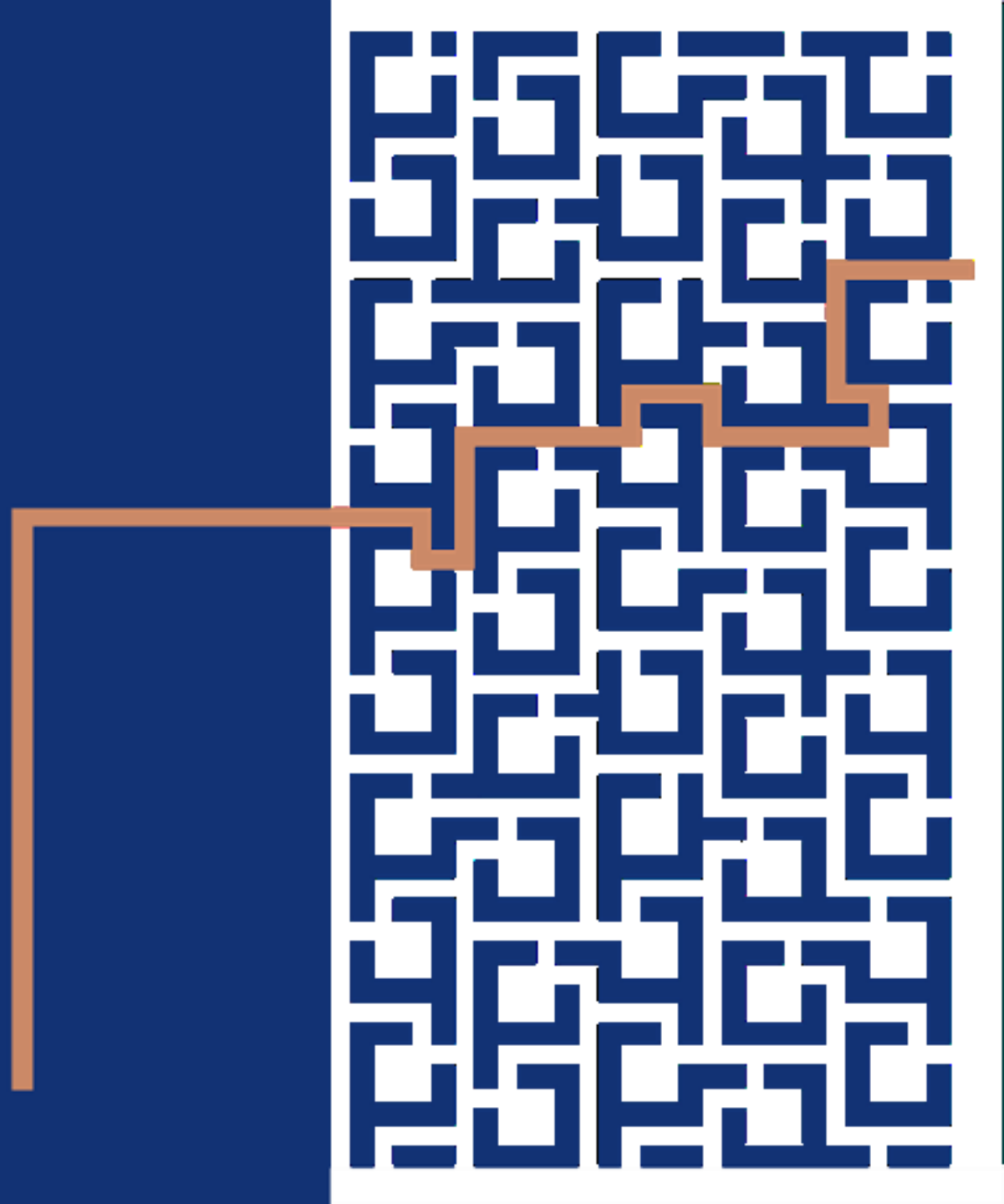
CEVRO VYSOKÁ ŠKOLA

5. PRÁVNÍ REGULACE UMĚLÉ INTELIGENCE, ŘÍZENÍ RIZIK UMĚLÉ INTELIGENCE V INFORMAČNÍCH SYSTÉMECH

Právo informačních a komunikačních technologií

29 | 10 | 2024

I. OBECNĚ O NAŘÍZENÍ O UMĚLÉ INTELIGENCI (AI ÁKT)



Proč máme v EU regulaci AI?

- AI může zasáhnout do základních práv člověka;
  - biais,
  - feedback loop,
  - confirmation biais…
- některá užití umělé inteligence nechceme
- pravidla musí být jednotná v celé EU
- podpora inovací v EU

Obecně o AI Aktu

- Počátky regulace AI na EU úrovni: 2016
- Návrh EK: 21. 4. 2021 | Rev. EP: 14. 6. 2023 | Podpis: 13. 6. 2024
- Publikace: 12. 7. 2024
- Vstup v platnost: **1. 8. 2024**
- Použitelnost: obecně od 2. 8. 2026
  - 2. 2. 2025 - obecná ustanovení, zakázané postupy
  - 2. 8. 2025 - oznamující orgány a oznámené subjekty, obecné modely AI, správa, zachovávání důvěrnosti, sankce a jejich stanovení
  - 2. 8. 2027 - specifické vysoce rizikové systémy dle přílohy I.

Obecně o AI Aktu II.

- Působnost: poskytovatelé, zavádějící subjekty (efekt v EU), dovozci a distributoři, výrobci produktů, zplnomocnění zástupci, dotčené osoby
- **Výjimky z působnosti:**
  - Národní bezpečnost, vojenské a obranné účely
  - Mezinárodní spolupráce v oblasti justice a vymáhání práva
  - AI výhradně pro účely vědeckého výzkumu a vývoje
  - Žádné činnosti výzkumu, testování či vývoje AI před uvedením na trh nebo do provozu; vyjma testování v reálných podmínkách
  - FO používající AI v rámci čistě osobní neprofesionální činnosti
  - AI zpřístupněné na základě svobodných licencí a licencí s otevřeným zdrojovým kódem (s výjimkou vysoce rizikových AI, zakázaných AI nebo AI s požadavky na transparentnost)

