Český rozhlas

Generální ředitel ČRo dne 1. září 2025
aktualizuje 3. novelou směrnici ve znění S*I*2023/001/GŘ//3

Pravidla pro užívání umělé inteligence v Českém rozhlasu

Účinnost od 15. ledna 2024, toto znění účinné od 1. září 2025

Čl. 1 Preambule

1. Podle zákona č. 484/1991 Sb., o Českém rozhlasu, (zejména § 3 odst. (1) písm. i) Český rozhlas (dále také jako „ČRo“) naplňuje veřejnou službu v oblasti rozhlasového vysílání rovněž tím, že vyvíjí činnost v oblastech nových vysílacích technologií a služeb. Dle Kodexu Českého rozhlasu (zejména čl. 2. 12) ČRo sleduje inovace v rozhlasovém vysílání ve smyslu technickém, žánrovém, obsahovém a formálním tak, aby bez závažného důvodu neopožďoval zavádění nových technologií, které mají přímý vliv na kvalitu služeb nabízených posluchačům. Do této oblasti spadá i zkoumání a využívání umělé inteligence (dále také jako „AI“).

2. AI umožňuje poskytovat veřejnosti více služeb a profesionálům ČRo pracovat efektivněji, ale představuje i nezanedbatelná rizika, která nelze s ohledem na specifické postavení ČRo jako média veřejné služby podceňovat.

3. V souladu se závazkem ČRo nabízet posluchačům služby v maximální možné kvalitě a s vědomím hodnot, na kterých je činnost ČRo postavena, považuje proto ČRo za nezbytné stanovit pravidla pro využití AI na činnosti ČRo.

4. Vývoj technologií AI nelze přesně předvídat. S vědomím překotného technologického rozvoje umělé inteligence a nastavení právního rámce fungování umělé inteligence, zejm. Nařízení Evropského Parlamentu a Rady (EU) 2024/1689 (Evropský akt o umělé inteligenci), se předpokládá, že bude operativně docházet ke změnám pravidel v této směrnici obsažených a jejich aktualizaci.

Čl. 2 Předmět a cíl směrnice

1. Předmětem této směrnice jsou pravidla pro využití umělé inteligence (dále také jako „směrnice AI“) v ČRo včetně nastavení procesu hodnocení rizik v případě systematického použití AI. Proces hodnocení rizik je postaven na definování rizik a možných dopadů, která s konkrétním způsobem využití AI souvisejí. Součástí procesu hodnocení rizik může být i návrh přijetí vhodných opatření, která případná rizika minimalizují.

2. Výklad pravidel v této směrnici AI uvedených a samotné využití AI v praxi je nutno činit způsobem, který minimalizuje rizika používání AI a bude v souladu s hlavními zásadami jejího využití dle této směrnice AI.

3. Cílem směrnice AI je nastavit srozumitelná pravidla pro používání AI nástrojů pro bezpečnou a užitečnou práci a zkoumání těchto nástrojů. Rizika v této směrnici AI uvedená demonstrují, v čem mohou spočívat problémy při využívání těchto nástrojů.

4. AI dělíme na dvě skupiny:

   a) analytickou (označovaná též negenerativní) AI

   b) generativní AI.

5. Směrnice AI se vztahuje na využití AI:

   a) na činnosti, které mají přímý vliv na obsah nabízený posluchačům a uživatelům digitálních platforem ČRo;

   b) na všechny ostatní činnosti, které mají charakter podpůrných služeb, např. IT služby, analytické a rešeršní práce, katalogizace a systematizace uzavřeného souboru dat, komunikace s posluchači a uživateli, právní, obchodní a ekonomické služby apod.

6. Zapojení nástrojů analytické AI při dodržení všech dosud platných zásad práce v ČRo nepředstavuje pro činnost ČRo významné riziko. Kreativní potenciál algoritmů a povaha tréninkových dat generativní AI ale významná rizika přinášejí a jsou popsány dále. S ohledem na provázanost a souběžné využití obou typů AI v jednotlivých nástrojích se rizika a zásady pro využití AI vztahují na oba typy AI, především však na generativní AI.

