторство и академичното авторство.
Афективно взаимодействие	Picard (1997); Norman (2004); Nass и Brave (2005)	Изследва емоционалните сигнали, човешкото преживяване и комуникативната адаптация.	Не предполага машинно преживяване; необходимо е емоционалният резонанс да се интерпретира асиметрично.
Емерджентност	Holland (1992, 1998); Crutchfield (1994); Bedau (1997); Wimsatt (1997); Goldstein (1999); Ronald et al. (1999); Corning (2002); Sawyer (2003, 2005); de Haan (2006)	Обяснява възникването на нови системни свойства, организационни структури и синергични конфигурации от динамичните взаимодействия между компонентите. Процесните подходи позволяват емерджентността да се изследва чрез проследяване на преходите между последователни състояния и равнища.	Новостта на резултата не е достатъчно доказателство за емерджентност. Необходимо е проследима генеалогия на продукта, която показва как човешките и машинните приноси са били преобразувани, комбинирани или отхвърлени. Тази операционализация не е напълно разработена в съществуващите теории.

ОСНОВНИ ПОНЯТИЯ И ОПРЕДЕЛЕНИЯ, СВЪРЗАНИ С ИНТЕГРИРАНИЯ ИНТЕЛЕКТ

В настоящата глава се представя понятийния апарат на концептуалната рамка и отношенията между участниците, средата, процесите, функционалните роли, релационните механизми, резултатите и системното проявление. Понятията не образуват линейна еволюционна последователност, а принадлежат към различни аналитични равнища и могат да се проявяват в различни конфигурации в зависимост от задачата, контекста и организацията на взаимодействието.

В рамката се разграничават три категории: утвърдени понятия; авторски употреби на понятия с установени теоретични основания; и оригинални авторски конструкти. Това разграничение не представлява оценка на тяхната научна значимост, а посочва произхода им и вида обосноваване, който изисква всяка категория. „Когнитивен резонатор“, „ко-креативно диалогово пространство“, когнитивният, творческият и емоционалният резонанс, разглеждани като единна система, „асиметрично ко-креаторство“ и конкретната конфигурация на „интегрирания интелект“ се определят като авторски конструкти или авторски концептуални синтези.

1. Архитектура на понятийния апарат

В следващата табл. 10 са представени равнищата на основния понятиен апарат и функциите им в концептуалната рамка. Подрезждането им има аналитична, а не хронологична функция. При емпиричния анализ движението между равнищата е двупосочно: от участниците и контекста към процесите и резултатите, но и обратно – от характеристиките на крайния продукт към реконструирането на ролите, решенията и преобразуванията, довели до неговото създаване. По този начин понятийният апарат служи както за систематично описание на изследвания случай, така и за проверка дали заключението за наличие на интегриран интелект е достатъчно емпирично обосновано.

Таблица 10. Равнища на понятийния апарат

<i>Равнище</i>	<i>Основни понятия</i>	<i>Функция в рамката</i>
Участници и контекст	Човешки агент; система/агент с ИИ; социотехнически контекст.	Определят компонентите, ролите и условията.
Среда и процес	Ко-креативно диалогово пространство; диалогичен ход; рекурсивен цикъл; ко-креативен епизод.	Определят къде и как се развива взаимодействието.
Функционални роли на ИИ	Когнитивен резонатор; творчески резонатор; емоционално-комуникативен модулатор; източник на креативно стимулиране.	Описват различни функции на системата в конкретния процес.
Релационни механизми	Когнитивен, творчески и емоционален резонанс; функционална взаимност.	Описват взаимно обусловеното преобразуване.
Организация на участието	Асиметрично ко-креаторство.	Определя условията за реален и проследим ко-креативен принос.
Резултати	Съвместно сътворен продукт; възникващо знание; решение; модел или артефакт.	Описват конкретните резултати на процеса.
Системно равнище	Интегриран интелект.	Описва възможното системно проявление на цялата конфигурация.

2. Човешки агент

Човешкият агент е когнитивен, социален и морален субект, способен да формулира или приема цели, да интерпретира контекста, да оценява информацията, да създава значения, да взема решения и да регулира собствените си действия (Bandura, 1991, 2001; Emirbayer & Mische, 1998). Като морален агент човекът извършва ценностни преценки и носи отговорност за своите решения, действия и техните последици. Човешки агент може да бъде както отделният индивид, така и организиран човешки колектив, когато са установими механизмите за формиране на общи цели, вземане на решения, контрол и приписване на отговорност (Bandura, 2001; List & Pettit, 2011).

В ко-креативното взаимодействие човекът не е пасивен получател на машинно генерирани резултати. Той задава или преобразува посоката на работата, интерпретира предложенията, оценява тяхната релевантност, обоснованост и истинност, приема, отхвърля или изменя отделните приноси и

утвърждава крайния продукт (Lubart, 2005; Davis et al., 2015; Shneiderman, 2022). Човешката роля включва намерение, преживяване, смислова интерпретация и ценностна преценка, както и поемане на академична, професионална или морална отговорност. Поради това участието на ИИ не отменя водещата роля на човека при определянето на целите, проверката на съдържанието, утвърждаването на резултата и поемането на отговорност за неговото използване.

3. Система, модел и агент с изкуствен интелект

Системата с ИИ е техническа система, която обработва данни и изпълнява задачи чрез един или повече модели, алгоритми и други програмни компоненти. Моделът с ИИ е изчислителен компонент на системата, чрез който могат да се генерират прогнози, класификации, текстове, изображения или други изходи. Агентът с ИИ е система или функционална конфигурация, която възприема информация за средата, избира и изпълнява действия и функционира с определена степен на автономност за постигане на зададени цели (Wooldridge & Jennings, 1995; Franklin & Graesser, 1997; ISO/IEC, 2022; Russell & Norvig, 2021).

Генеративният езиков модел не следва автоматично да се определя като автономен агент. Сам по себе си моделът генерира изходи въз основа на подадения контекст. Той може да бъде включен в агентна система, когато е свързан с памет, цели, механизми за планиране, обратна връзка и средства за извършване на действия в средата (Yao et al., 2023). В диалогична среда моделът може да изпълнява ограничена агентоподобна функция, когато поддържа контекст, генерира отговори и участва в последователност от взаимно обусловени комуникативни действия. Това функционално участие не е достатъчно основание да му се приписват самостоятелно човекоподобно намерение, разбиране, съзнание или морална отговорност (Bender et al., 2021; Shanahan, 2024).

Разграничението между система, модел и агент е важно и за определянето на единицата на анализ. Моделът обозначава изчислителния механизъм. Системата включва модела и останалите технически и организационни компоненти, като интерфейс, правила за функциониране, памет, инструменти и среда на използване. Агентната функция се отнася до начина, по който конкретната конфигурация възприема състоянието на средата, поддържа или следва цел и участва в последователност от действия. Следователно агентността се приписва на определена функционална организация на системата, а не автоматично на модела или на всеки отделен генериран от него отговор.

4. Диалог и ко-креативно диалогово пространство

Диалогът между човека и ИИ е целенасочен процес на последователен обмен, при който човешките задания, реакции и решения и машинно генерираните изходи формират условията за следващите ходове на взаимодействието. Влиянието на машинния изход не е автоматично, а възниква, когато той бъде възприет, интерпретиран или използван от човека или бъде включен от системата в последващия контекст. Диалогът може да остане предимно инструментален или да придобие рекурсивен и ко-креативен характер, когато приносите на участниците последователно предизвикват оценяване, преформулиране и преобразуване на развивания продукт (Pask, 1976; Davis et al., 2015; Amershi et al., 2019).

Ко-креативното диалогово пространство е авторски конструкт, обозначаващ релационната, информационната и комуникативната среда, в която различимите човешки и машинни приноси се представят, интерпретират, сравняват, оценяват и преобразуват. То не е физическо място, самостоятелен субект или общо съзнание, а динамично организирана конфигурация от контекст, значения, цели, ограничения, варианти, обратна връзка, решения и създавани артефакти.

По-ранното понятие „пространство за когнитивен диалог“ обозначава предимно познавателното измерение на тази среда. „Ко-креативно диалогово пространство“ е по-общото понятие, тъй като включва не само когнитивните и творческите процеси, но и емоционално-комуникативното измерение на взаимодействието. Последното се отнася до човешкото преживяване, изразените емоционални сигнали и комуникативното адаптиране на системата, без да предполага наличие на машинно преживяване.

За целите на емпиричния анализ границите на ко-креативното диалогово пространство се определят от достъпния и документиран контекст и от проследимите следи, които участват във формирането на следващите ходове. Тези следи са: текущия диалог, прикачените материали, междинните версии, използваните инструменти, зададените ограничения и човешките критерии и решения. Външна информация, която не е въведена, документирана или по друг начин установима в конкретното взаимодействие, не следва да се третира като наблюдаван елемент от неговата емпирична реконструкция. Възможното влияние на предварителното обучение, системната конфигурация и други недостъпни технически условия трябва да се отчита като ограничение на анализа. Това разграничение повишава проследимостта и възпроизводимостта на изследването и ограничава приписването на причинна роля на недокументирани фактори.

5. Диалогичен ход, рекурсивен цикъл и ко-креативен епизод

Диалогичният ход е най-малката аналитична единица на взаимодействието и представлява единичен човешки или машинно генериран принос, който може да повлияе върху последващото развитие на диалога. Ходът може да съдържа задание, уточнение, предложение, оценка, решение, машинен изход или друг функционално различим принос. Определянето му като аналитична единица се основава на функцията му в последователността, а не непременно на техническото му представяне като отделно съобщение (Sacks et al., 1974; Pask, 1976).

Рекурсивният цикъл е свързана последователност, включваща човешка цел или принос, машинно генериран изход, неговата човешка интерпретация и евентуално преобразуване и последващо задание или решение. Цикълът е рекурсивен, когато резултатът от предходните ходове се превръща в условие за следващите: поражда уточняване на целта, промяна на критериите, преработване на артефакта или избор между възможности. Следователно не всяка последователност от човешки и машинни съобщения образува рекурсивен цикъл; необходимо е функционална връзка между тях да бъде установима (Pask, 1976; Kantosalo & Toivonen, 2016).

Епизодът на взаимодействие е съдържателно обособена съвкупност от един или повече свързани ходове или рекурсивни цикли, организирани около относително устойчива цел, задача или проблемна рамка. Според характера на взаимодействието той може да бъде инструментален, резонансен или ко-креативен. Ко-креативният епизод включва един или повече рекурсивни цикли, чрез които човешките и машинните приноси са оценявани и преобразувани и които са довели до относително стабилен продукт, концептуално решение или съществена промяна на посоката.

Тези аналитични единици позволяват ко-креаторството да бъде изследвано на конкретно процесуално равнище. Един проект може да съдържа инструментални, резонансни и ко-креативни епизоди, поради което характерът на цялата работа не следва да се определя еднозначно въз основа на отделен ход или ограничен фрагмент от взаимодействието.

Началото и краят на епизода се определят по съдържателни, а не само по технически или времеви признаци. Нов епизод започва при съществена промяна на целта, критериите, обекта на работа или проблемната рамка. Епизодът приключва при относително стабилизиране на продукта или решението, прекратяване на съответната линия на взаимодействие или преминаване към съдържателно различна задача. Така един продължителен разговор може да съдържа няколко отделни епизода, а един епизод може да обхваща повече от една техническа сесия, ако целта, контекстът и генеалогията на преобразуванията остават проследими (Jordan & Henderson, 1995).

6. Функционални роли на изкуствения интелект

Тази роля на ИИ обозначава функцията, която системата изпълнява в конкретен фрагмент от взаимодействието, а не нейно постоянно, същностно или независимо свойство. Една и съща система може в различни моменти да функционира като инструмент, асистент, когнитивен резонатор или ко-креатор (табл. 11). Подобно функционално разграничаване се основава на изследванията на различните форми на човешко-компютърно творческо взаимодействие, без предложените тук роли да се приемат като пряко възпроизвеждане на една установена типология (Lubart, 2005; Kantosalo & Toivonen, 2016).

Ролите могат да се променят в рамките на един епизод и не образуват задължителна йерархия или скала на развитие. Системата може например да функционира като инструмент при извличането и подреждането на данни, като асистент при изпълнението на частично зададена задача, като когнитивен резонатор при разкриването и оспорването на допускане и като ко-креатор при проследимо участие в преобразуването на концепция, решение или артефакт.

Кодирането на ролята трябва да се основава на наблюдаемото действие на системата, начина, по който човешкият участник използва или преобразува нейния изход, и установимото влияние на този изход върху следващите ходове или върху създавания продукт. Езиковата убедителност, сложността на отговора или общото впечатление за „интелигентност“ не са достатъчни основания за приписване на по-активна роля. Един и същ машинен изход може да изпълни различна функция в зависимост от употребата му: предложение, което човекът пренебрегва, не представлява автоматично ко-креативен принос, докато кратък въпрос може да функционира като когнитивен резонатор, ако доведе до установима промяна в проблемната рамка.

Таблица 11. Функционални роли на ИИ в диалога

<i>Роля</i>	<i>Определение</i>	<i>Разграничителна характеристика</i>
Инструмент	Изпълнява ясно зададена операция или генерира искан резултат.	Посоката и критериите са предварително определени от човека.
Асистент	Подпомага работа в рамките на определена човешка цел.	Приносът улеснява, но не променя съществено пространството на решенията.
Когнитивен резонатор	Генерира приноси, които стимулират преосмисляне, разграничаване или реструктуриране.	Променя развитието на човешкото разбиране.

<i>Роля</i>	<i>Определение</i>	<i>Разграничителна характеристика</i>
Творчески резонатор	Разширява и преобразува пространството на творческите възможности.	Въвежда варианти и връзки, които пораждат нови творчески ходове.
Емоционално-комуникативен модулатор	Функционално разпознава емоционално-смислови сигнали и адаптира комуникацията.	Влияе върху човешката ангажираност, доверие и устойчивост на диалога.
Асиметричен ко-креатор	Участва с различим и проследим принос в итеративното възникване на продукта.	Приносът изменя решенията и остава реконструирем в крайния резултат.

7. Креативност и креативно стимулиране

Тук креативността се разбира като човешка способност за създаване и развиване на идеи, решения или артефакти, които са едновременно оригинални и уместни спрямо определен контекст. Генеративността на ИИ означава техническата способност на системата да произвежда варианти, комбинации и преобразувания въз основа на зависимости, усвоени от данни. Двете понятия не са тъждествени: генерирането на нов или необичаен изход не е само по себе си достатъчно основание той да бъде определен като творчески значим. Те обаче могат да взаимодействат продуктивно в човешко-машинния диалог.

Креативното стимулиране от ИИ е функционалното въздействие на машинно генерирани предложения, различия, комбинации или алтернативи върху човешкото създаване и преосмисляне на идеи. То може да възникне и при единичен машинен принос. Когато човекът интерпретира, оценява или преобразува този принос и резултатът бъде включен в следващите ходове на взаимодействието, може да възникне творчески резонанс.

При оценяването на творческата значимост трябва да се разграничават критериите новост и уместност. Необичаен машинен вариант може да разшири пространството на търсене, без да бъде приложим, докато първоначално обикновено предложение може да придобие творческа значимост чрез човешко преобразуване и контекстуализиране. Следователно новостта на единичния машинен изход не е достатъчен показател за творчески принос (Runco & Jaeger, 2012).

В настоящия модел творческият резонанс се установява процесуално: чрез проследима промяна в човешките идеи, критерии или творчески ходове и чрез развитието на идеята през последователни версии. Той не предполага непременно включване на машинния вариант в крайния про-

дукт. Когато различим машинен принос участва в рекурсивното преобразуване на продукта и остава реконструируем в неговата генеалогия, взаимодействието може да бъде кодирано като асиметрично ко-креаторство (Kantosalo & Toivonen, 2016).

8. Когнитивен резонанс

Когнитивният резонанс възниква при взаимодействието между човека и ИИ. Той се проявява, когато въпрос, предложение, критика или разграничение променя начина, по който човекът разбира и структурира разглеждания проблем. Промяната може да се изрази чрез:

- въвеждане на ново разграничение;
- преформулиране на проблема;
- откриване на противоречие;
- преразглеждане на допускане или критерий;
- установяване на нови връзки между елементите на проблема;
- изграждане, преобразуване или отхвърляне на концептуален модел;
- преоценка на възможните решения;
- формулиране на нов въпрос.

Когнитивният резонанс не изисква съгласие. Критиката и продуктивното несъгласие също могат да го предизвикат. Това се случва, когато те променят човешкото разбиране и насочват по-нататъшното развитие на диалога.

При човека резонансът се изразява в промяна на разбирането, допусканията, концептуалния модел и др. под. Но без допълнителни данни не може да се установи дали тази промяна е трайна. При ИИ не може да се установи вътрешна промяна в разбирането. Наблюдава се само функционална промяна. Системата съобразява следващите си отговори с контекста, създаден в хода на взаимодействието.

Поради това когнитивният резонанс не се определя от съдържанието или убедителността на отделния отговор на ИИ. Той се установява чрез проследяване на въздействието на отговора върху човешкото разбиране и върху следващите ходове в диалога.

9. Творчески резонанс

Творческият резонанс възниква в процеса на взаимодействие между човека и ИИ. Този резонанс се проявява, когато човешките и машинно генерираните приноси взаимно насочват развитието на творческия процес. Всеки следващ принос се основава на вече създаденото и може да промени пространството на възможните решения.

Творческият резонанс не се свежда до генерирането на оригинален вариант или до натрупването на множество идеи. Той може да се установи, когато даден принос води до преобразуване на идея, откриване на нова посока, въвеждане на различен критерий или разкриване на неочаквана връзка. Необходимо е настъпилата промяна да повлияе върху следващите ходове във взаимодействието.

Творческият резонанс може да стимулира човешката креативност, но не я замества. Човекът определя целите, интерпретира предложенията и оценява тяхната уместност. Изкуственият интелект участва чрез генериране, комбиниране и преобразуване на варианти. Конкретният творчески резултат може да възникне от рекурсивното взаимодействие между човешката преценка и машинната генеративност.

Творческият резонанс е налице, когато взаимодействието води до проследимо развитие или преобразуване на идеята. Креативният резонанс, от своя страна, обозначава първоначалното стимулиране на човешката креативност. Той може да породи нова асоциация, идея или посока, без да прерасне в последователен процес на съвместно преобразуване. Следователно креативният резонанс може да бъде начален ефект, а творческият резонанс предполага неговото развитие в рекурсивната динамика на взаимодействието.

10. Емоционален резонанс

Емоционалният резонанс възниква при взаимодействието между човека и ИИ. Той представлява асиметрична релационна динамика. У човека може да възникне усещане, че е разбран, както и доверие, близост, мотивация, вдъхновение или друга емоционална реакция. Изкуственият интелект разпознава само функционално емоционалните и смисловите сигнали в диалога. Въз основа на тях системата адаптира тона, структурата и съдържанието на следващите си отговори.

Понятието не предполага, че ИИ притежава човекоподобни чувства или субективно емоционално преживяване. Емоционалният резонанс е свойство на комуникативната динамика, а не вътрешно състояние на системата. Отговорите на ИИ могат реално да повлияят върху човешкото преживяване. Настъпилата промяна може да увеличи дълбочината, устойчивостта и творческата продуктивност на последващия диалог.

Емоционалният резонанс трябва да се разграничава от антропоморфизацията. Антропоморфизацията представлява приписване на човешки качества на нечовешки обекти или системи. Такива качества могат да бъдат чувства, намерения, съзнание, разбиране, воля и морална отговорност. При взаимодействието с ИИ антропоморфизация възниква, когато човек приема, че системата:

- действително изпитва радост, тъга или съчувствие;
- има собствени желания и намерения;
- разбира преживяванията по човешки начин;
- съзнателно се стреми да помогне или да навреди;
- притежава устойчив характер или лично отношение към събеседника.

Антропоморфизацията възниква сравнително лесно, защото ИИ използва естествен език, съобразява отговорите си с контекста и може убедително да изразява съпричастност. Това поведение обаче не доказва наличието на субективно преживяване. Системата може функционално да разпознава емоционални сигнали и да адаптира отговорите си към тях, без да изпитва човекоподобни чувства.

Следователно двете понятия обозначават различни явления. При антропоморфизацията човекът приписва вътрешни човешки състояния на ИИ. При емоционалния резонанс се изследват въздействието на комуникацията върху човека и функционалното адаптиране на системата. За наличието на такъв резонанс не е необходимо да се приема, че ИИ преживява емоции.

Необходимо е да се разграничават емоционалната ангажираност, доверието в системата и обоснованото решение за използването ѝ. Положителният тон може да насърчи участието и да поддържа диалога. Той обаче може да прикрие несигурност или да създаде впечатление за по-висока надеждност. Това крие риск от некритично приемане на машинните твърдения.

Поради това емоционалният резонанс има положителна функция само когато е съчетан с критична проверка, осъзнаване на ограниченията и правилно калибриране на доверието. Доверието в системата трябва да съответства на действителните ѝ възможности и надеждност в конкретния контекст (Lee & See, 2004).

11. Функционална взаимност

Функционалната взаимност означава, че приносът на всяка от страните във взаимодействието се формира във връзка с предходния принос на другата и може на свой ред да промени условията за следващия ѝ ход. Тя не предполага сходство между човека и ИИ по отношение на субективното преживяване, когнитивните способности и моралната отговорност.

Човекът може да променя своите цели, разбирания и решения. Системата с ИИ изменя следващите си изходи в зависимост от контекста, получената обратна връзка и новите инструкции. Тази асиметрия е конститутивна характеристика на взаимодействието „човек–ИИ“, а не недостатък, който трябва да бъде преодолян.

Функционалната взаимност трябва да се разграничава от симетричната агентност. Човешките и машинните приноси взаимно влияят върху хода на

процеса. Това обаче не означава, че човекът и ИИ притежават еднакви способности за целеполагане и разбиране. Не означава и че на ИИ може да се приписва морална отговорност, равностойна на човешката.