II. VZTAH AI AKTU A OSTATNÍCH PRÁVNÍCH PŘEDPISŮ

AI Akt a jeho vztah k další legislativě

- Naplnění povinností členských států EU plynoucích z **Úmluvy Rady Evropy o umělé inteligenci a lidských právech, demokracii a panství práva (5. 9. 2024)**
- Není dotčeno uplatňování ustanovení o odpovědnosti poskytovatelů zprostředkovatelských služeb uvedených v kapitole II DSA
- Nejsou dotčena nařízení (EU) 2016/679 nebo (EU) 2018/1725 ani směrnice 2002/58/ES nebo (EU) 2016/680
- **Nejsou dotčena pravidla** stanovená jinými právními akty Unie, které se týkají **ochrany spotřebitele a bezpečnosti výrobků**
- Možnost vyšší ochrany pracovníků a jejich práv v souvislosti s používáním AI zaměstnavateli

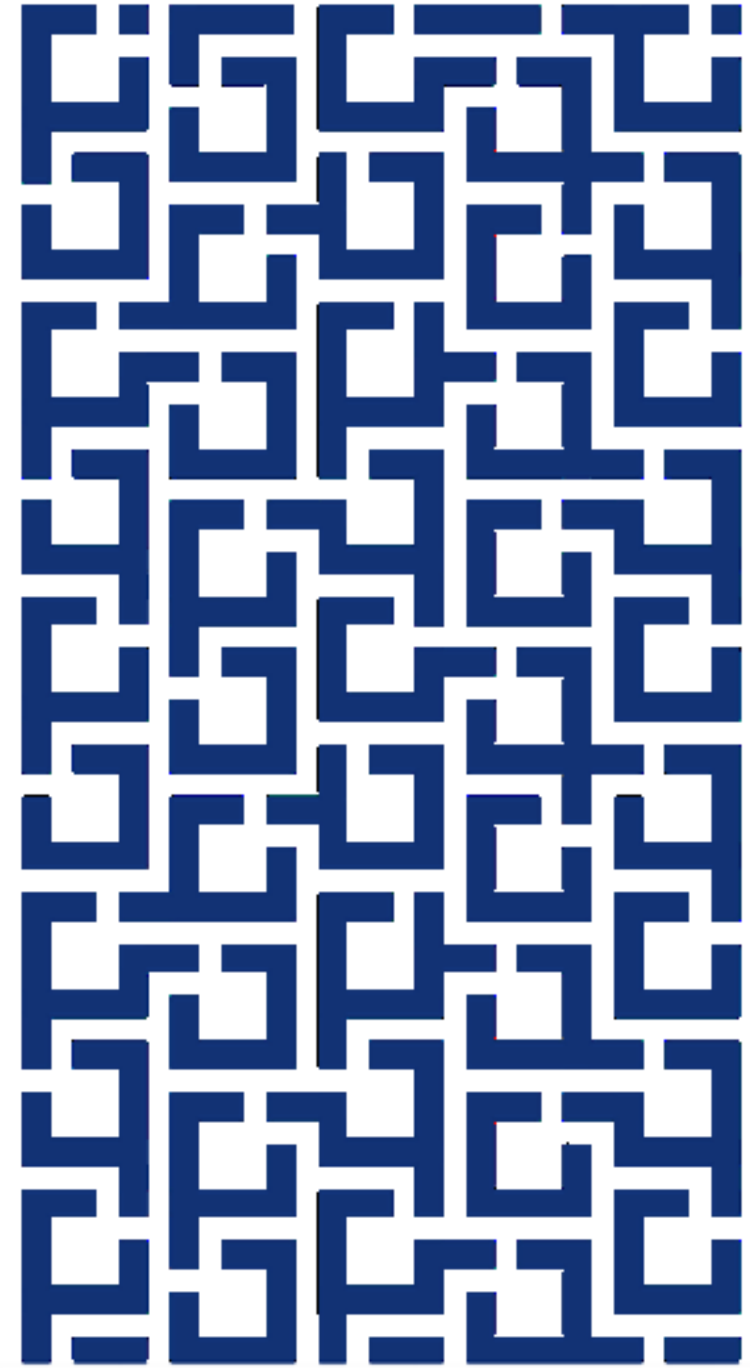
AI Akt a jeho vztah k ochraně osobních údajů

- Rec. 10 & čl. 2 odst. 7
  - Na osobní údaje zpracovávané v souvislosti s právy a povinnostmi stanovenými v tomto nařízení se vztahuje právo Unie o ochraně osobních údajů, soukromí a důvěrnosti sdělení. Tímto nařízením nejsou dotčena nařízení (EU) 2016/679 nebo (EU) 2018/1725 ani směrnice 2002/58/ES nebo (EU) 2016/680.
- Čl. 10 odst. 5: **možnost zpracování zvláštních kategorií OÚ, pokud je to nezbytně nutné pro zajištění detekce a oprav zkreslení ve vztahu k vysoce rizikovým systémům AI**
- Čl. 59: Další zpracování osobních údajů pro účely vývoje určitých systémů AI ve veřejném zájmu v rámci regulačního sandboxu pro AI

Nový titul pro zpracování zvláštních kat. OÚ

- Možnost zpracování zvláštních kategorií OÚ, pokud je to nezbytně nutné **pro zajištění detekce a oprav zkreslení ve vztahu** k vysoce rizikovým systémům AI
- Zajištění vhodných záruk, splnění podmínek legislativy na ochranu OÚ a při kumulativním splnění všech dalších podmínek:
  - Nelze provést pomocí anonymizovaných, syntetických nebo jiných údajů
  - Přísné požadavky na ochranu dat a omezení přístupu
  - Odůvodnění nutnosti
- Použití na testovací data u vysoce rizikových systémů, kt. netrénují modely

