

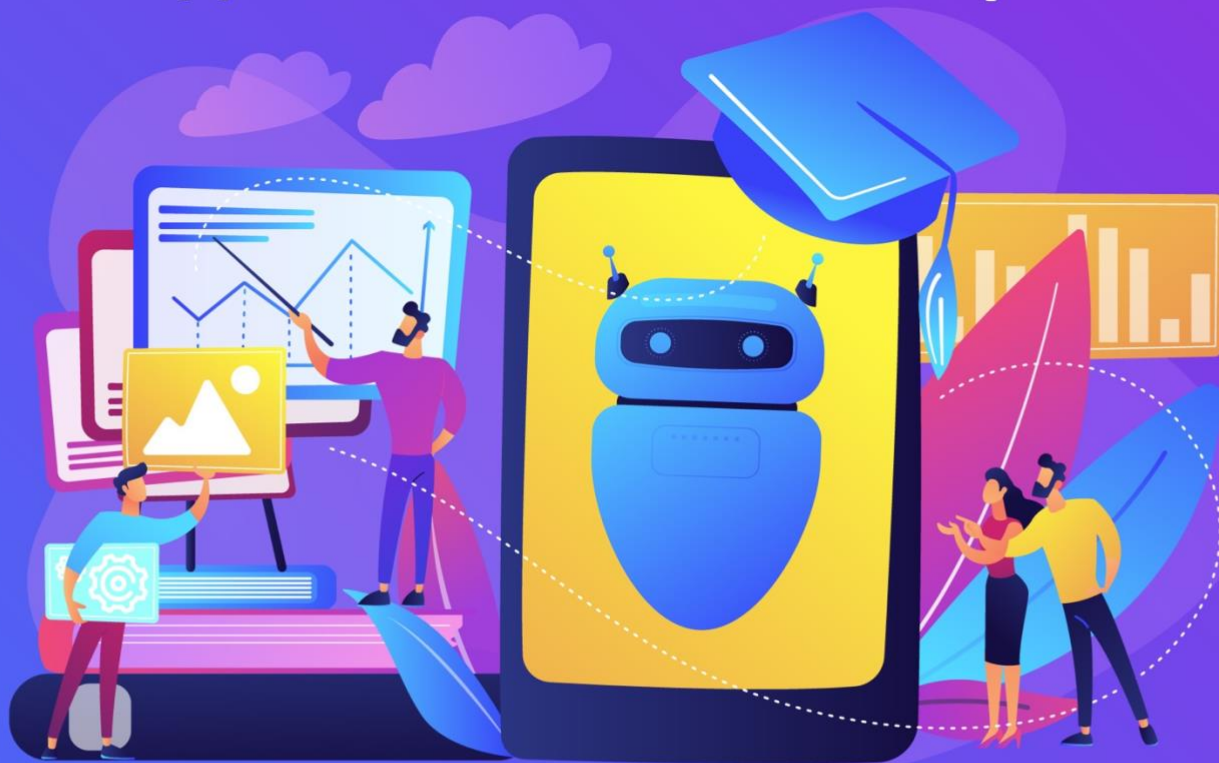


Navigating AI Regulations:
Practical Guide

Проектен Номер:
2024-2-DE02-KA210-VET-000287096

Учебна програма

AI-управлявана креативност: Разширено обучение за дигитални иноватори



Съфинансирано от
Европейския съюз

Учебна програма

AI-управлявана креативност: Разширено обучение за дигитални иноватори

Модул 1:

Разбиране на класификацията на рисковете при AI в креативните проекти

Navigating AI Regulations: Практически наръчник

Номер на проекта: 2024-2-DE02-KA210-VET-000287096

Финансирано от Европейския съюз. Изразените възгледи и мнения обаче принадлежат изцяло на техния(ите) автор(и) и не отразяват непременно възгледите и мненията на Европейския съюз или на Европейската изпълнителна агенция за образование и култура (EACEA). За тях не носи отговорност нито Европейският съюз, нито EACEA.



Съфинансирано от
Европейския съюз

Име на документа	Модул 1: Разбиране на класификацията на ИИ рисковете в креативните проекти
Проектна дейност	Дейност 2: Разработване на обучителната програма „AI-управлявана креативност: Разширено обучение за дигитални иноватори“
Тип ревизия	Финална
Дата на ревизия	30.06.2025
Автори	INI-Novation (Германия), Budakov Films (България)

Декларация за авторски права:



Този документ е защитен с лиценз Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International. Вие имате право да:

- Споделяте — копирате и разпространявате материала на всякакъв носител и във всякакъв формат
- Адаптирате — ремиксирате, преработвате и изграждате върху материала при следните условия:
- Признание на авторството — трябва да дадете подходящо признание, да предоставите връзка към лиценза и да посочите дали са направени промени. Можете да направите това по разумен начин, но не по начин, който предполага, че лицензодателят ви подкрепя или одобрява употребата.
- Некомерсиална употреба — не може да използвате материала за търговски цели.
- Споделяне при същите условия — ако ремиксирате, трансформирате или изграждате върху материала, трябва да разпространявате вашия принос под същия лиценз като оригинала.

Всяко неразрешено използване или възпроизвеждане на съдържанието на този обучителен модул ще се счита за нарушение на законите за авторско право и ще бъде обект на правни действия.



Съдържание

1. Относно учебната програма	5
2. Разкрийте пълния потенциал на обучението: съвети за участниците	6
2.1. Усвояване на ефективно събиране на съдържание	6
2.2. Определяне на собствен ритъм на самостоятелно обучение	5
2.3. Последователна структура, съобразена с компетенциите от EQF ниво 3	7
3. Композиция и представяне на обучителен модул 1 „Разбиране на класификацията на AI рисковете в креативните проекти“	8
3.1. Въведение	8
3.2. Цели и учебни задачи	9
3.3. Съдържание	10
3.4. Заключение	19
3.5. Практическо упражнение: Изследване на AI риска в креативен проект	24
3.6. Критерии за оценяване	25
3.7. Библиография	26
Приложение 1: Самооценъчен лист за надежден AI	27
За проекта	31

1. Относно учебната програма

Учебната програма „AI-управлявана креативност: Разширено обучение за дигитални иноватори“ е ключов резултат от съфинансирания по програма „Еразъм+“ проект „Навигиране в регулациите за ИИ: Практическо ръководство“ (Номер на проекта: 2024-2-DE02-KA210-VET-000287096). Адаптирана към динамично променящите се нужди на предприемачите в креативните индустрии, тази цялостна инициатива има за цел да предостави на участниците основни знания и умения за ефективно ориентиране в Закона за изкуствения интелект. Чрез разглеждане на критични аспекти като класификация на рисковете, съответствие с изискванията, прозрачност и управление на данните, програмата гарантира, че творческите професионалисти могат отговорно да интегрират ИИ в своите работни процеси, да спазват правните стандарти, да управляват данните етично и да изграждат доверие чрез прозрачни практики. В крайна сметка това обучение не само повишава осведомеността и разбирането за сложния регулаторен пейзаж, но и дава възможност на потребителите да използват технологиите на ИИ по иновативни начини, подобрявайки по този начин своята конкурентоспособност в дигиталната ера.

