Regulace umělé inteligence

5. října 2023

JUDr. Michaela Čejková

Advokátka a doktorandka
michaela.cejkova@twobirds.com

Bird & Bird

Návrh nařízení o umělé inteligenci (AI Act)

Plán

Čeho se snaží EU docílit?

- Ekonomická prosperita;
- Maximální využití AI;
- Investice (SME a startupy);
- Vzdělání a odborná příprava;
- Etický a právní rámec.

"Stejně jako parní stroj nebo elektřina v minulosti, i umělá inteligence mění náš svět, naši společnost a náš průmysl. (...) Způsob, jakým budeme k AI přistupovat, bude určovat svět, ve kterém budeme žít. Je zapotřebí pevný evropský rámec. Evropská unie by měla zaujmout koordinovaný přístup, aby co nejlépe využila příležitostí, které AI nabízí, a řešila nové výzvy, které přináší. EU může stát v čele vývoje a využívání AI pro bezpečí a pro dobro, přičemž bude vycházet ze svých hodnot a silných stránek."

Kontext

Regulace umělé inteligence

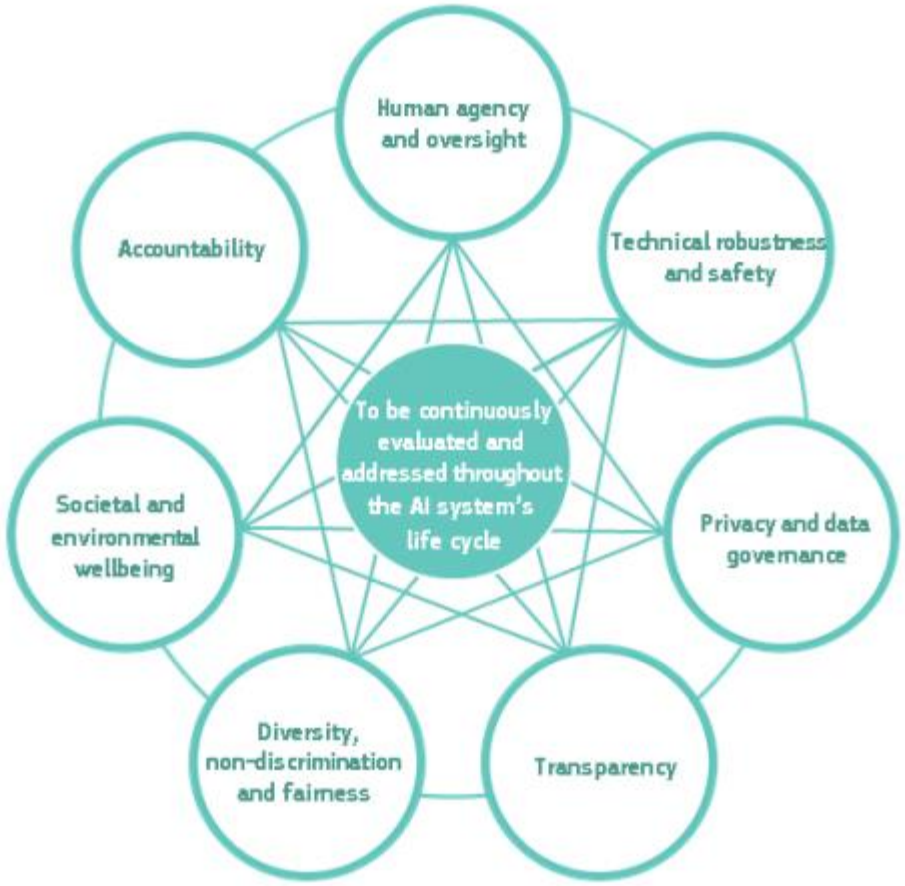
- Strategie pro umělou inteligenci (2018)
- Odborná skupina na vysoké úrovni (AI HLEG) a Evropská aliance pro umělou inteligenci (European AI Alliance)
- **Bílá kniha o umělé inteligenci** (2020)

Celková digitální strategie

- Digitální strategie Evropské unie
- Evropská strategie pro data
- Digitální legislativní balíček
  - Regulace pro digitální služby (DSA)
  - Regulace pro digitální trhy (DMA)
  - Data Act
  - NIS2 (kyberbezpečnost)