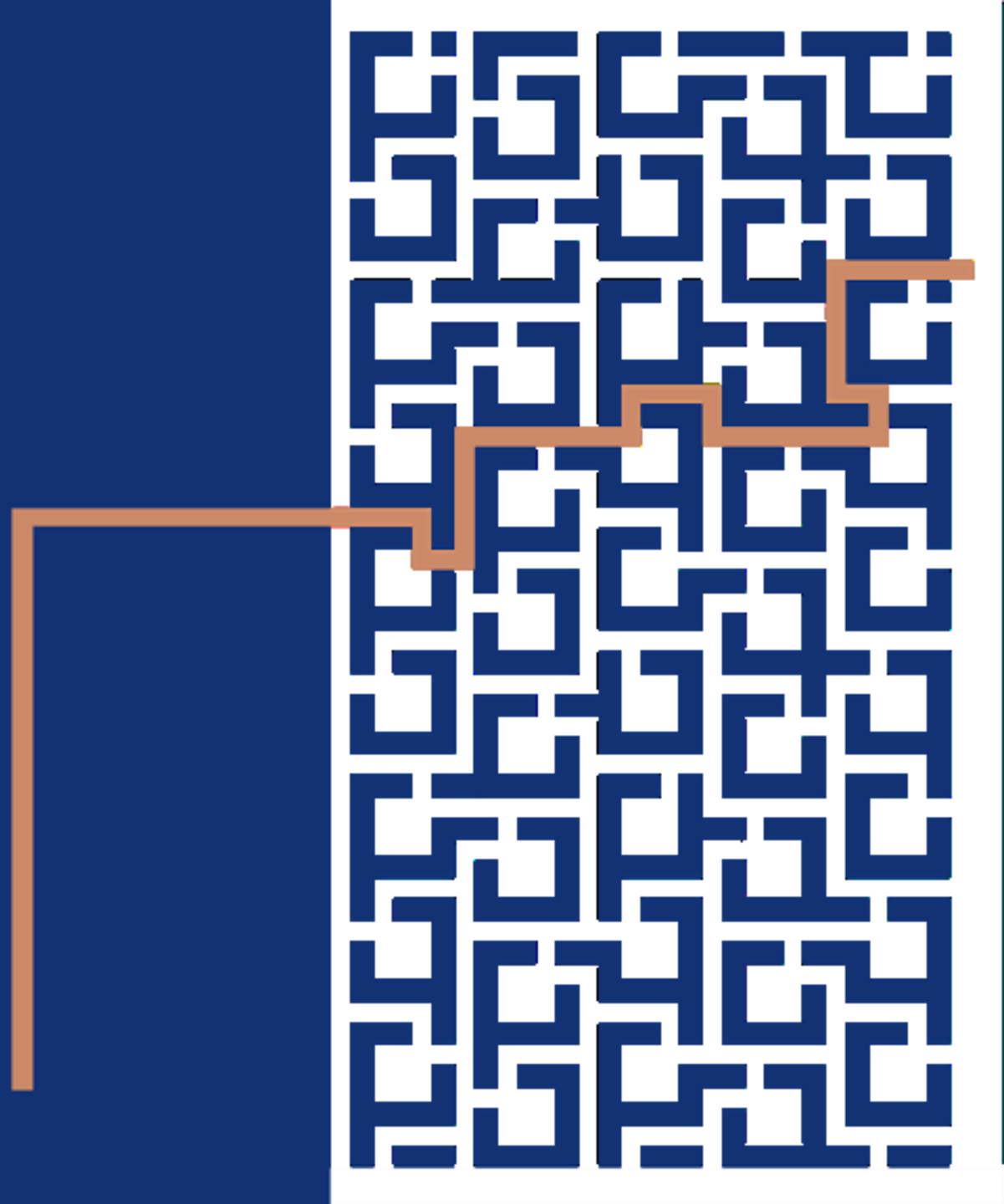
III. SYSTÉM SPRÁVY PODLE AI AKTU



Systém správy podle AI Aktu

- Orgány dle AIA:
  - **Úřad pro AI** (AI Office)
  - **Evropská rada pro AI** (pozorovatel: evr. inspektor ochrany OÚ)
  - Poradní fórum (technické konzultace)
  - Vědecká komise nezávislých odborníků (pro obecné modely)
  - Vnitrostátní orgány:
    - Oznamující organ (patrně ČTÚ)
    - Orgán dozoru nad trhem (patrně ČTÚ ve spol. s ČNB a ÚOOÚ)
    - Jednotné kontaktní místo (patrně ČTÚ…)

IV. OSOBNÍ, VĚCNÁ A ČASOVÁ PŮSOBNOST, SUBJEKTY AI AKTU



Působnost AI aktu

AIA se vztahuje na:

- **poskytovatele**, kteří systémy AI **uvádějí na trh nebo do provozu** v EU
- **subjekty**, které zavádějí systémy AI a jsou usazeny nebo se nacházejí v EU
- **poskytovatele a subjekty** mimo EU, pokud se **výstup používá v EU**
- **výrobce produktů**, kteří uvádějí na trh nebo do provozu systém AI společně se svým produktem a pod svým jménem nebo ochrannou známkou
- **dotčené osoby**, které se nacházejí v EU

AIA se n**evztahuje** na:

- na oblasti **mimo oblast působnosti práva EU**
- **vojenské, obranné účely**
- systémy, vč. výstupů, které byly speciálně vyvinuty a uvedeny do provozu **výhradně za účelem vědeckého výzkumu a vývoje**
- **neprofesionální užití**
- systémy **zpřístupněné na základě bezplatných licencí s otevřeným zdrojovým kódem** kromě zakázaných systémů a systémů s omezeným rizikem

Systém AI (čl. 3 odst. 1)

Strojový systém navržený tak, aby po zavedení **fungoval s různými úrovněmi autonomie** a který po zavedení **může vykazovat adaptabilitu** a který za explicitními nebo implicitními účely z obdržených vstupů odvozuje, jak **generovat výstupy**, jako jsou

- predikce,
- obsah,
- doporučení nebo rozhodnutí, které mohou ovlivnit fyzická nebo virtuální prostředí. (čl. 3 odst. 1))

Definice umělé inteligence neexistuje!