2. Разкрийте пълния потенциал на обучението: съвети за участниците

2.1. Ефективно събиране на съдържание

За да извлечете максимума от практическите обучителни модули, започнете с преглед на учебните цели, предоставени от консорциума. Помислете как тези цели се отнасят до конкретните предизвикателства, пред които се изправяте във всекидневната си работа, и идентифицирайте ключовите теми, които пряко влияят върху вашата професионална среда. След това направете връзка между практическите приложения, представени във всеки от петте модула, и реалните случаи от вашата дейност, като вземете предвид конкретните примери и казуси, включени в обучителните материали. Накрая, организирайте своите изводи с помощта на дигитални инструменти като приложения за водене на бележки или mind map карти, за да можете лесно да правите справки и да интегрирате тези концепции в ежедневната си практика за постигане на наистина ефективно обучение.

2.2. Задаване на собствен ритъм на учене

Нашите обучителни модули са проектирани с ясна структура, която включва следните части: **Въведение, Цели и задачи на обучението, Съдържание, Заключение, Практически упражнения и Критерии** за самооценка.

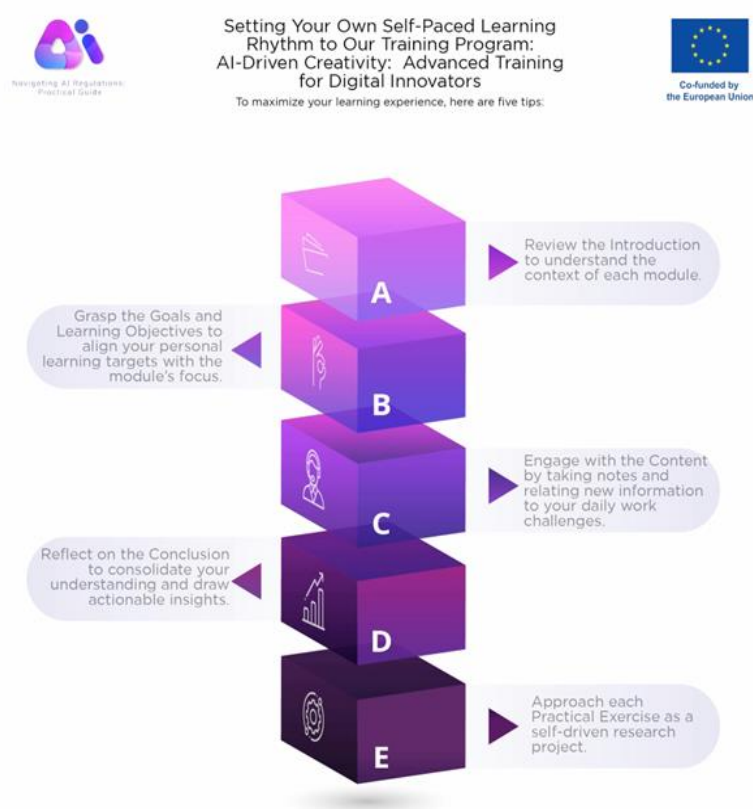
За да увеличите максимално учебния си опит, ето пет съвета:

- Първо, **прегледайте внимателно въведението**, за да разберете контекста на всеки модул.
 - Второ, разберете ясно **Целите и Задачите на Обучението**, за да съобразите личните си учебни цели с фокуса на модула.
 - Трето, запознайте се със **Съдържанието**, като си водите подробни бележки и свързвате новата информация с ежедневните си работни предизвикателства.
 - Четвърто, помислете върху **Заключението**, за да затвърдите разбирането си и да си извлечете практически прозрения.
- Пето, подхождайте към всяко **практическо упражнение** като към самостоятелен изследователски проект, като използвате критериите за оценка, за да измерите напредъка си и да определите областите за подобрене.

Накрая, за да сте сигурни, че сте усвоили уменията и знанията от всичките пет обучителни модула, силно ви препоръчваме да попълните финалния тест: Цялостна оценка на креативността, управлявана от изкуствен интелект. Финалната цялостна оценка не само затвърждава наученото от вас, но и ви мотивира да проучите всички части на програмата, и, моля, не се колебайте да се свържете с консорциума, ако имате някакви въпроси.

1.1. Последователна структура, съобразена с компетенциите от ниво 3 на EQF

Всеки обучителен модул е проектиран с последователна структура, съобразена с ниво 3 на ЕКР - установена по време на Дейност 1 и подробно описана в документа „Съгласуване на рамката на компетентностите: ниво 3 на ЕКР за ИИ и творчески практики“, който се основава на основните компетенции за креативност, задвижвана от ИИ: съобразяване с разпоредбите на Закона за ИИ.



3. Съставяне и представяне на учебния модул 1 „Разбиране на класификацията на риска, свързан с изкуствения интелект, в креативните проекти“

3.1. Въведение

Уважаеми обучаеми, в Модул 1 ще разгледаме как да разбираме категориите риск, свързани с изкуствения интелект, в съответствие с Европейския закон за изкуствения интелект, с особен акцент върху тяхното въздействие върху креативните индустрии като графичен дизайн, дигитален брандинг и уеб разработка. До края на този модул ще можете да:

- Идентифицирайте често срещаните рискове от злоупотреба с изкуствен интелект в областите на дигиталния дизайн, включително графичен дизайн, дигитален брандинг и уеб дизайн.
- Разберете как генерираните от изкуствен интелект пристрастия или грешки могат да повлияят на резултатите от дизайна и доверието на клиентите.

Ще се запознаете с инструмент за самооценка, който ще ви помогне да определите кога е необходима консултация с експерт. Чрез него ще научите практически стратегии за смекчаване на рисковете и осигуряване на отговорно използване на инструменти с изкуствен интелект в творческите работни процеси. В обобщение, Модул 1 предоставя фактически и теоретични знания, които ви насочват в управлението на предвидимите рискове в рамките на творческите работни процеси.

Гъвкавата и лесна за ползване структура ви позволява да напредвате със собствено темпо, като същевременно изграждате и усъвършенствате уменията си. Ако имате въпроси или се нуждаете от допълнителни разяснения, свържете се директно с нашия екип. Ние сме тук, за да подкрепим вашето обучение и да гарантираме, че можете да приложите обучението ефективно.

Ключови думи:

история на ИИ; европейски подход към ИИ; грамотност в областта на ИИ; прозрачност при използването на инструменти с ИИ; надеждност на инструментите с ИИ; управление на риска; класове риск, определени от Закона за ИИ; самооценка на прозрачността и рисковете

3.2. Цели и учебни задачи

Модул 1, „Разбиране на класификацията на риска, свързан с изкуствен интелект, в креативните проекти“, предоставя на креативните професионалисти фактически познания за класификациите на риска, свързан с изкуствен интелект, и уменията, необходими за оценка, категоризиране и смекчаване на рисковете, свързани с изкуствен интелект, в креативните проекти.

До края на модула обучаемите ще могат да идентифицират, класифицират и оценяват рисковете, свързани с изкуствен интелект, като осигуряват отговорно и съвместимо използване на изкуствен интелект в своите работни процеси.