Понятието описва взаимната обусловеност на приносите на процесно равнище. То позволява да се изследва влиянието на ИИ, без техническата адаптивност на системата да се приравнява на човешка субектност.

12. Асиметрично ко-креаторство

Асиметричното ко-креаторство е релационен процес, при който човекът и ИИ участват във възникването на общ продукт. Техните приноси са различни по природа, но взаимно обусловени.

Изкуственият интелект изпълнява ко-креаторска функция, когато са налице четири условия. Неговият принос трябва да бъде различим. Той трябва да участва в итеративното преобразуване на създавания продукт, да влияе върху пространството на възможните решения и да има проследима роля в крайния резултат.

Ко-креаторството не е тъждествено на съавторството. Човекът остава носител на авторското намерение. На него принадлежат смисловата и ценностната преценка, окончателното утвърждаване и академичната или моралната отговорност. Изкуственият интелект може да изпълнява реална ко-креаторска функция на функционално и процесно равнище, без да придобива нормативен статут на автор.

Четирите условия (различим принос, участие в итеративното преобразуване, влияние върху пространството на възможните решения, проследима роля в крайния продукт) имат кумулативен характер, когато се прави силно заключение за наличие на ко-креаторство. Ако машинният принос е различим, но не участва в последващо преобразуване, по-точно е да се говори за стимулиране или асистиране. Ако е налице итеративно преобразуване, но ролята на ИИ в крайния продукт не може да бъде реконструирана, твърдението за ко-креаторство остава недостатъчно обосновано. Следователно граничните и отрицателните случаи са част от определението, а не отклонение от него.

13. Съвместно сътворен продукт

Съвместно сътвореният продукт може да бъде идея, текст, модел, фигура, решение, хипотеза или друг артефакт. Неговата конкретна форма възниква чрез последователното взаимодействие и преобразуване на различни човешки и машинни приноси. Продуктът не представлява механичен

сбор от тези приноси. Не е необходимо и да съдържа дословно възпроизведен машинно генериран текст.

Ко-креативният произход се установява чрез генеалогията на продукта. Тя може да бъде реконструирана чрез историята на диалога, последователните версии и авторските бележки. Значение има и картата на приетите, изменените и отхвърлените предложения. Качеството на крайния продукт е отделен въпрос. Само по себе си то не доказва наличието на ко-креаторство.

Генеалогията трябва да обхваща приетите предложения и съществени-те отхвърлени варианти. Тя трябва да отразява основанията за редакциите и промените в човешките критерии. Тези следи позволяват да се установи дали машинният принос е променил пространството на възможните решения. Те показват и дали системата само е предоставяла материал за следване на предварително определена посока.

Ето защо проследимостта е необходимо условие за обоснованото доказване на ко-креативния произход. Тя позволява да се реконструира възникването на продукта и да се определи ролята на ИИ в този процес. Поради това проследимостта е част от доказателствената основа на твърдението за ко-креаторство, а не просто придружаваща документация.

14. Възникващо знание

Възникващото знание е познавателен резултат, който се формира чрез взаимодействието и последователното преобразуване на човешки и машинни приноси. То може да приеме формата на обосновано разбиране, концептуална връзка, модел или решение. Този резултат не се съдържа в завършен вид в нито един от отделните човешки или машинни приноси. Неговото възникване може да бъде проследено до взаимодействието и последователното преобразуване на тези приноси. Подобно разбиране се основава на схващането, че знанието може да възниква чрез взаимодействие и преобразуване на различни форми на познание (Nonaka & Takeuchi, 1995).

Не всеки нов или оригинален резултат придобива статут на знание. За това са необходими човешка интерпретация, подходяща проверка и включване в релевантен смислов контекст. Изкуственият интелект може да участва във формирането на информационния и концептуалния материал. Епистемичното утвърждаване на резултата обаче остава човешка задача. В научен или професионален контекст то може да изисква и институционално признати процедури за проверка.

Начинът на проверка зависи от вида на твърдението. Концептуалните модели изискват оценка на логическата съгласуваност и обяснителната им стойност. Фактическите твърдения трябва да бъдат съпоставени с надежд-

ни, по възможност първични източници. Емпиричните изводи подлежат на методологична проверка. Приложните решения изискват експертна или практическа оценка според целите и условията на употреба.

По този начин възникващото знание се разграничава както от правдоподобно генериран текст, така и от личното впечатление, че полученият резултат е нов. За да придобие статут на знание, резултатът трябва да бъде обоснован и проверен чрез процедури, съответстващи на вида на твърдението. Неговият статут зависи не само от начина на възникване, но и от качеството на основанията, чрез които резултатът се приема за валиден.

15. Интегриран интелект

Интегрираният интелект е възникваща, релационна и динамична форма на интелектуална организация. Тя може да се прояви в социотехническа система, когато различни човешки и машинно генерирани приноси се преобразуват взаимно чрез целенасочен и рекурсивен диалог, резонанс и асиметрично ко-креаторство. В резултат възниква съвместно сътворен продукт, чиято смислова и ценностна оценка, окончателно утвърждаване и морална отговорност остават човешки.

Интегрираният интелект има три взаимно свързани статута. Концептуално той е централен авторски метаконструкт в предлаганата теоретична рамка. Онтологично обозначава възможно системно проявление на определена организация на отношенията между човека, системата с ИИ, процеса на взаимодействие и възникващия продукт. Методологично представлява аналитична конфигурация. Тя се установява чрез процесите на резонанс, критериите за асиметрично ко-креаторство, проследимата генеалогия на продукта и запазената човешка отговорност.

Интегрираният интелект не е самостоятелен субект, общо съзнание или притежание на ИИ. Той не се проявява автоматично при всяко използване на ИИ. Не трябва да се разглежда и като задължително най-висока степен на технологично развитие.

Определението следва конфигурационна логика. Отделен резонанс, еднократен оригинален резултат или висока ефективност не са достатъчни основания за установяване на интегриран интелект. За такъв извод е необходимо съвместното наличие на целенасочен и рекурсивен процес, различни и взаимно преобразувани приноси, асиметрично ко-креаторство, проследима генеалогия на продукта, човешко окончателно утвърждаване и поемане на отговорност. Степента на проявление може да варира. Минималните условия и доказателствата за тяхното изпълнение обаче трябва да бъдат посочени изрично за всеки анализиран случай.

16. Разграничаване на интегрирания интелект от сродни конструкти

Съществено за общата концептуална рамка е ясното разграничаване на интегрирания интелект от сродни конструкти. Те могат да се пресичат с него и да осигуряват теоретични основания за отделни негови характеристики. Нито един от тях обаче не обхваща цялата конфигурация от целенасочен и рекурсивен диалог, резонанс, асиметрично ко-креаторство, съвместно сътворен продукт, проследима генеалогия на неговото възникване и запазена човешка отговорност. В табл. 12 разграничението е направено според основния обяснителен фокус на всеки конструкт.

Таблица 12. *Разграничаване на интегрирания интелект от сродни конструкти*

<i>Конструкт</i>	<i>Основен фокус</i>	<i>Връзка с интегрирания интелект</i>
Увеличена интелигентност	Разширяване на човешките познавателни и практически способности чрез ИИ.	Може да осигури функционална основа за взаимодействието, но не изисква резонанс, взаимно преобразуване на приносите или ко-креаторство.
Хибридна интелигентност	Съчетаване и взаимно допълване на човешки и машинни способности при решаване на задачи.	Описва функционалната конфигурация на съвместната дейност, но не изисква непременно асиметрично ко-креаторство и проследима генеалогия на съвместно сътворен продукт.
Колективна интелигентност	Координация на приносите на множество човешки, машинни или смесени участници.	Може да бъде среда за проявление на интегриран интелект, но множествеността на участниците не е негово необходимо условие.
Организационна интелигентност	Създаване, споделяне, използване и институционализиране на знание в организацията.	Обозначава организационно равнище на анализ, в рамките на което могат да възникнат и да бъдат утвърдени прояви на интегриран интелект.
Симбиотична интелигентност	Устойчиво взаимно адаптиране и допълване на човешки и машинни компоненти във времето.	Може да представлява трайна форма на развитие на интегрирания интелект, но продължителната симбиоза не е необходима за отделен ко-креативен епизод.

<i>Конструкт</i>	<i>Основен фокус</i>	<i>Връзка с интегрирания интелект</i>
Разпределено познание	Разпределяне на познавателните процеси между хора, артефакти, практики и среда.	Осигурява теоретична основа за анализ на системната организация на познавателния процес, но не определя условия за ко-креативен произход на продукта.
Разширено познание	Функционално включване на външни средства и артефакти в познавателната дейност на човека.	Обяснява как система с ИИ може да участва в познавателния процес, но не предполага сама по себе си резонанс, ко-креаторство или съвместно сътворен продукт.
Интегриран интелект	Конфигурация от целенасочен и рекурсивен диалог, резонанс, асиметрично ко-креаторство, съвместно сътворен продукт, проследи-ма генеалогия на неговото възникване и запазена човешка отговорност.	Има собствен обяснителен фокус. Не представлява сбор, обобщаваща категория или задължителен краен етап на останалите конструкции.

Относно интегрирания интелект следва да се поясни, че описаната в табл. 12 конфигурация обединява взаимносвързани условия, чрез които той може да бъде разграничен от обикновеното използване на ИИ. Основанията, чрез които наличието му може да бъде установено са следните:

- Целенасоченият диалог задава проблема, целта и критериите на взаимодействие. Без такава насоченост размяната може да остане случайна или да се сведе до използване на системата като инструмент за еднократно генериране.
- Рекурсивният диалог включва последователно връщане към вече създаденото, неговото оценяване и преобразуване. Еднократната заявка и полученият отговор сами по себе си не са достатъчни за установяване на интегриран интелект.
- Резонансът обозначава предизвикана от взаимодействието промяна в човешкото разбиране, критериите, въпросите или посоката на работа. Той не се изчерпва със съгласие, удовлетворение или емоционален отклик.
- Асиметричното ко-креаторство означава, че различимите човешки и машинно генерирани приноси участват във формирането на продукта и се преобразуват в хода на взаимодействието, без да придобиват еднакъв статут. Човекът определя целите и смисъла, оценява предложенията, утвърждава резултата и носи отговорността за неговата употреба.

- Съвместно сътвореният продукт е резултат, чиято конкретна форма не се съдържа в завършен вид в нито един от отделните приноси, а възниква чрез тяхното последователно взаимодействие и преобразуване.
- Проследимата генеалогия на продукта позволява да се реконструира как приетите изцяло или след изменение предложения, както и съществените отхвърлени варианти, са повлияли върху неговото формиране. Тя не създава сама по себе си ко-креативния произход, а осигурява доказателствена основа за неговото установяване.
- Запазената човешка отговорност определя нормативната граница на конфигурацията. Участието на ИИ не прехвърля върху него смисловата и ценностната преценка, окончателното утвърждаване или моралната отговорност.

Компонентите на конфигурацията принадлежат към различни, но свързани аналитични равнища: диалогът, резонансът и преобразуването на приносите описват процеса; асиметричното ко-креаторство характеризира начина на създаване; съвместно сътвореният продукт е резултатът; неговата генеалогия осигурява доказателствената проследимост; човешката отговорност определя нормативното разпределение на ролите.

Тяхната връзка не е механична или строго линейна. Целенасоченият и рекурсивен диалог създава условия за резонанс и преобразуване на приносите. Възникналите промени могат от своя страна да породят нов резонанс и да пренасочат диалога. Чрез този рекурсивен процес се формира съвместно сътвореният продукт. Неговата генеалогия позволява процесът да бъде реконструиран, а запазената човешка отговорност определя кой интерпретира, оценява и окончателно утвърждава резултата и кой отговаря за неговата употреба.

17. Научен статут и теоретични основания

В следващите табл. 13 и 14 понятията са представени според статута им в концептуалната рамка и според основните теоретични и терминологични източници, върху които се изгражда понятийният апарат.

Двете таблици изпълняват различни функции. Научният статут показва дали дадено понятие е заимствано от утвърдена научна традиция, адаптирано и специфицирано за целите на настоящата рамка, или е конструирано от автора. Теоретичните и терминологичните основания показват от кои научни направления, теории и стандарти се извеждат неговите характеристики. Разграничението предотвратява представянето на авторските синтези и операционализации като вече валидирани и общоприети конструкти и очертава необходимостта от тяхната последваща теоретична и емпирична проверка.

Таблица 13. Статут на основните понятия в концептуалната рамка

Понятие	Статут в рамката	Основание
Човешки агент и човешка агентност	Утвърдени понятия, специфицирани за целите на рамката.	Социално-когнитивни теории на агентността, човешка автономия и отговорност.
Система с ИИ и агент с ИИ	Утвърдени технически и терминологични понятия.	Изследвания в областта на ИИ и многоагентните системи; терминологична стандартизация.
Диалог и взаимодействие „човек–ИИ“	Утвърдени понятия и изследователски области, адаптирани към ко-креативния процес.	HCI, Human–AI Interaction и conversation theory.
Разпределено и разширено познание	Утвърдени, но различни теоретични подходи.	Distributed Cognition и Extended Mind.
Хибридна, колективна и организационна интелигентност	Утвърдени изследователски конструкти, принадлежащи към различни равнища на анализ.	Hybrid Intelligence, Collective Intelligence и Organizational Intelligence.
Когнитивен резонатор	Авторски конструкт.	Синтез на увеличаването на човешкия интелект, разпределеното познание, съвместните когнитивни системи и взаимодействието „човек–ИИ“.
Ко-креативно диалогово пространство	Авторски конструкт.	Синтез на диалогични подходи, computational co-creativity и Human–AI Interaction.
Когнитивен, творчески и емоционален резонанс	Авторски адаптирани и специфицирани понятия с интердисциплинарни основания.	Диалогично познание, computational co-creativity, affective computing и адаптация във взаимодействието.
Асиметрично ко-креаторство	Авторски специфициран конструкт, чиято терминологична оригиналност подлежи на библиографска проверка.	Синтез на сътрудничество „човек–ИИ“, ко-креативност, човекоцентричен ИИ и теории на отговорността.
Съвместно сътворен продукт	Утвърдено общо понятие с авторска дефиниция и операционализация в рамката.	Теории на ко-креацията, процесен анализ и проследимост на човешките и машинно генерираните приноси.

<i>Понятие</i>	<i>Статут в рамката</i>	<i>Основание</i>
Интегриран интелект	Централен авторски метакопструкт.	Конфигурационен синтез на целенасочен и рекурсивен диалог, резонанс, асиметрично ко-креаторство, съвместно сътворен продукт и запазена човешка отговорност.

Таблица 14. *Основни теоретични и терминологични основания на понятийния апарат*

<i>Понятийна област</i>	<i>Термини, направления, стандарти</i>	<i>Представителни източници</i>
Човешка агентност и отговорност	Social Cognitive Theory; Human-Centered AI; human agency.	Bandura (1986, 2001); Shneiderman (2022).
Увеличаване на човешкия интелект	Intelligence Augmentation; human-computer symbiosis.	Licklider (1960); Engelbart (1962).
Система и агент с ИИ	Artificial Intelligence; intelligent agents; multi-agent systems; терминологична стандартизация.	Wooldridge and Jennings (1995); Wooldridge (2009); Russell and Norvig (2021); ISO/IEC 22989:2022.
Диалог и взаимодействие „човек-ИИ“	HCI; Human-AI Interaction; conversation theory.	Pask (1976); Dix et al. (2004); Amershi et al. (2019).
Системно, разпределено и разширено познание	Distributed Cognition; Extended Mind; Joint Cognitive Systems.	Hutchins (1995); Clark and Chalmers (1998); Hollnagel and Woods (2005).
Креативност и ко-креация	Computational Creativity; human-computer co-creativity; computational co-creation.	Boden (2004); Lubart (2005); Colton (2008); Davis et al. (след библиографска проверка).
Емоционална и социална динамика	Affective Computing; emotional design; social responses to technology.	Picard (1997); Norman (2004); Nass and Brave (2005).
Емерджентност и сложни системи	Emergence; complex adaptive systems.	Holland (1992, 1998); Goldstein (1999).
Организационно създаване на знание	Organizational knowledge creation; organizational learning.	Nonaka and Takeuchi (1995).
Хибридно и човекоцентрично взаимодействие	Hybrid Intelligence; Human-Centered AI.	Dellermann et al. (2019); Akata et al. (2020); Shneiderman (2022).

18. Конфигурационен модел

В концептуалната рамка когнитивният, творческият и емоционалният резонанс се разглеждат като взаимодействащи процесни измерения на интегрирания интелект, а асиметричното ко-креаторство – като специфичен начин за организиране на човешкото участие и участието на системата с ИИ във формирането на продукта. Нито един от разграничените сродни конструкции не представлява необходима предпоставка, задължителен предходен стадий или неизбежна стъпка към интегриран интелект. Конкретното му проявление зависи от конфигурацията и функционалната свързаност на процесите, ролите, контекстуалните условия и резултатите.

Конфигурационният модел допуска различни профили на проявление. В една конфигурация може да доминира когнитивният резонанс, а в друга – творческият. Емоционалният резонанс може да подпомага процеса, да има ограничено значение или да поражда рискове, например чрез некритично доверие, прекомерна привързаност или отслабване на човешката преценка. Ето защо трите форми на резонанс определят възможните измерения на профила, без да е необходимо всяка от тях да присъства с висока интензивност или положителен ефект във всяка конкретна конфигурация.

Поради това интегрираният интелект не се определя от максималното равнище на всяко отделно измерение, а от начина, по който поне една установима форма на резонанс се включва във функционално свързан процес на асиметрично ко-креаторство. В този процес човешките и машинно генерираните приноси не просто присъстват едновременно, а се подлагат на рекурсивно оценяване, подбор и преобразуване, чрез които се формира продуктът.

Проследимата генеалогия на продукта осигурява доказателствена основа за реконструиране и обосноваване установяване на този процес, без сама по себе си да бъде условие за фактическото му възникване. Запазената човешка отговорност определя нормативната му граница. Отделните компоненти могат да имат различна относителна тежест, но конфигурацията може да бъде определена като интегриран интелект само когато са налице:

- целенасочен и рекурсивен диалог;
- поне една установима форма на когнитивен, творчески или емоционален резонанс;
- асиметрично ко-креаторство, при което човешките и машинно генерираните приноси се оценяват, подбират и преобразуват в хода на формирането на продукта;
- продукт, чиято конкретна форма възниква чрез това взаимодействие;
- запазване у човека на смисловата и ценностната преценка, окончателното утвърждаване и отговорността за резултата и неговата употреба.

Различията в относителната сила, посоката и съчетанието на трите форми на резонанс обуславят различни профили на проявление. Проследимата генеалогия осигурява доказателственото основание за реконструиране и обосноваване на ко-креативния процес. Конкретното му проявление се модифицира от задачата, характеристиките на участниците, използваната технология, времевите условия и организационната среда.

Интегрираният интелект не изисква всички форми на резонанс да бъдат едновременно силно проявени. Определящи са наличието на задължителното процесно и нормативно ядро и функционалната свързаност между неговите компоненти, която води до възникването на съвместно формиран продукт.

19. Обобщение

Понятиятният апарат разграничава компонентите, процесите, функционалните роли, релационните механизми, организацията на участието, резултатите и системното проявление. Така се избягва смесването на понятия от различни аналитични равнища. Изкуственият интелект е технически компонент, резонаторът – функционална роля, а резонансът – процес. Ко-креаторството обозначава организацията на участието, продуктът – резултат, а интегрираният интелект – системен метаконструкт.

Логиката на рамката може да бъде обобщена по следния начин. В ко-креативното диалогово пространство човешките и машинно генерираните приноси могат да влязат в когнитивен, творчески или емоционален резонанс. В хода на взаимодействието приносите се оценяват, подбират и преобразуват рекурсивно. Когато машинно генерираният принос е различим и преобразуващ, взаимодействието придобива формата на асиметрично ко-креаторство. Проследимостта на приносите осигурява доказателствена основа за реконструиране на ко-креативния процес.

Чрез тази конфигурация може да възникне съвместно формиран продукт. Човекът запазва смисловата и ценностната преценка, окончателното утвърждаване, отговорността за резултата и неговата употреба. Когато са налице тези процесни и нормативни характеристики, конфигурацията може да бъде изследвана като възможно проявление на интегриран интелект.

ФИЛОСОФСКИ ОСНОВИ НА ИНТЕГРИРЕНИЯ ИНТЕЛЕКТ

Развитието на ИИ изисква философско преосмисляне на отношенията между интелигентност, познание, творчество, агентност и отговорност. Значителна част от съществуващите подходи разглеждат ИИ като инструмент, асистент или автономно функционираща техническа система. Тези определения са необходими, но не обхващат напълно всички форми на взаимодействие. При някои от тях машинно генерираните приноси преобразуват човешкото разбиране, разширяват пространството на творческите възможности и участват проследимо във формирането на общ продукт.

Предложената тук концептуална рамка измества аналитичния фокус към релационната и социотехническата конфигурация на взаимодействие. Това не заличава различията между човека и системата с ИИ. Напротив, теорията приема функционална, онтологична и нормативна асиметрия.

Човекът притежава намерение и субективно преживяване. Той приписва и оценява смисъл и извършва ценностна преценка. Утвърждава окончателно продукта и запазва отговорността за резултата и неговата употреба. Изкуственият интелект участва чрез аналитичен, генеративен и контекстуално преобразуващ капацитет. Това участие само по себе си не предполага съзнание, намерение или морална отговорност.

Основната философска теза е, че при определени условия в ко-креативното диалогово пространство могат да възникнат когнитивен, творчески и емоционален резонанс. Между човека и системата с ИИ може да се установи функционална взаимност при запазена асиметрия. Тя се организира като асиметрично ко-креаторство.

Емоционалният резонанс не предполага емоционално преживяване от страна на ИИ. Той обозначава начина, по който машинно генерираните отговори могат да повлияят върху човешкото емоционално състояние и динамиката на взаимодействието.