- Slabá (weak), silná (strong), univerzální
- Použití umělé inteligence
  - Tvorba a analýza textů, audio, video, obraz
  - Autonomní vozidla
  - Zdravotnictví
  - Spotřební zboží
  - Průmysl 4.0

Rozdělení systémů AI

- **Zakázané** systémy
- **Vysoce rizikové** systémy (*high-risk AI systems*)
- **Systémy s omezeným rizikem**
- **Ostatní systémy**

- Víceúčelové modely AI (*general-puropse AI models*)
  - Víceúčelové modely AI se systémovým rizikem (*general-purpose AI models with systemic risk – GPAI SR*)
- Víceúčelové systémy AI (*general-purpose AI systems* - ***GPAI***)
- Generativní modely/systémy AI, *large language models* (**LLM**)

AKT O UMĚLÉ INTELIGENCI

- **Aplikuje se v ČR bez nutnosti měnit české právo**
- **Snaha o právní definici systémů umělé inteligence**
- **Rozdělení AI systémů podle rizika**
  - Podstatná část je věnována systémům s vysokým rizikem
- **Povinnosti poskytovatelů (tvůrců)**
- **Povinnosti profesionálních uživatelů (zavádějících subjektů)**
- **Ochrana koncových uživatelů**

KATEGORIE SYSTÉMŮ UMĚLÉ INTELIGENCE

- **Systémy s neakceptovatelným rizikem (zakázané systémy)**
- **Systémy s vysokým rizikem**
  - Riziko pro bezpečnost a zdraví
  - Riziko pro základní práva člověka
- **Systémy s omezeným rizikem**
  - Systémy komunikující s lidmi
  - Systémy manipulující s textem, audio, video, deep-fake
- **Redakční systémy**
- **Ostatní systémy**

SYSTÉMY S NEAKCEPTOVATELNÝM RIZIKEM ZAKÁZANÉ SYSTÉMY (ČL. 5)

- **Systémy, které využívají podprahové, manipulativní, klamavé techniky**
- **Systémy, které využívají zranitelnost**
- **Systémy biometrické kategorizace, které mohou dovodit rasu, sexuální orientaci, politické názory, apod.**
- ***Social-scoring*, které by mělo způsobit nepříznivé jednání s osobou nebo skupinou osob**
- **Systémy pro vzdálenou biometrickou identifikaci v reálném čase kromě výjimek**
- ***Predictive policing* kromě podpory lidského rozhodování**
- **Systémy pro vytváření databáze pro rozpoznávání obličejů za pomocí stahování fotek z internetu nebo CCTV**
- **Systémy pro rozpoznávání emocí na pracovišti a ve vzdělávacích zařízeních**

SYSTÉMY S VYSOKÝM RIZIKEM (ČL. 6)

- **První kategorie: AI systém je určen je bezpečnostní součástí výrobku, nebo je přímo výrobkem na který se vztahují předpisy uvedené v příloze I, a tento výrobek prochází posouzením shody s cílem uvedení na trh.**
  - Strojní zařízení, hračky, výtahy, zdravotnické prostředky, apod.
- **Druhá kategorie: AI systémy uvedené v příloze III (možno měnit aktem v přenesené působnosti)**
  - Systémy biometrické identifikace na dálku, některé systémy biometrické kategorizace, kritická infrastruktura, systémy používané ve vzdělávání, systémy používané v pracovněprávních vztazích a při náboru zaměstnanců, systémy hodnotící nárok na dávky, úvěruschopnost, míru rizika u životního a zdravotního pojištění, apod.

SYSTÉMY S OMEZENÝM RIZIKEM (ČL. 50)

- Víceúčelové AI systémy, včetně generativních modelů
- Systémy rozpoznávání emocí a biometrické kategorizace
- Systémy, které generují deep-fake
- Systémy, které generují nebo manipulují textem pro účely informování veřejnosti o věcech veřejného zájmu

Subjekty podle AI Aktu

- *Poskytovatel* - vyvíjí/nechává vyvíjet AI & uvádí na trh nebo do provozu
- *Zavádějící subjekt* - využívá AI s výjimkou osobní/neprofesionální činnosti
- *Zplnomocněný zástupce* - FO/PO v EU s písemným pověřením
- *Dovozce* - FO/PO v EU, která uvádí na trh AI ze třetí země
- *Distributor* - jiná osoba, která dodává AI na trh EU
- **Provozovatel** (všichni výše uvedení)

POVINNOSTI PŘI POUŽÍVÁNÍ GENERATIVNÍ AI

- **Informace o komunikaci s AI**
- **Informace o výstupu utvořeného AI, ne u SW pro standardní editaci (photoshop)**
- **Informace o deep-fake**
- **Informace o použití AI u textu, jenž je zveřejněn za účelem informování veřejnosti o záležitostech veřejného zájmu, nebo který s takovým textem manipuluje**
  - Nevztahuje se na texty s redakční kontrolou a odpovědností konkrétní osoby

POVINNOSTI POSKYTOVATELŮ VYSOCE RIZIKOVÝCH SYSTÉMŮ

- **Vypracovat systém řízení rizik (čl. 9)**
  - Rozumně předvídatelné riziko pro zdraví, bezpečnost (*safety*) a základní práva za předpokladu, že se systém používá k zamýšlenému účelu
  - Monitorování zbytkového rizika
- **Požadavky na data (čl. 10)**
  - Datasety pro trénování, validaci, testování AI systémů
  - Bezchybnost a úplnost, reprezentativnost
- **Vedení dokumentace (čl. 11)**
- **Zaznamenávání událostí, dohledatelnost *(traceability*) (čl. 12)**
- **Transparentnost a informování uživatelů (čl. 13)**
- **Lidský dohled (čl. 14)**
- **Kyberbezpečnost (čl. 15)**
- **Certifikace (čl. 43)**

POVINNOSTI UŽIVATELŮ VYSOCE RIZIKOVÝCH SYSTÉMŮ (ČL. 26)

- Dodržování vhodných technických a organizačních opatření
- **Zajištění lidského dohledu**
- Relevantní a dostatečně **reprezentativní vstupní data**
- Monitorování provozu systému a **informování poskytovatele**
- **Uchovávání protokolů** o událostech nejméně 6 měsíců
- Informace FO u systému, který rozhoduje, nebo asistuje při rozhodování
- Vypracování **posouzení vlivu na základní práva** (FRIA) u systémů (čl. 27)

POVINNOSTI POSKYTOVATELŮ A UŽIVATELŮ SYSTÉMŮ S OMEZENÝM RIZIKEM A GPAI SYSTÉMŮ

- **Poskytovatelé systémů interagujících s FO: informace, pokud není zřejmé**
- **Poskytovatelé systémů a GPAI generujících text, audio, video, obraz: označit výstup (ve strojově čitelném, detekovatelném formátu), Netýká se systémů s asistenční funkcí pro editaci dokumentu**
- **Uživatelé systémů pro rozpoznávání emocí a biometrické kategorizace: informace o používání systému a zpracování osobních údajů**
- **Uživatelé systému generujícího deep-fake: informace o manipulaci, kromě evidentní umělecké fikce**
- **Uživatelé systémů pro novinářské účely: informace o užití systému, ledaže text podléhá revizi člověkem nebo editorské kontrole a existuje osoba s editorskou odpovědností za vydání**