- **Разбиране на категориите риск**, свързани с изкуствения интелект, включително минимален, ограничен, висок и неприемлив риск, с фокус върху тяхното въздействие върху креативните индустрии като графичен дизайн, дигитален брандинг и уеб разработка.
- **Развиване на умения за решаване на проблеми**, за разпознаване и категоризиране на рисковете, свързани с изкуствения интелект, в творческите работни процеси, прилагайки примери от реалния свят от генерирането на съдържание, дизайна и медийното производство.
- **Придобиване на умения за оценка на риска**, свързан с изкуствения интелект, за да оценявате и класифицирате систематично рисковете, свързани с изкуствения интелект, съгласно рамката на Закона за изкуствения интелект.
- **Инструмент за самооценка**, за да анализирате самостоятелно рисковете, свързани с изкуствения интелект, да идентифицирате пропуски в знанията и да определите кога е необходима експертна консултация.
- **Практически примери**, които илюстрират различни нива на риск, свързан с изкуствен интелект (минимален, ограничен, висок) при генерирането на съдържание, дизайна и медийното производство.

Модул 1 служи като основа за програмата за обучение, подготвяйки обучаемите за по-задълбочено проучване на съответствието, прозрачността и управлението на данните в следващите модули.

3.3. Съдържание

Изкуственият интелект (ИИ) се отнася до разработването на машини, които могат да мислят като хора и да имитират действията им. Тази област използва различни

технологии, за да позволи на компютрите да изпълняват задачи, които обикновено изискват човешко познание, като разпознаване на изображения, разбиране на реч, вземане на решения и превод на езици.

Историята на развитието на изкуствения интелект е белязана от три така наречени „Големи вълни“:

- **Първата Голяма вълна** продължи от 1956 г. (когато е „въведен“ терминът „изкуствен интелект“) до началото на този век, когато първите графични процесори позволиха напредък в така наречените „процеси на дълбоко обучение“.
- **Втората Голяма вълна** продължи до 2017 г., когато бяха изобретени Трансформърс. Този период е белязан от компютрите, надминаващи хората в задачите за разпознаване на изображения, от пускането на пазара на „Alpha Go“ и „AlexNet“ и от това, че Google AI открива котки.
- Сега сме в **третата Голяма вълна**, белязана от приложения на генеративния изкуствен интелект, като например големи езикови модели и мултимодални системи.

ИИ грамотността (грамотност в областта на изкуствения интелект) е способността да се разбират, използват, наблюдават и критично да се разсъждава върху приложенията на ИИ. Терминът обикновено се отнася до преподаване на умения и знания на широката общественост, особено важни за хора, които не са специалисти в областта на ИИ, но взаимодействат с инструменти, задвижвани от ИИ, в ежедневието. ИИ грамотността се определя като: „набор от компетенции, които позволяват на хората критично да оценяват технологиите на ИИ; да комуникират и да си сътрудничат ефективно с ИИ; и да използват ИИ като инструмент онлайн, у дома и на работното място.“ [1]

Грамотността в областта на изкуствения интелект е основата на устойчива и иновативна работна сила. Грамотността осигурява предпазна мрежа, гарантираща, че потребителите на изкуствен интелект минимизират рисковете и работят отговорно. Това от своя страна дава възможност на творческите професионалисти, както и на мениджърите и служителите в творческите индустрии, да отключат пълния потенциал на изкуствения интелект и да доведат до значима промяна. Грамотността в областта на изкуствения интелект трябва да се фокусира върху развиването на умения за критична оценка на информацията и съдържанието, генерирани от изкуствен интелект, както и за идентифициране на дезинформация. Трябва също така да знаете как да използвате изкуствен интелект по безопасен, етичен и отговорен начин. Добра отправна точка е достъпът и работата с общи инструменти на изкуствен интелект на начално ниво, като същевременно развивате основни умения за подтикване. Освен това можете да проучите как работи технологията на изкуствения интелект, като дефинирате обща

терминология в областта на изкуствения интелект, опишете процеса и ролята на данните в процеса на разработване и обучение на изкуствен интелект.

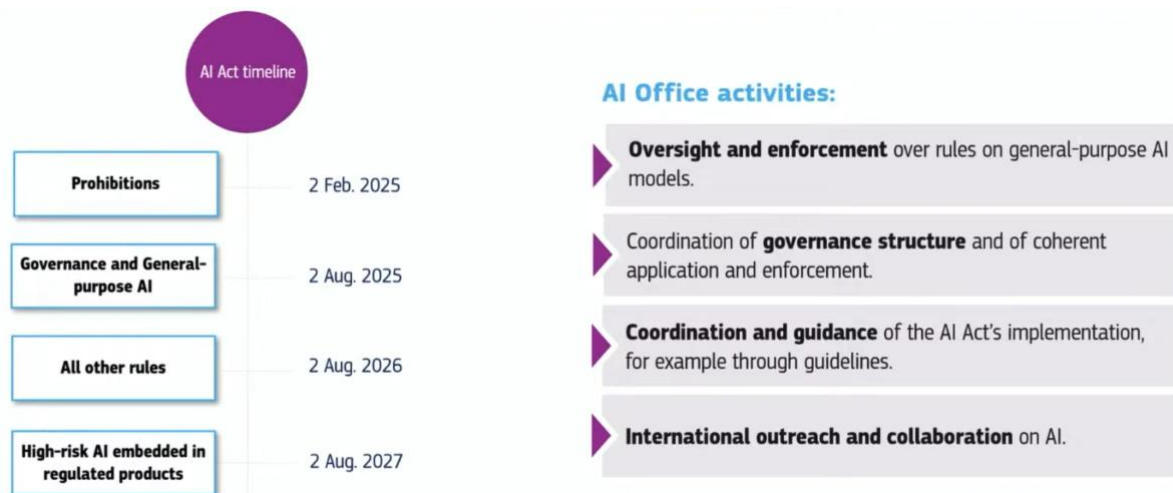
Грамотността в областта на изкуствения интелект трябва да се вземе предвид в подхода на Европейската комисия (ЕК), особено по отношение на съобщението „Изкуствен интелект за Европа“ от април 2018 г. [2], известно още като Стратегия на ЕС за изкуствен интелект. През април 2021 г. ЕК представи своя пакет за изкуствен интелект [3], който включва:

- Съобщение относно насърчаването на европейски подход към изкуствения интелект.
- Преглед на Координирания план за изкуствен интелект (с държавите членки на ЕС).
- Предложение за регламент, определящ хармонизирани правила за изкуствения интелект (Закон за изкуствения интелект), и свързана с него оценка на въздействието.