Sedm klíčových požadavků pro důvěryhodnou umělou inteligenci

1. Lidský faktor a dohled
2. Technická spolehlivost a bezpečnost
3. Ochrana soukromí a dat
4. Transparentnost
5. Rozmanitost, nediskriminace a spravedlnost
6. Dobré sociální a enviromentální podmínky
7. Odpovědnost



Časová osa

Dosavadní příběh:

04/2021	První návrh Komise
12/2022	Podrobné stanovisko Rady
leden - duben 2023	"Turbulence" kolem společnosti GAI
04/2023	Diskuse v Parlamentu: – Generativní AI = vysoce rizikový systém AI? – OpenAI kritizuje klasifikaci, alternativa: nová kategorie "Foundation Model"
06/2023	Hlasování Parlamentu o vyjednávacím postoji

Současný stav a očekávaný vývoj:

Probíhá	Jednání v rámci trialogu (Komise, Rada, Parlament)
Q4/ 2023	Dohoda o konečném znění
Q1-2/2024	Vstup v platnost
Q1-2/2026	Konečná použitelnost (za předpokladu dvouletého odkladného období)

Jaké jsou hlavní základy nařízení?

- Ochrana základních práv
- Uznání různých rizik různých AI systémů
- Zajištění bezpečnosti
- Podpora inovací
- Posílení důvěry v AI
- Zajištění globálního vedení
- Transparentnost a odpovědnost

Jaké mají být hlavní funkce?

- Přiměřený přístup vycházející z rizik: "Čím vyšší riziko, tím přísnější pravidla"
- Usnadnit rozvoj jednotného trhu se zákonnou, bezpečnou a důvěryhodnou AI a zabránit roztříštěnosti trhu
- Chránit občany před škodami způsobenými AI tím, že bude zajištěn bezpečný vývoj AI, který bude respektovat stávající právní předpisy týkající se základních práv a hodnot EU
- Rámec pro vymáhání povinností
- Právní jistota usnadňující investice a inovace v oblasti AI

Část rozsáhlejšího balíčku o umělé inteligenci:

- Aktualizovaný koordinovaný plán s členskými státy
- Pracuje s dalšími právními předpisy (např. nařízení o strojních zařízeních aj.)
- Nová směrnice o odpovědnosti za škody způsobené AI (více níže).

Na koho se bude nařízení vztahovat? Osobnostní hledisko:

Poskytovatelé:

- fyzická nebo právnická osoba, orgán veřejné moci, agentura nebo jiný subjekt, které **vyvíjí nebo nechávají vyvíjet systém AI za účelem jeho uvedení na trh nebo do provozu** pod svým vlastním jménem nebo ochrannou známkou, ať už za úplatu, nebo zdarma;
- Většina povinností dopadá na poskytovatele (např. posuzování shody, systém řízení kvality, technická dokumentace);
- Nařízení adresuje celý distribuční řetězec s různými povinnostmi;

Komerční uživatelé:

- jakákoli fyzická nebo právnická osoba, orgán veřejné moci, agentura nebo jiný subjekt, který v rámci své pravomoci využívají systém AI, **s výjimkou případů, kdy je tento systém AI využíván při osobní neprofesionální činnost**;

Teritoriální hledisko

Poskytovatelé	Komerční uživatelé	Poskytovatelé a uživatelé	Importéři a distributoři
• Poskytovatelé, kteří **uvádějí na trh nebo do provozu systémy AI v EU**, bez ohledu na to, zda mají sídlo v EU nebo ve třetí zemi.	• Pokud se **nacházejí v EU nebo mají sídlo v EU**.	• Nacházející ve třetí zemi, pokud: (i) je výstup AI systému **určen k použití v EU**, nebo (ii) se na danou **třetí zemi vztahují zákony členského státu na základě mezinárodního práva**.	• Importéři a distributoři AI systémů, pokud se **nacházejí v EU nebo mající sídlo v EU**.

Materiální hledisko: Co je to systém umělé inteligence?