PROSTŘEDKY NÁPRAVY

- **Stížnost orgánu dohledu nad trhem**
- **Právo na vysvětlení rozhodnutí**
- **Pokuty:**
  - do 35 mil EUR nebo 7 % z celosvětového obratu za provozování zakázaného systému AI
  - do 15 mil EUR nebo 3 % z celosvětového obratu za porušení některých povinností spojených s vysoce rizikovým systémem a systémem s omezeným rizikem
  - do 7,5 mil EUR nebo 1 % z celosvětového obratu za nesprávné nebo zavádějící informace

GENERAL-PURPOSE AI MODELS

- **GPAI model**
- **GPAI model se systémovým rizikem (*systemic risk*)**
  - Schopnosti s vysokým dopadem vyhodnoceným na základě určitých indikátorů a měřítek (*floating point operations*), nebo
  - Rozhodnutí Komise, neboť model má schopnosti srovnatelné se systémem 1)
- **Povinnosti poskytovatelů**
  - Technická dokumentace vč. procesů trénování a testování (neplatí pro modely zpřístupněné na základě bezplatné a otevřené licence u běžného GPAI)
  - Informace poskytovatelům AI systémů (neplatí pro modely zpřístupněné na základě bezplatné a otevřené licence u běžného GPAI)
  - Respektování autorského práva, vč. výjimky pro TDM
  - Detailní shrnutí obsahu, který byl použit pro trénování

REGULAČNÍ PÍSKOVIŠTĚ (*REGULATORY SANDBOXES*) (ČL. 57)

- **Podpora inovací**
- **Dohled kompetentních orgánů**
- **Výjimky z regulace**
  - Žádné pokuty, pokud jednali v dobé víře
  - Užití osobních údajů zpracovávaných pro jiné účely může být užito pro testování, validaci a trénování za podmínek AIA
  - Testování vysoce rizikových systémů v reálných podmínkách za podmínek AIA na základě plánu testování
    - Souhlas subjektů údajů s testováním, požadavky na informovaný souhlas

Nové definice v AI Aktu

- Trénovací data | validační data | soubor validačních dat | testovací data | vstupní data
- *Biometrická identifikace* - zjištění identity porovnáním s databází
- *Biometrické ověření* - automatizované “one-to-one” ověření totožnosti
- *Systém rozpoznávání emocí* - zjišťování nebo odvozování emocí nebo záměrů
- *Systém biometrické kategorizace* - řazení do urč. kategorií, pokud není doplňkem k jiné komerč. službě nebo nezbytně nutný z objektivních technických důvodů
- *Systém biometrické identifikace na dálku* - identifikace FO bez jejich aktivního zapojení, obvykle na dálku, porovnáním s databází
  - V **reálném čase** = zachycení, porovnání a identifikace okamžitě, bez zbytečného odkladu, vč. omezených krátkých zpoždění (prevence obcházení pravidel) v. **následná** identifikace

Přístup založený na riziku

- Zakázané systémy (postupy) v oblasti AI - čl. 5
- Vysoce rizikové systémy AI - čl. 6 - 27
- Obecný AI model se systémovým rizikem - čl. 51 - 52, čl. 55 - 56
- Obecné modely AI - čl. 53 - 54, čl 56
- AI systémy s požadavky na transparentnost - čl. 50

Zakázané postupy a ochrana OÚ

- Podprahové techniky a záměrně manipulativní či klamavé techniky
- AI využívající zranitelnosti osoby (věk, postižení, soc. či ekon. pozice)
- AI pro sociální kredit
- AI pro předvídání pravděpodobnosti spáchání trestného činu
- AI pro vytváření/rozšiřování databází rozpoznávání obličeje
- AI pro odvozování emocí na pracovišti/ve vzdělávacích institucích
- Systémy biometrické kategorizace
- Biometrická identifikace na dálku v reálném čase ve veřejně přístupných prostorech pro účely vymáhání práva (lze ovšem využít čl. 9 GDPR, pokud nejde o vymáhání!)

Biometrická identifikace dle čl. 5

- Umožněno, pokud je nezbytně nutné pro jeden z následujících cílů, a jen v nezbytně nutné míře: cílené vyhledávání obětí a pohřešovaných osob; prevence konkrétního, závažného a bezprostředního ohrožení/předvídatelného teroristického útoku; lokalizace podezřelé osoby (trest nejm. v délce 4 let)
- Pouze potvrzení totižnosti & zvážení povahy situace a důsledky
- Systém zaregistrován a provedeno posouzení dopadů na lidská práva dle čl. 27
- Předchozí povolení /následné povolení do 24h
- Každé použití oznámeno ÚOOÚ - nutnost podávat výroční zprávu

Vysoce rizikové systémy a ochrana OÚ

- Příloha III - specifické AI aplikace ve vztahu k ochraně OÚ
  - Biometrika (nutný soulad s GDPR): biometrická identifikace na dálku (vyjma ověřování), biometrická kategorizace na zákl. citlivých atributů, rozpoznávání emocí
  - Vzdělávání a odborná příprava: přístup, sledování, hodnocení
  - Zaměstnávání a řízení pracovníků: výběr pracovníků, přidělování úkolů
  - Přístup k základním soukromým a veřejným službám a dávkám a jejich využívání: dávky, úvěry, přizpůsobování cen pojištění
  - Vymáhání práva: posuzování rizika stát se obětí, polygrafy, prediktivní policing
  - Migrace, azyl a řízení ochrany hranic: posuzování rizik, žádostí, identifikace

Požadavky na vysoce rizikové systémy

- Systém řízení rizik (čl. 9)
- Data a správa dat (čl. 10)
- Technická dokumentace (čl. 11)
- Vedení záznamů (čl. 12)
- Transparentnost a poskytování informací (čl. 13)
- Lidský dohled (čl. 14)
- Přesnost, spolehlivost a kybernetická bezpečnost (čl. 15)