Този пакет е предназначен като първоначален и базов набор от разпоредби в областта на изкуствения интелект (ИИ). Член 4 от Закона за ИИ, който влезе в сила на 2 февруари 2025 г., постановява:

„Доставчиците и внедрителите на системи с изкуствен интелект предприемат мерки, за да осигурят, в най-добрата си степен, достатъчно ниво на грамотност в областта на изкуствения интелект на своя персонал и други лица, занимаващи се с експлоатацията и използването на системи с изкуствен интелект от тяхно име, като вземат предвид техните технически знания, опит, образование и обучение и контекста, в който ще се използват системите с изкуствен интелект, и като вземат предвид лицата или групите лица, върху които ще се използват системите с изкуствен интелект.“

Европейският закон за изкуствения интелект ще бъде въведен на етапи, както е илюстрирано на следната диаграма:



Фигура 2. Общ преглед на графика за пълното прилагане на Европейския закон за изкуствения интелект. Източник: Европейска служба за изкуствен интелект [4]

Законът за ИИ набляга на правилата, основани на риска, за системите с ИИ, които са определени като „машинни системи, проектирани да работят с различни нива на автономност и които могат да проявяват адаптивност след внедряване. Тези системи, за изрични или имплицитни цели, правят изводи от получените входни данни как да генерират резултати, като прогнози, съдържание, препоръки или решения, които могат да повлияят на физическа или виртуална среда“. Освен това, той изисква доставчиците и внедрителите на системи с ИИ да осигурят достатъчно ниво на грамотност в областта на ИИ.

Грамотността в областта на изкуствения интелект е разгледана в член 3 от Европейския закон за изкуствения интелект, както следва:

*AI literacy' means **skills, knowledge and understanding** that allow providers, deployers and affected persons, taking into account their respective rights and obligations in the context of this Regulation, **to make an informed deployment of AI systems**, as well as to gain awareness about the opportunities and risks of AI and possible harm it can cause.*

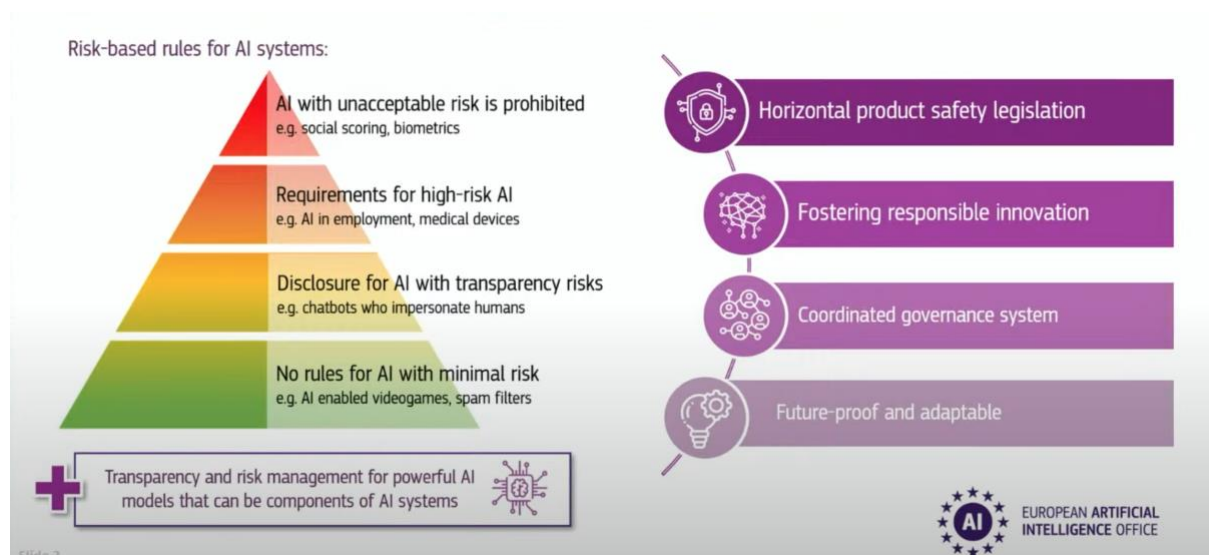
Следователно, за да отговорите на променящите се нужди на заинтересованите страни в творческите индустрии и да осигурите съответствие с Европейския закон за

изкуствения интелект, трябва да сте екипирани с основните знания и умения, за да се ориентирате ефективно в него. За да се справим с това, нашата Рамка за грамотност в областта на изкуствения интелект се състои от три взаимосвързани начина на ангажиране: **Разбиране, Оценка и Използване**. Рамката подчертава, че разбирането и оценяването на изкуствения интелект са от решаващо значение за вземането на информирани решения относно това дали и как да се използва изкуствен интелект в творческа среда.

Законът за изкуствения интелект осигурява прозрачност и управление на риска за мощни модели с изкуствен интелект. Прозрачността за всички модели с изкуствен интелект с общо предназначение се осигурява чрез:

- Техническа документация,
- Разкриване на информация на доставчици на системи с изкуствен интелект надолу по веригата,
- Спазване на законодателството на ЕС за авторското право, и
- Достатъчно подробно обобщение на данните за машинно обучение.

Управлението на риска е категоризирано в четири класа риск: „минимален“, „ограничен“, „висок“ и „неприемлив“. За всеки клас Законът очертава задължения като оценка на модела, оценка и смекчаване на системните рискове, докладване на сериозни инциденти и осигуряване на киберсигурност:



Фигура 3. Правила, основани на риска, за системи с изкуствен интелект. Източник: Европейска служба за изкуствен интелект [5]

Системите с изкуствен интелект, идентифицирани като високорискови, включват технологии с изкуствен интелект, като например системи, предназначени за използване

като полиграфи, инструменти за откриване на емоционалното състояние на физическо лице или системи, предназначени за оценка на индивидуалния риск, включително рискове за сигурността или здравето, породени от физическо лице.

These systems are subject to strict obligation, before being placed on the market, such as [6]:

- Адекватни системи за оценка и смекчаване на риска,
- Високо качество на наборите от данни за минимизиране на рисковете и предотвратяване на дискриминационни резултати,
- Регистриране на дейностите, за да се гарантира проследимост на резултатите,
- Подробна документация, предоставяща цялата необходима информация за системата и нейното предназначение, за да могат властите да оценят нейното съответствие,
- Ясна и адекватна информация за потребителя,
- Подходящи мерки за човешки надзор за минимизиране на риска,
- Високо ниво на надеждност, сигурност и точност.

Съгласно насоките за прозрачност и управление на риска, надеждният изкуствен интелект обхваща три основни аспекта:

- Законосъобразност - спазване на всички приложими закони и разпоредби.
- Етичност - придържане към етичните принципи и ценности.
- Надеждност - технически и социално последователн

Етичността в разработването на ИИ произтича от принципа, че системите с ИИ трябва да бъдат ориентирани към човека, като позволяват на индивидите и обществото да извършват стари и нови дейности по по-добри и по-прости начини, без да причиняват вреда или да ограничават индивидуалните и колективните свободи. Насоките определят четири основни принципа, на които системите с ИИ трябва да отговарят. Тези принципи, формиращи основата на надежден ИИ, са представени в Таблица 1 по-долу:

ПРИНЦИП	ОПРЕДЕЛЕНИЕ
Уважение към човешката автономия	Разработването, внедряването и използването на системи с изкуствен интелект трябва да гарантират пълното и ефективно самоопределение на хората, без никаква манипулация на индивидуалната воля, т.е. възможността за човешки избор и човешки надзор. Системите с изкуствен интелект трябва да „разширяват, допълват и овластяват човешките когнитивни, социални и културни способности“.
Предотвратяване на увреждане	Системите с изкуствен интелект трябва да избягват неблагоприятни последици за хората и/или всяко живо същество. Те трябва да запазват човешкото достойнство и неприкосновеност – както физически, така и