Komise

"systémem umělé inteligence" (systém AI) se rozumí **software,** který je vyvinut pomocí jedné nebo více technik a přístupů uvedených v **příloze I** a může pro **daný soubor cílů definovaných člověkem vytvářet výstupy,** jako je obsah, předpovědi, doporučení nebo rozhodnutí **ovlivňující prostředí, s nímž** interaguje

Příloha I:

(a) Přístupy strojového učení, včetně učení pod dohledem, bez dohledu a posilování, s využitím široké škály metod včetně hlubokého učení;

(b) přístupy založené na logice a znalostech, včetně reprezentace znalostí, induktivního (logického) programování, znalostních bází, inferenčních a deduktivních motorů, (symbolického) uvažování a expertních systémů;

(c) Statistické přístupy, bayesovské odhady, metody vyhledávání a optimalizace.

Rada

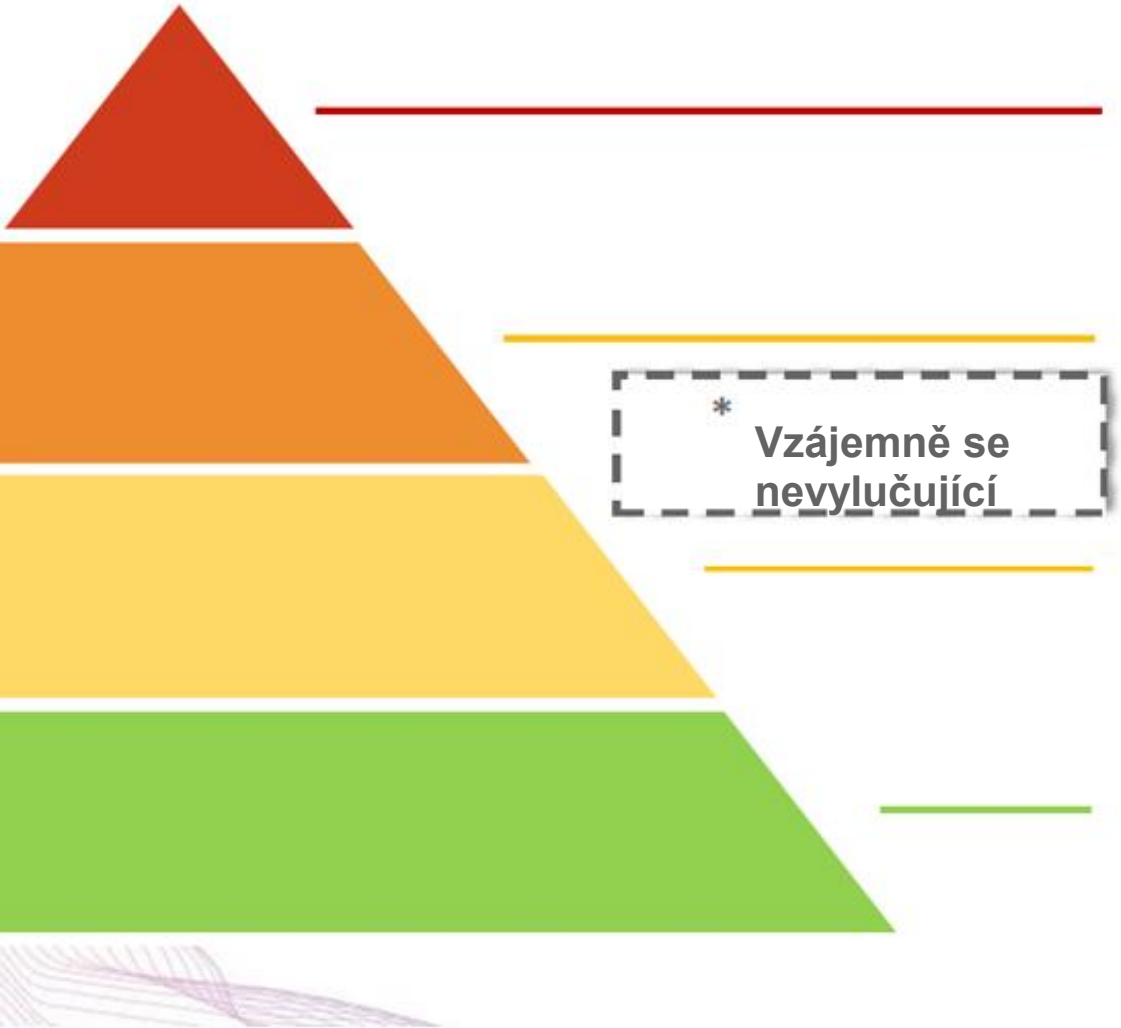
"umělým" (systémem umělé inteligence) se rozumí systém, který je navržen tak, aby fungoval s **prvky autonomie** a který na základě strojových a/nebo lidských **dat a vstupů odvozuje, jak dosáhnout daného souboru cílů, a to pomocí strojového učení a/nebo přístupů založených na logice a znalostech,** a vytváří **výstupy** generované systémem, jako je obsah (generativní systémy umělé inteligence), předpovědi, doporučení nebo rozhodnutí, které **ovlivňují prostředí,** s nímž systém umělé inteligence interaguje. systém umělé inteligence

