Průvodce pro úřady a další organizace veřejné správy

AI v kontextu
a systémově

Základní orientace
pro každodenní praxi

Profesionálně,
ne nahodile

Etické a odpovědné využívání AI
ve veřejné správě

Únor 2026

Etické a odpovědné využívání AI ve veřejné správě

Kolektiv autorů z organizací:

Sekce pro státní službu Úřadu vlády České republiky - garant

Český metrologický institut

Český telekomunikační úřad

Digitální a informační agentura

Filozofický ústav Akademie věd České republiky

Legální kód

Ministerstvo průmyslu a obchodu

Ministerstvo vnitra

Národní úřad pro kybernetickou a informační bezpečnost

prg.ai

Úřad vlády České republiky

1. vydání | únor 2026 | právní stav k únoru 2026

Obsah

Při přípravě tohoto Průvodce byla využita umělá inteligence, a to zejména pro úpravy textů a stylistickou práci s nimi a pro grafickou úpravu. Finální obsah a jeho interpretace jsou výsledkem práce autorského týmu.

Jak Průvodce používat

Průvodce nabízí praktická a obecně použitelná doporučení, jak využívat AI tak, aby byl respektován člověk, chráněno soukromí, hájen veřejný zájem, zajištěna spravedlnost a posilována důvěra mezi občanem a státem i demokratické hodnoty. Je míněn jako praktický podpůrný materiál pro zaměstnance, vedoucí pracovníky i metodiky a další aktéry, kteří se při své práci mohou setkat s technologiemi AI. Může pomoci při rozhodování o implementaci AI a nastavování procesů. Není to závazný předpis, ale opora pro organizace veřejné správy různého typu a velikosti. Vychází z respektu k různorodosti veřejné správy i jejích klientů. Samotné principy nestačí, pokud nezískají konkrétní podobu v každodenní praxi – tento Průvodce má pomoci, aby se tak skutečně stalo.

Průvodce nabízí srozumitelné vysvětlení pojmů, praktické pomůcky, otázky k zamyšlení při zavádění či využívání AI nástrojů a odkazy na právní rámec i další zdroje informací. Doporučujeme s ním pracovat jako s „živým“ dokumentem, který lze průběžně doplňovat podle vývoje technologií i situace ve Vaší organizaci či celé veřejné správě. Lze jej využít při tvorbě interních předpisů, při školení zaměstnanců či jako podklad pro hodnocení etických dopadů projektů s prvky AI – jako nástroj podpory pro odpovědné, bezpečné a veřejnosti prospěšné využívání AI.

Průvodce je rozdělen do tří vzájemně provázaných částí:

Část I

Zaměřuje se na vysvětlení toho, co je AI, a přibližuje její etický, lidskoprávní a odpovědnostní rozměr. Právě tyto aspekty by měly tvořit základní rámec našeho uvažování o AI a měly by být přítomny při jakémkoliv jejím využívání.

Část II

Věnuje se právním a bezpečnostním aspektům využívání AI v práci zaměstnanců veřejné správy. Ačkoliv se může její obsah na první pohled jevit jako odbornější, jde o souvislosti, které jsou pro odpovědné využívání AI zásadní. Bez jejich znalosti se vystavujeme zbytečným rizikům.

Část III

Primárně je určena těm, kteří mají v organizacích veřejné správy na starosti řízení a nastavování využívání AI, a proto klade větší důraz na pohled organizace jako celku. Její pročtení lze ovšem bezpochyby doporučit i ostatním uživatelům, neboť poskytuje širší kontext a podporuje komplexní porozumění této problematice.

Slovníček pojmů a zkratek

- **AI**

  Umělá inteligence (z anglického *artificial intelligence*)

  Lze na ni nahlížet jako na soubor metod, které umožňují systémům učit se ze vzorů v informacích, odhadovat nejlepší výstup, klasifikovat, generovat obsah nebo rozhodovat v nejistotě. Má schopnost se adaptovat podle nových vstupů.

  V Průvodci na ni budeme nahlížet primárně optikou tzv. generativní AI a velkých jazykových modelů.

- Algoritmus

  Přesný a jednoznačný postup řešení daného typu úlohy, složený z konečného počtu kroků.

  Sekvence instrukcí, které se provádějí v určitém pořadí, aby se dosáhlo požadovaného výstupu z daných vstupních informací.

- **Automatizace**

  Systematické převádění opakovatelných, předem určených činností z lidí na stroje či software podle pevně daných pravidel a postupů.

  Dělá přesně to, co je naprogramováno (pevný algoritmus); pokud vstup neodpovídá očekávání, selže nebo vyžaduje výjimku.

- **Bias**

  Neboli zkreslení, označuje odchylku výsledků, rozhodnutí nebo výstupů od objektivní či očekávané hodnoty. Může vznikat v důsledku způsobu sběru informací, jejich struktury, volby metod, nastavení modelu nebo lidského zásahu. Nejedná se nutně o vědomou předpojatost ani o systematickou chybu, ale o obecné zkreslení, které může mít různé příčiny a projevy.

  Může vznikat například přenesením lidských rozhodovacích vzorců do trénovacích informací (např. při učení nástroje pro výběr uchazečů o zaměstnání), což může vést k nežádoucímu znevýhodňování určitých skupin. Dalšími zdroji bias mohou být nevyvážené nebo neúplné vstupní informace, konstrukce algoritmu, nevhodné metriky hodnocení apod.

- **GDPR**

  Obecné nařízení o ochraně osobních údajů

  Norma chránící fyzické osoby a jejich osobní údaje. Má za cíl dát lidem větší kontrolu nad jejich osobními údaji a zároveň sjednotit a zmodernizovat pravidla pro jejich zpracování napříč Evropskou unií. Nařízení definuje základní práva občanů, povinnosti zpracovatelů údajů a zavádí sankce za porušení pravidel.

- **PDCA**

  Akronym pro metodu řízení *Plan-Do-Check-Act* (Plánuj-Proveď-Zkontroluj-Jednej), která slouží k neustálému zlepšování procesů a produktů.

  Tento cyklus je iterativní, tedy opakující se, a skládá se ze čtyř fází: nejprve se plánuje, poté se změny implementují, následuje kontrola výsledků, a nakonec se na základě zjištění buď implementuje trvalé řešení, nebo se cyklus zopakuje s novým plánem.

- **Principialismus**

  Teorie, která ve stručnosti říká, že jako východisko etického uvažování má posloužit ne nějaké ústřední jednotící pravidlo (typu kategorického imperativu nebo utilitaristického počítání kladů a záporů konkrétního jednání), ale sada principů, na kterých se v rámci dané disciplíny shodneme.

Symboly a jejich význam

Upozornění

Informace

Vhodný postup

Takto ne

K zamyšlení

Co najdete v přílohách

Kromě Průvodce, kterého právě máte před sebou, jsme pro Vás připravili také praktické pomůcky, které můžete ve Vaší organizaci využít v rámci implementace a řízení AI.

- **Ilustrativní "Základní přehled povinností při zavádění AI nástrojů"**

  Shrnuje povinnosti zavádějícího subjektu/povinné osoby v souvislosti s plánovaným zavedením AI nástroje. Vychází z materiálu Ministerstva vnitra jako zavádějícího subjektu/povinné osoby v souvislosti s přípravou projektu AI asistenta pro správní řízení, nicméně uváděné povinnosti jsou v zásadě obecně platné, a proto je přikládáme pro inspiraci.

- **Metodika analýzy vhodných případů využití AI (tzv. use-cases)**

  Vytvořena Ministerstvem průmyslu a obchodu jako reakce na potřebu organizací systematicky identifikovat oblasti, kde může AI přinést největší hodnotu. Součástí jsou také bezprostředně využitelné šablony.

- **Karty hodnocení dopadů implementace AI na etické principy**

  Zpracované dle doporučení UNESCO jako myšlenkový postup při hodnocení dopadu AI řešení na vybraný etický princip, normu či hodnotu. Obsahuje kartu hodnocení pozitivních a negativních dopadů implementace AI.

„Cesta k etickému a odpovědnému využívání AI vede přes osvětu, ne restrikce“

Poslání Průvodce

Východiskem vzniku tohoto Průvodce se stala profesionalita, kterou využívání AI může buďto umocňovat, nebo naopak ohrožovat. Být profesionálem neznamená jen dobře ovládat předpisy, postupy a nástroje, ale také používat je odpovědně a v souladu s etickými principy. Jako zaměstnanci veřejné správy nerozhodujeme jen sami za sebe – naše rozhodování a každodenní činnosti mají dopad na život občanů, fungování společnosti i důvěryhodnost státu. To platí i pro využívání AI. Patří k naší základní výbavě přemýšlet o tom, jak nás AI ovlivňuje, jak mění výstupy naší práce a jak tyto výstupy působí na jednotlivé lidi i na prostředí kolem nás.

Průvodce má sloužit hlavně Vám, lidem v organizacích veřejné správy, jako praktická pomůcka. Pomůže zorientovat se v rychle se rozvíjejícím světě AI a nabídne rámec pro její etické a odpovědné využívání ve veřejné správě. AI nás i naši práci změní, ale cílem je, aby tato změna byla k lepšímu. Naše úkoly a odpovědnost zůstávají, AI je „jen“ další nástroj, jak je plnit lépe, rychleji a efektivněji. Může nám ulevit od rutinních činností, podpořit kritické a systémové myšlení a pomoci nám tam, kde máme slabší místa – pokud si zároveň udržíme střízlivý nadhled a uvědomíme si její limity i vlastní odpovědnost.

Část I

Základy umělé inteligence a principy jejího využívání

AI jako pracovní nástroj, ne náhrada zaměstnance - o umělé inteligenci pro veřejnou správu

V poslední době je možné o AI slyšet v médiích, na nejrůznějších akcích i v rámci pracovních diskuzí jako o technologii, která „změní svět“ nebo „vezme lidem práci“. Realita ve veřejné správě je však mnohem střízlivější a praktičtější. AI není magie ani vševědoucí robot ze sci-fi filmů. Je to pracovní nástroj – podobně jako se kdysi staly nástrojem psací stroj, počítač nebo Excel.

Tato kapitola pomůže pochopit, jak tento nástroj funguje, kde Vám může ušetřit hodiny práce a kde byste naopak měli být maximálně obezřetní.