ПРИНЦИП	ОПРЕДЕЛЕНИЕ
	психически – и следва да бъдат проектирани така, че да избягват злонамерени и непреднамерени употреби, особено при работа с уязвими лица или в среда, характеризираща се с асиметрия на властта или информацията.
Справедливост	Обхватът, дизайнът и използването на системите с изкуствен интелект трябва да осигуряват справедливо разпределение на ползите и разходите и съгласуваност между средствата и целите, както и предотвратяване на несправедливи предразсъдъци, дискриминация и стигматизация. Процедурната справедливост изисква прозрачност в процесите на вземане на решения и възможност за „оспорване и търсене на ефективна защита срещу решения, взети от системите с изкуствен интелект и от хората, които ги управляват“.
Обяснимост	Системите с изкуствен интелект трябва да бъдат възможно най-прозрачни, като осигуряват ясна комуникация относно тяхната цел, особено на тези, които участват във внедряването на ИИ (напр. крайни потребители, целеви групи). Процедурната прозрачност може да срещне пречки, което изисква други мерки за обяснимост, като например проследимост, одитируемост и открита комуникация относно възможностите на системата.

Таблица 1. Принципи, формиращи основата на надежден изкуствен интелект [8]

Разбирането на тези принципи помага на учащите се да развият умения за оценка на риска, свързан с изкуствения интелект, което им позволява систематично да оценяват рисковете, свързани с изкуствения интелект, и да ги класифицират съгласно рамката на Закона за изкуствения интелект. Въз основа на гореспоменатите принципи, Насоките представят седем ключови изисквания за надежден изкуствен интелект. Тези изисквания подпомагат заинтересованите страни от творческите индустрии да развият умения за решаване на проблеми, разпознавайки и категоризирайки рисковете, свързани с изкуствения интелект, в творческите работни процеси и им дават възможност да прилагат примери от реалния свят, свързани с генериране на съдържание, дизайн и медийно производство. Таблица 2 по-долу обобщава изискванията:

ИЗИСКВАНИЯ	ОПРЕДЕЛЕНИЕ
Човешка активност и надзор	- Основни права: предполага необходимостта от извършване на оценка на въздействието върху основните права, когато съществува риск от нарушаване на основните права. - Човешка дейност: предполага необходимостта да се предоставят на потребителите необходимите знания и



ИЗИСКВАНИЯ	ОПРЕДЕЛЕНИЕ
	инструменти за разбиране и взаимодействие със системи с изкуствен интелект, вземане на информирани решения. ● Човешки надзор: трябва да се постигне чрез механизми за управление, като например „човек в цикъла“ (HITL), „човек в цикъла“ (HOTL) или „човек-командващ“ (HIC).
Техническа надеждност и безопасност	● Устойчивост на атаки и сигурност: Системите с ИИ трябва да бъдат защитени от атаки, които могат да използват системни уязвимости. ● Резервен план и обща безопасност: Системите с ИИ трябва да предвиждат резервен план, например, изискване за човешки оператор, преди да продължат процеса си, когато срещнат грешки. ● Точност: Системите с ИИ трябва да могат правилно да изпълняват действията си чрез класификации, прогнози, препоръки, решения. Това включва способността да се посочва вероятността от грешка, когато това е неизбежно. ● Надеждност и възпроизводимост: тези характеристики са необходими за проверка на системите с ИИ, т.е. за тестване и възпроизвеждане на резултати.
Поверителност и управление на данните	● Поверителност и защита на данните: Системите с изкуствен интелект трябва да гарантират поверителност и защита на данните през целия цикъл на системата, като същевременно гарантират, че събраните данни няма да бъдат използвани незаконно или несправедливо срещу потребителите. ● Качество и цялостност на данните: Системите с изкуствен интелект трябва да гарантират целостта на събраните данни и тяхното качество, особено когато става въпрос за самообучаващи се системи. ● Достъп до данни: протоколи за данни, определящи кой може да има достъп до данни и при какви обстоятелства са необходими, когато системите с изкуствен интелект обработват данни на лица.
Прозрачност	● Проследимост: всички процеси, характеризиращи процедурите на системата с изкуствен интелект, трябва



ИЗИСКВАНИЯ	ОПРЕДЕЛЕНИЕ
	да бъдат документирани във възможно най-голяма степен, за да се увеличи прозрачността. ● Обяснимост: този принцип включва както самия технически процес, така и решенията, взети от човека, които го характеризират; разбираемостта на системата с изкуствен интелект може да се сблъска с компромис с повишаването на нейната точност. ● Комуникация: хората имат право да бъдат информирани, когато взаимодействат със система с изкуствен интелект, вместо с друго човешко същество; те също така имат право да се откажат от взаимодействие с изкуствен интелект в полза на взаимодействието с човека.
Разнообразие, недискриминация и справедливост	● Избягване на несправедливи предубеждения: Системите с ИИ трябва да се справят с предубежденията, произтичащи от непълнота, неправилни модели на управление или други несъвършенства на предишни системи. ● Достъпност и универсален дизайн: Системите с ИИ трябва да бъдат ориентирани към потребителя и достъпни, независимо от възрастта, пола, способностите или характеристиките на потребителите. ● Участие на заинтересованите страни: разработването на системи с ИИ трябва да се характеризира с консултации със заинтересовани страни, които могат да бъдат пряко или косвено засегнати от самата система; заинтересованите страни могат също така да предоставят обратна връзка по време на целия жизнен цикъл на системата с ИИ.
Социално и екологично благополучие	● Устойчив и екологичен ИИ: Системите с ИИ трябва да се разработват, като се грижи за потреблението на ресурси и енергия през целия жизнен цикъл на ИИ. ● Социално въздействие: Развитието на системите с ИИ в ключови области, като образование, работа, грижи или развлечения, може да повлияе негативно на социалната активност и взаимоотношения; тези рискове трябва да бъдат адекватно разгледани.



ИЗИСКВАНИЯ	ОПРЕДЕЛЕНИЕ
	Общество и демокрация: Ефектите от системите с ИИ трябва да се оценяват и в обществен мащаб, а не само от индивидуална гледна точка, като се вземат предвид ефектите върху институциите, демокрацията и обществото като цяло.
Отчетност	- Одитируемост: Системите с ИИ трябва да могат да бъдат одитирани независимо, когато засягат основните права; в други случаи оценката на алгоритмите, данните и процесите на проектиране трябва да бъде балансирана със съществуващите права върху интелектуална собственост. - Минимизиране и докладване на отрицателните въздействия: трябва да се осигури сигурен начин за докладване на опасения относно разработването и внедряването на система с ИИ, също чрез разработване на оценка на въздействието по целия жизнен цикъл на ИИ. - Компромиси: компромисите между етичните принципи трябва да се разглеждат по рационален, методологичен и документиран начин, като се вземат предвид всички възможни променливи; когато не може да се постигне подходящо решение на конфликти, свързани с компромиси, с наличната технология, разработването на системи с ИИ трябва да бъде прекъснато по този начин. - Обезщетение: Системите с ИИ трябва да осигуряват начин за обезщетение, когато възникнат неблагоприятни въздействия, като се обръща специално внимание на уязвимите лица или групи.