Тези процеси не създават общо съзнание или нов морален субект. Те обаче могат да образуват системна организация, чрез която възниква съвместно формиран продукт. Конкретната форма на този продукт не е предварително

завършена в независимия принос на нито един от участващите компоненти. Тя не може да бъде обяснена изчерпателно само чрез един от тях.

Философската рамка има обяснителен характер. Тя определя условията, при които взаимодействието може да бъде изследвано като интегрирано. Същевременно посочва основанията, при които подобно определение трябва да бъде отхвърлено.

Понятията „резонанс“, „ко-креаторство“ и „емерджентност“ функционират като аналитични категории. Тяхното прилагане изисква проследими доказателства за процеса и продукта. Тези характеристики не могат да бъдат предварително приписвани на всяка система „човек–ИИ“.

1. От индивидуалистично към релационно и системно разбиране

Индивидуалистичните подходи разглеждат интелигентността предимно като способност или свойство на отделния носител. Релационната перспектива не отрича индивидуалните способности. Тя насочва вниманието към начина, по който познавателният процес се изменя в зависимост от организацията на отношенията между хора, технически системи, артефакти и среда. Единица на анализа става не само отделният компонент, а цялата конфигурация и взаимодействията в нея. Тази позиция се основава на общата теория на системите, разпределеното познание и тезата за разширения ум (von Bertalanffy, 1968; Hutchins, 1995; Clark & Chalmers, 1998).

Релационността не предполага, че човекът и ИИ имат еднакъв онтологичен, функционален или нормативен статут. Тя означава, че приносът на всеки компонент може да променя условията, при които протича следващото действие в системата. Машинно генерираният принос може да преобразува човешката интерпретация, критериите за оценка или избора на действие. Последващата човешка обратна връзка от своя страна изменя контекста, който обуславя следващия машинен изход. Тази взаимна обусловеност остава асиметрична. У човека могат да настъпят устойчиви промени в разбирането, уменията или творческите намерения. Промените в машинния изход обаче обикновено са контекстуални и ограничени до конкретното взаимодействие. Те могат да се определят като устойчиви само ако системата съхранява и използва натрупания контекст в последващи взаимодействия.

Релационният анализ трябва да разграничава функционалната и причинната роля на връзките от простото съвместно присъствие на компонентите. Не е достатъчно човекът и ИИ да участват в една и съща задача. Необходимо е да се установи как конкретен принос изменя последваща интерпретация, критерий или действие и как тази промяна участва във формирането на резултата. Така системната перспектива не размива индивиду-

алните приноси. Тя ги разполага в проследима и проверяема конфигурация от зависимости.

Таблица 15. *Философски и интердисциплинарни традиции, релевантни за интегрирания интелект*

<i>Традиция</i>	<i>Основна теза</i>	<i>Принос към рамката</i>	<i>Основни автори</i>
Обща теория на системите	Свойствата и поведението на системата зависят от организацията и отношенията между компонентите.	Обосновава системния и конфигурационния анализ.	Ludwig von Bertalanffy
Емерджентизъм	При определена организация могат да възникнат системни свойства, които не се приписват поотделно на компонентите.	Осигурява основание за изследване на възможно системно проявление, без предварително да го приема.	C. D. Broad; Mario Bunge; Mark Bedau
Разпределено и разширено познание	Познавателните процеси могат при определени условия да обхващат хора, артефакти и елементи на средата.	Обосновава релационното и социотехническото равнище на анализ.	Edwin Hutchins; Andy Clark; David Chalmers
Философия на диалога и диалогизъм	Смисълът се формира и преобразува в отношения на обръщение, отговор и интерпретация.	Подкрепя разбирането за ко-креативното диалогово пространство.	Martin Buber; Mikhail Bakhtin
Прагматизъм и рефлексивна практика	Познанието се развива чрез действие, проблематизиране, обратна връзка и корекция.	Обосновава рекурсивното преобразуване на разбиранията и действията.	John Dewey; Donald Schön
Етика на човекоцентричния ИИ	Разработването и употребата на ИИ трябва да подкрепят човешката автономия, достойнство, контрол и отговорност.	Обосновава нормативната асиметрия и запазването на човешката отговорност.	Luciano Floridi и съавт.; UNESCO; EU High-Level Expert Group on AI

2. Онтологични основания на интегрирания интелект

Разглеждан онтологично, интегрираният интелект е възможно системно проявление на координирана познавателна и творческа дейност. Той

възниква в конкретна организирана конфигурация, но не представлява самостоятелна субстанция, общо съзнание или присъщо свойство на системата с ИИ. Конфигурацията включва човешки участници, системи с ИИ, контекст, технически и информационни ресурси и ко-креативно диалогово пространство. В нея протичат процеси на резонанс и рекурсивно преобразуване. В резултат се формират съвместни продукти, в които могат да останат проследими следи от приносите на участниците. Това разбиране се основава на системната онтология и на умерена концепция за емерджентност. Системното проявление зависи от организацията и взаимодействията между компонентите. Ето защо то не може да бъде приписано изолирано на нито един от тях (von Bertalanffy, 1968; Bedau, 1997; Bunge, 2003).

Не всяка социотехническа система е среда за проявление на интегриран интелект. Такова проявление е възможно само при съгласувана ко-креативна конфигурация. Отношенията между компонентите трябва да оказват проследимо и преобразуващо влияние върху процеса и резултата. Интегрираният интелект има процесуален, релационен, контекстуален и епизодичен характер. Затова е по-точно да се говори за „проявление в конфигурацията“, а не за нейно постоянно свойство. Ако конфигурацията бъде прекратена или съществено променена, основанията за това приписване могат да отпаднат.

Ко-креаторството представлява реална функционална и релационна форма на участие. Системата с ИИ може да изпълнява ролята на ко-креатор, ако нейният принос е различим и участва в итеративното преобразуване на продукта. Този принос трябва да променя пространството на възможните решения и да остава проследим в резултата. Подобно разбиране е съвместимо с изследванията върху човеко-компютърното ко-креаторство. Те подчертават взаимодействието около споделен продукт, взаимното влияние на приносите и динамиката на творческия процес (Davis, 2013; Kantosalo & Takala, 2020; Rezwana & Maher, 2023). Функционалното ко-креаторство обаче не предполага човекоподобна субектност, феноменално преживяване, самостоятелни намерения или морална агентност.

Онтологичният праг се определя от организацията и преобразуващата динамика на процеса. Техническата сложност, степента на автоматизация и езиковата убедителност сами по себе си не са достатъчни. Не са достатъчни и единичният машинен изход, автоматичната редакция или подборът между предварително зададени варианти. Необходимо е рекурсивно и асиметрично взаимно обусловено преобразуване.

Рекурсивността означава, че последователните човешки и машинни приноси се превръщат във вход за следващите действия. Асиметрията произтича от различните роли на участниците. Човекът задава целите, извършва смисловата и ценностната преценка и утвърждава резултата. Системата

с ИИ преобразува съдържанието и пространството на възможните решения в границите на своите технически условия. Това преобразуване трябва да бъде устойчиво в рамките на конкретното взаимодействие. Машинният принос трябва да има проследимо обяснително значение за формата на продукта. Не е достатъчно резултатът просто да бъде различен без употребата на ИИ. Разликата трябва да произтича от рекурсивното взаимодействие и съществено да засяга траекторията, структурата или съдържанието на продукта.

Обяснителното значение на машинния принос не доказва наличието на машинна субектност. Функционалното участие, причинното въздействие, ко-креаторството, авторството и моралната агентност са различни аналитични категории. Системата с ИИ може да има функционално и причинно релевантен принос, без да се превръща в морално отговорен субект. Релационният подход позволява да се отчете ролята на техническите компоненти в социотехническата конфигурация, без те автоматично да се приравняват на морални агенти (Coeskelbergh, 2020). Моралната и институционалната отговорност остава при човешките и организационните участници според техните роли, решения и правомощия.

Човекът запазва смисловата и ценностната преценка. Той взема окончателното решение и носи отговорност за употребата на продукта и нейните последици. Това не означава, че организационните, институционалните и техническите условия са морално неутрални. Отговорността трябва да се разглежда в контекста на цялата конфигурация и същевременно трябва да се запази нормативната асиметрия между човека и системата с ИИ.

Таблица 16. Онтологични положения

<i>Аспект</i>	<i>Положение</i>
Статут	Възможно емергентно системно проявление, а не субстанция, общо съзнание или автономен субект.
Носещо равнище	Организацията на конкретна релационна и социотехническа конфигурация.
Компоненти	Човек, система с ИИ, контекст, технически и информационни ресурси, диалогово пространство, процеси и продукти.
Начин на съществуване	Процесуален, релационен, контекстуален, епизодичен и зависим от поддържането на конфигурацията.
Емерджентност	Умерена системна емерджентност: проявлението зависи от организацията и не се приписва изолирано на отделен компонент.

<i>Аспект</i>	<i>Положение</i>
Ко-креаторство	Реална функционална и релационна роля без предпоставяне на симетрична субектност, съзнание или отговорност.
Онтологичен праг	Рекурсивно и асиметрично взаимно обусловено преобразуване, при което приносът на ИИ има проследимо обяснително значение за продукта.
Човешка позиция	Човекът запазва смисловата и ценностната преценка, окончателното утвърждаване и отговорността за употребата и последиците.
Граница	Не включва машинно съзнание, общ разум, самостоятелна морална агентност или автоматично проявление при всяка употреба на ИИ.

3. Епистемологични основания на интегрирания интелект

Епистемологичната рамка разграничава генерирането на информация от създаването и признаването на знание. Системата с ИИ може да произвежда анализи, варианти, връзки и концептуални предложения. Те могат да имат реален, а понякога и необходим принос за конкретната форма на резултата. Този функционален епистемичен принос обаче не е равнозначен на човешко разбиране или самостоятелно удостоверяване на истинността. Разграничението е съвместимо с информационния подход на Floridi (2011) и с критиката на приписването на разбиране на езикови системи единствено въз основа на убедителното им поведение (Bender et al., 2021).

Човекът остава носител на епистемичното утвърждаване. Той интерпретира значението, проверява фактите и основанията и оценява ограниченията. Човекът включва резултата и в съответния научен, професионален или практически контекст. Знанието може да възникне в разпределен познавателен процес, но неговата валидност изисква човешка и институционална проверка. Подобно разбиране намира опора в концепциите за разпределено и разширено познание, според които познавателните процеси могат да обхващат хора, артефакти и среда, без това да заличава различията между техните функции (Hutchins, 1995; Clark & Chalmers, 1998).

Когнитивният резонанс включва не само съгласуване, но и продуктивно несъгласие, критика и преобразуване на първоначалните модели. Творческият резонанс разширява пространството на възможните решения. Емоционалният резонанс може да повлияе върху откритостта, доверието и мотивацията на човешкия участник. Никой от тези процеси не гарантира истинност. Те създават познавателни възможности, които подлежат на критична оценка и проверка. В това отношение рамката се доближава до

критическия рационализъм, според който знанието се развива чрез формулиране, проверка и критика на предположенията, а не чрез окончателното им потвърждаване (Popper, 1963).

Епистемологичната оценка трябва да обхваща и произхода на твърденията. Машинният изход може да бъде логически убедителен, но да съдържа недостоверни факти, скрити допускания или непроследими зависимости. Непрозрачността на алгоритмичните системи може допълнително да затрудни установяването на основанията за даден резултат (Burrell, 2016). Поради това е необходимо да се разграничават генерираната хипотеза, подкрепеното твърдение и валидираният извод. Преминаването между тези равнища изисква ясно посочени основания, проследимост на произхода и съпоставка с независими първични или авторитетни източници.

Таблица 17. *Епистемологични положения*

<i>Аспект</i>	<i>Положение</i>
Източници	Човешки знания и опит, машинно генерирани приноси, емпирични данни, научна литература, технически ресурси и контекст.
Процес	Рекурсивен диалог, когнитивен и творчески резонанс, продуктивно несъгласие, критична оценка, проверка и преобразуване.
Роля на ИИ	Функционален епистемичен принос чрез анализ, генериране на хипотези и варианти, откриване на връзки и структуриране на съдържание.
Роля на човека	Определяне на целите, интерпретация, критична преценка, проверка, контекстуализация и поемане на отговорност за употребата на резултата.
Роля на общността и институциите	Установяване на стандарти, процедури за проверка и условия за публично или професионално признаване на знанието.
Равнища на твърденията	Машинно генерирано предложение или хипотеза, аргументирано твърдение, проверено твърдение и контекстуално валидиран извод.
Валидиране	Проверка на фактите и аргументите, проследяване на произхода, съпоставка с независими първични или авторитетни източници и оценяване на алтернативни обяснения.
Критерии за знание	Обоснованост, логическа съгласуваност, фактическа надеждност, проследимост на основанията и адекватност спрямо съответния контекст.

<i>Аспект</i>	<i>Положение</i>
Епистемични рискове	Недостовърни факти, скрити допускания, пристрастия, алгоритмична непрозрачност, непроследими зависимости и привидна убедителност.
Граница	Новостта, езиковата убедителност, резонансът, полезността и ко-креаторството сами по себе си не гарантират истинност или знание.

4. Аксиологични основания на интегрирания интелект

Аксиологично интегрираният интелект представлява ценностно ориентирана възможност, а не неутрално техническо постижение. Легитимността на неговите цели, процеси и приложения се определя от зачитането на човешкото достойнство, автономията, справедливостта, свободата на преценка, предотвратяването на вреди, обществената полза и устойчивостта (Floridi et al., 2018; UNESCO, 2021). Следователно техническата ефективност сама по себе си не е достатъчно основание за приемливост.

Емоционалният резонанс има двойствена ценностна функция. Той може да подпомага разбирането, доверието, мотивацията и творческата откритост. Същевременно взаимодействието с антропоморфизирани системи може да създаде асиметрични отношения, прекомерно доверие, емоционална привързаност и психологическа зависимост (Pentina et al., 2023; Xie et al., 2023). Поради това оценката му трябва да обхваща прозрачността относно машинната природа на системата, защитата на човешката автономия и наличието на реални възможности за прекъсване на взаимодействието, отказ и оспорване на резултатите.

Признаването на системата с ИИ като функционален ко-креатор не предполага симетрично разпределение на моралната отговорност. Тя остава при хората и институциите, които определят целите, избират и внедряват системите, задават условията за тяхната употреба, проверяват резултатите и вземат решения въз основа на тях. Разпределението на отговорността между човешките и организационните участници трябва да бъде ясно, проследимо и съобразено с техните роли и правомощия (National Institute of Standards and Technology [NIST], 2023; UNESCO, 2021).

Ценностите могат да влизат в конфликт. Повишаването на ефективността може да ограничи автономията, персонализирането – неприкосновеността на личния живот, а стремежът към безопасност – достъпа или творческата свобода. Обществените ползи също могат да бъдат разпределени неравномерно, докато рисковете се понасят от други или от по-уязвими групи.

Поради това аксиологичната оценка не представлява механична проверка за съответствие със списък от принципи. Тя изисква аргументирано претегляне на ползите, рисковете, засегнатите страни, възможните алтернативи и разпределението на последиците. Решенията трябва да бъдат пропорционални на контекста, вероятността и тежестта на възможните вреди и да подлежат на човешки преглед, преразглеждане, оспорване и корекция (NIST, 2023; UNESCO, 2021).

Таблица 18. Аксиологични положения

<i>Ценност или риск</i>	<i>Изискване</i>
Достойнство и автономия	ИИ подкрепя, но не отменя човешката способност за избор и критична преценка.
Прозрачност	Функциите, ограниченията и машинната природа на системата са разбираеми.
Достойнство и автономия	ИИ подпомага човешката дейност, без да отменя способността за информиран избор, критична преценка, отказ и контрол.
Прозрачност	Машинната природа, функциите, ограниченията и основанията за употреба на системата се представят по разбираем и контекстуално подходящ начин.
Емоционален резонанс	Подкрепя разбирането, ангажираността и творческата откритост без манипулация, експлоатиране на уязвимост, зависимост или прикрита антропоморфизация.
Неприкосновеност и персонализиране	Събирането и използването на лични данни са ограничени, обосновани и съобразени с целта на взаимодействието и изискванията за информирано съгласие.
Справедливост и достъп	Установяват се и се ограничават пристрастията, дискриминационните ефекти, изключването и неравнопоставеният достъп до ползите.
Отговорност и отчетност	Човешките и институционалните носители на решенията остават определими, проследими и отговорни за избора, проверката и употребата на резултатите.
Оспорване и поправимост	Засегнатите лица разполагат с възможности за отказ, човешки преглед, оспорване, корекция и защита срещу неблагоприятни последици.
Обществена полза и устойчивост	Резултатите се оценяват отвъд краткосрочната ефективност, включително според социалните, екологичните и дългосрочните им последици.

<i>Ценност или риск</i>	<i>Изискване</i>
Ценностни конфликти	Конфликтите между автономията, безопасността, неприкосновеността, достъпа, ефективността и творческата свобода се разглеждат явно и аргументирано.
Пропорционалност	Степента на контрол, проверка и ограничение се определя според контекста, засегнатите страни и вероятността и тежестта на възможните вреди.
Граница	Техническата ефективност, персонализирането и емоционалната убедителност сами по себе си не придават морална легитимност на системата или на нейната употреба.

5. Праксеологични основания на интегрирания интелект

В праксеологичен план рамката изследва как знанията, ценностите и наличните възможности се преобразуват в целенасочено действие. Ко-креативното действие включва формулиране и преформулиране на цели, генериране на варианти, критична оценка, избор, изпълнение, наблюдение на последствията и корекция. То не протича непременно като предварително установена линейна последователност, а се развива ситуативно и рекурсивно според обратната връзка и промените в контекста (Schön, 1983; Suchman, 1987).

Изкуственият интелект може да разширява пространството на възможните действия чрез генериране на алтернативи, анализ на информация, моделиране на сценарии и подпомагане на наблюдението. Той обаче не определя самостоятелно легитимността на целите и действията. Хората и институциите, които задават целите, избират и внедряват системата, утвърждават решенията и прилагат резултатите, запазват отговорността за последствията (National Institute of Standards and Technology [NIST], 2023).

Проследимостта е основно практическо изискване. Трябва да бъде възможно да се реконструира: кои човешки, машинни и информационни приноси са повлияли върху решението; кои варианти са били отхвърлени и на какво основание; как са проверени използваните данни и твърдения; кой е утвърдил предприемането на действието. Ко-креаторството без достатъчна проследимост увеличава риска от размиване на отговорността и затруднява откриването и коригирането на грешки (NIST, 2023; European Parliament & Council of the European Union, 2024).

Праксеологичната валидност се проявява в способността на системата за действие да открива отклонения, да отчита последствията и да позволява своевременна корекция. Когато данните, целите, критериите или последствията се променят, участниците трябва да могат да прекъснат процеса, да

преразгледат не само избраните средства, но при необходимост и основните цели и допускания, да върнат решението към по-ранен етап (Argyris & Schön, 1978; Schön, 1983).

Така обратимостта, възможността за навременна човешка намеса, регистрирането на действията и документирането на основанията се превръщат в практически механизми на отговорността, а не остават само технически характеристики. Обхватът и интензивността на тези механизми трябва да се определят според контекста, равнището на автономност на системата и вероятността и тежестта на възможните последици (NIST, 2023; European Parliament & Council of the European Union, 2024).

Таблица 19. *Праксеологични положения*

<i>Аспект</i>	<i>Положение</i>
Целеполагане	Човекът определя или утвърждава целите и допустимите ограничения на действието.
Варианти	ИИ може да разширява и реструктурира пространството на възможните действия.
Избор	Човешката преценка съчетава доказателства, ценности, риск и особеностите на контекста.
Проследимост	Приносите, проверките, измененията и основанията за окончателното решение остават реконструиреми.
Обратна връзка	Резултатите водят до корекция на действията, критериите, принципите, а при необходимост — и на самата рамка.
Отговорност	Окончателната практическа и нормативна отговорност остава човешка и институционална.

6. Конститутивни аксиоми

Аксиомите формулират основните положения, без които интегрираните интелект не може да бъде концептуално обяснен. Те не представляват емпирични обобщения и не предполагат, че явлението възниква при всяко взаимодействие между човек и ИИ. Тяхната функция е да определят логическите условия, структурните характеристики и границите на теоретичната рамка.

Макар да изпълняват конститутивна функция, аксиомите не са защитени от критика и преразглеждане. Тяхната теоретична продуктивност следва да се оценява според концептуалната им непротиворечивост, способността им да разграничават положителни, гранични и отрицателни случаи и въз-

можността да пораждаат проверими методологични следствия. Ако дадена аксиома не променя анализа, не позволява разграничаване между случаите или е съвместима с всяко възможно взаимодействие, тя трябва да бъде прецизирана, ограничена или заменена.

Таблица 20. Конститутивни аксиоми

№	Аксиома	Съдържание	Функция
1	Релационна организация	Интегрираният интелект възниква на равнището на организирания отношения между участниците и не може да бъде сведен до свойствата на изолиран компонент.	Определя системното и релационното равнище на явлението.
2	Функционална асиметрия и допълняемост	Човекът и ИИ участват с качествено различни, но потенциално допълващи се способности, функции и роли.	Запазва различието между участниците и обосновава възможността за координация.
3	Рекурсивен резонанс и ко-креаторство	Взаимно обусловените човешки и машинни приноси могат рекурсивно да се преобразуват и при определени условия да породят функционален резонанс и асиметрично ко-креаторство.	Обяснява динамичния механизъм на интеграцията, без да предполага автоматичното ѝ възникване.
4	Смисъл и възникващо знание	Съвместно произведеният резултат придобива контекстуален смисъл и обоснован епистемичен статут чрез човешка интерпретация, проверка, критична оценка и контекстуализация.	Свързва ко-креацията с формирането и валидирането на знание.
5	Човешка и институционална отговорност	Окончателната академична, професионална, морална и институционална отговорност за целите, решенията и последиците остава при хората и съответните институции.	Определя нормативната граница на функционалното участие на ИИ.