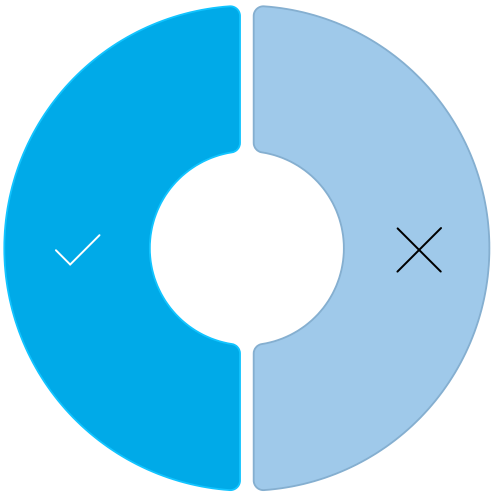
Co je (a není) AI

Pro potřeby běžné praxe veřejné správy se nejčastěji setkáváme s tzv. generativní AI (nástroje jako ChatGPT, Microsoft Copilot nebo Gemini), na kterou je položen důraz také v tomto Průvodci.

Co to je

Představme si AI nástroje jako extrémně výkonné našeptávače. Tyto nástroje „neví“ a „nemyslí“. Ony pouze na základě obrovského množství přečtených textů určují, jaké slovo či věta by měly následovat.

Jsou to mistři formy a generování jazyka.



Co to není

AI není vyhledávač faktů (jako Google). Nemá vědomí, morálku ani odpovědnost. Nerozumí obsahu textu, který generuje, stejně jako kalkulačka „nerozumí“ matematice – pouze zpracovává vzorce.

Aby mohla být AI využívána eticky, odpovědně i bezpečně, je nezbytně nutné znát její omezení a limity.

Klíčovým pravidlem je: **AI může být užitečným nástrojem pro tvorbu návrhů, avšak její výstupy je nutné v každé fázi kriticky posuzovat a nepřebírat je bez odborného lidského zhodnocení, zejména při rozhodování!**

K čemu AI využít

Práce s textem

Shrne dlouhé dokumenty, navrhne osnovu zprávy, přeformuluje složitý text do srozumitelné podoby nebo vytvoří první koncept e-mailu.

Kreativita a inspirace

Pokud máme „syndrom prázdného papíru“, AI během chvíle nabídne řadu nápadů, jak strukturovat projekt nebo materiál, či jaké otázky položit v dotazníku.

Překlady a korektury

Velmi rychle a kvalitně překládá úřední texty (i když vyžaduje kontrolu) a umí najít gramatické chyby či texty vhodně upravit.

Na co si dát pozor

Vymýšlení si

Smyšlený výstup je největší riziko. Protože AI jen „doplňuje slova“, klidně si s naprostou jistotou vymyslí neexistující zákon, citaci, událost nebo statistiku, aby věta zněla dobře.

Fakta, která AI uvádí, si vždy ověřujte!

Bias a zkreslení

AI byla trénována na informacích z internetu. Může tedy nezáměrně přebírat stereotypy a předsudky (např. genderové nebo rasové), které ve vstupech byly.

Absence kontextu

AI nezná vnitřní předpisy ani nepsaná pravidla komunikace s konkrétním útvarem či organizací, stejně jako přesné podmínky a kontext situace, kterou s její pomocí řešíme.

Existují situace, kdy je využití AI nejen nevhodné, ale může být i nelegální nebo eticky nepřípustné.

Citlivé a osobní údaje (GDPR)

Je nepřípustné do AI nástrojů nespravovaných Vaší organizací vkládat jména občanů, rodná čísla, adresy nebo neveřejné interní dokumenty. Jakmile jsou údaje vloženy, je nad nimi ztracena kontrola a mohou být použity k dalšímu trénování AI či k jiným skrytým účelům.

Rozhodování o občanech

AI nesmí sama rozhodovat o přidělení dávky, vydání povolení, sankci či jiných právech a povinnostech občanů. Nemůže ani nahradit princip správního uvážení. Finální rozhodnutí a odpovědnost musí vždy nést člověk. To se mimo jiné týká například i případného automatizovaného vedení správního řízení.

Nízká kvalita vstupů

Pokud nemáme ověřené a relevantní vstupy, AI z nich „nevykouzlí“ pravdu. Platí pravidlo "*Garbage In, Garbage Out (Vhodíš odpad, vypadne odpad)*".

Časový tlak bez kontroly

Pokud není čas si výstup AI přečíst a ověřit, neměl by se použít. Odeslat text vygenerovaný AI bez kontroly je neetické i neprofesionální.

Zlaté pravidlo

Na AI je potřeba pohlížet jako na velmi snaživého, ale nezkušeného stážistu. Svěříme mu úkol, ale jeho výstup si vždy pečlivě zkontrolujeme, než ho pošleme dál.
Podpis pod dokumentem je náš, ne stážisty.

Etika, lidská práva a odpovědnost ve využívání AI

AI se rychle stává běžnou součástí služeb pro občany i řízení organizací. Veřejná správa tak stojí před otázkou, jak tyto nástroje používat tak, aby pomáhaly a zároveň neoslabovaly její poslání – poskytovat veřejné služby, chránit práva lidí, zajišťovat dodržování principu rovného zacházení a udržovat důvěru občanů. U AI proto nejde jen o její funkčnost a efektivitu, ale i o to, zda její využití podporuje spravedlnost, rozmanitost, respekt k lidské důstojnosti a další hodnoty naší společnosti.

Etický rozměr využívání AI

Zákony určují, co je dovoleno a co zakázáno, ale samy o sobě nestačí. Etika, která studuje morální normy umožňující lidem žít a spolupracovat ve složitých společenstvích, není pouhým teoretickým cvičením, ale velmi užitečným nástrojem, který pomáhá společnosti fungovat a předcházet konfliktům. Proto je přínosné, ne-li nezbytné, nahlížet na AI z etického hlediska. To napomáhá předcházet nežádoucím jevům a udržovat rovnováhu mezi lidskými hodnotami a technologickým pokrokem.

> ⓘ Etika pomáhá odpovědět na otázku, co je správné a žádoucí. U AI to znamená přemýšlet o tom, jaké vstupy používá a jaké výstupy poskytuje, kdo nese odpovědnost za výsledná rozhodnutí, a kde už AI zapojit nechceme.
>
> I při formálně dostatečné lidské kontrole se mohou objevovat zkreslení vycházející z použitých kritérií nebo vstupů. K nim se přidávají i naše vlastní přirozené předsudky a zkratky v uvažování. U generativních systémů, které pracují s „nejpravděpodobnější kombinací prvků“, je toto riziko ještě výraznější. Proto je nutné pečlivě zvažovat, kde je využití AI vůbec vhodné, předem posuzovat rizika, nastavit jasná pravidla pro lidský dohled a kriticky hodnotit každý výstup.

Nejde jen o úsporu času a zrychlení agendy, ale také o to, aby rozhodování zůstalo srozumitelné, kontrolovatelné a nediskriminační. Pokud občané získají pocit, že AI je „černá skříňka“, která rozhoduje nespravedlivě, manipulativně, v rozporu s hodnotami společnosti nebo je možné ji snadno zneužít, ztratí důvěru nejen v technologii, ale i ve veřejnou správu.

Zaměstnanci veřejné správy už dnes podléhají etickým pravidlům a zodpovídají za své jednání vůči občanům. Stejné nároky musí platit i pro využívání AI – ta má lidský úsudek a odpovědnost doplňovat, nikoli nahrazovat. Tím, že veřejná správa vědomě otevírá otázky etiky a odpovědného využívání AI, ukazuje, že si uvědomuje svoji zvláštní roli a vyšší míru odpovědnosti.

Kontext lidských práv a inkluzivního přístupu

Lidská práva s etikou souvisejí, ale nejsou to totožné kategorie. Lidská práva a základní svobody jsou garantovány a chráněny ústavním pořádkem, zákony a mezinárodními úmluvami, které určují, co stát a jeho orgány dělat nesmí a co naopak musí zajistit. Etika říká, co je správné a žádoucí i tam, kde právo mlčí. U AI nástrojů ve veřejné správě proto potřebujeme obě perspektivy – právní pohled jako pevný rámec a etický pohled jako kompas, který pomáhá rozhodnout, jak nástroj využít s ohledem na důstojnost, spravedlnost, rozmanitost a důvěru občanů.

Lidská práva musí být respektována, chráněna a podporována ve všech fázích životního cyklu AI nástroje – od návrhu a vývoje, přes každodenní využívání až po jeho vyřazení z provozu. Moderní AI nástroje jsou schopné a flexibilní, nesou s sebou ovšem i nezanedbatelná rizika. I zdánlivě neškodná aplikace může poškodit jednotlivce nebo skupiny, prohloubit nerovnosti, oslabit důvěru ve veřejnou správu nebo zasáhnout do demokratických procesů.

⚠ Za diskriminaci se považuje nerovné zacházení ze zakázaných důvodů, které dopadá především na historicky znevýhodněné skupiny osob. Zcela nepřijatelné je využívání AI k represivním, diskriminačním, profilačním nebo jiným obdobně závažným postupům, které ohrožují lidskou důstojnost a další základní lidská práva.