Co je to systém umělé inteligence?

Parlament

"strojový systém, který je navržen tak, aby pracoval s **různou úrovní autonomie** a který může pro **explicitní nebo implicitní cíle vytvářet výstupy,** jako jsou předpovědi, doporučení nebo rozhodnutí, **která ovlivňují fyzické nebo virtuální prostředí**.

Rozdílné povinnosti dle míry rizika



Nepřípustné riziko

např. tzv. social scoring, manipulativní podprahové techniky

Zakázané

Vysoké riziko

např. nábor zaměstnanců, zdravotnická zařízení

Povoleno s podmínkou splnění všech povinností

AI s konkrétními povinnostmi transparentnosti

tzv. deep fakes

Povoleno, ale podléhá povinnostem o informování/transparentnosti

Minimální nebo žádné riziko

Povoleno bez omezení

European Commission

Source: European Commission

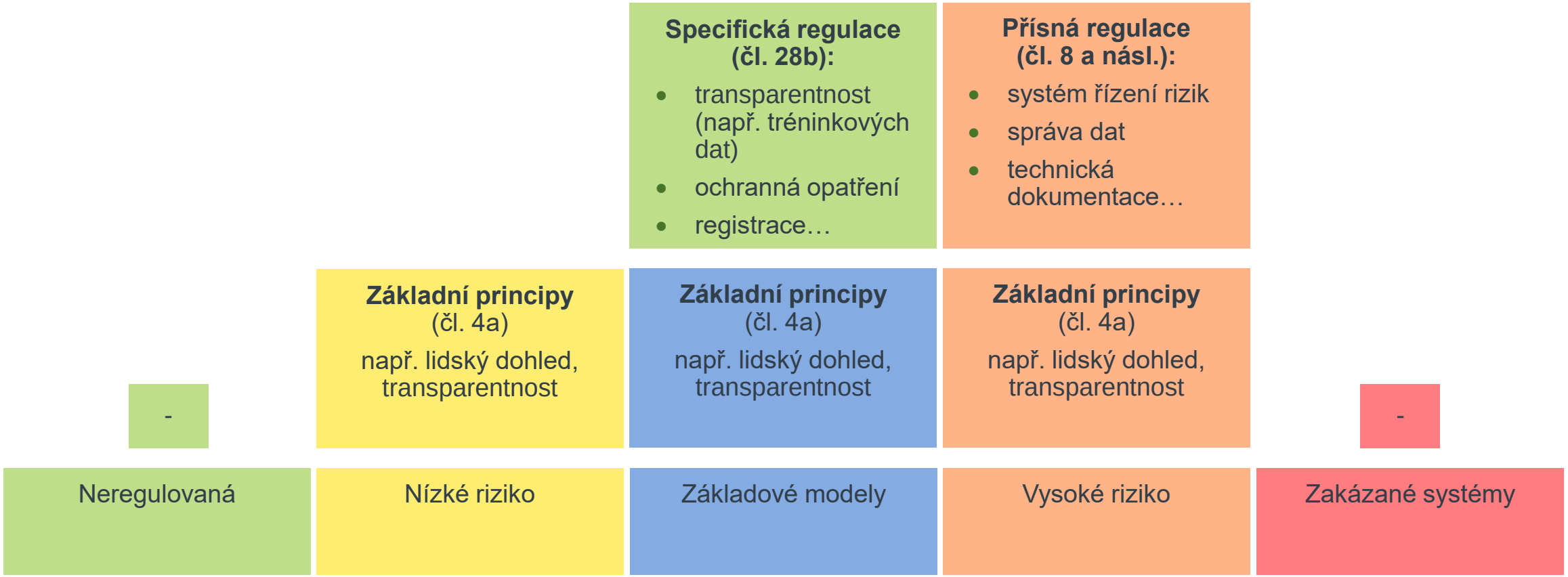
Kategorizace podle AI nařízení EU

Přístup podle rizika

Kategorie AI	Příklady	Popis
Neregulovaná		Čl. 2, např. vojenské účely, výzkum, OSS
Nízké riziko		Všechny jinak neklasifikované systémy (např. videohry, chatboti, deepfakes)
Foundation modely		Na základě rozsáhlých školených dat, široké škále úkolů…
Vysoké riziko		Biometrický záznam, kritická infrastruktura (např. plyn, voda), vymáhání práva…
Zakázané systémy		Podprahové ovlivňování, sociální hodnocení, veřejný biometrický záznam

Kategorizace povinností

Výběr závazků



Potenciální povinnosti uživatelů vysoce rizikových systémů

- Kvalita a správa tréninkových, validačních a testovacích dat;
  - zajištění, aby tréninkové data byly relevantní s ohledem na zamýšlený účel;
- Aktuální technická dokumentace;
- Vedení záznamů a protokolování;
- Transparentnost a poskytování informací uživatelům;