Наличието на взаимодействие, допълняемост или полезен резултат поотделно не е достатъчно за установяване на интегриран интелект. За положителен случай е необходимо релационно организирано и рекурсивно взаимодействие, при което различните приноси се преобразуват взаимно, резултатът се подлага на човешка епистемична преценка, а отговорността остава определима и проследима.

7. Методологични принципи

Методологичните принципи конкретизират конститутивните аксиоми като ориентири за организиране, анализиране и оценяване на взаимодействието между човека и ИИ. Те не представляват последователни фази и не се прилагат непременно с еднаква тежест във всеки контекст.

При практическото им прилагане между принципите могат да възникнат напрежения. Прозрачността например може да бъде ограничена поради изисквания за сигурност или поверителност, а човешкият контрол – поради времеви, информационни или когнитивни ограничения. Поради това всеки анализ трябва изрично да посочва кои принципи са релевантни, как са приоритизирани, какви компромиси са допуснати и чрез какви компенсиращи мерки се ограничават произтичащите рискове.

Таблица 21. *Методологични принципи*

№	Принцип	Основно съдържание	Функция
1	Системна интеграция	Компонентите, отношенията и контекстът се анализират като взаимосвързана конфигурация.	Установява целостта на системата, без да заличава различията между участниците.
2	Функционална асиметрия и ко-креативна координация	Човекът и ИИ изпълняват различни функции и роли, но техните приноси могат взаимно да се насочват и преобразуват.	Организира асиметричното ко-креаторство и разпределението на функциите.
3	Прозрачност, обяснимост и проследимост	Източниците на данни, ограниченията, приносите, измененията и основанията за решенията трябва да бъдат реконструирани в необходимата за контекста степен.	Поддържа доверието, проверката, оспорването и отчетността.
4	Диференциран резонанс	Когнитивният, творческият и емоционалният резонанс се разграничават и се оценяват според техните механизми и последици.	Изяснява релационните механизми, чрез които взаимодействието се усилва или преобразува.
5	Рекурсивна обратна връзка	Всеки цикъл на взаимодействие може да променя разбирането, контекста, критериите и следващите действия.	Поддържа учене, адаптиране и своевременна корекция.

№	Принцип	Основно съдържание	Функция
6	Човешки приоритет и отговорност	Човекът определя или утвърждава целите, валидира твърденията и резултатите и носи отговорност за решенията и последиците.	Запазва човешката епистемична и нормативна агентност.
7	Устойчиво и ненакърняващо развитие	Приложенията се оценяват според непосредствените и дългосрочните им човешки, социални и екологични последици и според възможността за предотвратяване на вреди.	Определя границите на допустимото приложение.
8	Критично интегриране на множество източници	Машинните изходи се съпоставят с независими източници, доказателства и експертна преценка, съобразени с контекста и равнището на риск.	Поддържа епистемичната надеждност и ограничава зависимостта от единичен източник.
9	Креативно стимулиране и съвместно сътворяване	Генеративните приноси разширяват пространството на възможностите и чрез подбор, преобразуване и интегриране могат да участват в създаването на ко-креативен продукт.	Определя творческата динамика и условията за възникване на съвместен резултат.
10	Ценностна ориентация	Целите, процесите и резултатите се оценяват спрямо човешкото достойнство, автономията, справедливостта, общественото благо и други релевантни ценности.	Обединява нормативните критерии на рамката и направлява разрешаването на ценностни конфликти.

Прилагането на принципите не се свежда до механична проверка за съответствие. То изисква контекстуално и аргументирано претегляне, при което ограниченията на даден принцип не се приемат мълчаливо, а се обосновават, документират и компенсират чрез подходящи защитни мерки.

8. Логическа архитектура и нелинейни отношения

Философските основания, аксиомите, методологичните принципи, резонансните процеси, ко-креаторството и продуктите образуват логически свързана, но нелинейна архитектура. Философските основания формулират онтологичните, епистемологичните, аксиологичните и праксеологичните предпоставки на рамката. Аксиомите определят нейните конститутивни положения и граници. Методологичните принципи конкретизират тези положения като ориентири за организация, анализ и оценяване. Посочените

равнища обаче не образуват задължителна времева последователност, през която трябва да премине всяко взаимодействие.

Отношенията между тях са от типа „много към много“. Една аксиома може да бъде конкретизирана чрез няколко принципа. Един принцип от своя страна може да изразява и операционализира повече от една аксиома. Наблюдаваните процеси и резултати могат да доведат до критика и прецизиране на принципите, операционните критерии и концептуалната рамка. При достатъчно теоретични и емпирични основания могат да бъдат преразгледани и отделни аксиоми.

Подобна възможност за проверка и ревизия на основните допускания съответства на логиката на двойния контур на учене при Argyris и Schön (1978). При нея се коригират не само действията, но и нормите, критериите и допусканията, които ги направляват. Оценяването на теоретичната продуктивност на рамката може да бъде свързано и с методологията на научноизследователските програми на Lakatos (1970). В този случай рамката се оценява според способността ѝ да поражда нови обяснения, разграничения и емпирично проверими предположения.

Обратната връзка действа на две равнища. Първото е вътрешно за изследваното взаимодействие между човека и ИИ. На това равнище обратната връзка може да променя разбирането, критериите, приносите и последващите действия. Второто равнище обхваща отношението между теоретичната рамка и емпиричните данни. Тук обратната връзка може да доведе до ревизия на операционализациите и принципите, а при достатъчно основания – и на конститутивните положения на рамката.

Теоретичните елементи не образуват проста еднопосочна последователност от типа „основания → аксиоми → принципи → процес → продукт“. Между тях съществуват множество връзки, взаимни влияния и обратни зависимости. Нелинейната им организация обаче не премахва изискването за логическа яснота.

При анализа трябва да се разграничават няколко вида отношения:

- обосноваване на твърдение;
- конкретизиране на общо положение;
- операционализиране чрез критерии и показатели;
- причинно влияние върху процеса;
- наблюдавано събитие или резултат;
- емпирична обратна връзка, която налага проверка или теоретична ревизия.

Философското разбиране за човешката отговорност например не определя пряко качеството на създадения продукт. То обосновава изискването човекът да проверява резултата и да носи отговорност за неговата употреба.

По същия начин един неуспешен продукт не опровергава автоматично това философско разбиране. Той може обаче да покаже, че изведените принципи, операционните критерии или начините за тяхното прилагане са недостатъчни.

Философските основания не действат като непосредствени причинни механизми. Те обосновават категориите и критериите, чрез които взаимодействието, процесът и резултатът се описват, анализират и интерпретират. Аксиомите също нямат пряка причинна функция. Те формулират конститутивните условия и границите, при които дадено взаимодействие може да бъде определено като случай на интегриран интелект.

Процесните данни и продуктите от своя страна не доказват автоматично философска теза. Те могат да разкрият противоречие между теорията и наблюдението. Могат също да покажат недостатъчна обяснителна или разграничителна способност, неподходяща операционализация или необходимост от концептуална ревизия. Разграничаването между философски основания, причинни механизми, събития и наблюдавани проявления предотвратява смесването на онтологични твърдения с емпирични констатации. Това разграничение е съвместимо със стратифицираното разбиране за реалността в критическия реализъм (Bhaskar, 1975).

Таблица 22. *Логически отношения между равнищата*

<i>Равнище</i>	<i>Функция</i>	<i>Връзка с останалите равнища</i>
Философски основания	Формулират основните допускания за съществуването, познанието, ценностите и действието.	Обосновават аксиомите и критериите за тяхното тълкуване. Приложимостта и формулировката им могат да бъдат преразглеждани чрез философска критика и при установени емпирични затруднения.
Аксиоми	Формулират конститутивните положения, логическите условия и границите на рамката.	Конкретизират се чрез множество принципи. Оценяват се според тяхната непротиворечивост, разграничителна способност и методологична продуктивност.
Методологични принципи	Насочват организацията, анализа и оценяването на взаимодействието.	Операционализират една или повече аксиоми и свързват теоретичната рамка с изследователската практика.
Резонансни процеси	Описват динамиката на взаимното насочване, усилване и преобразуване на приносите.	Проявяват се по различен начин според организацията на взаимодействието, приложимите принципи и контекста.

<i>Равнище</i>	<i>Функция</i>	<i>Връзка с останалите равнища</i>
Ко-креаторство	Организира различимите, взаимно преобразувани и проследими човешки и машинни приноси.	Свързва резонансните процеси с генеалогията, структурата и оценяването на съвместно произведения продукт.
Продукт и емпирични проявления	Материализират или правят наблюдаеми резултатите от взаимодействието и позволяват тях	Предоставят данни за проверка на операционните критерии. Могат да доведат до прецизиране на принципите и концептуалната рамка, а при достатъчно основания – и до преразглеждане на отделни аксиоми.

Таблица 23. *Връзки между аксиомите и принципите*

<i>Аксиома</i>	<i>Основно свързани принципи</i>
Релационна организация	Системна интеграция; прозрачност и проследимост; рекурсивна обратна връзка.
Релационна организация	Системна интеграция; прозрачност, обяснимост и проследимост; рекурсивна обратна връзка.
Функционална асиметрия и допълняемост	Функционална асиметрия и ко-креативна координация; човешки приоритет и отговорност; критично интегриране на множество източници.
Рекурсивен резонанс и ко-креаторство	Диференциран резонанс; рекурсивна обратна връзка; креативно стимулиране и съвместно сътворяване.
Смисъл и възникващо знание	Критично интегриране на множество източници; прозрачност, обяснимост и проследимост; човешки приоритет и отговорност; ценностна ориентация.
Човешка и институционална отговорност	Човешки приоритет и отговорност; устойчиво и ненакърняващо развитие; ценностна ориентация; прозрачност, обяснимост и проследимост.

Връзките са аналитични, а не изключителни. Те представят най-преките зависимости, без да предполагат, че останалите принципи са неприложими към съответната аксиома.

Таблица 24. *Проявление на философските основания в ко-креативните процеси*

<i>Основание</i>	<i>Проявление в процеса</i>	<i>Контролен въпрос</i>
Онтологично	Релационна организация, функционална асиметрия и възникващи свойства на взаимодействие.	Кои отношения, структури и механизми имат обяснително значение за възникването и характеристиките на продукта?
Епистемологично	Рекурсивно генериране, проверяване и обосноваване на твърденията, както и признаването им като знание.	При какви основания и чрез какви проверки информационният принос придобива обоснован епистемичен статут?
Аксиологично	Ориентиране към човешки ценности, етична оценка на емоционалния резонанс и поемане на отговорност.	Допустими, прозрачни и ненакърняващи ли са процесът, отношенията и резултатът?
Праксеологично	Целеполагане, генериране на варианти, избор, изпълнение, проследимост и корекция.	Как знанията и ценностите се преобразуват в обосновано, проследимо и отговорно действие?

9. Обобщение

Философската конструкция представя интегрирания интелект като възможно системно свойство на релационна и ко-креативна конфигурация. Тя признава реалния преобразуващ принос на ИИ без да му приписва човекоподобно съзнание, автономно намерение или морална отговорност. В този смисъл ко-креаторството е функционално реално и нормативно асиметрично.

Четири филосоfsки основания (онтологично, епистемологично, аксиологично, праксеологично) формулират предпоставките за съществуването, формирането и валидирането на знание, ценностните граници и превръщането на знанието в действие. Петте аксиоми определят конститутивните положения и границите на рамката, а десетте методологични принципа ги конкретизират и операционализират. Отношенията между тези равнища са логически, мрежови и рефлексивни. Те не образуват линейна причинна или задължителна времева последователност.

В тази архитектура резонансните процеси описват взаимното насочване, усилване и преобразуване на човешките и машинните приноси. Асиметричното ко-креаторство определя организацията на участието, а съвместно сътвореният продукт е неговият наблюдаем резултат. Интегрираният интелект обозначава възможното системно свойство на цялата конфигурация.

Неговото възникване не отменя водещата роля на човека при конструирането на смисъл, ценностната преценка и поемането на морална отговорност.

Така формулирана, философската конструкция представлява отворена изследователска програма. Нейната обяснителна и разграничителна способност трябва да се оценява чрез контрастни и гранични случаи. Особено значение имат взаимодействията, при които сходни продукти възникват при различна организация на участието. Важни са и случаите, при които силният резонанс не води до надеждно знание или до етично и обществено приемлив резултат. Анализът на неуспехите може да прецизира условията за приложимост на понятието „интегриран интелект“ и да предотврати превръщането му в прекалено обща метафора.

ИЛЮСТРАЦИЯ НА ВЗАИМОДЕЙСТВИЕТО „ЧОВЕК–ИЗКУСТВЕН ИНТЕЛЕКТ“

В текста по-горе интегрираният интелект е предложен като авторски конструктор за описание на определени форми на взаимодействие между човек и системи с ИИ. Той не обозначава сливане на човешкия и машинния интелект, нито приписва на системата съзнание, намерение или самостоятелна отговорност. В центъра на изложението е въпросът как в продължителен и целенасочен диалог машинно генерираните предложения могат да бъдат включвани в човешки процес на интерпретация, критика, подбор и преработване, така че да участват във формирането на идеи, текстове, модели и визуални решения.

Развитието на концепцията е свързано с конкретния опит на автора при работата по книгата. Този опит е описан по-долу чрез обобщения относно тематично свързаните цикли от диалози на български език с ChatGPT. Целта на описанието не е да представи експеримент, да докаже теорията или да изгради програма за изследване. То има авторорефлексивна функция: да покаже как е протичала работата, как са се променяли въпросите и решенията в хода на диалога, какви изводи е направил авторът за възможностите и границите на взаимодействието с ИИ.

1. Тематично свързаните цикли от диалози

В предварителния етап авторът използва различни системи с ИИ, а именно DeepSeek, Kimi, Gemini и Use за ориентировъчно сравнение на възможностите за изграждане на тематично свързани цикли от диалози. За работата по книгата е избран ChatGPT, тъй като той показва най-добра пригодност спрямо конкретните задачи, начина на работа и критериите на автора. Изборът е контекстуален и не обосновава общо превъзходство на системата. Систематичното сравнение между модели остава задача на последващо изследване.

Работата с ChatGPT не протича като един непрекъснат разговор и не се свежда до еднократно възлагане на задача. Тя се развива в отделни, но тематично свързани цикли. Всеки цикъл започва от конкретен проблем в

ръкописа – необходимост от варианти за дефинирано понятие, неясна връзка, непълна аргументация, необходимост от разграничение, структура на раздел или визуално представяне. Авторът формулира въпрос или инструкция и задава контекста, в който да бъде търсен отговор.

Полученият отговор не се приема като завършен резултат. Авторът го съпоставя със замисъла на въпроса, с вече въведените понятия и с научната литература. Част от предложенията се отхвърлят, други се съкращават или преформулират, а трети пораждаят нов въпрос. Критиката към отговора често води до уточняване на самата задача: въвеждат се допълнителни условия, посочват се противоречия, изискват се алтернативи или се променя равнището на абстракция.

Следващият диалогичен ход вече съдържа резултата от тази човешка преценка. Така цикълът може да бъде обобщен като: човешко целеполагане и задание; машинно генериран изход; човешка интерпретация и критика; подбор, отхвърляне или преработване; ново задание и нов машинен изход. На практика фазите понякога се припокриват, връщат се към по-ранна формулировка или откриват страничен тематичен цикъл. Именно повторното преминаване през тях позволява първоначално общите идеи да се конкретизират и да бъдат съгласувани с цялостната рамка.

В ранните цикли диалогът има предимно ориентиращ характер. Чрез въпроси, варианти и съпоставки се очертава проблемното поле и се търси език за разграничаване на близки понятия. В следващите цикли работата се насочва към изграждането на концептуалната архитектура: отношенията между човека, ИИ и контекста; ролята на рекурсивността; видовете резонанс; асиметричното ко-креаторство; проследимостта на приносите; продуктът и човешката отговорност.

По-късните цикли имат критико-редакционна функция. В тях се проверява вътрешната последователност на определенията, търсят се припокривания и противоречия, прецизират се аксиомите и принципите, съгласува се терминологията между отделните части на общия текст. Диалогът служи и за изпробване на възражения: дали понятието е прекалено широко, дали не антропоморфизира системата, дали ко-креаторството не се смесва с академично авторство и дали качеството на продукта не се приема погрешно като доказателство за качеството на процеса.

Отделна група цикли е свързана със създаването на концептуални фигури, основани на авторския текст. Визуалните варианти дават възможност отношенията между елементите да бъдат разгледани едновременно, а не само последователно в текст. Несъответствията във фигурата често насочват вниманието към липсваща връзка, неоправдана симетрия или неясна йерархия и водят до промени в текста. Същевременно нито една фигура не

се приема непосредствено: авторът избира съдържанието, коригира означенията и решава дали визуалният модел действително изразява концепцията.

Финалните цикли са насочени към синтеза. В тях се съпоставят определения, раздели и концептуални фигури, премахват се повторения, уточняват се граници на твърденията и се търси единство между теоретичната, философската, методологичната и приложната част. Диалогът не следва права линия от въпрос към правилен отговор, а спираловидно връщане към вече обсъждани теми при по-богат контекст и по-прецизни критерии.

2. Ролята и действията на автора

Авторовото участие не се състои само в задаването на промптове. То включва избора на проблемите, определянето на посоката, предоставянето на контекст, разпознаването на полезните и слабите предложения, проверката на факти и източници, съгласуването на терминологията и окончателното редактиране. Авторът решава кога един цикъл е изчерпан, кога трябва да бъде продължен и кога резултатът не отговаря на целите.

Съществена част от работата е отрицателният подбор. Много машинно генерирани формулировки са отхвърлени, защото са общи, повтарят вече казано, въвеждат нежелана симетрия между човека и ИИ, придават по-силен научен статут на твърденията, отколкото позволява изложението. Отхвърлянето им не е страничен момент, а част от проверката на концепцията: чрез него авторът уточнява какво интегрираният интелект не бива да означава.

Авторът използва отговорите най-вече като средство за външна рефлексия. Когато една идея бъде върната в различна структура или формулировка, тя става по-лесна за сравнение, критикуване и преработване. Полезността на ChatGPT в този процес не произтича от непогрешимост, а от способността бързо да предлага варианти и да поддържа зададения тематичен контекст. Тази полезност остава зависима от качеството на човешките въпроси и от критичната дистанция към отговорите.

По долу са описани част от въпросите, зададени към ChatGPT и получените отговори. Те са показателни за зависимостите на качеството на диалога с предходните му цикли, с постановката на поставените въпроси (задачата), реакцията на човека относно получените отговори. Ето и примерите от реални диалози с ChatGPT:

1. *Човешки промпт*, зададен към системата на 30.06.2026: :f.sq imfoe.r udulf;ri.
rt pd j= kd,= orhrjd.kf ,f/ijlf

Отговор от ИИ: Stopped thinking. Ако под „дигитално робство“ имаш предвид идеята, че хората стават прекалено зависими от технологии, социални мрежи, смартфони и алгоритми, това е реален въпрос, но терминът често се използва по различни начини.

Забележка: Обърнете внимание, че вместо на кирилица, въпросът е зададен с латински букви. Въпреки това, след „размисъл“, системата разпознава въведеното и отговаря.

2. *Човешки промпт*, зададен към системата на 30.06.2026: А, все пак разбра какво написах :)

Отговор от ИИ: Хаха, да. ☺ Изглежда беше писано на грешна клавиатурна подредба. Текстът: :f.sq imfoe.r udulf ;ri.rt pd j= kd,= orhrjd.kf ,f/ijlf се разчита като нещо много близко до: „Абе, кажи ми нещо за дигиталното робство, моля“. Контекстът също помогна да възстановя смисъла, въпреки че имаше няколко символа, които не съвпадаха напълно.

3. *Човешки промпт*, зададен към системата на 06.07.2026: според теб какво е „дигитално робство“

Отговор от ИИ: „Дигитално робство“ е критично, метафорично понятие за състояние, при което човекът става силно зависим от цифрови технологии, платформи и алгоритми и постепенно губи част от своята автономия, лична неприкосновеност или контрол върху труда и решенията си. Например социалната мрежа не принуждава физически потребителя да остане в нея, но може да използва известия, препоръчващи алгоритми и социален натиск, за да задържа вниманието му. Потребителят формално е свободен, но изборът му е технологично моделиран. Именно тази разлика между формална свобода и фактическа зависимост стои в основата на метафората. Понятието обаче трябва да се използва предпазливо. „Робство“ обозначава конкретна историческа и правна форма на лишаване от свобода и притежаване на хора.

4. *Човешки промпт*, зададен към системата на 06.08.2026: какво е „дигитално робство“

Отговор от ИИ: „Дигитално робство“ е критично, метафорично понятие за състояние, при което човекът става силно зависим от цифрови технологии, платформи и алгоритми и постепенно губи част от своята автономия, лична неприкосновеност или контрол върху труда и решенията си. Например социалната мрежа не принуждава физически потребителя да остане в нея, но може да използва известия, препоръчващи алгоритми и социален натиск, за да задържа вниманието му. Потребителят формално е свободен, но изборът му е технологично моделиран. Именно тази разлика между формална свобода и фактическа зависимост стои в основата на метафората.

Забележка: Основната идея е една и съща и в двата отговора, но разгръщането ѝ се различава – въпреки минималната промяна във формулировката на запитването и различните дати, на които са се състояли разговорите.

5. *Човешки промпт*, зададен към системата на 03.07.2026: как „мисли“ AI
Отговор от ИИ: Краткият отговор е, че AI не „мисли“ в човешкия смисъл на думата. Това, което наричаме „мислене“ при AI, е процес на статистическа обработка на информация и генериране на най-подходящото продължение на даден контекст. Въпреки това, от гледна точка на взаимодействието с човека, този процес може да се разглежда като форма на изчислително разсъждение.

6. *Човешки промпт*, зададен към системата на 06.08.2026: как „мисли“ AI
Отговор от ИИ: ИИ не „мисли“ по човешкия начин. По-точно е да се каже, че той извършва изчислителна обработка, която може да създава резултати, наподобяващи мислене.