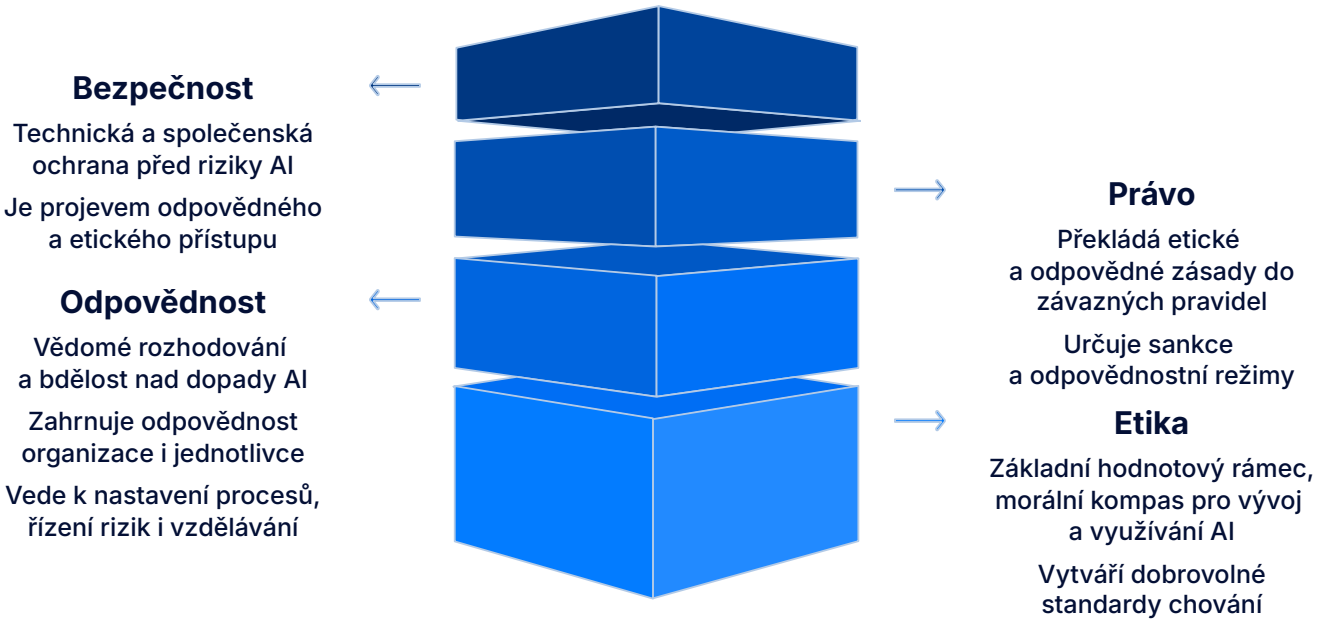
Obzvlášť rizikové se na mnoha zdokumentovaných případech ukázalo být nesprávné využití AI v oblastech týkajících se zdravotnictví, školství, justičního a správního rozhodování nebo bezpečnosti a policejního dohledu. AI je pouhý nástroj. Při nezodpovědném nebo nepozorném využití může napáchat značné škody v oblasti lidských práv, přičemž tyto škody mohou mít původ ve zdánlivě nevýznamném podcenění rizik, nedostatečném zapojení lidského faktoru či v podceněné průběžné kontrole.

Inkluzivním přístupem k využívání AI je myšleno to, že AI ani pravidla pro její využívání nesmí vytvářet nové bariéry a nerovnosti nebo prohlubovat ty stávající. AI nástroje by měly být navrhovány a využívány tak, aby byly, pokud možno, přístupné a prospěšné pro všechny, bez ohledu na osobní či sociální situaci.

⚠ Podpora využívání AI by neměla vést k vyloučení těch, kteří ji z různých důvodů neumí, nemohou nebo nechtějí využívat. Neméně důležité je také zamezení nevhodného sociálního hodnocení těch, kteří AI využívají – neměli bychom se setkat s jejich označováním za méně pracovité, méně kompetentní či lenivější.

Stejně inkluzivně by mělo být nastaveno i vzdělávání v oblasti AI gramotnosti – s respektem k různým výchozím podmínkám zaměstnanců.

Vztah etiky, odpovědnosti, práva a bezpečnosti ve využívání AI lze zachytit např. takto:



Ohled na životní prostředí

Součástí odpovědného přístupu je také ohled na životní prostředí. AI nástroje jsou energeticky náročné, a proto by se měly používat tam, kde skutečně přinášejí přidanou hodnotu – zvyšují efektivitu, kvalitu služeb nebo dostupnost informací. V některých agendách se jejich využití nemusí vyplatit ekonomicky ani environmentálně, a může znamenat jen zbytečné zatížení infrastruktury i životního prostředí.

Při vývoji nebo využití AI by organizace i jednotliví zaměstnanci měli zvažovat nejen funkčnost a přínos, ale i energetickou náročnost, dopady na životní prostředí a připravenost digitální infrastruktury.

Etika a nevyužívání AI

Debata o AI ve veřejné správě se většinou soustředí na souvislosti jejího využívání. Méně si všímáme etiky jejího nevyužití – situací, kdy organizace veřejné správy technologii nezavede, i když existují bezpečné a ověřené nástroje, které by mohly zlepšit služby občanům nebo podmínky pro zaměstnance. Pokud lze díky AI zjednodušit rutinní agendu, ušetřit čas a peníze a přesunout lidi tam, kde je potřeba jejich úsudek a empatie, pak dlouhodobé přehlížení těchto možností není jen technické rozhodnutí, ale i etická volba – občané i zaměstnanci nesou náklady v podobě zbytečných průtahů, výdajů, chyb či nerovností, kterým šlo předejít.

Zároveň ale platí, že neimplementovat AI může být v některých případech správné a odpovědné – pokud chybí důkazy o přínosu, hrozí nepřiměřená rizika, nejsou kapacity pro bezpečný provoz nebo je právní rámec nejasný.

> Eticky přiměřený postup stojí na proporcionalitě a opatrnosti. Tento Průvodce nechce AI ani nekriticky prosazovat, ani odmítat. Nabízí způsob, jak o ní racionálně uvažovat.

Veřejnosprávní organizace, které se ani nepokusí rozumně využít dostupný potenciál, mohou ztrácet důvěru a legitimitu. Ty, které inovují přiměřeně, s jasnými pojistkami a vysvětlením, naopak důvěru veřejnosti posilují.

Minimální standard pro etické a odpovědné využívání AI

Tato část Průvodce nastavuje minimální standard pro etické a odpovědné využívání AI ve veřejné správě a vymezuje základní principy, které by měly řídit každodenní práci s AI nástroji.

Cílem této části je vysvětlit soulad využívání AI s obecnými etickými principy při specifickém přihlédnutí k nárokům a potřebám veřejné správy. Než se dostaneme k detailům, je důležité pochopit některá základní východiska etického uvažování. Platí, že etická pravidla a normy nejsou neměnné – některé jsou s námi dlouho (nekrást nebo být slušný), jiné se za posledních několik desetiletí zásadně proměnily (postoj k domácímu násilí, vztah k různým menšinám). Pozoruhodné je, že významným posunem prošla i aplikovaná etika, která se věnuje pravidlům správného jednání v jednotlivých oborech lidské činnosti (příkladem může být, kromě námi probírané etiky využívání AI ve veřejné správě, také biomedicínská etika nebo etika pomáhajících profesí).

Principialismus jako základní přístup

Znamená, že v etice nevycházíme z jednoho univerzálního pravidla, ale z několika základních principů, na kterých se v dané oblasti shodneme. Ty pak používáme jako vodítko při rozhodování v konkrétních situacích.

Důsledky principialismu

- **Principy jako ideál, který se vyvíjí** - určují směr, ale jejich konkrétní obsah se může v čase měnit. Díky tomu je lze uplatnit i na nové, dosud neřešené situace.
- **Důraz na kontext** - význam principu se vždy posuzuje podle konkrétní situace – například to, co znamená „chránit soukromí“, závisí na čase, místě, roli aktérů a účelu. Uplatnění principu proto vyžaduje dobrou znalost daného kontextu.
- **Vyvažování mezi principy** - principy si mohou navzájem konkurovat (např. transparentnost vs. ochrana soukromí). Etické rozhodování pak spočívá v tom, jak je v konkrétním případě rozumně vyvážit, ne slepě upřednostnit jeden z nich.

> **Řízení se principy není jednoduchý proces, ve kterém je vše předem jasné. Etické problémy jsou často mimořádně komplexní, a tak nejen jejich řešení, ale i pozdější odůvodnění ve světle kritiky (nebo chvály) není jednoduché.**

Pro zaměstnance státní správy by tato problematika neměla být překvapivá novinka – Etický kodex státních zaměstnanců z října roku 2023 vychází ze základních principů správného jednání.

V dalším textu budeme na Etický kodex státních zaměstnanců občas odkazovat, abychom ukázali, jak je etické využívání AI svázáno s obecnějšími pravidly správného jednání zaměstnance (nejen) ve státní správě.

Nyní již k samotným principům, kterými se řídí etika v oblasti AI (pokud by Vás blíže zajímalo, z čeho vycházejí, přelistujte na seznam inspirativních materiálů na konci Průvodce). O některých principech je řeč v dalších částech Průvodce, zde nastíníme, co se za nimi skrývá.

Principy etického a odpovědného využívání AI

Ochrana soukromí

Ochrana soukromí je ve veřejné správě samozřejmý požadavek, u AI ale zvlášť upozorňuje na riziko úniku údajů z nástrojů, do kterých je zadáváme – ať už jde o velké jazykové modely, nebo menší lokální nástroje. Zaměstnanci by vždy měli vědět, kam se zadané údaje dostanou a kdo k nim může mít přístup. Zároveň by se neměli ptát AI nástrojů na osobní údaje svoje, kolegů nebo občanů, ani je do nich přímo zadávat.

> **Takový přístup je v souladu s prvním bodem Etického kodexu státních zaměstnanců.**

Transparentnost

U AI nemusí vždy znamenat rozumět vnitřnímu fungování systému, což je často nereálné. V našem pojetí jde hlavně o to, aby bylo vždy zřejmé, co vytvořil zaměstnanec a kde mu pomohla nebo ho dokonce nahradila AI. Transparentnost, a s ní se pojící vysvětlitelnost původu výsledků a rozhodnutí, musí být přiměřená kontextu a dopadu. Všichni máme nicméně právo vědět, kdo je autorem konkrétního výstupu a jak AI ovlivnila dané rozhodnutí.

Odpovědné využívání

Zahrnuje řadu kroků, které spolu souvisejí, i když to na první pohled nemusí být zřejmé. Není nutné rozumět vnitřní architektuře AI, důležité je vědět, jak s ní pracovat v praxi a kde hledat spolehlivé informace. Základem je využívat AI jen v oblastech, kde je to povolené a vhodné. Podstatné je znát právní i etické limity. Zcela vyloučeno je využívat AI pro protiprávní úkoly či manipulativně - např. k působení na emoce lidí, k podbízení se nebo k šíření nepravdivých informací (včetně deepfake obsahu). Podobně by AI neměly být zadávány ani úkoly zjevně neetické. Velké jazykové modely mohou vytvářet přesvědčivé, ale nepravdivé výstupy („halucinace“), proto je nepřípustné spoléhat se na ně bez pečlivého ověření. Nelze se tvářit, že za výstupy neneseme odpovědnost. Tu má vždy člověk, ne AI nástroj.