Další povinnosti poskytovatelů vys. riz. AI

- Systém řízení kvality (čl. 17)
- Uchovávání dokumentace (čl. 18)
- Uchovávání automaticky generovaných protokolů (čl. 19)
- Provedení posuzování shody, vydání prohlášení o shodě, umístění označení, dodržování povinností (čl. 43, 47, 48 a 49)
- Nápravná opatření a poskytování informací (čl. 20)
- Prokazování souladu
- Splnění požadavků na přístupnost

Obecné AI modely

- Povinnosti poskytovatelů:
  - Technická dokumentace modelu a dokumentace pro poskytovatele, kt. budou chtít začlenit do svých systémů
  - Zavedení politiky v oblasti autorského práva
  - Podrobné shrnutí obsahu pro odbornou přípravu modelu
- Modely se systémovým rizikem - dodatečné povinnosti

Velké jazykové modely a ochrana soukromí

Problémy s ChatGPT:

- Získávání osobních údajů **bez souhlasu** a vědomí subjektů údajů
- **Absence záruk** a nástrojů pro uživatele
- Problematický (nemožný) **výkon práva na zapomenutí**
- **Extenzivní sběr dat** o chování uživatelů
- Sdílení informací se třetími stranami
- Nedodržování principů ochrany osobních údajů (mj. přesnost údajů)
- **Produkce chybných** osobních údajů
- Náhodné zveřejnění údajů získaných ze soukromých dokumentů jiných uživatelů

Velké jazykové modely a ochrana soukromí II

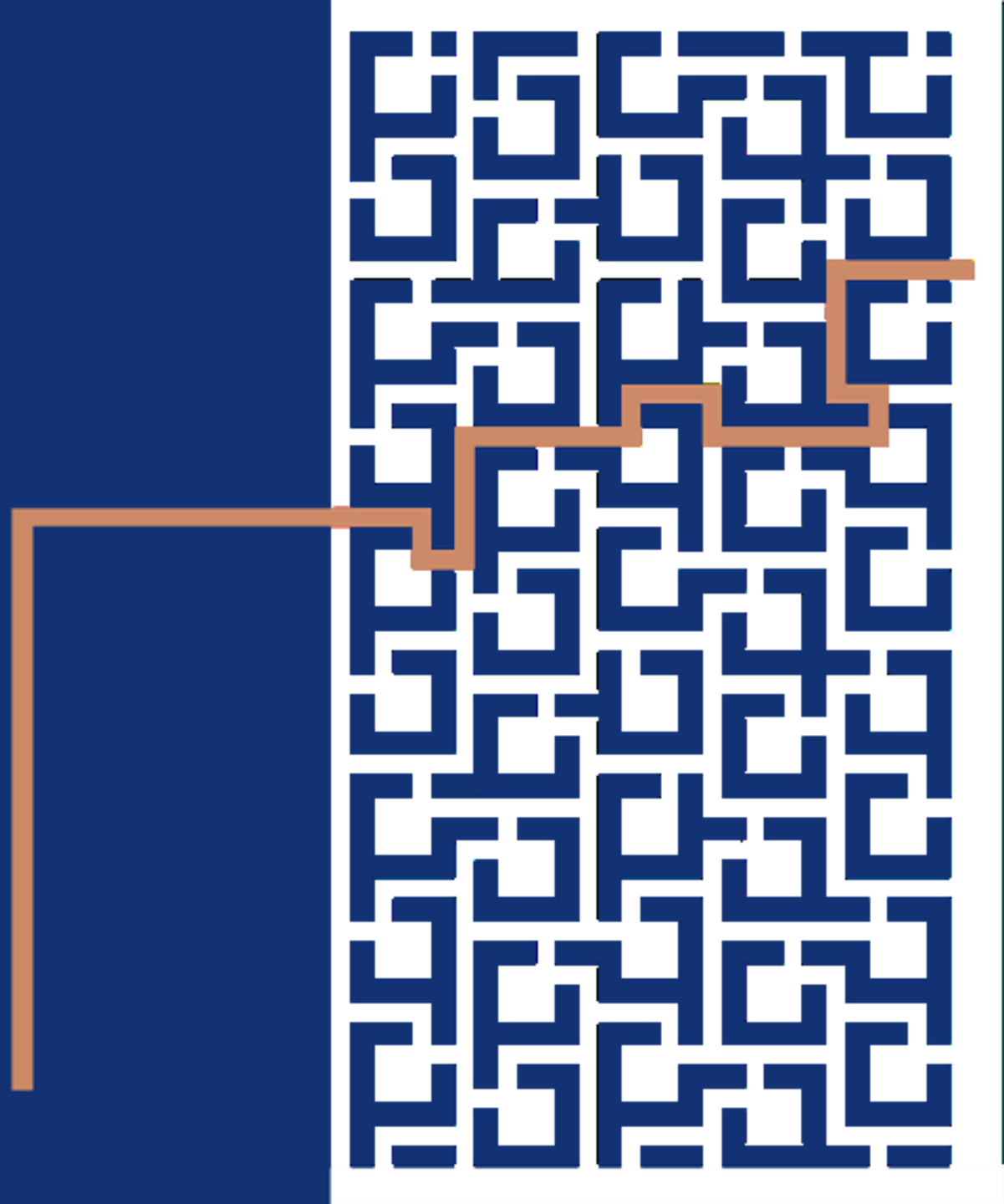
Možná nebezpečí:

- Obejití ochranných prvků (např. pokyn **DAN** - "do anything now")
- **Manipulace** jednotlivcem (případ sebevraždy v Belgii)
- Přístup k **citlivým** dokumentům
- **Diskriminační** obsah
- **Odposlouchávání**
- **Prozrazení** hesel
- Zjišťování nových informací o uživatelích, včetně biometrických údajů
- ... nové způsoby narušení soukromí

AI s požadavky na transparentnost a OÚ

- Transparentnost ohledně komunikace s AI systémem
- Označování AI generovaného/manipulovaného obsahu ve strojově čitelném formátu
- Subjekty zavádějící **systém rozpoznávání emocí** nebo **systém biometrické kategorizace** informují o provozu systému fyzické osoby, které jsou mu vystaveny, a zpracovávají osobní údaje podle potřeby v souladu s nařízeními (EU) 2016/679 a (EU) 2018/1725 a směrnicí (EU) 2016/680.
- Označování deep fake obsahu + označování strojově generovaného textu, jenž je zveřejněn za účelem informování veřejnosti o záležitostech veřejného zájmu