Table 2. Requirements for trustworthy AI [8]

За да се гарантира пълно съответствие с насоките за прозрачност и управление на риска, обучаемите могат да извършат оценка въз основа на „Списъка за оценка на надежден изкуствен интелект“ (ALTAI), представен от „Експертната група на високо ниво по изкуствен интелект“ (AI HLEG) през юли 2020 г. [9]. Листът за самооценка на надежден изкуствен интелект се основава на следните седем ключови изисквания и е представен в Приложение 1 към този документ:

- Човешка активност и надзор,
- Техническа надеждност и безопасност,
- Поверителност и управление на данните,
- Прозрачност,
- Разнообразие, недискриминация и справедливост,
- Екологично и обществено благополучие, и
- Отговорност.

Самооценката осигурява своевременно идентифициране на потенциални проблеми и рискове при използването/прилагането на система с изкуствен интелект, както и на възможни мерки за предотвратяване/сметчане на последиците за справяне с установените проблеми.

След приключване на оценката, обучаемите трябва да отделят време, за да разберат по-добре най-важните въпроси, повдигнати в оценката, и да се уверят, че те са обмислени и интегрирани в техните операции. Инструментите подпомагат независимия и насочен анализ на рисковете, свързани с изкуствения интелект, като помагат на обучаемите да идентифицират пропуските си в знанията. Освен това, обучаемите ще бъдат насочвани да определят и решат дали е необходима експертна консултация.

Моля, вижте допълнителни насоки в практическите примери по-долу, които илюстрират различни нива на риск, свързан с изкуствен интелект (минимален, ограничен, висок) при генерирането на съдържание, дизайна и медийното производство.

3.4. Заключение

Този модул въвежда основите на Европейския закон за изкуствения интелект и предоставя кратък преглед на международните развития в областта на изкуствения интелект. Той разглежда съобщението **на Европейската комисия от април 2018 г. относно „Изкуствен интелект за Европа“**. С фокус върху членове 3 и 4 от Европейския закон за изкуствения интелект, този модул помага на учащите да разберат и да осигурят достатъчно ниво на грамотност в областта на изкуствения интелект. В този контекст **прозрачността и управлението на риска** при използването на модели на изкуствен интелект са от първостепенно значение. Това включва разбирането на концепцията за надежден изкуствен интелект, както и на системата за класификация на риска, свързан с изкуствения интелект, като се фокусира върху това как творческите професионалисти могат да оценяват и категоризират рисковете, свързани с инструментите на изкуствения интелект в своите проекти.

Ще се научите да оценявате рисковете и да съгласувате практиките си с регулаторните рамки. Като част от ключовите съображения относно използването на изкуствен

интелект, модулът описва процеса на провеждане на самооценка на риска, използвайки така **наречения инструмент ALTAI**.

Препоръчваме ви да попълните теста, свързан с този модул, за да проверите напредъка си и да затвърдите разбирането си.

Практически примери

Ето два (2) практически примера за това как може да ви бъде оказана подкрепа при прилагането на разпоредбите във вашите ежедневни обстоятелства:

Практически пример 1: Използвайте онлайн инструмент за удобно откриване на риск, свързан с изкуствен интелект

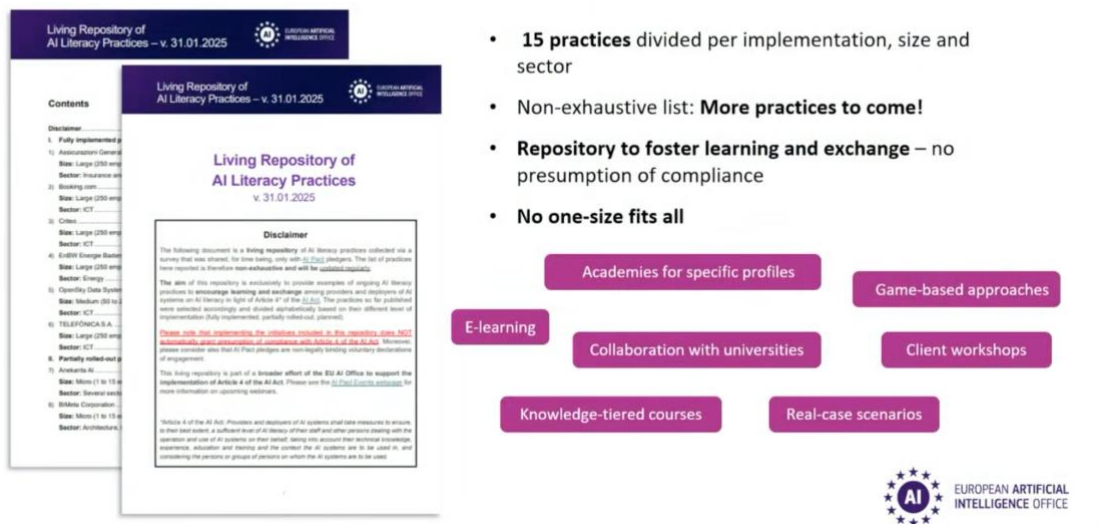
CopyLeaks (<https://copyleaks.com/de/>) предлага лесен за употреба интерфейс за проверка на генерирано от изкуствен интелект съдържание. Можете да използвате демо версия на CopyLeaks, като просто поставите или въведете текст, за да получите вероятностен резултат от изкуствен интелект в рамките на секунди. Инструментът предоставя и ефективно разширение за браузър, което ви позволява да проверявате онлайн страници за съдържание, написано от изкуствен интелект (без ръчно копиране и поставяне).

Практически пример 2: Открийте как организациите се справят с грамотността в областта на изкуствения интелект

Член 4 от Закона за изкуствения интелект (ИИ) изисква доставчиците и внедрителите на системи с ИИ да осигурят достатъчно ниво на грамотност в областта на ИИ сред своя персонал и всяко лице, което използва системите от тяхно име. Службата на ЕС за ИИ е

събрала някои от текущите практики с цел създаване на живо хранилище, което да предоставя примери за текущи практики за грамотност в областта на ИИ:

A living repository of practices



- **15 practices** divided per implementation, size and sector
- Non-exhaustive list: **More practices to come!**
- **Repository to foster learning and exchange** – no presumption of compliance
- **No one-size fits all**

Academies for specific profiles

Game-based approaches

E-learning

Collaboration with universities

Client workshops

Knowledge-tiered courses

Real-case scenarios

AI EUROPEAN ARTIFICIAL INTELLIGENCE OFFICE

Фигура 4. Компилация от практики за грамотност в областта на изкуствения интелект - 28 март 2025 г.
Източник: Европейска служба за изкуствен интелект [10]

Живото хранилище е част от по-широки усилия на Службата на ЕС за ИИ в подкрепа на прилагането на член 4 от Закона за ИИ. То има за цел да насърчи обучението и обмена на знания между доставчиците и внедрителите на системи с ИИ. Можете да изтеглите хранилището тук: <https://ec.europa.eu/newsroom/dae/redirection/document/11220>. Използвайте го, за да задълбочите основните си знания за грамотност в областта на ИИ.

Препоръчваме ви да използвате някои от тези примери за насоки и подкрепа при използването на обучението за изкуствен интелект в рамките на вашите ежедневни работни предизвикателства.