Férovost a nediskriminace

Jsou přirozenou součástí toho, jak bychom měli jednat s lidmi obecně. U AI je to ovšem zvlášť citlivé, protože systémy vycházejí z trénovacích vstupů z minulosti. Ty často obsahují skryté i otevřené formy diskriminace, vycházející z nerovnoměrného zastoupení různých (zejm. historicky marginalizovaných) skupin společnosti. Výstupy AI tak mohou být zkreslené, i když jí výslovně „zakážeme“ diskriminovat. Proto je v celé veřejné správě nutný velmi přísný dohled nad tím, jak je AI využívána.

> **Tento princip přímo navazuje na požadavek nestrannosti v Etickém kodexu státních zaměstnanců.**

Bezpečnost

Nástroje AI jsou podobně jako jiné programy zranitelné – mohou obsahovat slabiny, které útočníci zneužijí k získání osobních údajů, informací nebo tajemství. Dalším rizikem je, že řadu nástrojů vyvíjejí soukromé firmy nebo subjekty z nedemokratických zemí, u kterých neznáme jejich skutečné záměry. Proto je rozumné využívat jen takové AI nástroje, které doporučila sama organizace nebo ověřený správce IT, a nezkoušet náhodné nové služby, i když vypadají na první pohled velmi lákavě.

Zachování lidské autonomie a rozhodování

Znamená to, že konečné rozhodnutí musí vždy příslušet člověku, nikoli AI. Ta může pomoci např. při zpracování podkladů, ale odpovědnost za výsledek nese konkrétní zaměstnanec veřejné správy. Tento princip nelze naplnit pouhým formálním potvrzováním výstupů AI ani nastavením hranic, ve kterých člověk rozhodnutí bez dalšího posouzení jen přebírá. Pouze vůči člověku se může klient odvolat, vznést námitku či žádat vysvětlení postupu.

Rozhodování AI navíc často nelze plně vysvětlit, zatímco lidé mají právo vědět, proč bylo v jejich věci rozhodnuto konkrétním způsobem. Proto není role člověka ve finálním rozhodování pouze formální, ale skutečně nezastupitelná.

> **Tento požadavek odpovídá také zásadě odbornosti obsažené v Etickém kodexu státních zaměstnanců.**

Soulad s právem

Soulad s právem vlastně není jen etickým principem, ale základním předpokladem fungování veřejné správy. V oblasti AI to znamená využívat pouze schválené nástroje a počítat s tím, že jejich výstupy mohou být chybné či smyšlené. Zvláštní pozornost je třeba věnovat autorským právům. Ačkoli tvůrci nástrojů často tvrdí, že výsledky patří uživateli, vycházejí z trénovacích vstupů, jejichž původ není vždy jasný. Je proto nutné dbát na to, aby výstupy AI neporušovaly práva autorů (např. při tvorbě obrázků v konkrétním výtvarném stylu). Právních a etických problémů v této oblasti je mnoho, proto je opatrnost a střídmost zcela na místě.

Část II

Právní a bezpečnostní aspekty využívání AI ve veřejné správě

Právní aspekty - přehled pro praxi

Tato část Průvodce poskytuje základní přehled nejdůležitějších legislativních souvislostí problematiky využívání AI organizacemi veřejné správy.

Akt o umělé inteligenci pro organizace veřejné správy

Akt o AI je evropské nařízení, které určuje, které AI nástroje jsou zakázané, které „vysoce rizikové", a na které jsou kladeny mírnější povinnosti. Pracuje s tzv. rizikovým přístupem - čím vyšší riziko pro práva, bezpečí či důstojnost osob, tím přísnější pravidla. Ukládá povinnosti těm, kdo AI vyvíjejí, dodávají nebo využívají (vč. organizací veřejné správy) a zavádí dohled a sankce (podobně jako GDPR; klíčová úloha zde bude podle české úpravy náležet Českému telekomunikačnímu úřadu).

> ⓘ Akt o AI již vstoupil v platnost, ale jeho jednotlivá ustanovení se začínají uplatňovat postupně (v letech 2025–2027). Ne všechny povinnosti tedy platí okamžitě a ve stejném rozsahu. Např. od února 2025 platí část vztahující se k zakázaným praktikám AI (sociální hodnocení, biometrická identifikace v reálném čase, zneužívání zranitelnosti osob aj.). Poslední část Aktu, která vejde v platnost, je část týkající se vysoce rizikových AI nástrojů. U těch začnou pravidla platit od poloviny roku 2026 a v určitých případech až v roce 2027.

Pro organizace veřejné správy neznamená Akt o AI „jen něco pro IT", ale změnu nákupu a zavádění nástrojů a technologií s AI; nové povinnosti dokumentace, posuzování rizik, školení; nové procesy pro transparentnost (nejen) vůči občanům.

Kdy se organizace veřejné správy Aktem o AI musí řídit

Organizace veřejné správy je:

- typicky „uživatel" AI systému – když nástroj s AI využívá v agendě (analýzy, rozhodování, triáž žádostí...),
- někdy i „zadavatel / specifikátor" – když v zakázce požaduje konkrétní funkce AI,
- spíše výjimečně „poskytovatel" – pokud si nechá vyvinout vlastní systém a poskytuje ho dalším subjektům.

> **Jakmile organizace využívá nástroj s AI a ovlivňuje práva lidí nebo rozhodování, Akt se jí nějakým způsobem týká. Rozsah povinností se však liší podle toho, do jaké rizikové kategorie konkrétní AI nástroj spadá.**

Na každodenní práci v organizaci veřejné správy má Akt o AI dopad v několika rovinách:

- Nesmí docházet k vytváření klamavého dojmu, že komunikace probíhá výhradně s člověkem, pokud je část komunikace generována automatizovaně. Současně musí být zajištěna možnost obrátit se na lidského zaměstnance. Hlavní je transparentnost.

- AI nenahrazuje rozhodovací odpovědnost zaměstnanců. Zaměstnanci jsou povinni kontrolovat a posuzovat relevanci a správnost výstupů AI a zohledňovat konkrétní okolnosti situace.
- Na organizace veřejné správy dopadne povinnost vést záznamy o tom, kde se AI používá, k čemu a s jakými vstupy. Cílem je zajistit institucionální přehled a řídit rizika, nikoli vytvářet administrativní zátěž v podobě individuálních „deníčků“ jednotlivých zaměstnanců.
- U vysoce rizikových systémů bude třeba evidovat další záznamy. S tím se pojí i nezbytnost posuzování rizik při zavádění AI, dopadů na práva občanů, diskriminaci, spravedlnost a transparentnost. I když se s nimi běžný zaměstnanec veřejné správy nesetká, je dobré vědět, že tyto povinnosti existují.
- V organizacích bude nezbytné určit role a odpovědnosti a uzpůsobit zadávání veřejných zakázek i koncepce smluv.

Z pohledu zaměstnance se Akt o AI projeví hlavně takto:

1. **Bude vědět, kde v agendě pracuje s AI**
   – žádná „skrytá“ AI v pozadí bez informovanosti.
2. **Nebude brát výstupy AI jako pravdu, ale jako pomůcku**
   – vždy zapojí vlastní úsudek a vždy musí učinit vlastní rozhodnutí.
3. **Bude mít nastaveny postupy práce a bude se účastnit vzdělávání**
   – jak AI využívat, jak kontrolovat výsledky a posuzovat rizika
4. **Přibude administrativa (záznamy, checkboxy, šablony)**
   – cílem je mít „stopu“, kde byla AI využita a kdo to schválil.

Ochrana osobních údajů

Využívání AI s sebou často nese i zpracování osobních údajů, jak jej vymezuje GDPR. GDPR chápe zpracování osobních údajů velmi široce a je jím téměř jakékoliv nakládání s osobními údaji (výslovně hovoří například o shromáždění, zaznamenání, uspořádání, strukturování, uložení, pozměnění, vyhledání, nahlédnutí, seřazení, zkombinování, výmazu apod.).

Ke zpracování dochází nejen v případech, kdy je AI využívána cíleně pro práci s údaji, ale i v situacích, kdy plní podpůrnou funkci v procesech, které primárně s osobními údaji nepracují. Zde totiž obvykle dochází alespoň k zaznamenání aktivit uživatelů nebo zaměstnanců, k analýze chování uživatelů systému či k dalším operacím, které mohou mít povahu zpracování osobních údajů.

Z pohledu GDPR představuje využití AI zpravidla nový způsob zpracování osobních údajů. Proto bychom měli v každém případě posoudit, jaká rizika takové využití přináší, měli bychom být schopni doložit, že využití systémů AI je v souladu s GDPR, a že jsou zavedena přiměřená technická a organizační opatření k ochraně osobních údajů.

Na jaké zásady a žádoucí postupy dbát v kontextu souladu s GDPR?

Minimalizace údajů | **Přiměřenost** | **Transparentnost**

Při zavádění nebo testování systémů AI je vhodné provést posouzení vlivu na ochranu osobních údajů (DPIA). To lze doporučit zejména tam, kde dochází k rozsáhlému zpracování osobních údajů, využívání inovativních technologií nebo tam, kde by mohlo dojít k významnému zásahu do práv občanů.

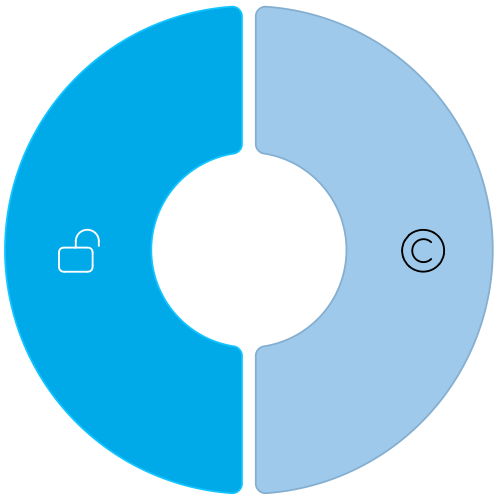
Autorské právo a licencování

Využití systémů AI rovněž přináší řadu otázek a rizik v oblasti autorského práva. Především ve vztahu k autorským dílům, která jsou do nástroje při využívání vkládána nebo na kterých jsou nástroje trénovány. Otázky mohou vzbuzovat i autorská práva k samotným výstupům, které nástroje AI generují.

Autorské právo a obecně právo duševního vlastnictví chrání celou řadu děl a jiných výsledků. Vložení chráněných předmětů do nástrojů AI může tato práva porušovat.

