Microsoft

Umělá inteligence

ochrana dat, soukromí, bezpečnost a autorské právo

Dalibor Kačmář
National Technology Officer
Microsoft ČR a SR

Co je generativní AI?

Generativní AI

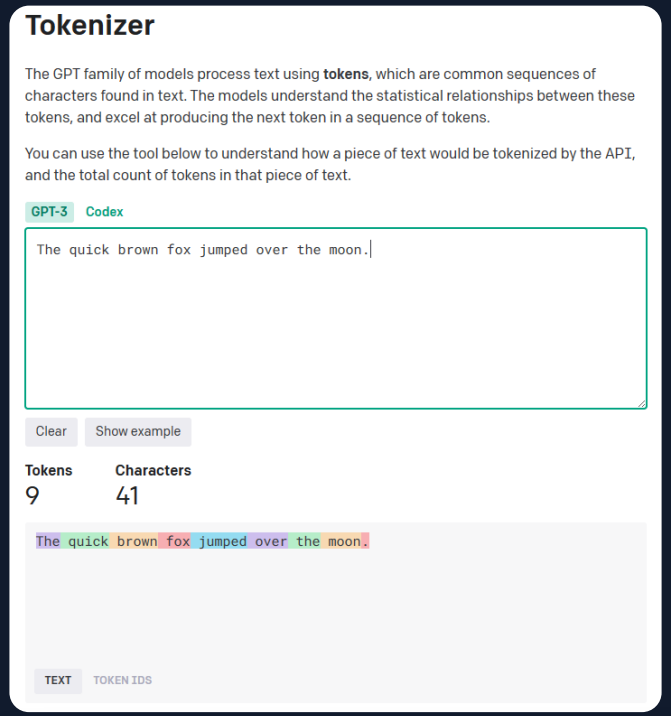
Jde o AI model, který je vybudován na základě analýzy velkého množství dat, aby predikoval jaká sekvence bude následovat. Uživatelé mohou použít také systémy pro „generování“ nebo vytváření textů, obrázků nebo videa na základě vstupu v přirozeném jazyce.

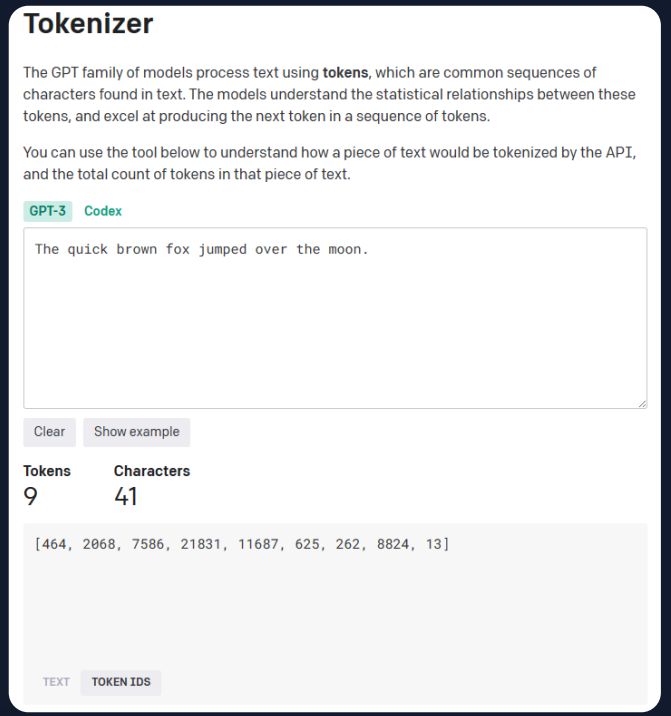
Příklad generativní AI

- Jazyk – GPT-4, LLaMa, Kosmos-1
- Obrázky – DALL-E, Midjourney, Stable Diffusion
- Kód – Codex, GitHub Copilot

Jak generativní AI funguje

- AI Model e trénuje analýzou velkého množství dat.
- Trénovací data jsou vstupem do neuronové sítě jako "tokeny", kde síť přijímá vstup jako čísla.
- Síť se sama přizpůsobuje, dokud výstup neukáže, že se model naučil vzory a může na základě těchto vzorů vytvářet předpovědi.
- Tento proces se nazývá „self-supervised learning" a může odvodit znalosti i v případě, že data nemají „labels" vytvořené člověkem.
- Jakmile je model vytvořen, k provádění svých úkolů už nepoužívá trénovací data.