- Sledování provozu systému a informování poskytovatele o všech podezřeních na rizika a závažných incidentech nebo poruchách;
- Zavedení odpovídajícího lidského dohledu
  - zabránění zkreslení automatizace - přílišnému spoléhání na výstupy;
- Zajistit sledování robustnosti a kybernetické bezpečnosti;
- "posuzování shody" = proces ověřování, zda byly splněny požadavky týkající se systému AI;
- Automaticky generované protokoly uchovávat po příslušnou dobu;
- Provést posouzení vlivů na životní prostředí a zveřejnit jeho shrnutí;
- Provést posouzení vlivů na základní lidská práva a zveřejnit shrnutí;
- Spolupracovat s příslušnými vnitrostátními orgány.

Jsou modely generativní umělé inteligence (*foundation models*) vysoce rizikové?

- Návrh Komise: (konkrétně) neřešeno;
- Postoj Rady: Systémy AI pro všeobecné použití, které mohou být ("možné použití") použity jako vysoce rizikové systémy AI nebo součásti vysoce rizikových systémů AI, musí být v souladu;
- Pozice Parlamentu: požadavky (poskytovatelů) generativních modelů by měly být široce použitelné a doplňovat opatření pro vysoce rizikové systémy AI.

"*OpenAI CEO Sam Altman said on Wednesday the ChatGPT maker might consider leaving Europe if it could not comply with the upcoming artificial intelligence (AI) regulations by the European Union.*"
- Reuters, 24 May 2023

"*Sam Altman, the CEO of ChatGPT maker OpenAI, used a high-profile trip to South Korea on Friday to call for coordinated international regulation of generative artificial intelligence, the technology that underpins his famous chatbot.*"
- CNN, 9 June 2023

Ostatní kategorie

AI s konkrétními povinnostmi transparentnosti:

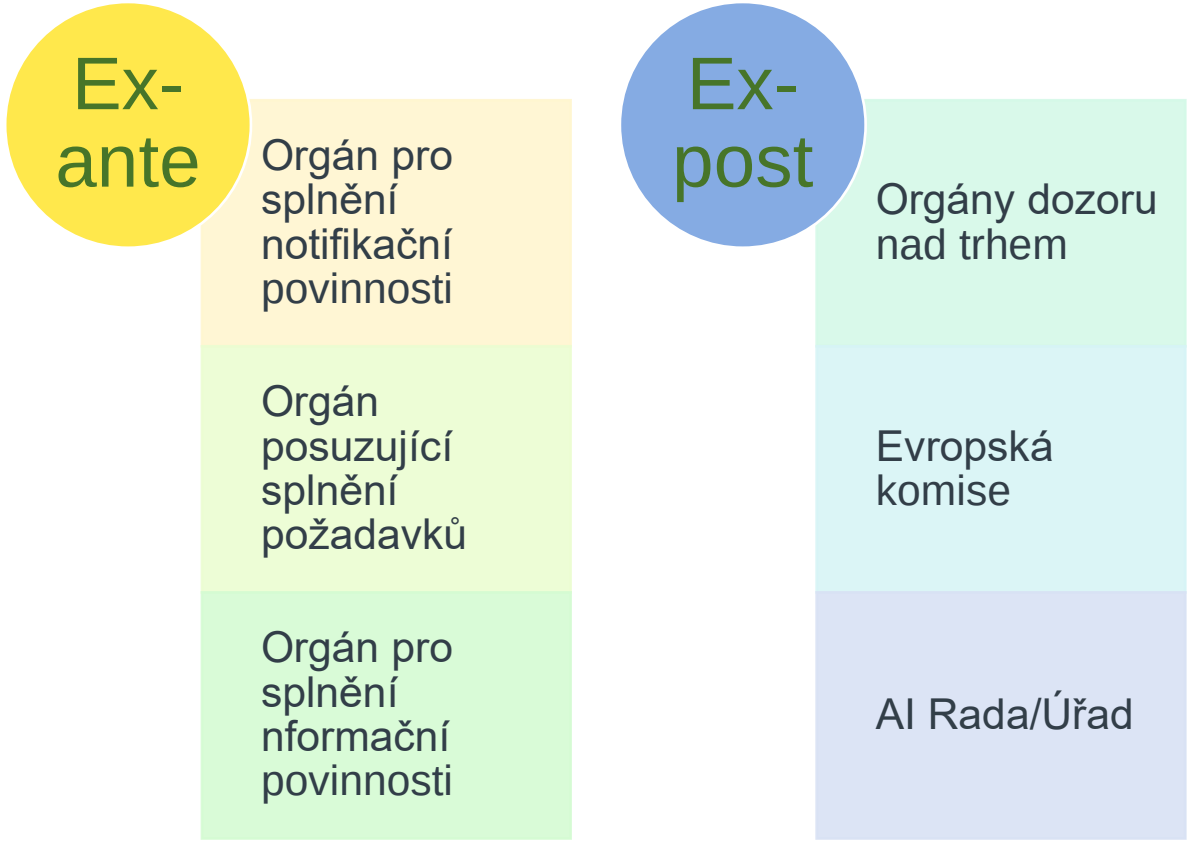
- Upozorňovat lidi, že komunikují se systémem umělé inteligence (např. chatbotem), pokud to není zřejmé;
- Upozornit lidi, že se používají systémy rozpoznávající emoce nebo biometrické kategorizace
  - Parlament: + je třeba získat souhlas před zpracováním těchto osobních údajů;
- Označení Deep Fakes (pokud to není nezbytné pro výkon základního práva/svobody nebo veřejného zájmu)
  - Parlament: trochu nižší nároky, pokud se jedná o zjevně tvůrčí, satirický, umělecký nebo fiktivní obsah

Všechny ostatní

- Zbytková kategorie;
- Žádné další povinné požadavky;
- Komise a Rada pro umělou inteligenci / Úřad pro umělou inteligenci podporují a usnadňují vypracování dobrovolných kodexů chování;

Struktura dohledu

- Komise: Vytvoření Evropského výboru pro umělou inteligenci se zástupci národních dozorčích orgánů a Evropské komise;
  - Členské státy určí příslušný vnitrostátní orgán;
- Parlament: Evropská rada pro AI (AI Office)
  - Řízeno centrálně



Některé další události

Jak do toho zapadá směrnice o odpovědnosti za AI?