Výstupy, které jsou plně nebo převážně generovány AI bez tvůrčího vkladu člověka, zpravidla nebudou autorským dílem ve smyslu autorského zákona.

To znamená, že na takové výstupy se autorskoprávní ochrana obvykle nevztahuje.



Pokud se na tvorbě výstupu podílí člověk (např. formuluje tvůrčí pokyny, provádí výběr, uspořádání nebo zásadní úpravy generovaného obsahu), může být takový výstup teoreticky za určitých podmínek považován za dílo chráněné autorským právem.

Pokud je nicméně výstup připraven v rámci úřední činnosti, ochrana podle autorského zákona se na něj vztahovat nebude. Půjde totiž o úřední dílo, které není podle zákona autorským právem chráněno.

Právní rizika vznikají také při používání chráněných děl a jiných výsledků pro trénink AI. Pokud organizace vyvíjí nebo upravuje vlastní nástroj, je potřeba zajistit, aby použití vstupů bylo v souladu se zákonem a právy „vlastníků“ chráněných předmětů.

Při využívání nástrojů třetí strany je nezbytné klást důraz na to, aby poskytovatel disponoval všemi potřebnými právy ke vstupům, na nichž je nástroj založen a trénován.

Práce s dokumenty s využitím AI

Zavádění nástrojů AI do činnosti organizací veřejné správy se dotýká také oblasti spisové služby. AI může ve spisové službě výrazně pomoci – například při třídění, vyhledávání, rozpoznávání typů dokumentů nebo navrhování spisových znaků a skartačních lhůt. I v těchto případech ale platí, že se musíme řídit zákonem o archivnictví a spisové službě, prováděcí vyhláškou a dalšími předpisy.

> ⚠ Automatizované zpracování dokumentů nesmí narušit klíčový prvek spisové služby - možnost doložit původ dokumentu, jeho úplnost, neměnnost a správné evidování.

AI proto může fungovat jen jako podpůrný nástroj, ani zde jí nelze svěřit rozhodování. Konečné zařazení dokumentu, nastavení skartační lhůty a další klíčové kroky musí vždy potvrdit oprávněná osoba. Pro oblast archivace je zásadní, aby dokumenty zpracované AI byly uchovány v dlouhodobě čitelných formátech a aby každá úprava provedená AI byla zaznamenána v metadatech (kdy, jakým způsobem, jaká změna). Tím se zachovává princip důvěryhodného elektronického úložiště a přezkoumatelnosti všech úkonů.

Bezpečná práce, ověřování kvality a kontrola výstupů

Práce v bezpečném režimu

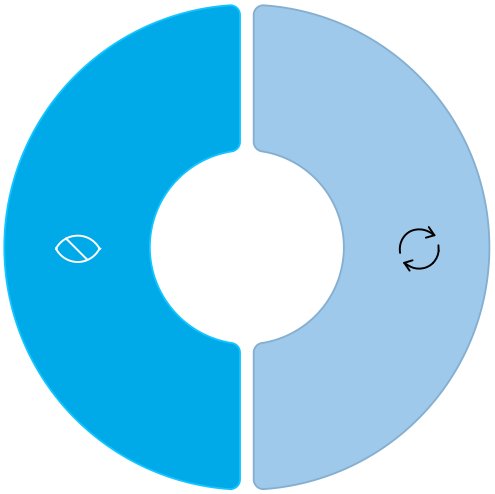
AI technologie nám mohou výrazně pomoci, ale jejich využívání je vždy spojené s problematikou bezpečné práce. Než začneme, položme si otázku: "Jaké údaje potřebujeme zpracovat a proč? Máme jasná pravidla pro jejich správu?". Bezpečnost nezačíná u technologie, ale u nás – uživatelů. Musíme rozumět údajům, se kterými pracujeme, a až poté je svěřit AI nástrojům.

Některé údaje podléhají zákonným pravidlům. **Například osobní údaje lze zpracovávat jen zákonným způsobem a k legitimním účelům.**

Pokud pracujeme s osobními údaji, je žádoucí je před použitím v AI nástroji upravit, aby nebylo možné identifikovat konkrétní osoby. To lze udělat dvěma způsoby, anonymizací nebo pseudonymizací.

Anonymizace

Odstranění všech údajů (identifikátorů), podle kterých by šlo určit totožnost člověka.



Pseudonymizace

Nahrazení identifikátorů kódy nebo jinými znaky, které umožňují dočasné propojení s původní osobou, ale jen za přísně kontrolovaných podmínek.

Jaké jsou nejběžnější situace při využívání AI?

AI nástroj provozovaný přímo organizací

Ze strany organizace je zajištěno, že k AI nástroji a vkládaným údajům nemají přístup externí osoby.

Údaje zůstávají v interním IT prostředí organizace, nikdo zvenku k nim nemá přístup. Tato varianta je pro organizaci jednodušší a bezpečnější.

AI nástroj poskytovaný jinou firmou ("AI jako služba")

Nástroj vložené údaje zpracovává na serverech poskytovatele.

Údaje se zpracovávají v IT prostředí poskytovatele, často i mimo Českou republiku. V takovém případě je nutné pečlivě poskytovatele vybírat a ověřit:

- jeho spolehlivost,
- rozsah přístupu k údajům,
- podmínky zpracování (kde a jak se údaje ukládají, jak jsou chráněny, zda je poskytovatel může využít pro jiné účely),
- právní roli poskytovatele (zpracovatel vs. správce osobních údajů).

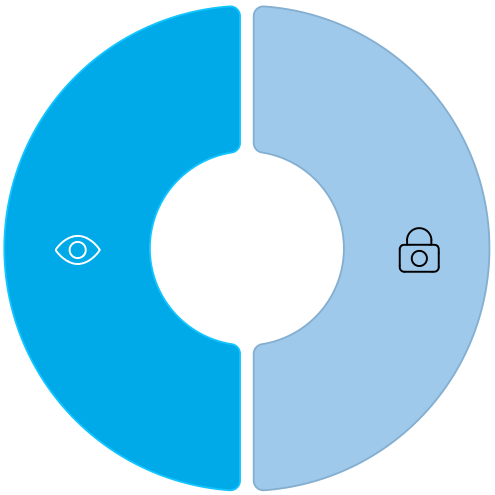
⚠ U neplacených verzí AI nástrojů bývá běžné, že poskytovatel může vložené údaje využít k jejich dalšímu tréninku. Ani placená verze však nemusí vždy automaticky znamenat bezpečnost a vylučovat sdílení údajů k dalším účelům.

Klasifikace údajů

Základem bezpečné práce je jasné rozdělení údajů podle citlivosti. Organizace by měla mít rámec, který rozlišuje:

Otevřené údaje

Lze je volně používat, sdílet a zpracovávat (např. informace zveřejňované podle zákona o svobodném přístupu k informacím). Tyto údaje jsou pro AI velmi cenné.



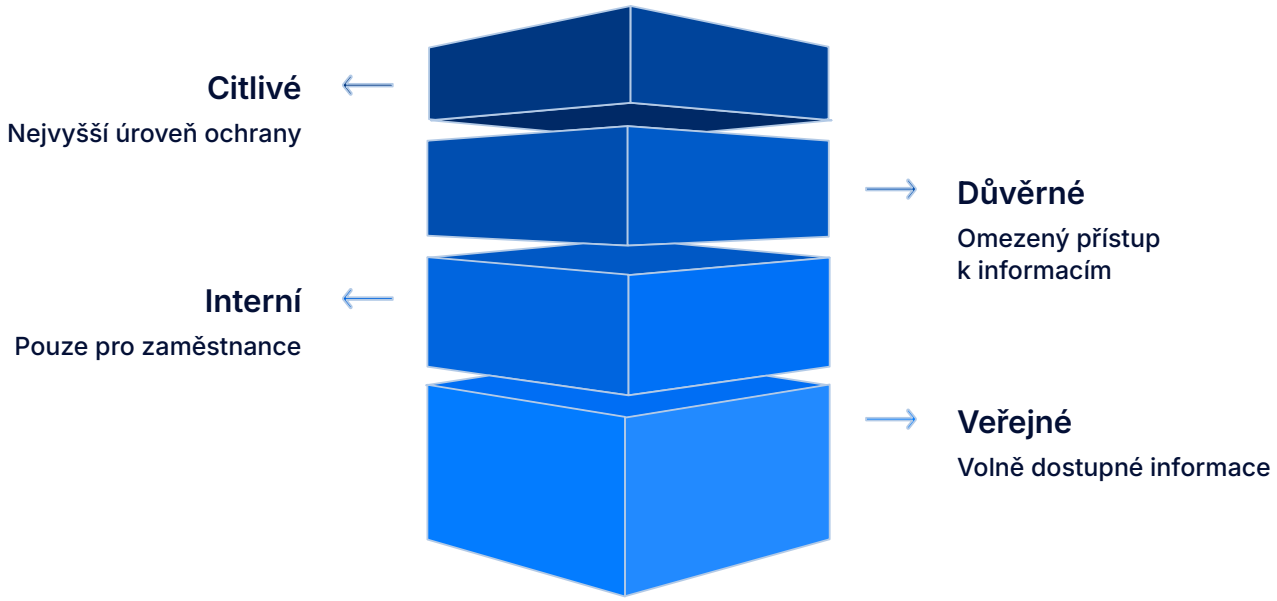
Neveřejné údaje

Chráněné ze zákona (osobní údaje, utajované informace, obchodní tajemství, bezpečnostní zájmy státu). Jejich využití musí být vždy zdůvodněné, přiměřené a dokumentované.

Adekvátně strukturovaná podrobnější klasifikace pomůže nastavit správnou úroveň zabezpečení a rozhodnout, zda je využití údajů v konkrétním AI nástroji vhodné.

ⓘ **Přílišné množství** různých **režimů pro práci** s údaji v různých nástrojích AI může být pro **uživatele matoucí.**

Příklad možné klasifikace:



U každého datového úložiště, aplikace a nástroje AI je potřeba předem určit, jaké kategorie údajů mohou zpracovávat a v jakém režimu.

Otázky, které je dobré si položit před využitím AI:

Jaké údaje opravdu potřebujeme nástroji svěřit a proč právě ty?

Každý údaj navíc je další zodpovědností.

Jak kvalitní a smysluplné údaje máme?