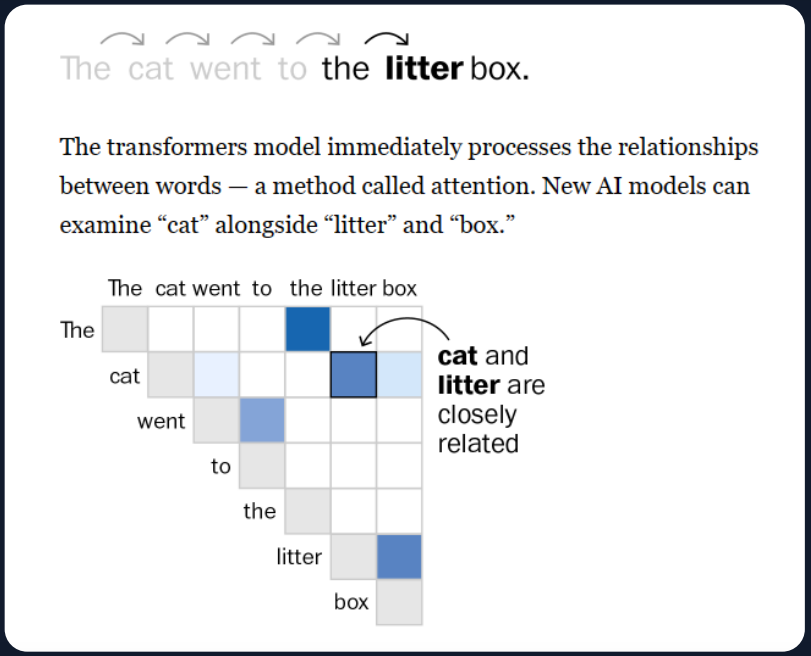




OpenAI API, https://platform.openai.com/tokenizer

Jak generativní AI funguje:
Velké jazykové modely

- Podíváme-li se na velké jazykové modely (LLM) například:
  - Jazykové modely udávají pravděpodobnost dalšího slova v posloupnosti slov.
  - Generativní AI předpovídá další slovo na základě toho, co bylo předtím.
- LLM se dívají na naše používání jazyka v masivním měřítku.



[How AI like ChatGPT and Dall-E got frighteningly good so quickly - Washington Post]

Co generativní AI je a co není

je …

- ✓ Prediktivní algoritmus, který provádí matematiku
- ✓ Nástroj používaný lidmi, „copilot“
- ✓ Proces, jehož cílem je učinit znalosti dostupnějšími a užitečnějšími

není …

- × Kopírovací stroj
- × Databáze
- × Schopnost tvorby úsudku nebo lidského poznání

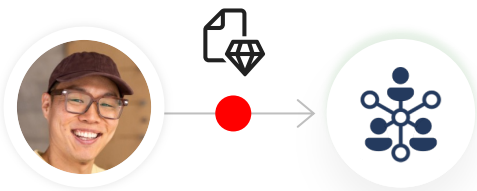
Copilot for M365

Bezpečnostní obavy spojené s použitím AI

Nedostatečné porozumění funkce AI aplikace může vyústit v problému s bezpečností a souladem

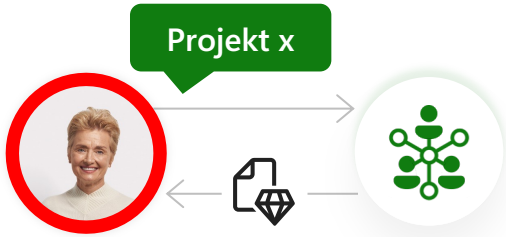
1

Únik dat:
Uživatelé mohou neúmyslně vyzradit citlivá data prostřednictvím AI aplikací