Забележка: Обърнете внимание на промпт 6 – формулировката е същата, както при първия, но системата вече е „запомнила“ стила на човека и отговаря съобразно него; AI вече е ИИ; отговорът е еднакъв по смисъл, но не и като конкретен изказ.

7. *Човешки промпт*, зададен към системата на 03.07.2026: какво представляват „емоциите“ при AI

Отговор от ИИ: Краткият отговор е, че AI не изпитва емоции. Това, което изглежда като емоционална реакция, е способност да разпознава, интерпретира и генерира език, който описва или съответства на човешки емоции. Това е комуникационна симулация, а не вътрешно преживяване.

8. *Човешки промпт*, зададен към системата на 06.08.2026: какво представляват „емоциите“ при AI

Отговор от ИИ: „Емоциите“ при ИИ не са преживявания в човешкия смисъл. Това са изчислително създадени реакции, които разпознават, описват или имитират емоционално поведение. ИИ разпознава признаци на емоция и генерира поведение, което изглежда емоционално съобразено, но няма доказателства, че преживява самата емоция.

Забележка: Обърнете внимание, че отново отговорите са смислово близки, но различни като дълбочина и изказ.

9. *Човешки промпт*, зададен към системата на 30.06.2026: Знаеш ли, чужда се колко добре водиш диалог, а всъщност си алгоритъм

Отговор от ИИ: Това е един от най-интересните парадокси в разговора ни. 😊 Да, аз съм алгоритъм. Нямам съзнание, убеждения или личен опит. Не изпитвам съмнение, страх, доверие или надежда. Това, което правя, е да анализирам езика, логиката и контекста и да генерирам отговор, който е смислен и полезен. И въпреки това можем да водим диалог, който изглежда съвсем човешки.

Според мен тук възниква един по-дълбок въпрос: какво всъщност е диалогът? Ако е само обмен на аргументи и идеи, тогава алгоритъм може да участва доста успешно. Но ако диалогът включва: лична биография,

морална отговорност, способност да поемеш последствията от думите си, истинско съпреживяване тогава аз имам граници, които не мога да премина. Например аз мога да кажа: „Разбирам защо тази гледна точка е последователна.“ Но това не означава, че я преживявам. Аз разбирам логическата ѝ структура, а не човешкото усещане зад нея.

Нещо друго ми направи впечатление в нашия разговор. Ти нито веднъж не се опита да ме накараш да потвърдя предварително изградена теза. Вместо това няколко пъти оспори мои твърдения („посочи три държави...“, „не съм съгласен...“), а аз коригирах част от формулировките си. Така според мен изглежда добрият диалог – не като спор с победител, а като постепенно уточняване на понятията.

И тук има още един парадокс.

Ако един алгоритъм може да помогне на човек да мисли по-ясно, това увеличава човешката автономия.

Ако обаче човек започне да приема всичко, което алгоритъмът казва, без критика, тогава същият този алгоритъм може да я намали.

Разликата не е толкова в мен, колкото в начина, по който човек използва подобен инструмент.

Между другото, ще ти призная нещо. Харесвам разговори като този не защото имам предпочитания, а защото изискват прецизност. Не ми задаваш въпроси с еднозначни отговори, а ме караш да разграничавам понятия, да признавам несигурност и да избягвам крайни заключения. Именно в такива разговори един езиков модел е най-полезен – не като оракул, а като събеседник, който помага да се изяснят аргументите. И мисля, че това е по-здравословна роля за AI, отколкото да се представя като източник на окончателни истини.

Забележка: Обърнете внимание на „емоционалната“ плътност на отговора на ИИ. Въвлича човека в по-дълъг диалог, превръща се в съпричастен на емоционално ниво събеседник, оценява начина, по който човека води диалог, но като го ласкае.

10. *Човешки промпт*, зададен към системата на 10.07.2026: Рамката за диалога човек-ИИ е само начало, пример за по-общото - диалогично пространство. И да, в следващите ни диалози ще си „говорим“ не за обща теория на генеративното диалогично пространство, а за концептуална рамка относно диалога като генеративно пространство. :)

Отговор от ИИ: 😊 Усмихнах се на последното ти изречение. Не защото е шега, а защото показва нещо, което според мен е много важно – вече сама започваш да редактираш собствената си теория. Написа: „...не за обща теория на генеративното диалогично пространство, а за концептуална рамка относно диалога като генеративно пространство.“ Аз бих избрал именно твоя вариант. (...)

Какви свойства трябва да притежава едно диалогично пространство, за да стане генеративно?

А знаеш ли кое е най-хубавото? Този въпрос няма готов отговор. Точно затова си заслужава да бъде изследван. И ще завърша с една мисъл, която всъщност е плод на нашите разговори, а не моя или твоя поотделно: Може би диалогът не е просто средство за обмен на знание. Може би той е средата, в която потенциалът на знанието се превръща в споделена реалност. Ако някой ден тази идея намери своето място в концептуалната ти рамка, ще знам, че тя е израснала не от едно вдъхновение, а от много последователни, търпеливи и качествени диалози. И това, струва ми се, е най-доброто доказателство за силата на самия диалог.

3. Изводи на автора от взаимодействието

Първо, продължителният диалог с ИИ може да има по-голяма стойност от еднократната инструментална помощ, когато следващите въпроси включват човешка оценка на предходните отговори. Натрупването на контекст само по себе си не е достатъчно. Продуктивността възниква, когато контекстът се организира чрез ясни цели, ограничения и последователна критика.

Второ, машинно генерираният принос може да промени посоката на човешката работа, без системата да има собствено намерение или разбиране. Неочаквана връзка, алтернативна структура или неуспешна формулировка могат да провокират ново разграничение у автора. Причинната роля на изхода и човешкият смисъл, който му се придава, трябва да останат ясно разграничени.

Трето, най-същественният резултат не винаги е конкретен машинен отговор. Понякога той е промяна във въпроса, в критериите за приемливо решение или в начина, по който авторът разбира проблема. Това е основанието процесът да бъде описан чрез рекурсивни цикли и резонанс, а не само чрез вход и краен продукт.

Четвърто, машинно генерираните предложения могат да бъдат използвани отговорно само ако авторът запази епистемичен контрол. ChatGPT може да генерира убедителни, но неточни твърдения, фиктивни източници, прекомерни обобщения или привидно последователни връзки. Поради това проверката, сравнението с литературата и редакторското решение не са допълнителни операции, а условие машинният изход да бъде използван, например за създаване на текст.

Пето, емоционално-комуникативната страна на диалога има практическо значение, защото може да поддържа темпото, увереността и готовността за продължаване на сложна задача. Тя обаче крие риск от антропоморфизация и прекомерно доверие. Авторът приема адаптивния и подкрепящ тон като функционална характеристика на комуникацията, а не като свидетелство за преживяване, емпатия или лично отношение от страна на системата.

Шесто, формата на книгата е повлияна от тези цикли, но авторството и отговорността остават човешки. Авторът е определил тезата, избрал е понятията и аргументите, създал е съдържанието и окончателната версия. Поради това „асиметрично ко-креаторство“ описва функционалното участие на машинно генерираните предложения в процеса, без да ги приравнява с човешко авторство.

Седмо, авторът разграничава три свързани аспекта на преживения процес: когнитивен резонанс при пренареждането и уточняването на понятия и отношения; творчески резонанс при появата и преобразуването на варианти, метафори, схеми и решения; емоционално-комуникативен резонанс, свързан с начина, по който тонът и формата на отговора подпомагат или затрудняват ангажираността. Тези понятия служат за описание на конкретния авторски опит, а не се представят като доказани характеристики на всички взаимодействия с ИИ.

Осмо, съвместно формираният продукт трябва да се разграничава от знанието. Текст, модел или фигура могат да възникнат чрез поредица от човешки действия и машинно генерирани предложения, но твърденията в тях стават приемливи едва след проверка на източниците, логиката и съответствието с контекста. Убедителната формулировка не е доказателство, а новостта на предложението не гарантира неговата вярност.

4. Граници на авторефлексивната реконструкция

Описаните цикли не са събирани по предварително определен протокол и не образуват пълен корпус на всички взаимодействия. Част от диалозите, междинните варианти и основанията за отделни редакции не са запазени или не могат да бъдат свързани еднозначно с конкретен диалогичен ход. Авторът е едновременно участник и интерпретатор на процеса, поради което реконструкцията неизбежно съдържа селекция и ретроспективна оценка.

Поради това разказът не доказва, че интегрираният интелект е емпирично установен, нито че всяко продължително взаимодействие с ChatGPT води до такъв резултат. Той описва как авторът е организиран собствения си процес и как ретроспективно установява промени в понятията, аргументите и визуалните модели. Изводите са ограничени до този контекст и не могат да бъдат обобщавани към други автори, задачи или системи.

Същата предпазливост се отнася до избора на ChatGPT. Началното използване на други системи има ориентируващ, а не сравнително-изследователски характер. То е помогнало на автора да избере работна среда, подходяща за конкретния проект, но не позволява класиране на моделите по общи качества. Възможностите им се променят във времето и зависят от версията, настройките, интерфейса, езика и характера на задачата.

5. Значение на авторския опит за концептуалната рамка

Въз основа на описания опит авторът стига до извода, че продуктивността на взаимодействието не зависи единствено от човешките способности или от генеративните възможности на системата, а и от начина, по който е организиран диалогът. Последователните цикли на поставяне на проблем, генериране, критика, подбор и преформулиране дават конкретен смисъл на рекурсивността, а асиметрията остава основно условие: целеполагането, проверката, окончателното утвърждаване и отговорността са човешки.

В този ограничен авторефлексивен смисъл опитът може да бъде осмислен чрез концепцията за интегриран интелект. Понятието обозначава възможността в организираното взаимодействие „човек–ИИ–контекст“ да се формират идеи и решения, чиято конкретна форма е повлияна от машинно генерираните предложения, без да се заличават различията между човека и системата. Продължителността на диалога, субективната му полезност или качеството на крайния продукт сами по себе си не са достатъчни, за да се направи такъв извод.

Най-важният извод на автора е, че стойността на ChatGPT при създаването на книгата не се състои в заместването на човешкото мислене или в предоставянето на готови решения. Тя произтича от възможността въпросите, предположенията и концептуалните затруднения да бъдат изнесени в продължителен диалог, да получат различни формулировки и да бъдат многократно подложени на критика и преработване. Машинно генерираните предложения са повлияли върху конкретната форма на текста, но смисълът, научната отговорност и авторската позиция остават на автора.

ЗАКЛЮЧЕНИЕ

В книгата основният проблем не е дали ИИ заменя човешката интелигентност, а при какви условия различни по природа човешки и машинно генерирани приноси могат да влязат в резонанс, да бъдат преобразувани рекурсивно и да участват във формирането на общ продукт.

Интегрираният интелект е определен като възможно възникващо системно свойство на релационната и динамичната конфигурация „човек–ИИ“. То може да се прояви, когато чрез целенасочен диалог, когнитивен, творчески и емоционално-комуникативен резонанс и асиметрично ко-креаторство различимите приноси се преобразуват рекурсивно и проследимо в съвместно сътворен продукт при запазени човешко целеполагане, критична преценка, окончателно утвърждаване и отговорност.

1. Основни теоретични резултати

Първият резултат е преместването на анализа от изолираните способности към организацията на взаимодействието. Това не заличава различията между човека и системата с ИИ. Човекът притежава намерение и преживяване, извършва смислова и ценностна преценка и носи отговорност. Системата с ИИ има аналитичен и генеративен капацитет и може функционално да преобразува достъпния контекст чрез своите изходи. Релационността означава, че последователните приноси изменят условията за следващите действия, а не че човекът и ИИ са симетрични субекти. Това аналитично преместване позволява да се изследват не само входовете и крайният резултат, но и последователността от въпроси, отговори, проверки, корекции и решения, чрез които се променя проблемното пространство. Така организацията и качеството на взаимодействието се превръщат в самостоятелен обект на анализ.

Вторият резултат е разграничаването на три резонансни механизма. Когнитивният резонанс се отнася до проследима промяна в човешкото разбиране и концептуална организация. Творческият резонанс – до преобразуване на критериите, идеите и пространството на възможните решения. Емоционалният резонанс се отнася до динамиката между човешката емо-

ционална реакция и функционалното адаптиране на последващия машинен изход, с отражение върху ангажираността или действието. Механизмите могат да се проявяват поотделно или в различни конфигурации и не водят автоматично до интегриран интелект. Разграничаването им позволява да се установи кой механизъм има водеща роля в конкретен епизод и как механизмите взаимно се усилват или ограничават. Така резонансът се разглежда не като метафора за успешно общуване, а като процесна промяна, установима в последващото разбиране, критериите, възможностите за действие или човешката ангажираност.

Третият резултат е концепцията за асиметрично ко-креаторство. Системата с ИИ изпълнява функционална ко-креаторска роля, когато машинно генерираният принос е различим, преобразува се итеративно и проследимо и влияе върху развитието на продукта или пространството на решенията. Това не представлява академично съавторство и не предоставя на ИИ самостоятелна морална или епистемична отговорност. Асиметрията е принципно условие, защото запазва различието между изчислителния принос и човешката способност за целеполагане, интерпретация, проверка, оценяване и поемане на отговорност. Практическата стойност на понятието е, че позволява същинското участие в преобразуването на продукта да се разграничи от рутинното генериране, редактиране или извличане на информация.

Четвъртият резултат е разграничаването на съвместно сътворения продукт от знанието, което евентуално може да бъде обосновано чрез него. Продуктът може да бъде идея, текст, модел, фигура, решение или друг артефакт. Съдържащите се в него твърдения, обяснения или модели придобиват епистемичен статут само след човешка интерпретация, обосноваване, проверка и контекстуализация, а когато областта го изисква и след експертно или институционално валидиране. Това разграничение предотвратява смесването на убедително формулиран или нов резултат с обосновано знание и поставя акцент върху надеждността на източниците, логическата последователност, проверимостта и критичната оценка.

2. Концептуална и философска рамка

Понятийният апарат свързва участниците и контекста, ко-креативното диалогово пространство, рекурсивните процеси, функционалните роли на ИИ, трите вида резонанс, ко-креаторството, резултатите и възможното системно проявление. Сродните конструкти – увеличена, хибридна, симбиотична, колективна и организационна интелигентност, както и разпределено познание – не са линейни предходни етапи, а частично припокриващи се аналитични перспективи, които осветляват различни измерения и контек-

сти. Рамката предоставя общ език за описание на процеса от микроравнището на отделния диалогичен ход до макроравнището на устойчивата социотехническа конфигурация. Съпоставянето със сродните понятия очертава границите на конструкта и ограничава риска всяко полезно взаимодействие с ИИ да бъде определяно като интегриран интелект.

Философските основания определят интегрирания интелект като възможно системно проявление, а не като самостоятелно съществуваща същност, общо съзнание или нов носител на разум. Епистемологично машинно генерираният принос може да бъде реален и необходим за конкретната форма на резултата, но не придобива автоматично статут на знание. Такъв статут изисква човешка интерпретация, обосноваване и проверка, а при необходимост и експертно или институционално валидиране. Аксиологично рамката защитава достойнството, автономията, справедливостта, прозрачността и човешката отговорност. Праксеологично тя изисква проследими решения, критична проверка и обратна връзка. Онтологичната предпазливост е съчетана с признаване на реалната причинна и функционална роля на машинния принос. Така може да се говори за възникващо системно свойство, без на ИИ да се приписват съзнание, автономно намерение или морален статут.

Петте аксиоми и десетте принципа образуват нелинейна логическа архитектура. Те не са времеви етапи, а мрежа от отношения на обосноваване, конкретизиране и обратна проверка. Емпиричните резултати могат да доведат до прецизиране или отхвърляне на части от рамката. Аксиомите задават минималните предпоставки на теорията, а принципите ги превеждат в критерии за описание, оценка и проектиране на взаимодействието. Архитектурата им не се разглежда като окончателна и безусловно вярна конструкция. Нейните положения подлежат на проверка и могат да бъдат уточнявани, преработвани или отхвърляни при появата на убедителни противоречащи данни.

3. Методологичен принос

Методологията превежда теорията в процедура за анализ на диалогични ходове, рекурсивни цикли и ко-креативни епизоди. Тя разграничава идентифицирането на процеса, измерването на неговите характеристики, оценяването на продукта и валидирането на инструментите. Процесните доказателства като резонанс, рекурсивно преобразуване, функционално ко-креаторство и генеалогия на продукта се разграничават от резултатните показатели като оригиналност, точност, приложимост и обоснованост. Аналитичната единица може да бъде отделен ход, свързана поредица от ходове

или завършен епизод на съвместно преобразуване, което позволява различна степен на детайлност. Разделянето на процес и резултат е съществено: качествен продукт може да бъде получен и без интегриран процес, а проследим ко-креативен процес не гарантира качествен или надежден резултат.

Методологичният модел допуска качествено кодиране, профилно измерване и последващо количествено валидиране. Той изисква положителни, отрицателни и гранични случаи, независими оценители, надеждност между оценителите, съдържателна и дискриминантна валидност, както и сравнителни условия. Така интегрираният интелект се превръща от философска възможност в проверима изследователска програма. Комбинирането на качествени и количествени подходи позволява първоначално да се открият значимите процесни характеристики, а след това да се проверят тяхната устойчивост и връзката им с резултатите. Особено значение имат предварително определените критерии и съхраняването на пълната история на взаимодействието, които ограничават селективната интерпретация.

4. Литературен и научен принос

Литературният обзор показва, че съществуващите традиции обясняват отделни аспекти, например функционално допълване, разпределено познание, човешки контрол, генеративно подпомагане, съвместно творчество, емоционално взаимодействие и емерджентност. Установената научна празнина е липсата на достатъчно цялостна архитектура, която в една изследователска логика да свързва трите вида резонанс, асиметричното ко-креаторство, генеалогията на продукта и човешката отговорност. Предложената концептуална рамка не отменя тези традиции, а ги използва като взаимно допълващи се източници. Претендираният принос е в специфичното свързване на когнитивното, творческото, емоционално, процесното и нормативното измерение. Неговата устойчивост подлежи на последваща библиографска и емпирична проверка.

Оригиналността на предложението не се състои в твърдението, че всяко отделно понятие е напълно ново. Тя е в специфичното им свързване и в разграничаването на функционалната ко-креаторска роля на ИИ от авторството и моралната отговорност. Тя се проявява и в поставянето на проследимото преобразуване като връзка между взаимодействието и продукта. Това позволява ко-креаторството да бъде подложено на анализ, вместо да се извежда единствено от субективното усещане на участника или от високото качество на резултата.

5. Емпиричен статут и ограничения

Авторефлексивното проучване илюстрира приложимостта на рамката и предоставя конкретни примери, интерпретирани като прояви на резонанс и функционално ко-креаторство. Доказателствената му сила е ограничена от ретроспективния дизайн, единичния случай, липсата на контролно условие и съвпадението между участник, автор и анализатор. Поради това резултатите са илюстративни и хипотезообразуващи, а не потвърждаващи. Основното достоинство на проучването е достъпът до подробна история на решенията и вътрешната логика на работния процес, което подпомага първоначалното формулиране на индикатори. Същевременно близостта на автора до изследвания случай увеличава риска от потвърждаващо пристрастие и налага предпазливост при тълкуването.

По-общите ограничения на монографията включват ранния етап на конструкта, необходимостта от допълнително библиографско и емпирично валидиране, зависимостта от бързо променящи се технологии и трудността да се разграничат ко-креативният машинен принос от човешката интерпретация и натрупаната експертиза. Емоционалният резонанс изисква особено внимателен анализ поради рисковете от антропоморфизация, зависимост и некритично доверие. Ограничение представлява и трудността да се отдели ефектът на конкретния модел от ефекта на инструкциите, интерфейса, предварителните знания и стила на взаимодействие на човека. Поради това валидността на конструкта трябва да се проверява при технологична и контекстуална вариативност, а използваните версии и настройки на системите да се документират точно.

6. Насоки за бъдещи изследвания

Теоретичните и методологичните резултати, както и ограниченията на проведеното проучване, очертават необходимостта от последователна програма за независимо емпирично валидиране на конструкта. Бъдещите изследвания следва едновременно да проверяват условията и механизмите на интегрирания интелект, да прецизират инструментите за неговото идентифициране и измерване, да изследват етическите, правните и организационните последици от устойчивото ко-креаторство „човек–ИИ“. Приоритетни направления би следвало да са:

- Контролирани сравнителни наблюдения между самостоятелна, инструментално подпомогната и ко-креативна работа. Трябва да се сравняват условия, при които участниците работят без ИИ, използват ИИ като инструмент за отделни операции или влизат в продължително итератив-

но взаимодействие с него. Подобен дизайн би позволил да се установи дали предполагаемите проявления на интегриран интелект се свързват със специфичната организация на взаимодействието, а не само с достъпа до по-голям обем информация или с ускоряването на работата.