Pokud pracujeme s nekvalitními, neúplnými, zastaralými či zkreslenými vstupy, dostaneme i takový výstup AI.

Jsme schopni mluvit o využití údajů srozumitelně?

Srozumitelnost, vysvětlitelnost a jednoduchost je potřeba aplikovat vůči všem aktérům.

Kdo nese odpovědnost?

Odpovědnost leží na lidech a musí být konkrétní. Kdo má právo finálně rozhodnout? Jak jsou nastavené kontrolní procesy?

Jak jsme připraveni na změny?

Regulace se mění, stejně tak i nároky na správu údajů a fungování AI nástrojů i způsoby jejich využití, někdy velmi rychle. Je potřeba být připraven na změny regulatorní i technologické.

Na co si dát pozor ve smlouvách?

Je nutné věnovat pozornost smluvním podmínkám – co přesně může nástroj dělat s našimi údaji. Už při zadávání zakázky je vhodné využít standardizované smluvní doložky, které zajistí otevřenost a férovost.

Kontrola a průběžné ověřování

Bezpečnost, správnost a přesnost údajů musí být zajištěna jak u nástrojů provozovaných organizací, tak u těch od jiných poskytovatelů.

Kromě technických opatření (šifrování, pseudonymizace, řízení přístupů) je nutné, aby správci AI nástroje pravidelně kontrolovali:

- zda nástroj stále funguje v souladu s původním účelem,
- zda se nezměnily podmínky využívání nebo spolehlivost poskytovatele,
- zda nástroj neprovádí operace nad rámec povoleného použití.

Odpovědnost za výstupy AI

⚠ Výstupům z jakéhokoliv nástroje AI je vždy potřeba věnovat náležitou pozornost. Za výstupy i za rozhodnutí, která jsou na nich založená, nese v první řadě odpovědnost ten, kdo s nástrojem pracuje.

Tato odpovědnost je jasně daná a nelze ji přenést na samotný AI nástroj. To ale neznamená, že bychom nemohli AI využívat – ani správní řád použití AI nezakazuje. Pořád však platí, že úkony v řízení provádí oprávněné úřední osoby a rozhodnutí musí obsahovat jejich identifikační údaje a podpis. Zaměstnanec tedy nemůže „přenechat odpovědnost“ AI a musí výstupy vždy ověřit.

Pokud kvůli neověřenému výstupu AI vznikne někomu škoda, stát (nebo jiný příslušný orgán) odpovídá za škodu způsobenou nezákonným rozhodnutím nebo nesprávným postupem. Náhrada škody může být následně požadována po zaměstnanci, jehož činností byla škoda způsobena.
V některých případech může stát škodu vymáhat i po poskytovateli AI nástroje.

Jak správně kontrolovat výstupy AI

Porovnat výstup AI se spisem a dostupnými podklady

Kontrolujte, zda výstup odpovídá skutečnostem v řízení, neobsahuje věcné chyby a správně pracuje s právními předpisy a jinými zdroji.

Posoudit logiku a vnitřní konzistenci výstupu

Ověřujte, že závěry odpovídají uvedeným skutečnostem a jednotlivé části nejsou v rozporu se zadáním ani s interními dokumenty a jinými skutečnostmi.

Konfrontovat další zdroje

Vyhledávejte relevantní literaturu, zákony a materiály, podle kterých lze platnost uváděných skutečností ověřit.

> **Pokud máte pochybnosti, výstup upravte, přepočítejte, ověřte v právních předpisech nebo se poraďte s kolegy či nadřízeným. Při ověřování postupujte tak, jako by šlo o podklad od stážisty - tedy jako o pomůcku, nikoli neomylnou pravdu.**

Součástí správného postupu je i to, že jako zaměstnanci veřejné správy víte, kdy výstup AI nepoužít vůbec – například pokud je nekvalitní nebo v rozporu s právem.

U složitějších věcí je vhodné projít text větu po větě, přizpůsobit jej konkrétnímu případu, doplnit vlastní úvahy a převzít odpovědnost za konečnou podobu výstupu (jak ukládá i platná legislativa). Je vhodné si poznamenat, jak byl nástroj použit – k čemu sloužil, co bylo upraveno člověkem. To pomůže při kontrole i při případné obhajobě, že zaměstnanec nejednal mechanicky, ale výstupy aktivně ověřoval.

Kombinujte kritické myšlení s ověřováním informací. AI je nástroj, nikoli autorita. Každé tvrzení, které má vliv na rozhodnutí, ověřte z více nezávislých zdrojů. Důležité je **rozumět danému tématu** tak, abyste v rámci své práce s AI byli schopni **validovat** její **výstupy**.

Transparentní informování o využití AI

Organizace veřejné správy by měly jasně označovat výstupy, které byly vytvořeny nebo výrazně ovlivněny AI (texty, grafy, analýzy aj.).

Uvedení, kdy a k čemu byla AI využita:

- zajišťuje přezkoumatelnost rozhodnutí
- pomáhá ověřit soulad se zákonem
- posiluje důvěryhodnost

Označování má dvě roviny:

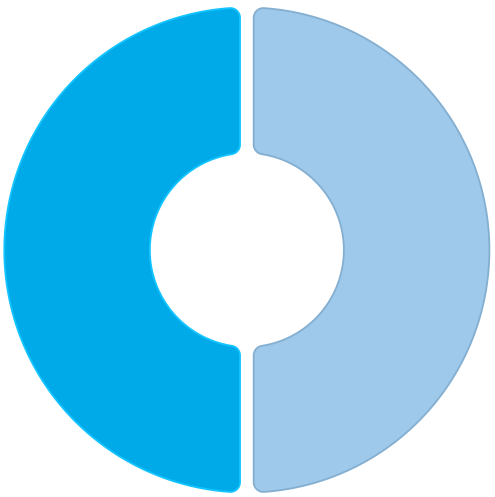
Označování obsahu zpracovaného AI pro strojové čtení (zahrnuje technické řešení pro různé typy obsahu i metriky využití AI). To je primárně úkol pro poskytovatele AI řešení.

Jasné a viditelné grafické označení pro uživatele tak, aby bylo jasné, že obsah zcela vytvořila AI, nebo se na něm podílela. Ve veřejné správě se jedná o informování ve veřejném zájmu a odpovědnost nese jak původce, tak i každý zaměstnanec.

Obsah plně generovaný AI

Vytvořený zcela bez lidského autentického zásahu do tvorby výstupu či jeho úprav, včetně automatizovaných úkonů, rozhodování apod.

Obsah nemusel vůbec projít lidskou supervizí.



Obsah tvořený s asistencí AI

AI významně ovlivnila obsah, který by mohl být považován za vytvořený člověkem.

Za významné ovlivnění lze považovat přidání či úpravu obsahu, přepis, sumarizaci, změnu kontextu či významu, překlad apod.

Všímejte si a označujte

Kontrolu, zda jsou veškerý obsah a výstupy AI vhodně a zřetelně označeny, by měl provádět každý, kdo přichází s AI a jejími výstupy do kontaktu v průběhu celého „uživatelského řetězce“.

AI výstupy často vypadají realisticky a přesvědčivě, ale uživatel musí vždy vědět, že se na nich AI podílela.

Varovné signály

Jak jsme si již vysvětlili, AI nástroje při tvorbě výstupů ve skutečnosti nerozumí dané problematice, ale pracují s algoritmy a skládají text podle pravděpodobnostních vzorců. I proto si ve výstupech musíme všímat varovných signálů, které naznačují, že výstup není spolehlivý, přesný nebo vhodný jako podklad pro rozhodování.

Neurčitost a vágnost odpovědí

AI se vyhýbá jasným formulacím a používá nejasné výrazy typu „**pravděpodobně**“ nebo „**někteří se domnívají**“. Takové fráze naznačují, že nástroj může odpovídat na základě neúplných či nejasně definovaných, možná i protichůdných, údajů.

Přílišná sebejistota modelu

AI může působit velmi jistě, i když nemá dostatek informací. Tón odpovědi může znít autoritativně, ale to není záruka správnosti. AI může např. tvrdit, že „**toto je jednoznačně nejlepší řešení**“ nebo „**určitě platí, že…**“, aniž by uvedla jiné možnosti nebo vysvětlení.

Přehnaná snaha vyhovět uživateli

AI nástroje bývají nastaveny tak, aby působily vstřícně a přizpůsobovaly se očekáváním uživatele. Takový postup může usnadnit komunikaci, ale v praxi někdy vede k tomu, že AI přejímá chybné nebo nevhodné předpoklady uživatele, místo aby je korigovala.

Zastaralé nebo smyšlené informace

Pokud AI nemá přístup k aktuálním údajům, může čerpat ze zastaralých zdrojů nebo si informace „vymýšlet“. Typickým příkladem je tzv. halucinace – situace, kdy AI uvede neexistující údaje, vymyslí odkazy nebo citace, které vypadají věrohodně, ale nejsou pravdivé.

Překrucování kontextu

AI může odpověď zasadit do jiného kontextu, než byl původní záměr. To se stává, když nástroj nemá dostatečně specifikované zadání, špatně ho pochopí nebo doplní domněnky, které nejsou správné. Proto je nutné vždy ověřit, zda odpověď odpovídá zadání.

Antropomorfizace

AI nástroje jsou stavěny tak, aby v nás vyvolaly dojem, že komunikujeme s velice inteligentní a úslužnou, empatickou a skutečně myslící bytostí. To může být matoucí i nebezpečné, protože uživatelé mohou mít tendenci se na výstupy spoléhat více, než by si zasluhovaly, a svěřovat AI i informace, které mohou být citlivé.

Organizace by měla mít jasně a srozumitelně popsané podmínky pro zavádění a využívání AI nástrojů.

Mělo by platit, že zaměstnanci:

- používají jen schválené nástroje a pracují s nimi určeným způsobem,
- vědí, co je automatizováno či zpracováno s využitím AI, jaké údaje AI zpracovává a jak se kontrolují vstupy i výstupy,
- jsou pravidelně informováni o novinkách či změnách a školeni v etickém a bezpečném využívání AI.