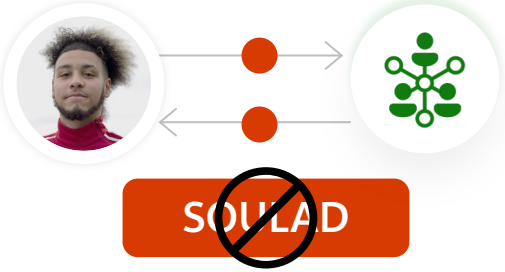
2

Přílišné sdílení dat:
Uživatelé mohou získat přístup k citlivým datům prostřednictvím AI aplikací, k nimž nenají autorizaci



3

Použití porušující soulad:
Uživatelé použijí AI aplikace pro vytváření neetického nebo jiného rizikového obsahu



Sdílené odpovědnosti za bezpečnost při použití AI

Přístup k produkčnímu prostředí

Nastavuje řízení přístupu, které striktně omezuje stálý přístup k datům zákazníka nebo produkčnímu prostředí

Ochrana dat

Šifrování uložených i přenášených dat průmyslovými standardy (BitLocker, HSM, TLS atd.)

Přístup k AI aplikacím

Nastavení politik přístupu s využitím řešení pro správu identit a přístupu

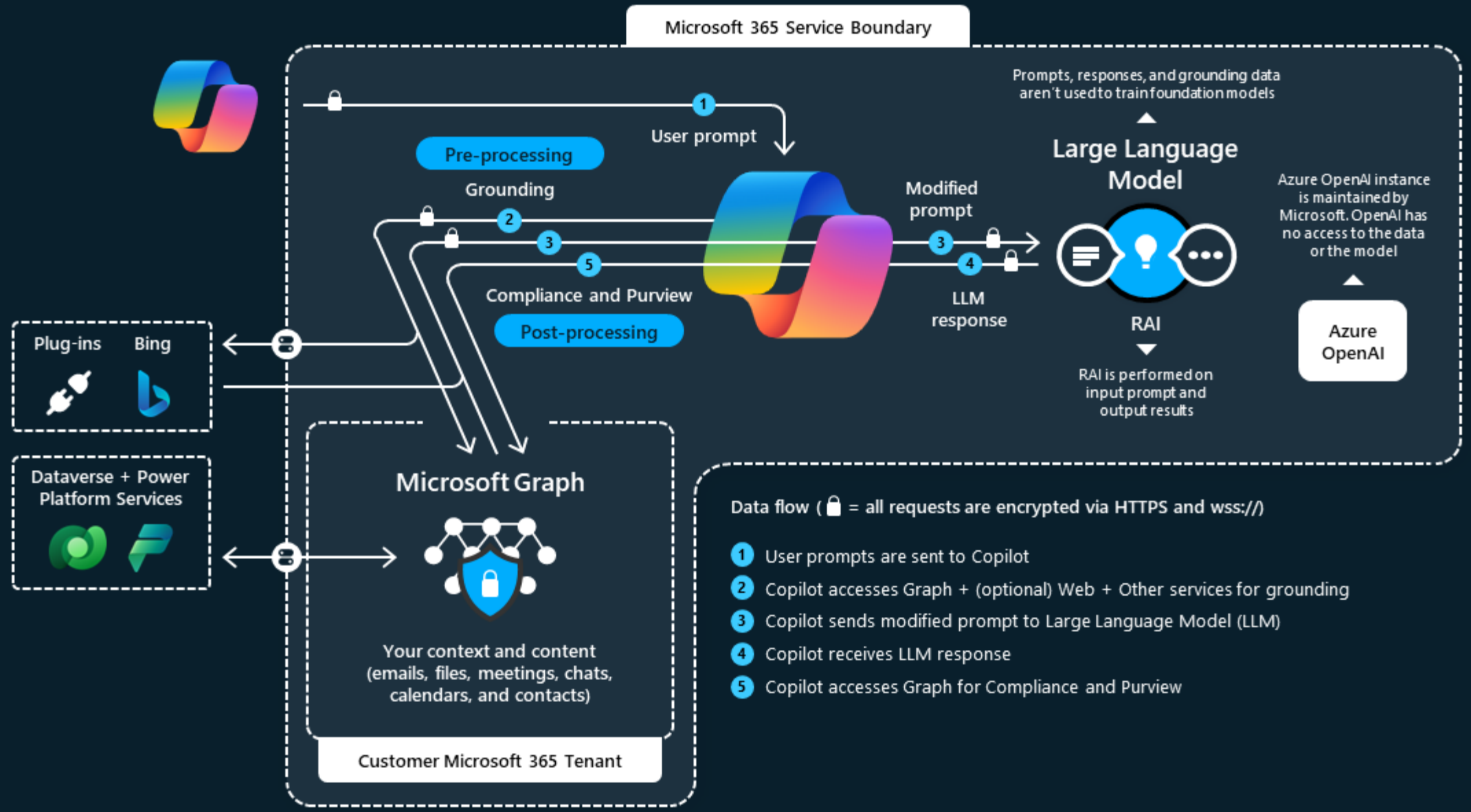
Ochrana a správa dat vkládaných a vytvářených aplikacemi AI