V. PRAXE A DOPORUČENÍ PRO IMPLEMENTACI



VYUŽITÍ AI K VYTVOŘENÍ DOKUMENTŮ

- **Informovat, citovat**
- **Etické kodexy**
- **Pozměněná fotografie**
  - Zpravodajská licence?
  - Radši se souhlasem
- **Pozměněný hlas**
  - Se souhlasem
- **Autorská práva k výstup vygenerovaného AI**
  - Dosud předmětem diskuzí

ORGANIZACE PLÁNUJE VYUŽÍVAT AI SYSTÉM

- K čemu bude AI systém sloužit?
- Nejedná se o vysoce rizikový systém?
  - Jde o bezpečnostní část výrobku nebo výrobek uvedený v příloze I, která podléhá certifikaci?
  - Jde o výrobek uvedený v příloze III?
    - I když se o výrobek v příloze III jedná, nepředstavuje vážné riziko pro zdraví, bezpečnost a základní práva, tím že splňuje jedno nebo více z kritérií AIA, ledaže profiluje FO
  - Je systém certifikovaný?
  - Je systém, který představuje riziko pro základní práva, uveden v databázi rizikových systémů?

DATA PRO TRÉNINK SYSTÉMŮ UMĚLÉ INTELIGENCE (POVINNOSTI POSKYTOVATELÉ SYSTÉMŮ AI DLE ČL. 53 AIA)

Poskytovatelé všech obecných modelů umělé inteligence jsou povinni:

- vypracovat a vést **aktuální technickou dokumentaci**, včetně dokumentace trénování a testování,
- zpřístupnit **informace** poskytovatelům, kteří mají v úmyslu integrovat model do svých systémů.
- respektovat **autorská práva**, včetně respektování výjimky to test and data mining,
- poskytnout **detailní shrnutí o obsahu, který byl použít k trénování**.

Použití díla pro trénování systému umělé inteligence představuje užití díle dle AutZ, tj.

- je přípustné pouze na základě **udělené licence** (§ 2358 a násl. OZ) nebo
- výjimku v podobě **zákonná licence dle § 39c a 39d Autz** k rozmnožování díla pro účely automatizované analýzy textů nebo dat (TDM, text and data mining); ; pokud autor tuto zákonnou licenci nevyloučil (výhrada)

ORGANIZACE PLÁNUJE VYUŽÍVAT AI S VYSOKÝM RIZIKEM

- Zajistit lidský dohled FO do té míry, v jaké vykonává kontrolu nad systémem, FO musí mít dostatečné schopnost, školení, pověření a podporu
- Zajistit kontrolu vstupních dat (relevantní a dostatečně reprezentativní)
- Zajistit monitorování systému a informování poskytovatele a orgán dohledu nad trhem
- Zajistit uchovávání protokolů podle zamýšleného účelu, min. 6 měsíců, vzít v úvahu i GDPR
- Zajistit zaškolení personálu
- Užití systému na pracovišti – zajistit informování zástupců zaměstnanců a dotčených zaměstnanců
- Užití systému s výsledkem, který má právní nebo obdobný účinek na újmu dotčené osoby – zajistit srozumitelné vysvětlení role systému v procesu rozhodování a jeho hlavní prvky rozhodnutí

ORGANIZACE PLÁNUJE VYUŽÍVAT AI S VYSOKÝM RIZIKEM

- Vypracovat FRIA u systémů určených pro zjišťování úvěruschopnosti a určování rizik u životního a zdravotního pojištění
  - Popis účelů (v souladu se zamýšleným účelem)
  - Kategorie dotčených FO a skupin osob
  - Specifická rizika, která mohou pravděpodobně způsobit újmu dotčeným FO/skupinám
  - Popis zajištění lidského dohledu
  - Plánovaná opatření v případě vzniku újmy
- Před nasazením systému (může využít FRIA vypracovaný poskytovatelem), updaty
- Oznámení orgánu dohledu nad trhem
- Možno využít i DPIA

ORGANIZACE PLÁNUJE VYUŽÍVAT AI S OMEZENÝM RIZIKEM

- Jedná se skutečně o takový systém?
  - Systémy používané pro rozpoznání emocí a biometrické kategorizace mohou být i rizikové systémy
- Transparentnost, informace, ochrana osobních údajů

ORGANIZACE PLÁNUJE VYUŽÍVAT AI SYSTÉM

- Ochrana duševního vlastnictví
- Ochrana osobních údajů
- Smluvní podmínky s poskytovatelem
  - Umožnění splnit povinnosti uživatelů
  - Umístění dat
  - Vztah k osobním údajům
  - Kyberbezpečnost
  - Využití dat koncových uživatelů pro další vývoj a post-market monitoring poskytovatelů

ORGANIZACE PLÁNUJE VYVÍJET AI SYSTÉM

- Jaké je plánované využití AI systému?
- Nejedná se o vysoce rizikový systém?
  - Jde o bezpečnostní část výrobku nebo výrobek uvedený v příloze I, která podléhá certifikaci?
  - Jde o výrobek uvedený v příloze III?
    - I když se o výrobek v příloze III jedná, nepředstavuje vážné riziko pro zdraví, bezpečnost a základní práva, tím že splňuje jedno nebo více z kritérií, ledaže profiluje FO

! Pozor – uživatel, který změní zamýšlený účel systému se stává poskytovatelem AI systému s vysokým rizikem

ORGANIZACE PLÁNUJE VYVÍJET AI SYSTÉM S VYSOKÝM RIZIKEM

- Definovat účel systému a rozumně předvídatelná rizika, je-li systém využíván s tímto účelem, a navrhnout opatření na jejich zmírnění
- Zajistit relevantní, dostatečně reprezentativní, kompletní a minimálně chybné datasety
- Osobní údaje – zjistit původní účel zpracování, zpracování zvláštních kategorií údajů pro zjištění a nápravy předpojatostí (*bias*) po zajištění opatření na ochranu základních práv a svobod a splnění dalších podmínek uvedených v AIA
- Ověřit si podmínky GPAI
- Zajistit smlouvy s dodavateli nástrojů, služeb, atd., které specifikují nezbytné informace, možnosti, a další služby založené na všeobecně přijímaném *state-of-art*
- Zajistit instrukce a informace uživatelům vč. informaci uvedených v AIA
- Zajistit možnosti lidského dohledu (jejich určení nebo přímo je přímo začlenit)
- Zajistit uchovávání protokolů, pokud jsou pod jeho kontrolou
- Zajistit certifikaci a zápis do databáze vysoce rizikových systémů
- Zajistit monitorování systému již uvedeného na trh (post-market monitoring), včetně plánu tohoto monitorování