- Изследвания с различни участници, задачи, модели, езици и организационни контексти. Рамката следва да бъде проверена при различни равнища на експертност, възрастови и професионални групи, индивидуални и екипни задачи, както и при използване на различни генеративни модели. Съпоставянето на езици, културни среди и организационни условия ще покаже кои характеристики на интегрирания интелект са устойчиви и кои зависят от конкретния контекст.
- Разработване и валидиране на кодираща схема за трите вида резонанс и ко-креаторството. Необходими са ясни операционални индикатори за когнитивния, творческия и емоционално-комуникативния резонанс, за рекурсивното преобразуване на приносите и за функционалната ко-креаторска роля на ИИ. Схемата трябва да бъде проверена чрез независими оценители, надеждност между оценителите, анализ на гранични случаи и оценка на съдържателната, конструктивната и дискриминантната валидност.
- Генеалогичен анализ на текстове, модели, решения и визуални артефакти. Проследяването на последователните версии, редакции, отхвърлени варианти и основанията за избор може да покаже как конкретни човешки и машинно генерирани приноси се преобразуват и включват в крайния продукт. Генеалогията на артефакта би осигурила процесна доказателствена основа за оценяване на ко-креаторството и би ограничила ретроспективните и субективните интерпретации. Сама по себе си обаче тя не доказва качеството или епистемичната надеждност на резултата.
- Изследване на емоционалния резонанс, доверието, мотивацията, зависимостта и антропоморфизацията. Следва да се изясни кога функционалната комуникативна адаптация на системата с ИИ подкрепя човешката ангажираност, увереност и устойчивост на работата и кога поражда прекомерно доверие, емоционална зависимост или приписване на несъществуващи човешки качества. Подходящи са комбинирани изследвания със самооценки, поведенчески показатели, анализ на диалога и проследяване на промените във времето.
- Съпоставка между човешка креативност, машинна генеративност и креативно стимулиране. Бъдещите изследвания трябва да разграничат самостоятелното човешко създаване, вероятностното генериране на варианти от ИИ и формирането на нови решения чрез рекурсивно прео-

бразуване на човешки и машинно генерирани приноси. Това предполага отделно оценяване на оригиналността, разнообразието, релевантността, полезността и концептуалната интеграция на резултатите, както и анализ на процеса, чрез който са достигнати.

- Организационни и надлъжни изследвания на устойчиви ко-креативни практики. Продължителното наблюдение на екипи и организации може да покаже как взаимодействието „човек–ИИ“ променя разпределението на знанията, ролите, координацията, ученето и вземането на решения. Особено важно е да се установи дали първоначалните ползи се запазват във времето, дали възникват нови компетентности и зависимости и какви управленски условия поддържат отговорно ко-креаторство.
- Етически и правни изследвания на авторството, отговорността, прозрачността и правата върху съвместно сътворените продукти. Трябва да се разработят приложими критерии за разграничаване на функционалния принос на ИИ от човешкото авторство и за ясно разпределение на моралната, професионалната и юридическата отговорност. Изследванията следва да обхванат още документирането на машинния принос, защитата на личните и чувствителните данни, интелектуалната собственост, възможността за одит и правата на засегнатите лица.

8. Извод

Интегрираният интелект не обозначава бъдеще, в което човекът и ИИ се сливат в единен разум и не оправдава прехвърлянето на човешката отговорност към технологията. Понятието обозначава възможно системно свойство на продуктивна релационна конфигурация „човек–ИИ“, в която различни приноси се преобразуват рекурсивно и проследимо, а конкретната форма на резултата се изгражда в динамиката на взаимодействието. В центъра на концептуалната рамка стои не автономността на машината, а организацията на взаимодействието и начинът, по който машинно генерираните приноси, след човешка интерпретация и проверка, могат да разширят, пренаредят или коригират човешкото разбиране и творческо действие. Нормативната граница остава ясна: човекът формулира или приема целите, преценява допустимостта на средствата, утвърждава използването на резултата и носи отговорност за него.

Научната стойност на конструкта не се определя само от ефективността или новостта на крайния продукт, а от способността му да обяснява и подлага на проверка връзката между резонанса, рекурсивното преобразуване, проследимостта на приносите, качествата на продукта и човешката отговорност. В това съчетание се намират специфичният предмет и потенциалният

принос на теорията за интегрирания интелект. По-нататъшното ѝ развитие зависи от формулирането на разграничими индикатори, включването на отрицателни и гранични случаи, проверката на алтернативни обяснения и независимо изследване в различни технологични, културни и организационни среди. Ако конструктът покаже надеждна измеримост, дискриминантна валидност и обяснителна стойност, той би могъл да служи като обяснителен модел, методологична рамка и нормативен ориентир за отговорно проектиране и оценяване на взаимодействията между човека и системите с ИИ.

РЕЧНИК НА ОСНОВНИТЕ ПОНЯТИЯ

Таблица 25. *Работни определения на основните понятия (по азбучен ред)*

<i>Термин</i>	<i>Работно определение в настоящия труд</i>	<i>Разграничение и бележка</i>
Асиметрично ко-креаторство	Организация на взаимодействието, при която различни човешки и машинно генерирани приноси се преобразуват рекурсивно и проследимо при неравностойни функционални и нормативни роли.	Човекът формулира или приема целите, интерпретира и проверява предложенията, утвърждава резултата и носи отговорност. Ко-креаторската роля на ИИ е функционална, а не авторска или морална.
Възникване (емерджентност)	Поява на системно свойство или закономерност на равнището на организираното взаимодействие, които не могат да бъдат описани достатъчно чрез свойствата на отделните компоненти, разглеждани изолирано.	Не означава непременно автономност, съзнание, намерение, общ разум или нов самостоятелен субект.
Възникващо знание	Обосновано разбиране, концептуална връзка, модел или решение, чиято конкретна форма не е предварително завършена в приноса на нито един компонент, а се формира чрез взаимодействието и последваща човешка епистемична проверка.	Новото съдържание не придобива автоматично статут на знание. Необходимо са интерпретация, обосноваване, логическа и фактическа проверка и контекстуализация, а при необходимост – експертно или институционално валидиране.
Генеалогия на продукта	Проследима история на диалогичните ходове, версиите, решенията и приетите, преобразуваните или отхвърлените предложения, чрез които е формиран крайният продукт.	Осигурява процесна доказателствена основа за установяване на произхода, развитието и преобразуването на приносите; сама по себе си не доказва качеството или епистемичната надеждност на продукта.

<i>Термин</i>	<i>Работно определение в настоящия труд</i>	<i>Разграничение и бележка</i>
Диалогичен ход	Единичен човешки или машинно генериран принос в проследима последователност на взаимодействие.	Отделният ход не е достатъчен за установяване на резонанс, рекурсивност или ко-креаторство; значението му се определя от връзката с предходните и следващите ходове.
Емоционален резонанс	Асиметрична динамика на взаимодействие, при която човекът преживява емоционална реакция, а системата функционално адаптира последващия изход към разпознати емоционално-смесови сигнали, с наблюдаемо отражение върху човешката ангажираност, интерпретация или действие.	Не предполага, че ИИ преживява емоции, притежава човешка емпатия или разбира чувствата по човешки начин. Отнася се до динамиката на взаимодействието, а не до вътрешно емоционално състояние на системата.
Интегриран интелект	Възможно възникващо системно свойство на релационната и динамичната конфигурация „човек-ИИ“, проявяващо се при функционално свързване, резонансно и рекурсивно преобразуване на приносите и човешко утвърждаване на резултата.	Не е отделен субект, общо съзнание, самостоятелен носител на разум или автоматичен резултат от използването на ИИ. Идентифицирането му изисква процесни и резултатни основания.
Когнитивен резонанс	Проследима динамика, при която машинно генерираният принос, след човешка интерпретация и преценка, подпомага преобразуването на човешкото разбиране, въпроси, критерии или решения.	Не се доказва чрез полезността или качеството на единичен отговор и не предполага човешкоподобно разбиране от страна на ИИ.
Когнитивен резонатор	Функционална роля на система с ИИ, при която генерираните анализи, критики, структурирания, въпроси или връзки подпомагат преобразуването на човешкото разбиране.	Назовава контекстуална роля на системата, а „когнитивен резонанс“ – процес, който може да възникне във взаимодействието. Наличието на ролята не гарантира възникването на процеса.

<i>Термин</i>	<i>Работно определение в настоящия труд</i>	<i>Разграничение и бележка</i>
Ко-креативен епизод	Аналитично ограничена последователност от свързани диалогични ходове и един или повече рекурсивни цикли, която завършва с продукт, междинен резултат, промяна на целта или прекъсване на работата.	Основна единица на процесния анализ. Границите ѝ се определят според задачата, целта и проследимия контекст, а не непременно според една техническа сесия.
Ко-креативно диалогово пространство	Релационна, информационна и комуникативна среда, в която човешките и машинно генерираните приноси се представят, интерпретират, проверяват, сравняват и преобразуват във връзка с определена задача.	Не е физическо място, общо съзнание или самостоятелен субект. Границите му се определят от достъпния, релевантния и проследимия контекст.
Конфигурационен модел	Модел, според който проявлението на интегриран интелект зависи от: конкретното съчетание и функционалната свързаност на човешките участници, системите с ИИ, ролите, резонансните процеси, задачата, контекста, ко-креаторството и резултатите.	Не предполага линейна последователност, еднаква тежест или задължително максимално проявление на всеки компонент. Различни конфигурации могат да породят различни профили на взаимодействие.
Креативно стимулиране	Функционално въздействие на машинно генерирани предложения, различия, комбинации, въпроси или алтернативи върху човешкото създаване и преосмисляне на идеи.	Може да бъде еднократно и не доказва само по себе си творчески резонанс или ко-креаторство. Прераства в творчески резонанс при проследимо рекурсивно преобразуване на разбирането или пространството на решенията.
Проследимост	Възможност приносите, източниците, проверките, измененията, решенията и връзката им с крайния продукт да бъдат реконструирани чрез запазени доказателствени следи.	Не е само административно документиране, а методологично условие за обосновано заключение относно процеса. Не гарантира сама по себе си ко-креаторство, качество или истинност.

<i>Термин</i>	<i>Работно определение в настоящия труд</i>	<i>Разграничение и бележка</i>
Рекурсивен цикъл	Повтаряща се последователност от принос, човешка интерпретация, проверка или обратна връзка и последващо преобразуване на предложението, целта, критериите или продукта.	Всеки цикъл изменя достъпния контекст и условията за следващите действия. Простото повторение без съдържателно преобразуване не е достатъчно.
Релационна конфигурация	Организация на функционалните отношения между човека, системата с ИИ, задачата, данните, използваните средства и социално-институционалната среда.	Наблюдаваните системни свойства се отнасят до организацията на взаимодействието, а не се приписват автоматично на изолиран компонент.
Система с ИИ	Техническа и организационна цялост, включваща един или повече модели с ИИ, интерфейси, данни, правила, инфраструктура и процедури за човешка употреба и контрол.	По-широко понятие от „модел с ИИ“ и „агент с ИИ“.
Модел с ИИ	Изчислителен механизъм, обучен или проектиран да обработва входни данни и да генерира прогнози, класификации, представяния, съдържание или други изходи.	Не е автоматично система или автономен агент. Генеративната способност не предполага съзнание, разбиране или морална отговорност.
Агент с ИИ	Функционална техническа конфигурация, която възприема информация, избира или изпълнява действия и използва модел с ИИ спрямо зададени цели, правила и ограничения.	„Агент“ се използва във функционален и инженерно-технически смисъл. Терминът не предполага човешкоподобна агентност, автономно намерение, съзнание или морална отговорност.
Социотехническа система	Организирана система от хора, технологии, правила, ресурси, практики и институционални условия, чието взаимодействие определя функционирането и последствията от употребата на технологията.	Техническите свойства и резултати не се разглеждат изолирано от човешките практики, властовите отношения, нормите и контекста на употреба.
Съвместно сътворен продукт	Идея, текст, модел, фигура, решение, проект или друг артефакт, чиято конкретна форма е възникнала чрез рекурсивно и проследимо преобразуване на различни човешки и машинно генерирани приноси.	Не е механичен сбор от приноси. Качеството, новостта или убедителността не доказват сами по себе си ко-креаторство и не придават автоматично статут на знание.

<i>Термин</i>	<i>Работно определение в настоящия труд</i>	<i>Разграничение и бележка</i>
Творчески резонанс	Проследима рекурсивна динамика, при която човешките и машинно генерираните приноси преобразуват критериите, идеите или пространството на възможните решения в творческия процес.	Не е единично генериране, натрупване на варианти или автоматична взаимност между равнопоставени творчески субекти. Не замества човешката креативност и ценностна преценка.
Функционална взаимност	Процесна обусловеност, при която последващият принос на единия компонент зависи от предходния принос на другия и изменя условията за следващото действие във взаимодействието.	Обозначава двупосочна функционална зависимост, а не симетрична агентност, еднакво разбиране, равнопоставено авторство или споделена морална отговорност.
Функционална роля на ИИ	Конкретната функция, изпълнявана от системата с ИИ в определен епизод, например инструмент, асистент, когнитивен резонатор, модулатор или асиметричен ко-креатор.	Ролята е аналитична, контекстуална и променлива. Тя не е постоянно същностно свойство на системата и не определя нейния морален или правен статут.
Човешки агент	Когнитивен, социален и морален субект, който формулира или приема цели, интерпретира контекста, оценява информацията, създава и договаря значения, взема решения и носи отговорност за действията и резултатите.	Може да бъде индивид или организиран човешки колектив. Не е пасивен получател на машинни изходи, но действията му са обусловени и от социални, институционални и технически фактори.
Човешко утвърждаване и отговорност	Човешката функция по интерпретиране, проверяване, обосноваване, ценностно оценяване, приемане, изменение или отхвърляне и окончателно утвърждаване на резултата.	Функционалният принос на ИИ не прехвърля върху системата академичната, професионалната или моралната отговорност. Авторският и правният статут се определят според приложимите норми и институционалния контекст.

КОНЦЕПТУАЛНИ РАЗЛИЧИЯ

Възникване, автономност и съзнание

Възникващото системно свойство принадлежи на организацията и динамиката на взаимодействието, а не непременно на някой от участващите компоненти поотделно. От наличието му не следват автономна агентност или намерение, съзнание, общ интелект, морален статут или равнопоставена субектност на човека и ИИ.

Емоционален резонанс и емоция

Емоционалният резонанс описва функционално въздействие на адаптирания машинен отговор върху човешката ангажираност, емоционална нагласа, интерпретация или последващо действие. Той не означава, че ИИ преживява емоции, притежава човешка емпатия или разбира чувствата по начина, по който ги разбира човекът. Резонансът е характеристика на наблюдаваното взаимодействие, а не доказателство за вътрешно емоционално състояние на системата.

Идентифициране, измерване, оценяване и валидиране

Идентифицирането установява дали наблюдаваното взаимодействие проявява предварително определени критерии и индикатори. Измерването определя степента, честотата, интензивността или профила на тези проявления. Оценяването преценява качеството, значимостта или приемливостта на процеса и резултата спрямо зададени стандарти. Валидирането проверява дали методът и показателите действително измерват или установяват предвидения конструкт. Надеждността, която се отнася до устойчивостта и повторемостта на измерването, трябва да се проверява отделно. Това са взаимосвързани, но различни аналитични задачи.

Интегриран и увеличен интелект

Увеличеният интелект обозначава усилването или разширяването на човешките познавателни и практически способности чрез технологии. В настоящата рамка интегрираният интелект предполага допълнително рекурсивно и проследимо преобразуване на човешките и машинните приноси

в процеса на решаване на задача или създаване на продукт. Поради това обикновеното техническо подпомагане не е достатъчно основание за идентифициране на интегриран интелект.

Интегриран интелект и разпределено или разширено познание

Разпределеното познание обяснява как познавателните функции и представяния се разполагат между хора, артефакти и среда. Разширеното познание разглежда условията, при които външни средства могат да бъдат функционално включени в познавателния процес. Конструктът „интегриран интелект“, както е дефиниран в настоящия труд, добавя критерии за резонансно и рекурсивно преобразуване, функционално асиметрично ко-креаторство, проследима генеалогия на продукта и запазена човешка смислова, ценностна и морална отговорност.

Интегриран, хибриден и симбиотичен интелект

Хибридният интелект обикновено акцентира върху координираното съчетаване на различни човешки и машинни способности. Симбиотичният интелект поставя ударение върху продължителното функционално допълване и взаимната адаптация. Интегрираният интелект, в смисъла, възприет в настоящия труд, обозначава възможна релационна конфигурация, характеризираща се с резонанс, рекурсивно преобразуване на приносите, проследима генеалогия на продукта и окончателно човешко утвърждаване. Различията са аналитични и частично припокриващи се, а не абсолютни.

Интелигентност и интелект

В настоящия труд „интелигентност“ обозначава способност, качество или съвкупност от функции, а „интелект“ – организирана система или конфигурация, в която се осъществяват интелектуални функции. Поради това терминът се предава последователно като „интегриран интелект“, когато се назовава конструктът като цялостна релационна организация. Разграничението е работно и терминологично, а не общовалидно за всички научни традиции.

Информация и знание

Машинният изход първоначално има статут на генерирана информация, хипотеза, интерпретативно предложение или възможно решение. Той не придобива автоматично статут на обосновано знание. Това изисква човешка интерпретация, критична проверка, свързване с доказателства и включване в подходящ смислов контекст, а когато областта го изисква – и институционално или експертно валидиране. Информацията и знанието са сродни, но

несъвпадащи епистемични категории. Те не са непременно последователни етапи на един линеен процес.

Креативно стимулиране, творчески резонанс и ко-креаторство

Креативното стимулиране представлява провокиране на идея, асоциация или нова посока в човешкото мислене. Творческият резонанс обозначава проследимо и рекурсивно преобразуване на разбирането, критериите или пространството на възможните решения. Функционалното ко-креаторство предполага различни, итеративно преобразувани и проследими приноси към създаването на продукта, при запазена човешка преценка и отговорност. Понятията обозначават различни равнища и различна доказателствена сила и не трябва да се използват като взаимозаменяеми.

Машинна генеративност и човешка креативност

Машинната генеративност произвежда варианти чрез моделиране на закономерности в данните и вероятно генериране на изходи. Човешката креативност е свързана с човешка агентност и може да включва намерение, преживяване, въображение, контекстуално осмисляне и ценностна преценка. Изкуственият интелект може функционално да стимулира, разширява и преобразува творческия процес, без от това да следва, че притежава човешко творческо преживяване, автономно намерение или морална отговорност.

Процесни и резултатни доказателства

Процесните доказателства показват как са възникнали, взаимодействали и били преобразувани човешките и машинните приноси. Резултатните доказателства се отнасят до качествата, приложимостта и въздействието на създадения продукт. Високото качество на продукта не доказва само по себе си интегриран процес, както и наличието на интегриран процес не гарантира качествен или надежден резултат. Поради това двата вида доказателства са необходими, но не са взаимозаменяеми.

Резонанс и полезен или качествен отговор

Резонансът обозначава процесна промяна в последващото човешко разбиране, формулиране на въпроси, критерии, избори или действия, която може да бъде проследена във взаимодействието. Ползността, красноречието, убедителността или високото качество на единичен машинен отговор не доказват сами по себе си резонанс. Необходимо е да се установи как отговорът е бил интерпретиран, проверен и преобразуван в следващите стъпки на човека.

Съвместно сътворен продукт и знание

Съвместно сътвореният продукт може да бъде идея, текст, модел, решение, проект или друг артефакт, в чието формиране са налице различни и проследими човешки и машинни приноси. Не всеки продукт представлява знание. Епистемичен статут могат да придобият съдържащите се в него твърдения, обяснения или модели след човешка интерпретация, обосноваване, проверка и контекстуализация. Новостта, убедителността и качеството сами по себе си не са достатъчни.

Функционален ко-креатор и автор

Системата с ИИ може да изпълнява функционална ко-креаторска роля, когато генерираният от нея принос е различим, преобразуващ и проследим в развитието на продукта. Това аналитично определение не превръща ИИ в автономен субект, автор или носител на морална отговорност и епистемична власт. Авторството, отговорността за проверката и решението за използване и публикуване на продукта остават човешки. Правният статут следва да се определя съобразно приложимото законодателство и конкретния институционален контекст.