Aplikace řízení bezpečnosti a souladu na data v AI interakcích

Odpovědnost uživatele při použití AI

Realizace školení bezpečnosti použití AI v rámci celé organizace

Microsoft Copilot for Microsoft 365 architecture



Data flow (= all requests are encrypted via HTTPS and wss://)

1. User prompts are sent to Copilot
2. Copilot accesses Graph + (optional) Web + Other services for grounding
3. Copilot sends modified prompt to Large Language Model (LLM)
4. Copilot receives LLM response
5. Copilot accesses Graph for Compliance and Purview

Nejčastější dotazy:

Ochrana dat a osobních údajů v AI

Podmínky pro použití a uložení/zpracování dat

Které T&Cs jsou aplikovatelné na Copilot for M365 službu?

Od března, 2024, Copilot for M365 Service patří mezi Microsoft **Office 365 Core Online Service** v Podmínkách pro produkty. Vztahují se na něj všechny závazky pro ochranu dat a soukromí v Microsoft Products and Services Data Protection Addendum.

Je služba poskytována v EU?

Ano.

Služba Microsoft Copilot for M365 je službou, kterou se vztahuje závazek EU datové hranice

Zpracování dat

Jaká data Copilot používá a zpracovává?

Služba pracuje primárně s daty zákazníka uloženými v

- Službách M365 (Exchange Online, SharePoint Online, OneDrive, Teams, atd.), včetně sémantického indexu
- Data obsažená v Microsoft Graph

Služba může dále zpracovávat data nacházející se mimo hranice služby Copilotu

- Web content plugin – využívá data Bingu a službu Microsoft Copilot (ex Bing Chat Enterprise)
- Datové konektory - externí datové zdroje

Použití uvedených zdrojů je konfigurovatelné a plně pod kontrolou zákazníka

Přístup provozovatele k datům

Bude mít Microsoft přístup k datům zákazníků, promptům a odpovědím?

Ochrana dat zákazníka je identická jako u ostatních služeb Microsoft 365. Smluvně ošetřena v Podmínkách pro produkty a Dodatku pro ochranu osobních údajů.
Mezi zákaznická data se patří i prompty a generované odpovědi.

Na rozdíl od Azure OpenAI se Abuse Monitoring provádí pouze v reálném čase a nedává Microsoftu možnost pro automatický nebo personální přezkum.

Microsoft 365 data nejsou sbírána nebo ukládána ve službě Azure OpenAI.

Microsoft Copilot for M365 – ochrana dat

Jak je nakládáno s daty uživatele a klasifikovaným obsahem?

Model oprávnění brání nechtěnému úniku informací mezi uživateli, skupinami a tenanty.

Sémantický index respektuje identitou stanovené hranice => proces groundingu přistupuje pouze k obsahu dostupnému uživateli.

Jsou respektována Sensitivy Labels a odpovídající šifrování, které jsou aplikovány přes Microsoft Purview Information Protection.