Část III

Řízení a implementace AI v organizacích veřejné správy

Řízení AI v organizacích veřejné správy

Ačkoliv se AI ve veřejné správě začíná zavádět už relativně široce, převládají spíše ad-hoc projekty bez jednotného rámce, což vede k roztříštěnosti, duplicitám a horšímu sdílení dobré praxe. Dobře nastavené řízení tomu předchází, propojuje AI řešení s cíli veřejné správy a jejích organizací, nastavuje priority, metriky, role a odpovědnosti, pravidla pro bezpečnost, právo i etiku a dělá z dílčích projektů řízené portfolio, které lze vyhodnocovat, škálovat a zlepšovat. Koncepční přístup umožňuje sdílení standardů, opakované využívání řešení, interoperabilitu systémů, snižuje náklady a rizika, zrychluje zavádění změn a posiluje důvěryhodnost. Řízení AI tak není „zátěž navíc“, ale praktický nástroj, který dává využívání AI směr.

Co může organizace udělat už nyní?

Inspirovat, sdílet a komunikovat

Lidé v organizacích možná nevědí, jak je možné využít AI komplexněji, a další o tom nikomu neříkají.

Oceňovat

Vyzdvihnutí vhodného, odpovědného využívání AI má význam. Kdo zná někoho používajícího AI, má větší pravděpodobnost, že to bude dělat také.

Vyhodnotit dopad a přijmout výsledky

Znalost, co funguje, podložená důkazy může posílit přijetí. Součástí je uznat, co nefungovalo, což vytváří pocit bezpečí pro další experimenty.

Působit na hlavní faktory, které ovlivňují přijetí a využívání AI:

Motivace

Vidím jasný a žádoucí důvod pro využití AI? Obecné, běžně uváděné výhody mohou být odtržené od skutečných úkolů zaměstnance.

Schopnosti

Cítím se schopen využívat AI efektivně, mám sebejistotu? Problémy spočívají více v lidech a procesech než technice.

Důvěra

Je AI v souladu s mými hodnotami? Představuje hrozbu pro mou identitu?

Podle provedených studií se ukazuje, že zaměstnanci veřejné správy sice AI využívají, ale spíše nárazově a pro jednotlivé úkony než systematicky a např. i s kombinací více nástrojů tak, aby se AI stala integrálním nástrojem každodenní práce. Tento proces má svůj vývoj, který vyžaduje vzdělávání, osvojování dovedností, pravidel i posilování vlastního sebevědomí.

Pro vedení organizací z toho plyne několik výzev:

- Brát AI dovednosti jako nový typ kompetence zaměstnanců. Zajišťovat a podporovat kontinuální rozvoj a cílené učení zaměstnanců pro posilování sebevědomí a aktualizaci jejich dovedností.
- Vytvářet rámec a pravidla pro využívání AI (i s rozlišením různých rolí).
- Vyvíjet a zlepšovat AI nástroje i jejich využití ve spolupráci se zaměstnanci jako odborníky na věcné agendy a zajišťovat propojení technologie s problémy, které mají řešit.

Proces implementace etického a odpovědného využívání AI v organizaci veřejné správy

Zavádění AI je vždy do určité míry experiment, ale v prostředí veřejné správy musí mít i experiment jasná pravidla. Nestačí jednotlivé nápady, je potřeba vědět, proč daný nástroj zavádíme, jaká rizika s sebou nese a jak poznáme, že nám skutečně pomáhá. Proto má smysl řídit implementaci AI jako cyklus PDCA – stanovit cíle a metriky, pilotně ověřit řešení, průběžně kontrolovat dopady (vč. právních, etických) a podle výsledků upravovat praxi. Součástí takového rámce jsou interní předpisy, vzdělávání, řízení rizik, jasný postup pro řešení problémů i otevřená komunikace dovnitř organizace i směrem k veřejnosti. Výsledkem není méně svobody v rozhodování, ale méně improvizace, větší odpovědnost, lepší ochrana údajů a rychlejší cesta k dosahování přínosů AI.

Při zavádění je možné využít regulační sandboxy pro pilotování AI nástrojů pro veřejný sektor. Tato kontrolovaná reálná testovací prostředí pomáhají identifikovat technické/etické problémy a ověřit přijetí uživateli.

Vyplatí se investovat do robustní infrastruktury, včetně interních cloudových platforem či výpočetních klastrů.

Postup implementace AI při zohlednění PDCA přístupu může vypadat následovně:

Identifikace případu využití

- Určení záměru a definice cílů
- Vyhodnocení připravenosti
- Nastavení procesu implementace

Interní předpisy a definice rolí

- Stanovení odpovědností
- Zavedení pravidel a postupů
- Popis rolí
- Soulad s legislativou

Komunikace a transparentnost

- Otevřenost vůči zaměstnancům
- Transparentnost vůči veřejnosti
- Interní dialog a zpětná vazba

Úprava dle konceptu PDCA

- Nastavení strategie
- Implementace AI řešení
- Sledování výsledků a rizik
- Úpravy procesů a standardů

Analýza rizik a jejich zmírňování

- Identifikace rizik
- Predikce dopadů
- Nastavení opatření
- Záznam a průběžná aktualizace

Vzdělávání a osvěta

- Školení zaměstnanců
- Praktické zásady
- Rozvoj digitálních kompetencí
- Podpora kultury důvěry

Měření přínosů a plán řešení incidentů

- Stanovení metrik a indikátorů
- Hodnocení dopadů
- Postupy pro řešení incidentů
- Sdílení poznatků

Řízení etického a odpovědného využívání AI by mělo zahrnovat vyhodnocení dopadů využívání AI nástrojů na etické principy i východiska základních práv, které jsou řešeny v jiné části Průvodce. Předmětem této části textu je poskytnout rámec pro toto hodnocení v podobě návodných otázek, který je ovšem vždy potřeba uzpůsobit konkrétním podmínkám a situaci organizace i charakteru AI nástroje.

Odpovědi na otázky by měly zasahovat celý životní cyklus AI (naznačen v obrázku). Je důležité se k nim při postupu AI životním cyklem opakovaně a progresivně vracet, a revidovat hodnocení dopadů na etické i právní principy.

Jaké otázky si tedy jako organizace klást při hodnocení etických dopadů implementace AI nástroje?

Strategický záměr a přínos

- Jaký AI nástroj zavádíme a jaký cíl tím sledujeme?
- Jaký problém řeší a jak zapadá do strategie a priorit organizace i celé veřejné správy?
- Jaké metriky úspěchu a očekávané přínosy nastavíme?
- Je rozsah nasazení přiměřený a časově realistický?

Uživatelé a jejich kompetence

- Kdo bude s AI interagovat (zaměstnanec, příjemce výstupu, správce)?
- Jaká je jejich úroveň kompetencí a jak zajistíme školení/podporu?
- Existuje možnost AI nepoužívat nebo namísto AI komunikovat s člověkem?
- Jsou aktéři transparentně informováni, že pracují s AI či jejími výstupy?

Procesní integrace a rozsah

- Které procesy/agendy budou dotčeny? Jedná se o zásah do kritických procesů/agend?
- Zasahuje nástroj celou organizaci, nebo pouze část/specifickou agendu?
- Jak AI navazuje na existující procesy a rozhodovací pravomoci?
- S jakými systémy bude AI spolupracovat a jak je řízena návaznost, propojenost, toky informací?
- Nahrazuje AI jiný IT systém/člověka/přináší novou funkci/doplňuje stávající?

Informace, kvalita a ochrana

- K jakým údajům má AI přístup? Jaký je zdroj a způsob vkládání/správy?
- Jsou údaje klasifikovány dle citlivosti a je rozsah vkládaných údajů omezen na nezbytné minimum?
- Jak je zajištěna kvalita, ověřitelnost a sledování bias ve vstupech?
- Kdo má kontrolu nad údaji v celém životním cyklu AI a jak detekujeme únik/zneužití?

Model, transparentnost a kontrolovatelnost

- Je znám původ nástroje/model (včetně dokumentace), jeho limity a předpoklady?
- Je míra vysvětlitelnosti a vysledovatelnosti odpovídající riziku?
- Jaké jsou přístupové role, auditní role a jak je nastaven dohled nad změnami?
- Je nastaven proces validace/testování před nasazením i po něm?

Přiměřenost a volba řešení

- Byly zvažovány ne-AI alternativy a proč nepostačují?
- Proč volíme právě toto AI řešení (varianta, poskytovatel, vývoj/nákup)?
- Jaké jsou známé limity, rizika a jak je eliminujeme?
- Jaké jsou náklady (vč. nákladů obětované příležitosti) oproti alternativám a jak to ovlivní rozhodnutí o volbě řešení?

Minimalizace újmy a základní práva

- Hrozí využití k sociálnímu hodnocení, sledování v reálném čase či zásahu do základních lidských práv? Jaké jsou pojistky?
- Mohou být dopady využití AI nezvratitelné či zásadní? Jak zajistíme lidskou kontrolu?
- Může AI nepřiměřeně zasahovat do autonomie osob nebo vytvářet diskriminaci?
- Jak je řešena možnost námitky, opravného prostředku a přezkoumatelnost?

Stakeholdeři a participace

- Které skupiny budou nejvíce ovlivněny a jaké mají potřeby/obavy?
- Jak budou zapojeny napříč životním cyklem AI (návrh, test, provoz) a s jakým cílem?
- Jaká je jejich role ve schvalování/využívání a jaký mají reálný vliv?
- Proběhla cílená komunikace a bylo poskytnuto srozumitelné vysvětlení?

Právo, bezpečnost a soulad s požadavky

- Byl proveden právní rozbor (ochrana osobních údajů, sektorové normy, veřejné zakázky apod.)?
- Jak naplníme bezpečnostní požadavky?
- Jsou definovány odpovědnosti a dohledové mechanismy pro celý životní cyklus AI?
- Je zajištěna auditovatelnost (evidence rozhodnutí, kontrolní body, důkazní stopa)?

Nasazení, provoz a zlepšování (PDCA)

- Je popsán postup nasazení, kritéria „kam až AI sahá“ a plán ukončení?
- Jsou nastaveny metriky výkonu/kvality?
- Existuje plán řízení rizik a incidentů?
- Jak probíhá průběžné přezkoumání a zlepšování (PDCA, opětovná validace)?

Role a odpovědnosti

Určení jasných rolí v rámci řízení AI je důležité pro zajištění jejího vhodného využívání. Každá role má své specifické hlavní úkoly a odpovědnosti. Jasným určením základních rolí se předchází nejasnostem v kompetencích, zmatkům, zajišťuje se transparentnost a kontrolovatelnost a posiluje se i důvěra v AI řešení. Jasně definované role navíc podporují spolupráci, efektivní řízení rizik a dlouhodobou udržitelnost využívání AI.