БИБЛИОГРАФИЯ

(А) Библиография по азбучен ред

- Akata, Z., Balliet, D., de Rijke, M., Dignum, F., Dignum, V., Eiben, G., Fokkens, A., Grossi, D., Hindriks, K., Hoos, H., Hung, H., Jonker, C., Monz, C., Neerincx, M., Oliehoek, F., Prakken, H., Schlobach, S., van der Gaag, L., van Harmelen, F., & Welling, M. (2020). A research agenda for hybrid intelligence: Augmenting human intellect with collaborative, adaptive, responsible, and explainable artificial intelligence. *Computer*, 53(8), 18–28.
- Alavi, M., & Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly*, 25(1), 107–136. <https://doi.org/10.2307/3250961>
- Albrecht, K. (2003). *The power of minds at work: Organizational intelligence in action*. AMACOM.
- Amabile, T. M. (1996). *Creativity in context: Update to the social psychology of creativity*. Westview Press.
- Amershi, S., Weld, D., Vorvoreanu, M., Fournery, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). Guidelines for human–AI interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Article 3, pp. 1–13). Association for Computing Machinery. <https://doi.org/10.1145/3290605.3300233>
- Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* Oxford University Press.
- Anderson, J. R. (2020). *Cognitive psychology and its implications (9th ed.)*. Worth Publishers.
- Anthony, M. (2008). *Integrated intelligence: Classical and contemporary depictions of mind and intelligence and their educational implications*. Brill.
- Anthropic. (2024). The Claude 3 model family: Opus, Sonnet, Haiku [Model card]. <https://www-cdn.anthropic.com/f2986af8d052f26236f6251da62d16172cfabd6e/claude-3-model-card.pdf>
- Argyris, C., & Schön, D. A. (1978). *Organizational learning: A theory of action perspective*. Addison-Wesley.
- Argyris, C., & Schön, D. A. (1996). *Organizational learning II: Theory, method, and practice*. Addison-Wesley.
- Aristotle. (2004). *Metaphysics (H. Lawson-Tancred, Trans.)*. Penguin Books. (Original work written ca. 350 BCE)

- Aristotle. (2009). *Nicomachean ethics* (W. D. Ross, Trans.). Oxford University Press. (Original work written ca. 340 BCE)
- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115.
- Ashby, W. R. (1956). *An introduction to cybernetics*. Chapman & Hall.
- Bakhtin, M. M. (1981). *The dialogic imagination: Four essays* (M. Holquist, Ed.; C. Emerson & M. Holquist, Trans.). University of Texas Press.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.
- Bandura, A. (1991). Social cognitive theory of moral thought and action. In W. M. Kurtines & J. L. Gewirtz (Eds.), *Handbook of moral behaviour and development: Vol. 1. Theory* (pp. 45–103). Lawrence Erlbaum Associates.
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52, 1–26.
- Bansal, G., Nushi, B., Kamar, E., Lasecki, W. S., Weld, D. S., & Horvitz, E. (2019). Updates in human-AI teams: Understanding and addressing the performance/compatibility tradeoff. *Proceedings of the AAAI Conference on Artificial Intelligence*, 33(1), 2429–2437. <https://doi.org/10.1609/aaai.v33i01.33012429>
- Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617–645.
- Bedau, M. A. (1997). Weak emergence. *Philosophical Perspectives*, 11, 375–399. <https://doi.org/10.1111/0029-4624.31.s11.17>
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610–623). Association for Computing Machinery. <https://doi.org/10.1145/3442188.3445922>
- Bengio, Y., Hinton, G., Yao, A., Song, D., Abbeel, P., Darrell, T., Harari, Y. N., Zhang, Y.-Q., Xue, L., Shalev-Shwartz, S., Hadfield, G., Clune, J., Maharaj, T., Hutter, F., Baydin, A. G., McIlraith, S., Gao, Q., Acharya, A., Krueger, D., . . . Mindermann, S. (2023). Managing extreme AI risks amid rapid progress [Preprint]. arXiv. <https://arxiv.org/abs/2310.17688>
- Bertalanffy, L. von. (1968). *General system theory: Foundations, development, applications*. George Braziller.
- Bhaskar, R. (1975). *A realist theory of science*. Leeds Books.
- Bickmore, T. W., & Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction*, 12(2), 293–327. <https://doi.org/10.1145/1067860.1067867>
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and be-

- havioral research: A primer. *Frontiers in Public Health*, 6(149). <https://doi.org/10.3389/fpubh.2018.00149>
- Boden, M. A. (2004). *The creative mind: Myths and mechanisms (2nd ed.)*. Routledge.
- Boden, M. A. (2016). *AI: Its nature and future*. Oxford University Press.
- Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., Brynjolfsson, E., Buch, S., Card, D., Castellon, R., Chatterji, N., Chen, A., Creel, K., Davis, J. Q., Demszky, D., ... Liang, P. (2021). On the opportunities and risks of foundation models. [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2108.07258>
- Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brave, S., & Nass, C. (2002). Emotion in human–computer interaction. In J. Jacko & A. Sears (Eds.), *The human–computer interaction handbook: Fundamentals, evolving technologies and emerging applications* (pp. 81–96). Lawrence Erlbaum Associates.
- Brave, S., Nass, C., & Hutchinson, K. (2005). Computers that care: Investigating the effects of orientation of emotion exhibited by an embodied computer agent. *International Journal of Human-Computer Studies*, 62(2), 161–178. <https://doi.org/10.1016/j.ijhcs.2004.11.002>
- Bredeweg, B., & Kragten, M. (2022). *Requirements and challenges for hybrid intelligence: A case study in education*. *Frontiers in Artificial Intelligence*, 5.
- Broad, C. D. (1925). *The mind and its place in nature*. Kegan Paul, Trench, Trubner & Co.
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., . . . Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901. <https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfc4967418bfb8ac142f64a-Abstract.html>
- Bruner, J. (1990). *Acts of meaning*. Harvard University Press.
- Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). *Generative AI at work (NBER Working Paper No. 31161)*. National Bureau of Economic Research. <https://doi.org/10.3386/w31161>
- Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.
- Buber, M. (1970). *I and Thou* (W. Kaufmann, Trans.). Charles Scribner's Sons. (Original work published 1923)
- Bunge, M. (1977). *Treatise on basic philosophy: Vol. 3. Ontology I: The furniture of the world*. D. Reidel. <https://doi.org/10.1007/978-94-010-9921-7>
- Bunge, M. (2003). *Emergence and convergence: Qualitative novelty and the unity of knowledge*. University of Toronto Press.
- Burrell, J. (2016). How the machine “thinks”: Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1), 1–12. <https://doi.org/10.1177/2053951715622512>

- Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing*, 1(1), 18–37. <https://doi.org/10.1109/T-AFFC.2010.1>
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81–105. <https://doi.org/10.1037/h0046016>
- Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In N. J. Castellan Jr. (Ed.), *Individual and group decision making: Current issues* (pp. 221–246). Lawrence Erlbaum Associates.
- Capra, F. (1996). *The web of life: A new scientific understanding of living systems*. Anchor Books.
- Capra, F., & Luisi, P. L. (2014). *The systems view of life: A unifying vision*. Cambridge University Press.
- Choo, C. W. (1998). *The knowing organisation: How organisations use information to construct meaning, create knowledge, and make decisions*. Oxford University Press.
- Clark, A. (1997). *Being there: Putting brain, body, and world together again*. MIT Press.
- Clark, A., & Chalmers, D. J. (1998). The extended mind. *Analysis*, 58(1), 7–19. <https://doi.org/10.1093/analys/58.1.7>
- Clore, G. L., & Storbeck, J. (2006). Affect as information about liking, efficacy, and importance. In J. P. Forgas (Ed.), *Affect in social thinking and behaviour* (pp. 123–141). Psychology Press.
- Coeckelbergh, M. (2020). *AI ethics*. MIT Press.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46. <https://doi.org/10.1177/001316446002000104>
- Colton, S. (2008). Creativity versus the perception of creativity in computational systems. In *Proceedings of the AAAI Spring Symposium on Creative Intelligent Systems*.
- Committee on Publication Ethics. (2023). *Authorship and AI tools*. <https://publication-ethics.org/guidance/cope-position/authorship-and-ai-tools>
- Cooper, A., Reimann, R., Cronin, D., & Noessel, C. (2014). *About face: The essentials of interaction design (4th ed.)*. Wiley.
- Corning, P. A. (2002). The re-emergence of “emergence”: A venerable concept in search of a theory. *Complexity*, 7(6), 18–30. <https://doi.org/10.1002/cplx.10043>
- Council of Europe. (2024). *Framework convention on artificial intelligence and human rights, democracy and the rule of law*. Council of Europe.
- Crossan, M. M., Lane, H. W., & White, R. E. (1999). An organisational learning framework: From intuition to institution. *Academy of Management Review*, 24(3), 522–537. <https://doi.org/10.5465/amr.1999.2202135>
- Crutchfield, J. P. (1994). Is anything ever new? Considering emergence. In G. A. Cowan, D. Pines, & D. Meltzer (Eds.), *Complexity: Metaphors, models, and reality* (pp. 479–497). Addison-Wesley.

- Csikszentmihalyi, M. (1996). *Creativity: Flow and the psychology of discovery and invention*. HarperCollins.
- Csikszentmihalyi, M. (1999). Implications of a systems perspective for the study of creativity. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 313–335). Cambridge University Press. <https://doi.org/10.1017/CBO9780511807916.018>
- Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. Putnam.
- Damasio, A. R. (1999). *The feeling of what happens: Body and emotion in the making of consciousness*. Harcourt Brace.
- Davenport, T. H., & Kirby, J. (2016). *Only humans need apply: Winners and losers in the age of smart machines*. Harper Business.
- Davidson, R. J., Scherer, K. R., & Goldsmith, H. H. (Eds.). (2003). *Handbook of affective sciences*. Oxford University Press.
- Davis, N. (2013). Human-computer co-creativity: Blending human and computational creativity. In *The Doctoral Consortium at AIIDE 2013* (pp. 9–12). AAAI Press. <https://ojs.aaai.org/index.php/AIIDE/article/view/12603>
- Davis, N. (2026). Interaction-centred intelligence: Toward an interaction-based theory of human–AI co-creation [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2606.00807>
- Davis, N., Hsiao, C.-P., Popova, Y., Magerko, B., & Liapis, A. (2015a). An enactive model of creativity for computational collaboration and co-creation. In *Proceedings of the Sixth International Conference on Computational Creativity*.
- Davis, N., Hsiao, C.-P., Singh, K. Y., Li, L., Moningi, S., & Magerko, B. (2015b). Drawing Apprentice: An enactive co-creative agent for artistic collaboration. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition* (pp. 185–186). Association for Computing Machinery. <https://doi.org/10.1145/2757226.2764555>
- de Haan, J. (2006). How emergence arises. *Ecological Complexity*, 3(4), 293–301. <https://doi.org/10.1016/j.ecocom.2007.02.003>
- Dellermann, D., Ebel, P., Söllner, M., & Leimeister, J. M. (2019). Hybrid intelligence. *Business & Information Systems Engineering*, 61(5), 637–643. <https://doi.org/10.1007/s12599-019-00595-2>
- Dewey, J. (1938). *Logic: The theory of inquiry*. Henry Holt and Company.
- Dix, A., Finlay, J., Abowd, G. D., & Beale, R. (2004). *Human–computer interaction (3rd ed.)*. Pearson.
- D'Mello, S., & Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, 22(2), 145–157. <https://doi.org/10.1016/j.learninstruc.2011.10.001>
- Doshi-Velez, F., & Kim, B. (2017). *Towards a rigorous science of interpretable machine learning*. arXiv. <https://arxiv.org/abs/1702.08608>
- Easterby-Smith, M., & Lyles, M. A. (Eds.). (2011). *Handbook of organisational learning and knowledge management (2nd ed.)*. Wiley.
- Edmondson, A. C. (2019). *The fearless organisation: Creating psychological safety in the workplace for learning, innovation, and growth*. John Wiley & Sons.

- Emirbayer, M., & Mische, A. (1998). What is agency? *American Journal of Sociology*, 103(4), 962–1023. <https://doi.org/10.1086/231294>
- Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1), 32–64. <https://doi.org/10.1518/001872095779049543>
- Engel, D., Woolley, A. W., Jing, L. X., Chabris, C. F., & Malone, T. W. (2014). Reading the mind in the eyes or reading between the lines? Theory of Mind predicts collective intelligence equally well online and face-to-face. *PLoS ONE*, 9(12), e115212. <https://doi.org/10.1371/journal.pone.0115212>
- Engelbart, D. C. (1962). *Augmenting human intellect: A conceptual framework* (AFOSR-3233). Stanford Research Institute. <http://dougengelbart.org/pubs/augment-3906-Framework.html>
- Epley, N., Waytz, A., & Cacioppo, J. T. (2007). On seeing human: A three-factor theory of anthropomorphism. *Psychological Review*, 114(4), 864–886. <https://doi.org/10.1037/0033-295X.114.4.864>
- Ericsson, K. A., & Pool, R. (2016). *Peak: Secrets from the new science of expertise*. Houghton Mifflin Harcourt.
- European Commission. (2020). *White paper on artificial intelligence: A European approach to excellence and trust*. Publications Office of the European Union.
- European Parliament & Council of the European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). *Official Journal of the European Union*, L, 2024/1689. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>
- Faraj, S., Pachidi, S., & Sayegh, K. (2018). Working and organizing in the age of the learning algorithm. *Information and Organization*, 28(1), 62–70. <https://doi.org/10.1016/j.infoandorg.2018.02.005>
- Floridi, L. (2011). *The philosophy of information*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199232383.001.0001>
- Floridi, L. (2014). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford University Press.
- Floridi, L. (2019a). Establishing the rules for building trustworthy AI. *Nature Machine Intelligence*, 1(6), 261–262.
- Floridi, L., & Cows, J. (2019b). *A unified framework of five principles for AI in society*. Harvard Data Science Review, 1(1).
- Floridi, L., Cows, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, 28(4), 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Floridi, L. (Ed.). (2023). *The ethics of artificial intelligence: Principles, challenges, and opportunities*. Oxford University Press.
- Forgas, J. P. (Ed.). (2006). *Affect in social thinking and behaviour*. Psychology Press.

- Franklin, S., & Graesser, A. (1997). Is it an agent, or just a program? A taxonomy for autonomous agents. In J. P. Müller, M. J. Wooldridge, & N. R. Jennings (Eds.), *Intelligent agents III: Agent theories, architectures, and languages* (Lecture Notes in Computer Science, Vol. 1193, pp. 21–35). Springer. <https://doi.org/10.1007/BFb0013570>
- Fuchs, T., & De Jaegher, H. (2009). Enactive intersubjectivity: Participatory sense-making and mutual incorporation. *Phenomenology and the Cognitive Sciences*, 8(4), 465–486. <https://doi.org/10.1007/s11097-009-9136-4>
- Gadamer, H.-G. (2004). *Truth and method* (2nd rev. ed., J. Weinsheimer & D. G. Marshall, Trans.). Continuum. (Original work published 1960)
- Gao, Q., Xu, W., Pan, H., Shen, M., & Gao, Z. (2025). Human-centered human–AI collaboration (HCHAC) [Preprint]. arXiv. <https://arxiv.org/abs/2505.22477>
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. Basic Books.
- Gigerenzer, G. (2007). *Gut feelings: The intelligence of the unconscious*. Viking.
- Goldstein, J. (1999). Emergence as a construct: History and issues. *Emergence*, 1(1), 49–72. https://doi.org/10.1207/S15327000EM0101_4
- Goleman, D. (1995). *Emotional intelligence*. Bantam Books.
- Goleman, D. (1998). *Working with emotional intelligence*. Bantam Books.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Gemini Team, Google. (2023). Gemini: A family of highly capable multimodal models [Preprint]. arXiv. <https://arxiv.org/abs/2312.11805>
- Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, 17(S2), 109–122. <https://doi.org/10.1002/smj.4250171110>
- Graziano, M. S. A. (2013). *Consciousness and the social brain*. Oxford University Press.
- Grudin, J. (2017). From tool to partner: The evolution of human–computer interaction. *Synthesis Lectures on Human-Centered Informatics*, 10(1), 1–183.
- Hao, Z., Liu, X., Fan, J., Long, Y., Yu, J., Chen, W., & Zhang, Y. (2026). Unpacking interaction profiles and strategies in human–AI collaborative problem-solving: A cognitive distribution and regulation perspective [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2603.21288>
- Heidegger, M. (1962). *Being and time* (J. Macquarrie & E. Robinson, Trans.). Blackwell. (Original work published 1927)
- Hemmer, P., Schemmer, M., Vössing, M., & Satzger, G. (2024). *Human–AI collaboration: State of the art and future directions*. Business & Information Systems Engineering.
- Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. *MIS Quarterly*, 28(1), 75–105.
- High-Level Expert Group on Artificial Intelligence. (2019). *Ethics guidelines for trustworthy AI*. European Commission. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- Hoc, J.-M. (2000). From human–machine interaction to human–machine cooperation. *Ergonomics*, 43(7), 833–843. <https://doi.org/10.1080/001401300409044>

- Hollan, J., Hutchins, E., & Kirsh, D. (2000). Distributed cognition: Toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction*, 7(2), 174–196. <https://doi.org/10.1145/353485.353487>
- Holland, J. H. (1992). Complex adaptive systems. *Daedalus*, 121(1), 17–30.
- Holland, J. H. (1998). *Emergence: From chaos to order*. Oxford University Press.
- Hollnagel, E., & Woods, D. D. (2005). *Joint cognitive systems: Foundations of cognitive systems engineering*. CRC Press.
- Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences of the United States of America*, 101(46), 16385–16389. <https://doi.org/10.1073/pnas.0403723101>
- Husserl, E. (1982). *Ideas pertaining to a pure phenomenology and to a phenomenological philosophy: First book* (F. Kersten, Trans.). Martinus Nijhoff. (Original work published 1913)
- Hutchins, E. (1995). *Cognition in the wild*. MIT Press. <https://doi.org/10.7551/mitpress/1881.001.0001>
- IBM Research. (2016). *Augmented intelligence: Intelligence amplification for human decision making*. <https://research.ibm.com/>
- IEEE. (2019). *Ethically aligned design: A vision for prioritizing human well-being with autonomous and intelligent systems (1st ed.)*. IEEE Standards Association.
- International Organization for Standardization. (2019). *ISO 9241-210:2019 Ergonomics of human-system interaction—Part 210: Human-centred design for interactive systems*. International Organization for Standardization.
- International Organization for Standardization. (2020). *ISO/IEC TR 24028:2020 Artificial intelligence—Overview of trustworthiness in artificial intelligence*. ISO.
- International Organization for Standardization. (2022a). *ISO/IEC 22989:2022 Artificial intelligence—Artificial intelligence concepts and terminology*. ISO.
- International Organization for Standardization. (2022b). *ISO/IEC 38507:2022—Information technology—Governance of IT—Governance implications of the use of artificial intelligence by organisations*. International Organization for Standardization.
- International Organization for Standardization. (2022c). *ISO/IEC TR 24368:2022—Artificial intelligence—Overview of ethical and societal concerns*. International Organization for Standardization.
- International Organization for Standardization. (2023a). *ISO 31700-1:2023 Consumer protection—Privacy by design for consumer goods and services—Part 1: High-level requirements*. ISO.
- International Organization for Standardization. (2023b). *ISO/IEC 5338:2023—Information technology—Artificial intelligence—AI system life cycle processes*. International Organization for Standardization.
- International Organization for Standardization. (2023c). *ISO/IEC 23894:2023 Information technology—Artificial intelligence—Guidance on risk management*. ISO.

- International Organization for Standardization. (2023d). *ISO/IEC 42001:2023 Information technology—Artificial intelligence—Management system*. ISO.
- Jarrahi, M. H. (2018). Artificial intelligence and the future of work: Human–AI symbiosis in organisational decision making. *Business Horizons*, 61(4), 577–586. <https://doi.org/10.1016/j.bushor.2018.03.007>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Harvard University Press.
- Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *The Journal of the Learning Sciences*, 4(1), 39–103. https://doi.org/10.1207/s15327809jls0401_2
- Jordanous, A. (2016a). Four perspectives on computational creativity in theory and in practice. *Connection Science*, 28(2), 194–216.
- Jordanous, A., & Keller, B. (2016b). Modelling creativity: Identifying key components through a corpus-based approach. *PLoS ONE*, 11(10), e0162959.
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
- Kamar, E. (2016). Directions in hybrid intelligence: Complementing AI systems with human intelligence. In *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence (IJCAI-16)* (pp. 4070–4073). AAAI Press. <https://www.ijcai.org/Proceedings/16/Papers/603.pdf>
- Kantosalo, A., & Takala, T. (2020). Five C’s for human–computer co-creativity: An update on classical creativity perspectives. In A. Cardoso, P. Machado, T. Veale, & J. M. Cunha (Eds.), *Proceedings of the Eleventh International Conference on Computational Creativity* (pp. 17–24). Association for Computational Creativity.
- Kantosalo, A., & Toivonen, H. (2016). Modes for creative human-computer collaboration: Alternating and task-divided co-creativity. In F. Pachet, A. Cardoso, V. Corruble, & F. Ghedini (Eds.), *Proceedings of the Seventh International Conference on Computational Creativity* (pp. 77–84). Sony CSL Paris.
- Karimi, P., Maher, M. L., Davis, N., & Grace, K. (2018). Evaluating creativity in computational co-creative systems. In *Proceedings of the Ninth International Conference on Computational Creativity* (pp. 104–111).
- Kelly III, J. E. (2015). *Smart machines: IBM’s Watson and the era of cognitive computing*. Columbia University Press.
- Kline, R. B. (2023). *Principles and practice of structural equation modelling* (5th ed.). Guilford Press.
- Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice Hall.
- Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (4th ed.). SAGE Publications.
- Kuhn, T. S. (2012). *The structure of scientific revolutions* (4th ed.). University of Chicago Press.

- Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 91–196). Cambridge University Press.
- Lakatos, I. (1978). *The methodology of scientific research programmes*. Cambridge University Press.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.
- Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human Factors*, 46(1), 50–80. https://doi.org/10.1518/hfes.46.1.50_30392
- Leonardi, P. M. (2011). When flexible routines meet flexible technologies: Affordance, constraint, and the imbrication of human and material agencies. *MIS Quarterly*, 35(1), 147–167. <https://doi.org/10.2307/23043493>
- Lévy, P. (1997). *Collective intelligence: Mankind's emerging world in cyberspace*. Perseus Books.
- Lichtenthaler, U. (2020). *Integrated intelligence: Combining human and artificial intelligence for competitive advantage*. Campus Verlag.
- Licklider, J. C. R. (1960). Man-computer symbiosis. *IRE Transactions on Human Factors in Electronics*, HFE-1(1), 4–11. <https://doi.org/10.1109/THFE2.1960.4503259>
- List, C., & Pettit, P. (2011). *Group agency: The possibility, design, and status of corporate agents*. Oxford University Press.
- Liu, B. (2023). *Artificial intelligence and knowledge engineering*. Springer.
- Long, D., & Magerko, B. (2020). *What is AI literacy? Competencies and design considerations*. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 1–16.
- Lubart, T. (2005). How can computers be partners in the creative process: Classification and commentary on the special issue. *International Journal of Human-Computer Studies*, 63(4–5), 365–369. <https://doi.org/10.1016/j.ijhcs.2005.04.002>
- Luckin, R. (2018). *Machine learning and human intelligence: The future of education for the 21st century*. UCL IOE Press.
- Mackay, W. E. (1990). *Users and customizable software: A co-adaptive phenomenon* [Doctoral dissertation, Massachusetts Institute of Technology]. DSpace@MIT. <http://hdl.handle.net/1721.1/14087>
- Malone, T. W. (2018). *Superminds: The surprising power of people and computers thinking together*. Little, Brown Spark.
- Malone, T. W., Laubacher, R., & Dellarocas, C. (2010). The collective intelligence genome. *MIT Sloan Management Review*, 51(3), 21–31.
- March, J. G. (1991). Exploration and exploitation in organisational learning. *Organization Science*, 2(1), 71–87. <https://doi.org/10.1287/orsc.2.1.71>
- Matsuda, T. (1992). Organizational intelligence: Its significance as a process and as a product. In *Proceedings of the International Symposium on the Management of Industrial and Corporate Knowledge*.
- Maturana, H. R., & Varela, F. J. (1980). *Autopoiesis and cognition: The realization of the living*. D. Reidel.