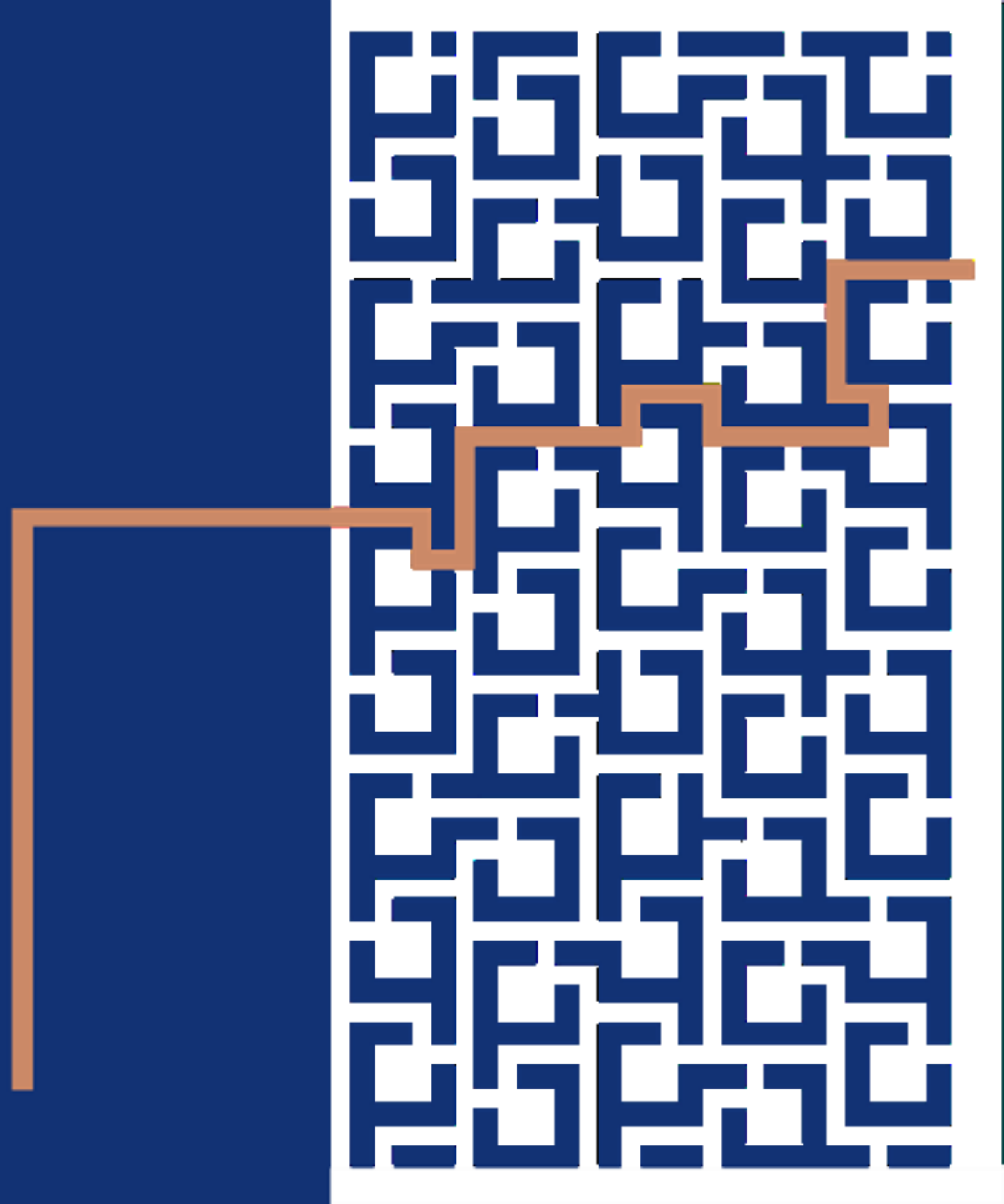
ORGANIZACE PLÁNUJE VYVÍJET AI SYSTÉM S OMEZENÝM RIZIKEM

- Zajistit, aby systém, který interaguje s FO, umožňoval, aby tyto osoby byly informovány o komunikaci s AI systémem, pokud to není zřejmé
- Zajistit, aby systémy vytvářející obsah označovaly obsah generovaný AI ve strojově čitelném formátu
  - Netýká se editorů dokumentů

ORGANIZACE PLÁNUJE VYVÍJET AI SYSTÉM

- Využití duševního vlastnictví
  - Výjimka TDM v AZ
- Využití dat k trénování a testování
  - Osobní údaje nebo syntetická data
  - Testování je také zpracování osobních údajů
- Podmínky GDAI
- Smluvní podmínky s uživateli
  - Dodání dat k trénovaní a testování
  - Dodávání dat k post-market monitoringu
  - Kde budou data/osobní údaje uloženy
  - Smluvní zajištění povinností uživatelů

VI. ZÁVĚREČNÁ SHRNUTÍ



Implementace AI Aktu

- Určení/vytvoření příslušného orgánu dohledu a určení jednotného kontaktního místa
- Změny na úrovni vnitrostátních předpisů
  - Pravidla pro AI systémy pro účely národní bezpečnosti
  - Povolení využití biometriké identifikace na dálku v reálném čase a nastavení procesních pravidel reportování
  - Procesní pravidla a nastavení kompetencí
  - Zavedení správních sankcí
  - Zavedení možnosti podat stížnost
  - ...
- Česká republika jako **AI Friendly Country** - klíčová role veřejných orgánů!



DISKUSE?

DĚKUJI
ZA POZORNOST

Jungmannova 17 | 110 00 Praha 1
| Česká republika

+420 221 506 700
www.cevro.cz