Systém hlavních rolí v řízení AI se může mezi organizacemi různit, nicméně je vždy vhodné jasně určit alespoň následující role:

U každého člověka v dané roli se musíme ptát, jaká je jeho **znalost problematiky** AI ve vztahu k jeho odpovědnostem, jaké jsou jeho **kompetence**, jak byl na danou roli **připraven a vybaven**, zda má vše, **co** pro výkon role **potřebuje**.

Manažer

Osoba odpovědná za strategická rozhodnutí o tom, zda a jak bude AI v organizaci využívána.

Typické úkoly a odpovědnosti:

- Stanovuje cíle a strategii využívání AI.
- Zajišťuje soulad s právem a etickými zásadami.
- Přiděluje odpovědnosti, zdroje a určuje priority.

> **Bez jasné rozhodovací autority hrozí nekoordinovanost, nedostatečná odpovědnost a ztráta důvěry v AI řešení.**

AI uživatel

Zaměstnanec, který aktivně využívá nástroje nebo systémy založené na AI při výkonu své agendy.

Typické úkoly a odpovědnosti:

- Využívá AI v souladu s interními pravidly a etickými zásadami.
- Vyhodnocuje výstupy AI a uplatňuje vlastní úsudek a kritické myšlení.
- Hlásí nesrovnalosti, chyby či rizika spojená s využíváním AI.

> **Klíčový pro zajištění „lidské kontroly“ nad technologií a pro realistickou zpětnou vazbu, jak AI skutečně pomáhá v praxi.**

Vlastník věcného procesu

Osoba odpovědná za konkrétní agendu nebo proces, do kterého je AI implementována.

Typické úkoly a odpovědnosti:

- Definuje cíle využití AI v daném procesu, aby se AI nemíjela účinkem.
- Určuje požadavky na kvalitu a správnost výstupů.
- Spolupracuje s IT a poskytovateli na správném nastavení nástroje.

> **Zajišťuje, že AI řešení skutečně odpovídá potřebám organizace a podporuje věcnou agendu, nikoliv naopak.**

Bezpečnostní / IT podpora

Technický garant zajišťující, že nástroje AI jsou provozovány bezpečně a v souladu s interními, právními i bezpečnostními a technologickými požadavky.

Typické úkoly a odpovědnosti:

- Spravuje přístupová práva, údaje a technickou infrastrukturu, zajišťuje spolehlivost.
- Monitoruje kybernetická rizika a zabezpečení AI.
- Poskytuje technickou podporu uživatelům.

> **Bez této role nelze zaručit ochranu údajů, dostupnost ani důvěryhodnost AI řešení.**

Dodavatel / vývojář

Externí nebo interní partner, který navrhuje, vyvíjí nebo dodává AI řešení pro organizaci.

Typické úkoly a odpovědnosti:

- Zajišťuje transparentnost algoritmů, dokumentaci a testování.
- Spolupracuje s organizací při vyhodnocování dopadů AI.
- Dodržuje bezpečnostní a etické standardy stanovené organizací.

> **Jeho přístup a kvalita práce mají přímý vliv na důvěru, funkčnost a udržitelnost AI.**

AI mentor / zkušený uživatel

Interní ambasador, který pomáhá kolegům s orientací v AI nástrojích a rozvíjí digitální kompetence a know-how organizace.

Typické úkoly a odpovědnosti:

- Rozvíjí a metodicky vede ostatní uživatele.
- Sleduje trendy a dobrou praxi v oblasti AI.
- Působí jako spojka mezi uživateli, vedením a IT.

> **Podporuje kulturu etického a odpovědného využívání AI, zvyšuje důvěru zaměstnanců a pomáhá překonávat nechuť ke změnám.**

Kde hledat další ověřené informace a inspiraci

Pokud se chcete o tématu AI dozvědět více, doporučujeme tyto zdroje, které jsou srozumitelné a důvěryhodné:

- **Elements of AI (elementsofai.cz)**

  Celosvětově uznávaný online kurz v češtině, který vysvětluje základy AI pro naprosté laiky. Je zdarma.

- **Portál AI dětem (aidetem.cz)**

  Ačkoliv je určen primárně pro školství, nabízí jedny z nejsrozumitelnějších materiálů a příruček o generativní AI, které jsou aplikovatelné i pro zaměstnance veřejné správy.

- Využít můžete podpory **Ministerstva průmyslu a obchodu**, které zabezpečí sběr otázek prostřednictvím online dotazníku, na jehož základě poskytne odpovědi na nejčastěji se opakující dotazy, zejm. v oblasti veřejných zakázek, analýzy vhodných případů využití AI, nastavení základních vnitřních pravidel pro bezpečné zavedení a provoz AI, vč. doporučení cílové AI architektury a sdílení ověřených postupů a zkušeností z realizovaných AI projektů ve veřejné správě.

  Dotazník je dostupný zde: **https://forms.office.com/e/Wrd2xL9N5F**

- Na národní úrovni zejména **NÚKIB** (nukib.gov.cz) zveřejňuje seznamy aplikací, jež zcela jistě nejsou bezpečné a v případě pochybností je možné si na jeho webových stránkách daný nástroj či aplikaci ověřit.

- Evropská komise spustila službu **AI Act Service Desk** a **Jednotnou informační platformu** na podporu implementace Aktu o AI, jakožto svůj primární komunikační kanál v této oblasti.

  Vydala také **Pokyny Komise k zakázaným postupům v oblasti umělé inteligence podle nařízení (EU) 2024/1689 (Akt o AI)**, které si můžete stáhnout zde: Komise zveřejňuje pokyny k zakázaným praktikám umělé inteligence, jak jsou definovány v aktu o umělé inteligenci. | Utváření digitální budoucnosti Evropy

  Připravují se také Pokyny pro vysoce rizikové systémy.

  Obecně doporučujeme se s Pokyny Komise v případě zájmu seznámit, neboť jsou velmi přínosné.

Legislativní okénko

Nařízení (EU) 2024/1689, kterým se stanoví harmonizovaná pravidla pro umělou inteligenci

Nařízení Evropského parlamentu a Rady (EU) 2016/679 ze dne 27. dubna 2016 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů a o zrušení směrnice 95/46/ES (obecné nařízení o ochraně osobních údajů)

Rámcová úmluva Rady Evropy o umělé inteligenci a lidských právech, demokracii a právním státě 2024/225.

Zákon č. 106/1999 Sb., o svobodném přístupu k informacím

Zákon č. 500/2004 Sb., správní řád

Zákon č. 499/2004 Sb., o archivnictví a spisové službě a o změně některých zákonů

Zákon č. 365/2000 Sb., o informačních systémech veřejné správy

Zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů

Inspirativní zdroje (nejen) k AI

AI Verify Testing Framework. AI Verify Foundation. Dostupné z: **aiverifyfoundation.sg/what-is-ai-verify/**.

Apolitical.co - celosvětová platforma pro sdílení, vzdělávání a networking pro úředníky a veřejný sektor. Je v angličtině a dostupná po jednoduché registraci.

Ethical Impact Assessment: A Tool of the Recommendation on the Ethics of Artificial Intelligence (2023). UNESCO. Dostupné z: **https://www.unesco.org/en/artificial-intelligence/recommendation-ethics**.

Guidelines for the Responsible Use of Artificial Intelligence in the Public Service (2025). Department of Public Expenditure, Infrastructure, Public Service Reform and Digitalisation. Irsko. Dostupné z: **www.gov.ie/en/department-of-public-expenditure-infrastructure-public-service-reform-and-digitalisation/publications/guidelines-for-the-responsible-use-of-ai-in-the-public-service/**.

Príbaň Žolnerčíková, V. et al. Příručka pro společnosti a výzkumné organizace zpracovávající velká data za účelem vývoje umělé inteligence. Ústav státu a práva, Akademie věd ČR.

Recommendation on the Ethics of Artificial Intelligence (2021). UNESCO. Dostupné z: **https://www.unesco.org/en/artificial-intelligence/recommendation-ethics**.

Revised Recommendation of the Council on Artificial Intelligence (2024). OECD. Dostupné z: one.oecd.org/document/C/MIN(2024)16/FINAL/en/pdf.

Umělá inteligence (2021). Úřad vlády ČR. Dostupné z: Umělá inteligence | Vláda České republiky

Understanding Responsibilities in AI Practices (2024). Department of Customer Service. Austrálie, Nový Jižní Wales. Dostupné z: www.aigl.blog/content/files/2025/04/Understanding-Responsibilities-in-AI-Practices.pdf.

Autorský kolektiv a odborná spolupráce

Tento materiál vznikl ve spolupráci níže uvedených osob. Uvedení jmen vyjadřuje podíl na tvorbě textů a konzultacích; pořadí nemusí odrážet míru přispění. Děkujeme všem, kteří se podíleli na přípravě tohoto materiálu.

Autoři:

Juraj Hvorecký (Filozofický ústav Akademie věd České republiky) - odborný garant

Lukáš Kalenský (sekce pro státní službu, Úřad vlády České republiky) - garant

Jana Novosáková (sekce pro státní službu Úřadu vlády České republiky) - garant, manažer projektu

Šimon Svoboda (Právnická fakulta Masarykovy univerzity)

Jan Chmelíček (sekce pro státní službu, Úřad vlády České republiky)

Filip Šmakal (oddělení evropské digitální agendy, Úřad vlády České republiky)

Lukáš Flek (Digitální informační agentura)

Kateřina Kellerová (Národní úřad pro kybernetickou a informační bezpečnost)

Adam Jareš (Legální kód)

Spolupracující osoby:

Emma Lachová (Ministerstvo průmyslu a obchodu)

Jakub Kortus (Ministerstvo průmyslu a obchodu)

Tomáš Machovec (Ministerstvo průmyslu a obchodu)

Ivana Kudláčková (Český telekomunikační úřad)

Filip Hartmann (prg.ai)

Michal Řezáč (Český metrologický institut)

Kristián Malina (Ministerstvo vnitra)

Poděkování patří také všem dalším participujícím osobám, které se na vzniku materiálu podílely a přispěly svou expertízou, konzultací či názorem.

Digitálně podepsal
Tünde Bartha
2. 4. 2026