Освен това, не се колебайте да се свържете с авторите на този обучителен модул за допълнителна подкрепа или за да започнете дискусия. Участието в диалог спомага за по-дълбоко разбиране чрез практическо проучване и споделяне на опит.

3.5. Практическо упражнение: Проучване на риска, свързан с изкуствения интелект, във вашия творчески проект.

Практическото упражнение е предназначено да даде възможност на обучаемите да се включат в самостоятелно, практическо изследване, което е пряко свързано с основните теми на модула.

Задача:

AI in Practice: Classifying and Evaluating Risk in Your Creative Work

В тази практическа и саморефлексивна задача сте поканени да проучите как изкуственият интелект (ИИ) е – или би могъл да бъде – част от вашия собствен творчески или културен проект. Въз основа на наученото в Модул 1 за Закона на ЕС за ИИ и класификацията на риска, ще анализирате един проект, ще оцените потенциалното му ниво на риск и ще разсъждавате върху ключови аспекти като прозрачност, етично използване на данни и въздействие върху потребителите.

Изисквания:

- Изберете проект:** Изберете един от вашите съществуващи или планирани творчески проекти, който включва (или би могъл да включва) инструменти или системи с изкуствен интелект - например използване на генеративен изкуствен интелект във визуалните изкуства, композирането на музика, автоматизацията на дизайна, разказването на истории и др.
- Опишете проекта накратко:** Очертайте целите, творческия процес и как се използва или е предвидено да се използва изкуственият интелект.
- Попълнете самооценката:** Използвайки петте области за оценка - използване на изкуствен интелект, класификация на риска, етика на данните, въздействие върху заинтересованите страни и правна осведоменост, помислете върху проекта си, като отговорите честно и обмислено на въпросите, предоставени по-долу:

Област	Насочващи въпроси
Използване на AI	Уверете се, че сте наясно как и къде се използва изкуствен интелект във вашата творческа работа: ○ Ясно ли съм определил кои компоненти на моя проект включват изкуствен интелект?



	○ Кои инструменти или системи в моя проект използват изкуствен интелект (напр. генеративен дизайн, обработка на изображения, генериране на музика, чатботове)? ○ Какви решения или резултати са повлияни от тези системи с изкуствен интелект?
Класификация на риска	Разберете как вашият проект се отразява на категориите риск за ИИ в ЕС: ○ Моят проект използва ли ИИ по начин, който би могъл да повлияе на човешките права, безопасността или вземането на решения? ○ Използва ли се ИИ в биометричен анализ, решения за заетост или публично оценяване? Ако да, може ли да е високорисков? ○ Наясно ли съм какво се квалифицира като забранено, високорисково, с ограничен риск или с минимален риск съгласно закона? ○ Дали е система с ИИ с общо предназначение с потенциал за злоупотреба?
Етика на данните	Започнете да вграждате етично мислене и прозрачност в творческия си процес. ○ Събирам ли и използвам ли данни по начин, който е законен, етичен и прозрачен? ○ Информирам ли потребителите/аудиторията, че в тази работа е използван изкуствен интелект? ○ Документирам ли съм откъде идват данните, особено ако е включено генерирано от изкуствен интелект съдържание? ○ Би ли разбрал обикновен човек как изкуственият интелект е допринесъл за крайния резултат?
Въздействие върху потребителите	Предвиждайте как вашата творческа работа, подобрена с изкуствен интелект, ще се отрази на другите. ○ Обмислил ли съм как различните заинтересовани страни взаимодействат с изкуствения интелект в моя проект или са засегнати от него? ○ Обмислил ли съм евентуални рискове от вреда или нежелани последици?



Правно съзнание	Запознайте учащите се с начина на мислене за съответствие, изискван съгласно Закона за изкуствения интелект. ○ Необходима ли ми е оценка на въздействието върху защитата на данните? ○ Разбирам ли отговорностите си съгласно Закона на ЕС за изкуствения интелект дори на основно ниво? ○ Обмислил ли съм как да направя използването на изкуствения интелект от моя страна по-прозрачно?

4. **Обобщете ключовите прозрения:** Завършете размисъла си с 2–3 параграфа, обобщаващи:

- а. Какво научихте за риска и етиката, свързани с изкуствения интелект, във вашия контекст и
- б. Всички промени, които може да направите в проекта си в резултат на това

Бъдете честни и любопитни - не става въпрос за това да имате „правилния“ отговор, а за това да се научите как да мислите критично за ИИ в собствената си практика.

- **Не прекалявайте с техническите подробности** - фокусирайте се върху предназначението, а не върху кода зад инструментите.
- **Върнете се към съдържанието на Модул 1** - особено категориите риск и етичните съображения на Закона на ЕС за изкуствения интелект.
- **Визуални обучаващи:** Използвайте диаграми, скици или ментални карти, за да изразите как изкуственият интелект взаимодейства с различните етапи от вашия творчески процес.
- **Творчески мислещи:** Отнасяйте се към това като към дизайнерско предизвикателство - как можете да подобрите проекта си, като същевременно уважавате потребителите, законите и ценностите?

Резултати от обучението:

Чрез изпълнението на тази задача, учащите ще могат да:

- Да идентифицират използването на ИИ в собствената си творческа практика.
- Да приложат категориите за класификация на риска по Закона на ЕС за ИИ към пример от реалния свят.
- Да разсъждават критично върху етичните, правните и социалните последици от интеграцията на ИИ в творческата работа.
- Да демонстрират основна регулаторна грамотност, съобразена с ниво 3 на ЕКР.

- Да развият по-голяма осведоменост за това как да подобрят прозрачността, отчетността и доверието на потребителите в собствените си творчески проекти.

3.6. Критерии за оценка

1. Идентифициране и разбиране на употребата на ИИ в творчески контексти

- Обучаващият се демонстрира ясна способност да идентифицира къде ИИ се използва или би могъл да се използва в неговите творчески или предприемачески проекти.
- Описанията на употребата на ИИ са конкретни, релевантни за практиката на обучаващия се (напр. генеративно изкуство, автоматизация на дизайна, интерактивни медии) и показват основно разбиране за ролята на ИИ.
- Обучаващият се разсъждава върху функцията и целта на ИИ в работния процес, като подчертава области, в които ИИ допринася за генерирането на съдържание, вземането на решения или взаимодействието с потребителя.

2. Прилагане на класификацията на риска от ИИ съгласно Закона на ЕС за ИИ:

- Обучаващият се е в състояние предварително да класифицира системите с изкуствен интелект, включени в неговия проект, използвайки категориите риск по Закона на ЕС за изкуствения интелект: Минимален риск, Ограничен риск, Висок риск или Забранен.
- Обосновката за класификацията отразява разбирането на критериите, използвани в Закона за изкуствения интелект, като например контекст на употреба, потенциално въздействие върху отделни лица или специфични за сектора опасения.
- Оценката на риска е съобразена с реалните условия на проекта и може да включва размисъл върху несигурността или сложността при класификацията.