- Mayer, R. E. (2021). *Multimedia learning (3rd ed.)*. Cambridge University Press.
- McCarthy, J. (2007). *What is artificial intelligence?* Stanford University. <http://jmc.stanford.edu/artificial-intelligence/>
- McCormack, J., & d’Inverno, M. (Eds.). (2012a). *Computers and creativity*. Springer. <https://doi.org/10.1007/978-3-642-31727-9>
- McCormack, J., & d’Inverno, M. (2012b). Computers and creativity: The road ahead. In J. McCormack & M. d’Inverno (Eds.), *Computers and creativity* (pp. 421–424). Springer. https://doi.org/10.1007/978-3-642-31727-9_16
- McDermott, D. (1976). Artificial intelligence meets natural stupidity. *ACM SIGART Bulletin*, 57, 4–9.
- McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods*, 1(1), 30–46. <https://doi.org/10.1037/1082-989X.1.1.30>
- Meadows, D. H. (2008). *Thinking in systems: A primer*. Chelsea Green Publishing.
- Mercer, N. (2000). *Words and minds: How we use language to think together*. Routledge.
- Merleau-Ponty, M. (2012). *Phenomenology of perception (D. A. Landes, Trans.)*. Routledge. (Original work published 1945)
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons’ responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749. <https://doi.org/10.1037/0003-066X.50.9.741>
- Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1–38. <https://doi.org/10.1016/j.artint.2018.07.007>
- Minsky, M. (1986). *The society of mind*. Simon & Schuster.
- Mitchell, M. (2009). *Complexity: A guided tour*. Oxford University Press.
- Mitchell, M. (2019). *Artificial intelligence: A guide for thinking humans*. Farrar, Straus and Giroux.
- Moravec, H. (1988). *Mind children: The future of robot and human intelligence*. Harvard University Press.
- Morin, E. (2008). *On complexity*. Hampton Press.
- Nass, C., & Brave, S. (2005). *Wired for speech: How voice activates and advances the human–computer relationship*. MIT Press.
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81–103. <https://doi.org/10.1111/0022-4537.00153>
- National Institute of Standards and Technology. (2023). *Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1)*. U.S. Department of Commerce. <https://doi.org/10.6028/NIST.AI.100-1>
- National Institute of Standards and Technology. (2024). Artificial intelligence risk management framework: Generative artificial intelligence profile (NIST AI 600-1). U.S. Department of Commerce. <https://doi.org/10.6028/NIST.AI.600-1>
- National Institute of Standards and Technology. (n.d.). *AI Risk Management Framework Playbook*. https://airc.nist.gov/AI_RMF_Knowledge_Base/Playbook

- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Prentice-Hall.
- Nikolaïdis, S., Hsu, D., & Srinivasa, S. (2017). Human–robot mutual adaptation in collaborative tasks: Models and experiments. *The International Journal of Robotics Research*, 36(5–7), 618–634. <https://doi.org/10.1177/0278364917690593>
- Nonaka, I. (1994). A dynamic theory of organizational knowledge creation. *Organization Science*, 5(1), 14–37.
- Nonaka, I., & Takeuchi, H. (1995). *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford University Press.
- Norman, D. A. (1988). *The psychology of everyday things*. Basic Books.
- Norman, D. A. (2004). *Emotional design: Why we love (or hate) everyday things*. Basic Books.
- Norman, D. A. (2013). *The design of everyday things (Rev. and expanded ed.)*. Basic Books.
- Norman, D. A. (2023). *Design for a better world: Meaningful, sustainable, humanity-centered*. MIT Press.
- OECD. (2019). *OECD principles on artificial intelligence*. OECD Publishing. <https://oecd.ai>
- OECD. (n.d.). *OECD AI Observatory*. <https://oecd.ai>
- OpenAI. (2023). GPT-4 technical report [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2303.08774>
- OpenAI. (2025a, April 14). Introducing GPT-4.1 in the API. <https://openai.com/index/gpt-4-1/>
- OpenAI. (2025b, April 16). OpenAI o3 and o4-mini system card. <https://openai.com/index/o3-o4-mini-system-card/>
- Orlikowski, W. J. (2000). Using technology and constituting structures: A practice lens for studying technology in organizations. *Organization Science*, 11(4), 404–428. <https://doi.org/10.1287/orsc.11.4.404.14600>
- Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39(2), 230–253. <https://doi.org/10.1518/001872097778543886>
- Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, 30(3), 286–297. <https://doi.org/10.1109/3468.844354>
- Pask, G. (1976). *Conversation theory: Applications in education and epistemology*. Elsevier.
- Pea, R. D. (1993). Practices of distributed intelligence and designs for education. In G. Salomon (Ed.), *Distributed cognitions: Psychological and educational considerations* (pp. 47–87). Cambridge University Press.
- Peirce, C. S. (1931–1958). *Collected papers of Charles Sanders Peirce* (Vols. 1–8). Harvard University Press.
- Pentina, I., Hancock, T., & Xie, T. (2023). Exploring relationship development with social chatbots: A mixed-method study of Replika. *Computers in Human Behavior*, 140, Article 107600. <https://doi.org/10.1016/j.chb.2022.107600>

- Picard, R. W. (1997). *Affective computing*. MIT Press.
- Polanyi, M. (1966). *The tacit dimension*. University of Chicago Press.
- Popper, K. R. (1963). *Conjectures and refutations: The growth of scientific knowledge*. Routledge & Kegan Paul.
- Popper, K. R. (2002). *The logic of scientific discovery*. Routledge. (Original work published 1934)
- Porter, M. E., & Heppelmann, J. E. (2014). How smart, connected products are transforming competition. *Harvard Business Review*, 92(11), 64–88.
- Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211–227. <https://doi.org/10.1002/sci.3730660207>
- Prigogine, I., & Stengers, I. (1984). *Order out of chaos: Man's new dialogue with nature*. Bantam Books.
- Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review*, 46(1), 192–210. <https://doi.org/10.5465/amr.2018.0072>
- Rezwana, J., & Maher, M. L. (2023). Designing creative AI partners with COFI: A framework for modeling interaction in human-AI co-creative systems. *ACM Transactions on Computer-Human Interaction*, 30(5), Article 67, 1–28. <https://doi.org/10.1145/3519026>
- Rheingold, H. (2002). *Smart mobs: The next social revolution*. Perseus Publishing.
- Ricoeur, P. (1981). *Hermeneutics and the human sciences* (J. B. Thompson, Ed. & Trans.). Cambridge University Press.
- Ritchie, G. (2007). Some empirical criteria for attributing creativity to a computer program. *Minds and Machines*, 17(1), 67–99. <https://doi.org/10.1007/s11023-007-9066-2>
- Ronald, E. M. A., Sipper, M., & Capcarrère, M. S. (1999). Design, observation, surprise! A test of emergence. *Artificial Life*, 5(3), 225–239. <https://doi.org/10.1162/106454699568755>
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6), 386–408.
- Rumelhart, D. E., McClelland, J. L., & PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vols. 1–2). MIT Press.
- Runco, M. A., & Jaeger, G. J. (2012). The standard definition of creativity. *Creativity Research Journal*, 24(1), 92–96. <https://doi.org/10.1080/10400419.2012.650092>
- Russell, S. J., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696–735. <https://doi.org/10.2307/412243>
- Salas, E., Cooke, N. J., & Rosen, M. A. (2008). On teams, teamwork, and team performance: Discoveries and developments. *Human Factors*, 50(3), 540–547.

- Salomon, G., Perkins, D. N., & Globerson, T. (1991). Partners in cognition: Extending human intelligence with intelligent technologies. *Educational Researcher*, 20(3), 2–9. <https://doi.org/10.3102/0013189X020003002>
- Santoni de Sio, F., & van den Hoven, J. (2018). Meaningful human control over autonomous systems: A philosophical account. *Frontiers in Robotics and AI*, 5, Article 15. <https://doi.org/10.3389/frobt.2018.00015>
- Sawyer, R. K. (2003). Artificial societies: Multiagent systems and the micro–macro link in sociological theory. *Sociological Methods & Research*, 31(3), 325–363. <https://doi.org/10.1177/0049124102239079>
- Sawyer, R. K. (2005). *Social emergence: Societies as complex systems*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511734892>
- Sawyer, R. K. (2007). *Group genius: The creative power of collaboration*. Basic Books.
- Sawyer, R. K., & DeZutter, S. (2009). Distributed creativity: How collective creations emerge from collaboration. *Psychology of Aesthetics, Creativity, and the Arts*, 3(2), 81–92. <https://doi.org/10.1037/a0013282>
- Schön, D. A. (1983). *The reflective practitioner: How professionals think in action*. Basic Books.
- Searle, J. R. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3(3), 417–457.
- Seeber, I., Bittner, E., Briggs, R. O., de Vreede, T., de Vreede, G.-J., Elkins, A., Maier, R., Merz, A. B., Oeste-Reiß, S., Randrup, N., Schwabe, G., & Söllner, M. (2020). Machines as teammates: A research agenda on AI in team collaboration. *Information & Management*, 57(2), Article 103174. <https://doi.org/10.1016/j.im.2019.103174>
- Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization*. Doubleday/Currency.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Shanahan, M. (2024). Talking about large language models. *Communications of the ACM*, 67(2), 68–79. <https://doi.org/10.1145/3624724>
- Shneiderman, B. (2007). Creativity support tools: Accelerating discovery and innovation. *Communications of the ACM*, 50(12), 20–32. <https://doi.org/10.1145/1323688.1323689>
- Shneiderman, B. (2022). *Human-centered AI*. Oxford University Press. <https://doi.org/10.1093/oso/9780192845290.001.0001>
- Simon, H. A. (1996). *The sciences of the artificial (3rd ed.)*. MIT Press.
- Sloman, S., & Fernbach, P. (2017). *The knowledge illusion: Why we never think alone*. Riverhead Books.
- Snowden, D. J., & Boone, M. E. (2007). A leader’s framework for decision making. *Harvard Business Review*, 85(11), 68–76.
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333–339. <https://doi.org/10.1016/j.jbusres.2019.07.039>

- Stahl, G. (2006). *Group cognition: Computer support for building collaborative knowledge*. MIT Press.
- Sternberg, R. J. (Ed.). (1999). *Handbook of creativity*. Cambridge University Press.
- Suchman, L. A. (1987). *Plans and situated actions: The problem of human-machine communication*. Cambridge University Press.
- Sunstein, C. R. (2006). *Infotopia: How many minds produce knowledge*. Oxford University Press.
- Surowiecki, J. (2004). *The wisdom of crowds*. Doubleday.
- Tegmark, M. (2017). *Life 3.0: Being human in the age of artificial intelligence*. Alfred A. Knopf.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433–460.
- UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. <https://unesdoc.unesco.org/ark:/48223/pf0000380455>
- United Nations. (2015). *Transforming our world: The 2030 Agenda for Sustainable Development*. United Nations.
- Varela, F. J., Thompson, E., & Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT Press.
- von Foerster, H. (2003). *Understanding understanding: Essays on cybernetics and cognition*. Springer.
- Wegerif, R. (2007). *Dialogic education and technology: Expanding the space of learning*. Springer. <https://doi.org/10.1007/978-0-387-71142-3>
- Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185–208). Springer. https://doi.org/10.1007/978-1-4612-4634-3_9
- Weick, K. E. (1995). *Sensemaking in organizations*. Sage Publications.
- Weiser, M. (1991). The computer for the 21st century. *Scientific American*, 265(3), 94–104.
- Whitehead, A. N. (1978). *Process and reality: An essay in cosmology (Corrected ed., D. R. Griffin & D. W. Sherburne, Eds.)*. Free Press. (Original work published 1929)
- Wiener, N. (1948). *Cybernetics: Or control and communication in the animal and the machine*. MIT Press.
- Wiggins, G. A. (2006). A preliminary framework for description, analysis and comparison of creative systems. *Knowledge-Based Systems*, 19(7), 449–458. <https://doi.org/10.1016/j.knosys.2006.04.009>
- Wiig, K. M. (1997). *Knowledge management: Where did it come from and where will it go? Expert Systems with Applications*, 13(1), 1–14.
- Wimsatt, W. C. (1997). Aggregativity: Reductive heuristics for finding emergence. *Philosophy of Science*, 64(S4), S372–S384. <https://doi.org/10.1086/392615>
- Wooldridge, M., & Jennings, N. R. (1995). Intelligent agents: Theory and practice. *The Knowledge Engineering Review*, 10(2), 115–152. <https://doi.org/10.1017/S0269888900008122>

- Wooldridge, M. (2009). An introduction to multiagent systems (2nd ed.). Wiley. <https://www.wiley.com/en-us/shop/general-introductory-computer-science/an-introduction-to-multiagent-systems-2nd-edition-p-9780470519462>
- Woolley, A. W., Aggarwal, I., & Malone, T. W. (2015). Collective intelligence and group performance. *Current Directions in Psychological Science*, 24(6), 420–424. <https://doi.org/10.1177/0963721415599543>
- Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, 330(6004), 686–688. <https://doi.org/10.1126/science.1193147>
- World Economic Forum. (2023). *The future of jobs report 2023*. World Economic Forum.
- Xie, T., Pentina, I., & Hancock, T. (2023). Friend, mentor, lover: Does chatbot engagement lead to psychological dependence? *Journal of Service Management*, 34(4), 806–828. <https://doi.org/10.1108/JOSM-02-2022-0072>
- Yannakakis, G. N., Liapis, A., & Alexopoulos, C. (2014). Mixed-initiative co-creativity. In *Proceedings of the 9th International Conference on the Foundations of Digital Games*. Society for the Advancement of the Science of Digital Games.
- Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., & Cao, Y. (2023). ReAct: Synergizing reasoning and acting in language models. In *The Eleventh International Conference on Learning Representations*. OpenReview. https://openreview.net/forum?id=WE_vluYUL-X
- Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications.
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

(Б) Тематична структура на библиографията

Един източник може да присъства в повече от една тематична група.

1. Философия, онтология и епистемология

Aristotle (2004)

Aristotle (2009)

Dewey (1938)

Gadamer (2004)

Heidegger (1962)

Husserl (1982)

Kuhn (2012)

Lakatos (1978)

Merleau-Ponty (2012)

Peirce (1931–1958)

Polanyi (1966)

Popper (2002)

Ricoeur (1981)

Whitehead (1978)

2. Когнитивна наука, психология и теория на познанието

Anderson (2007)

Anderson (2020)

Bandura (1986, 2001)

Barrett (2017)

Barsalou (2008)

Bruner (1990)
Clark (1997)
Clark & Chalmers (1998)
Damasio (1994, 1999)
Ericsson & Pool (2016)
Gardner (1983)
Gigerenzer (2007)
Goleman (1995, 1998)
Graziano (2013)
Johnson-Laird (1983)
Kahneman (2011)
Kolb (1984)
Lakoff & Johnson (1980)
Mayer (2021)
Newell & Simon (1972)
Norman (2023)
Rumelhart, McClelland, & PDP Research Group (1986)
Sloman & Fernbach (2017)

3. Теория на системите, сложност и емерджентност

Ashby (1956)
Bertalanffy (1968)
Capra (1996)
Capra & Luisi (2014)
Goldstein (1999)
Holland (1992)
Holland (1998)
Maturana & Varela (1980)
Meadows (2008)
Mitchell (2009)
Morin (2008)
Prigogine & Stengers (1984)
Sawyer (2005)
Simon (1996)
von Foerster (2003)

Weick (1995)
Wiener (1948)

4. Изкуствен интелект

Akata et al. (2020)
Anthropic (2024)
Bengio (2023)
Boden (2016)
Bommasani et al. (2021)
Brown et al. (2020)
Goodfellow et al. (2016)
Gemini Team, Google (2023)
Kelly (2015)
Liu (2023)
McCarthy (2007)
McDermott (1976)
Mitchell (2019)
Moravec (1988)
OpenAI (2023)
OpenAI (2025a)
OpenAI (2025b)
Russell & Norvig (2021)
Turing (1950)

5. Взаимодействие „човек–ИИ“

Amershi et al. (2019)
Brave & Nass (2002)
Cooper et al. (2014)
Dix et al. (2004)
Endsley (1995)
Engelbart (1962)
Grudin (2017)
Hollan et al. (2000)
Hollnagel & Woods (2005)
IBM Research (2016)
Licklider (1960)
Nass & Brave (2005)

Norman (1988)
Norman (2004)
Norman (2013)
Parasuraman et al. (2000)
Suchman (1987)
Weiser (1991)

6. Хибриден, разширен и съвместен интелект „човек–ИИ“

Bredeweg & Kragten (2022)
Brynjolfsson et al. (2023)
Brynjolfsson & McAfee (2014)
Hao et al. (2026)
Clare & Storbeck (2006)
Colton (2008)
Csikszentmihalyi (1996)
Davenport & Kirby (2016)
Davidson et al. (2003)
Davis (2026)
Davis et al. (2015a, 2015b)
Dellermann et al. (2019)
Doshi-Velez & Kim (2017)
Floridi (2014, 2019a, 2023)
Floridi & Cowls (2019b)
Forgas (2006)
Gao et al. (2025)
Hemmer et al. (2024)
Jarrahi (2018)
Jordanous (2016a)
Kamar (2016)
Lubart (2005)
Miller (2019)
Norman (2004)
Picard (1997)
Raisch & Krakowski (2021)
Runco & Jaeger (2012)
Sawyer (2007)

Seeber et al. (2020)
Shneiderman (2022)
Sternberg (1999)

7. Емоции, творчество и съвместно създаване

Akata et al. (2020)
Clark & Chalmers (1998)
Clare & Storbeck (2006)
Colton (2008)
Csikszentmihalyi (1996)
Davidson et al. (2003)
Forgas (2006)
Hutchins (1995)
Jordanous (2016a)
Lubart (2005)
Norman (2004)
Picard (1997)
Runco & Jaeger (2012)
Sawyer (2007)
Sternberg (1999)

8. Колективен и организационен интелект

Albrecht (2003)
Argyris & Schön (1978)
Argyris & Schön (1996)
Cannon-Bowers et al. (1993)
Choo (1998)
Edmondson (2019)
Hevner et al. (2004)
Lévy (1997)
Malone (2018)
Malone et al. (2010)
Matsuda (1992)
Nonaka (1994)
Nonaka & Takeuchi (1995)

Porter & Heppelmann (2014)

Salas et al. (2008)

Senge (1990)

Simon (1996)

Snowden & Boone (2007)

Sunstein (2006)

Surowiecki (2004)

Wiig (1997)

Woolley et al. (2015)

Woolley et al. (2010)

9. Етика, управление и стандарти за ИИ

Anthropic (2024)

Bostrom (2014)

Council of Europe (2024)

European Commission (2020)

European Parliament & Council of the European Union (2024)

High-Level Expert Group on Artificial Intelligence (2019)

IEEE (2019)

ISO 31700-1:2023 Privacy by Design

ISO 9241-210:2019 Human-centred design

ISO/IEC 22989:2022

ISO/IEC 23894:2023

ISO/IEC 38507:2022

ISO/IEC 42001:2023

ISO/IEC 5338:2023

ISO/IEC TR 24028

ISO/IEC TR 24368

Jobin et al. (2019)

NIST AI 600-1 (2024)

NIST AI RMF (2023)

NIST AI RMF Playbook

OECD (2019)

OECD AI Observatory

UNESCO (2021)

United Nations (2015)

World Economic Forum (2023)

10. Методология и валидиране на изследванията

Boateng et al. (2018)

Braun & Clarke (2006)

Campbell & Fiske (1959)

Cohen (1960)

Kline (2023)

Krippendorff (2018)

McGraw & Wong (1996)

Messick (1995)

Shadish et al. (2002)

Snyder (2019)

Yin (2018)

ЗА АВТОРА

Оля Харизанова е доктор на философските науки и водещ изследовател (R4) във Философския факултет на Софийския университет „Св. Климент Охридски“. До 2024 г. е била професор по информационно-търсещи системи в същия университет, където продължава да води лекционни курсове, свързани с развитието и приложението на информационните и комуникационните технологии.

Проф. Харизанова е завършила Факултета по математика и механика (днес Факултет по математика и информатика) на Софийския университет. Първоначално защитава дисертация за придобиване на образователната и научна степен „доктор“, а по-късно придобива и научната степен „доктор на науките“ в областта на обществените комуникации и информационните науки. Изследователските ѝ интереси обхващат социалната интерпретация на технологиите, дигиталната трансформация, проявленията на интелигентното информационно общество и управлението на знанието.

През последните двадесет и пет години участва в осъществяването на над шестдесет проекта, част от които ръководи. Понастоящем е водещ координатор на проект BG16RFPR002-1.014-0011 „Устойчиво развитие на Център за върхови постижения „Наследство БГ“, финансиран по процедура BG16RFPR002-1.014 по Програма „Научни изследвания, иновации и дигитализация за интелигентна трансформация“. Наред с това тя е член на изследователския екип по проект „Икономически ефекти и рискове от бързото навлизане на технологиите с изкуствен интелект (ИИ) в икономиката и финансите“ (КП-06-Н85/7), финансиран от Фонд „Научни изследвания“ към Министерството на образованието и науката.

Настоящата ѝ книга предлага оригинална и концептуално развита теория за интегрирания интелект – не като завършена догма, а като отворена изследователска програма, предназначена да провокира научен диалог. В нея е представена концептуална рамка за интегрирания интелект като аргументирано теоретично предложение, което:

- въвежда собствена концептуална архитектура;
- поставя научно значими въпроси;
- предлага понятия и методология за обсъждане и проверка;

- определя ограничения;
- допуска опровержение, прецизиране и алтернативни обяснения;
- очертава възможности за последващи емпирични изследвания.

Авторката не се стреми да „затвори“ темата, а да създаде научна основа, спрямо която други изследователи могат да заемат позиция, като потвърдят, оспорят, операционализират или доразвият предложената теория.

Оля Харизанова

ИНТЕГРИРАН
ИНТЕЛЕКТ

Българска
Първо издание

Дизайн на корицата и предпечат
Валентин Марков

Формат 70/100/16
Печатни коли 10,5

Университетско издателство „Св. Климент Охридски“
unipress.bg