Nejčastější dotazy:

AI a autorská práva

Časté dotazy

- Umožňuje autorský zákon trénovat modely umělé inteligence na dílech chráněných autorským právem?
- Jsou výstupy generované AI chránitelné autorskými právy?
- Porušuje uživatel AI autorská práva?

Náš přístup

Níže jsou uvedené principy, které Microsoft používá při přemýšlení o AI a autorských právech

1. nástroje a uživatelé AI musí respektovat autorská práva
2. veřejnost má právo používat technologie k rozvíjení znalostí na základě děl chráněných autorským právem
3. nástroje AI musí být přínosem pro širokou společnost, nikoli úzkou

Duševní vlastnictví

Kdo je vlastníkem promptu / vstupních dat?

Prompty zákazníků odeslané do AI služby splňují definici "zákaznických dat" v Microsoft's Data Protection Addendum. Společnost Microsoft nemá žádná vlastnická práva k těmto datům.

Dodatek o ochraně osobních údajů stanoví, že:

„Platí ustanovení mezi smluvními stranami, že si zákazník zachová všechna práva, duševní vlastnictví a zájem týkající se údajů zákazníka a dat odborných služeb. Společnost Microsoft nezískává k údajům o zákazníkovi nebo datům odborných služeb žádná práva s výjimkou práv, která jsou společnosti Microsoft udělena v tomto oddílu.“

Duševní vlastnictví

Kdo vlastní výstupní obsah / odpovědi generované AI?

Microsoft ne.

Obsah generovaný umělou inteligencí, tj. výstup nebo odpovědi, není ve vlastnictví Microsoftu a Microsoft si nebude nárokovat žádná vlastnická práva. To, zda je zákazník vlastní, bude záviset na řadě faktorů, které společnost Microsoft nemůže ovlivnit, včetně platných zákonů.

Nejčastější dotazy:

Odpovědnost

AI služby – odpovědnost použití

Jsou všechny AI služby použitelné bez ohledu na potenciální rizika?

Některé služby (nebo jejich verze) mohou vyžadovat registraci a podléhají omezením přístupu a užití na základě kritérií oprávněnosti a používání stanovených společností Microsoft.

- Schválení použití pro deklarované scénáře a podmínky způsobilosti
- Microsoft může přehodnotit oprávněnost použití
- Při neplnění podmínek nebo uvedení nesprávných informací – ukončení služby, bez odkladu
- V případě, že zákazník již nesplňuje kritéria oprávněnosti - 12 měsíční výpovědní lhůta

Na následující služby Azure AI se vztahují podmínky omezeného přístupu (viz Product Terms):

- Azure AI Speech, převod textu na řeč, funkce Custom Neural Voice
- Azure AI Vision Face API
- Azure AI Vision celebrity recognition
- Služba ověření mluvčího ve službě Azure AI
- Applied AI ve službě Azure Video Indexer
- Azure OpenAI
- Azure OpenAI (filtrování upraveného obsahu/monitorování nesprávného použití)

AI služby - Odpovědnost

Bude společnost Microsoft bránit zákazníka, pokud výstupní obsah porušuje autorská práva třetích stran?