3. Размисъл върху прозрачността, етиката на данните и въздействието върху заинтересованите страни:

- Обучаващият се демонстрира осъзнаване на изискванията за прозрачност при използването на инструменти с изкуствен интелект в рамките на творческата работа (напр. разкриване на генерирани от изкуствен интелект елементи).
- Етичните съображения, като например извличане на данни, съгласие, потенциални пристрастия и очаквания на потребителите, се отчитат във връзка с проекта.

- Обсъжда се въздействието върху заинтересованите страни (аудитория, клиенти, сътрудници), показвайки способността на обучаващия се да предвижда или разсъждава върху това как изкуственият интелект може да повлияе на другите в практически или етичен смисъл.

4. Доказателство за компетентност по ниво 3 на EQF в областта на регулаторната осведоменост

- Обучаващият се показва зараждаща се автономност при прилагането на регулаторни концепции от Закона на ЕС за изкуствения интелект в собствения си творчески контекст.
- Самооценката разкрива основно, но точно разбиране на отговорностите, свързани с използването на изкуствен интелект (напр. оповестяване, отчетност, смекчаване на риска).
- Рефлексията на обучаващия се показва готовност да действа съгласно насоките, да адаптира работните процеси и да отговаря на регулаторните очаквания по практичен начин.

3.7. Библиография

[1] Long, D., & Magerko, B. (2020). What Is AI Literacy? Competencies and Design Considerations. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1-16). Association for Computing Machinery.

[2] European Commission, Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions COM (2018) 237 final, Artificial Intelligence for Europe, 25.04.2018

[3] European Commission, European Approach to Artificial Intelligence, Link: <https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence>, last time accessed on 20th of April 2025

[4] European Artificial Intelligence Office (cnect-aioffie@ec.europa.eu), Third AI Pact webinar on AI literacy; 20.02.2025

[5] European Artificial Intelligence Office (cnect-aioffie@ec.europa.eu), Third AI Pact webinar on AI literacy; 20.02.2025

[6] European Commission, AI Act.

Link: <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>, last time accessed on 20th of April 2025

[7] AI HLEG, Ethics Guidelines for Trustworthy AI, 2019

[8] AI HLEG, Ethics Guidelines for Trustworthy AI, 2019

[9] ALTAI, The Assessment List for Trustworthy Artificial Intelligence. Link: <https://altai.insight-centre.org>, last time accessed on 20th of April 2025

[10] European Artificial Intelligence Office (eu-ai-pact@ec.europa.eu), Third AI Pact webinar on AI literacy; 20.02.2025

Приложение 1: Лист за самооценка на надеждността на AI



Лист за самооценка

Документът включва разделите от ALTAI, където трябва да се предостави информация за оценка на пълното съответствие с насоките за прозрачност и управление на риска, свързани с ИИ (ако е приложимо и необходимо):

1. Човешка дейност и надзор
Включително основни права, човешка дейност и човешки надзор
2. Техническа надеждност и безопасност
Включително устойчивост на атаки и сигурност, резервен план и обща безопасност, точност, надеждност и възпроизводимост
3. Поверителност и управление на данните
Включително зачитане на поверителността, качеството и целостта на данните, както и достъпа до данни
4. Прозрачност
Включително проследимост, обяснимост и комуникация
5. Разнообразие, недискриминация и справедливост
Включително избягване на несправедливи предразсъдъци, достъпност и универсален дизайн, както и участие на заинтересованите страни
6. Социално и екологично благополучие
Включително устойчивост и екологичност, социално въздействие, общество и демокрация
7. Отговорност
Включително одитируемост, минимизиране и докладване на отрицателното въздействие, компромиси и обезщетения.



Самооценката беше проведена от:		Самооценката беше одобрена от:		
Име				
Фамилия				
Институция				
Дата				
		yes	no	n/a
1	Човешка дейност и надзор			
1.1	Автономия на човека			☐
	Допълнителни обяснения:			
1.2	Човешко наблюдение			☐
	Допълнителни обяснения:			
2	Техническа надеждност и безопасност			
2.1	Обща безопасност			☐
	Допълнителни обяснения:			
2.2	Точност			☐
	Допълнителни обяснения:			
2.3	Надеждност, резервни планове и възпроизводимост			☐
	Допълнителни обяснения:			
3	Поверителност и управление на данните			



3.1	Поверителност и управление на данните	☐
	Допълнителни обяснения:	
4	Прозрачност	
4.1	Проследимост	☐
	Допълнителни обяснения:	
4.2	Обяснимост	☐
	Допълнителни обяснения:	
4.3	Комуникация	☐
	Допълнителни обяснения:	
5	Разнообразие, недискриминация и справедливост	☐
5.1	Избягване на предубеждения	☐
	Допълнителни обяснения:	
5.2	Достъпност и универсален дизайн	☐
	Допълнителни обяснения:	
5.3	Участие на заинтересованите страни	☐
	Допълнителни обяснения:	
6	Социално и екологично благополучие	



6.1	Екологично благополучие	☐
	Допълнителни обяснения:	
6.2	Въздействие върху работата и уменията	☐
	Допълнителни обяснения:	
7	Отчетност	
7.1	Одитируемост	☐
	Допълнителни обяснения:	
7.2	Управление на риска	☐
	Допълнителни обяснения:	
Коментари		

За Проекта

Съфинансираният от Еразъм+ проект „Навигиране в регулациите за изкуствения интелект: Практическо ръководство“ (номер на проекта: 2024-2-DE02-KA210-VET-000287096) има за цел да преодолее критичните пропуски в знанията за изкуствения интелект, цифровите умения и осведомеността за политиките на ЕС сред обучители и фрилансъри в творческата индустрия. Основан на задълбочен анализ на нуждите, проведен по време на етапа на подготовка, проектът възприема целенасочен подход в подкрепа на дигиталната трансформация на този динамичен сектор.

Проектът има три основни цели

- Подобряване на компетентността в областта на изкуствения интелект и използването на данни: Чрез предоставяне на персонализирана обучителна програма на 57 участници, проектът ще подобри разбирането на разпоредбите на Закона за изкуствения интелект, включително класификацията на риска, съответствието, прозрачността и управлението на данните. Тези знания ще дадат възможност на обучителите да насочват младите предприемачи в използването на изкуствения интелект за бизнес иновации, като същевременно се спазват регулаторните стандарти.
- Подобряване на дигиталните умения за изкуствен интелект в творчеството: Участниците ще придобият умения за работа с инструменти, задвижвани от изкуствен интелект, анализ на данни и грамотност в областта на изкуствения интелект, което ще им позволи да интегрират авангардни технологии в творческите процеси. Тази цел е фокусирана върху насърчаване на иновациите, подобряване на творческите работни процеси и изграждане на дигитална устойчивост в сектор
- Повишаване на осведомеността относно политиките на ЕС и Закона за изкуствения интелект: Чрез повишаване на запознатостта с политиките и етичните рамки на ЕС, проектът ще гарантира, че участниците работят отговорно и в съответствие със Закона за изкуствения интелект, насърчавайки доверието и устойчивия растеж в творческата индустрия.

Резултатите от проекта ще допринесат пряко за осигуряването на инструменти и знания на обучителите и фрилансърите, за да се справят успешно в бъдеще, основано на изкуствен интелект, като същевременно се съобразят с регулаторните и етични стандарти на ЕС.