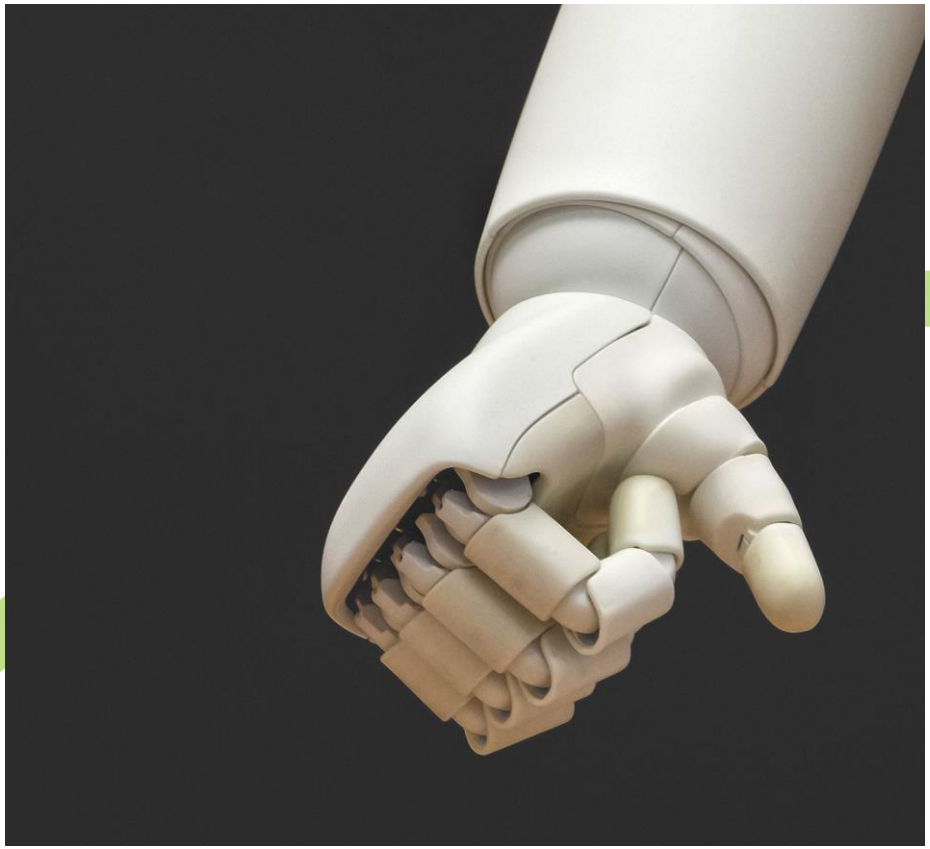
- Navrhla Evropská komise: 28. září 2022
- Zajišťuje, aby osoby poškozené umělou inteligencí měly stejnou úroveň ochrany jako osoby poškozené jinými technologiemi
- Sjednocení definic
- Právo na přístup k důkazům od poskytovatelů vysoce rizikových AI systémů

Vyvratitelná domněnka příčinné souvislosti mezi chybou a škodným následkem

Právo na přístup k důkazům od vysoce rizikových poskytovatelů UI

Různé cesty k regulaci

- EU: Nařízení o umělé inteligenci (maximální harmonizace) a aktualizace režimu občanskoprávní odpovědnosti
- Čína: Nařízení o deepfakes a doporučujících systémech a návrh nařízení pro generativní umělou inteligenci
- Brazílie: Návrh zákona vytváří rámec pro regulaci umělé inteligence, ale pokrok je pomalý
- USA: Návrh zákona v Massachusetts zaměřený na obecnou umělou inteligenci, zákon o odpovědnosti algoritmů, návrh zákona o právech AI, financování a politické pokyny
- Austrálie: v současné době zkoumá, jak dosáhnout bezpečné a zodpovědné umělé inteligence (otevřená konzultace)
- Indie: Vyloučila nutnost přijmout právní předpisy týkající se umělé inteligence, ale je možné, že se na něčem pracuje



Děkuji

twobirds.com

Abu Dhabí ● Amsterdam ● Bratislava ● Brusel ● Budapešť ● Casablanca ● Dubaj ● Dublin ● Düsseldorf ● Frankfurt● Haag ● Hamburk ● Helsinky ● Hongkong ● Kodaň ● Londýn ● Lucemburk ● Lyon ● Madrid ● Milán ● Mnichov ● Peking ● Paříž ● Praha ● Řím ● San Francisco ● Šanghaj ● Singapur ● Stockholm ● Sydney ● Varšava