Ano, Customer Copyright Commitment – součást Product Terms

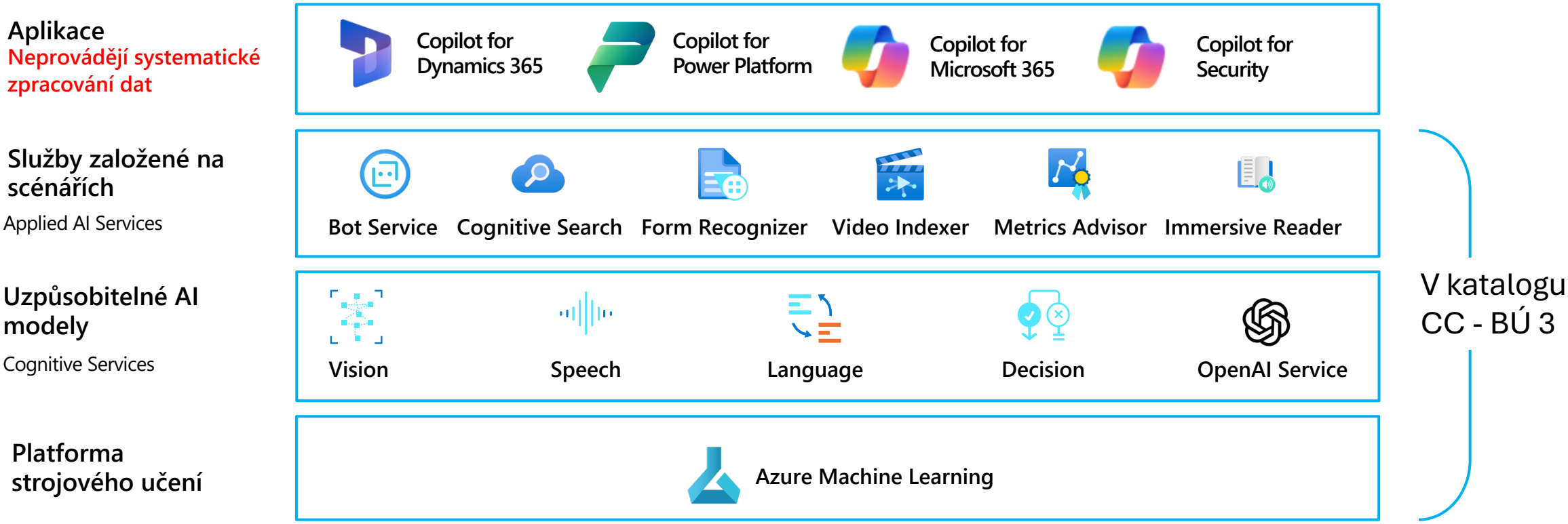
Pokud bude zákazník obviněn z porušení autorských práv, převezmeme odpovědnost za potenciální právní rizika, která s tím souvisí.

Konkrétně platí, že pokud třetí strana zažaluje komerčního zákazníka za porušení autorských práv při používání **krytých služeb** společnosti Microsoft nebo výstupu, který generují, budeme zákazníka hájit a uhradíme částku všech nepříznivých rozsudků nebo vyrovnání, která vyplynou ze soudního sporu, pokud zákazník použil doporučení a filtry obsahu, které jsme zabudovali do našich produktů.

Pozor ale na podmínky! Content filtry a metaprompty, vědomé zneužití, práva na vlastní data

Mýtus 1: Moderní služby umělé inteligence nemůžeme ve veřejné správě používat

- Všechny Microsoft AI služby, které se používají pro vytváření informačních systémů veřejné správy jsou schváleny v BÚ3 „vysoká“ v katalogu cloud computingu (podle 316/2021 Sb.)
- Služby Copilot nenaplňují definici ISVS a nespadají do povinností podle 365/2000 Sb. (ZoISVS)



Mýtus 2: Naše data jsou používána k trénování AI, nemáme nad daty kontrolu

Azure Open AI

Nasazeno ve vaší Azure subskripci, vámi zabezpečené a svázané s vašimi datovými sadami a aplikacemi

Zabezpečení na firemní úrovni Role-based access control (RBAC)

Možnost šifrovat data pomocí customer-managed key

Vestavěná zodpovědná AI pro detekci a zmírnění škodlivého používání

Bezpečné propojení přes Azure Virtual Network s privátním koncovým bodem

Copilot for Microsoft 365

Používá a ukládá data v M365 tenantu, včetně sémantického indexu a Microsoft grafu

Respektuje přístupová práva a klasifikaci informací (Purview Information Protection)

Dotazy a odpovědi jsou pouze vaše data, zůstávají výhradně uchovány ve vašem tenantu a nepoužívají se k trénování

Možnost omezení zpracování dotazů pouze nad vlastními daty (bez internetových zdrojů)

Služby jsou provozovány se zárukami Evropské datové hranice (EUDB)


Microsoft