Čl. 3 Výklad pojmů pro účely směrnice AI

1. **Umělou inteligencí (AI)** se rozumí počítačová technologie založená na struktuře tzv. umělých neuronových sítí, která analyzuje a klasifikuje existující data, vyhledává v nich objekty nebo vzory. Může sloužit pro automatizaci procesů a vytváření nových dat. Algoritmy umělé inteligence nejsou naprogramované do posledního detailu jako běžné počítačové programy, ale učí se z tréninkových dat (zpravidla se jedná o data z internetu).

2. **Analytická (negenerativní) AI** slouží k rozpoznávání skrytých vzorů, k analýze obrázků (například vyhledání obrázků, na kterých je kočka), rozpoznávání textů v obrázcích, k převodu textu do mluveného slova a naopak, k indexaci a doporučování souvisejícího obsahu podle definovaných algoritmů.

3. **Generativní AI** umožňuje vytvářet nový obsah (např. texty, obrázky, zvuk, hudbu, videa nebo počítačové kódy), a to na základě požadavků jednotlivých uživatelů. Tyto požadavky lidé zadávají umělé inteligenci většinou prostřednictvím tzv. textových promptů (příkazů či pokynů). Prompty by měly být formulovány co nejpřesněji, aby byly výsledky co nejlepší. Generativní AI funguje na principu pravděpodobnosti, neboť se snaží předpovědět, jaký prvek (např. slovo) se bude s největší pravděpodobností vyskytovat v blízkosti předchozího prvku na základě analýzy tréninkových dat. Kromě generování textů, grafiky či hudby dokáže umělá inteligence také provádět analýzu generovaného obsahu nebo kombinovat kreativní práci s texty a obrázky apod. I když funguje na základě složitých umělých neuronových sítí a matematických algoritmů, přesto může vytvářet chybné nebo zkreslené výstupy (tzv. halucinovat).

4. Seznam nejčastějších aplikací a nástrojů, které jsou založeny na technologii generativní AI, je k dispozici na intranetu zde. Tento seznam není vyčerpávající a slouží pouze k základnímu přehledu o tom, které nástroje jsou založeny na modelu generativní AI. Použití těchto nástrojů podléhá plně všem pravidlům v této směrnici uvedeným.

Čl. 4 Hlavní zásady použití AI

1. Uživatelé mají povinnost seznámit se s obchodními podmínkami konkrétního nástroje a tyto podmínky dodržovat.

2. AI nástroje se v ČRo využívá zejména pro zrychlení příprav, zpřesnění nebo jiné zkvalitnění původního obsahu. Žádná zodpovědnost nemůže být z principu přenesena na AI. AI také není ze žurnalistického hlediska samostatným zdrojem informací.

3. ČRo při využití AI dbá na dodržení následujících zásad:

   - **Transparentnost** – Tato zásada je důležitá pro budování důvěry mezi ČRo, posluchači a uživateli služeb ČRo. ČRo je povinen informovat publikum o tom, že konkrétní obsahové položky byly vytvořeny za pomoci nástrojů generativní AI, zejména pokud jsou použity k vytvoření obsahu, který je prezentován jako autorský. Syntetický obsah musí být vždy označen ve strojově čitelném formátu a zjistitelný jako uměle vytvořený, případně uměle manipulovaný.

   - **Osobní odpovědnost** – Pokud zaměstnanec ČRo použije při své práci nástroje AI, je za výsledky práce zodpovědný stejně jako v případě využití jiných technologických prostředků typu vyhledávání informací na internetu. Současně si musí být vědom základních podmínek fungování AI a rizik spojených s použitím nástroje umělé inteligence v rozsahu této směrnice AI a dalších aktuálních sdělení na intranetu. Zaměstnanci musí vždy ověřovat a provádět kontrolu kvality.

   - **Společenská odpovědnost při využití AI** – V ČRo platí i při používání nástrojů AI zásada vždy jednat v nejlepším zájmu veřejnosti. ČRo musí zkoumat, jak tyto nástroje využít pro posílení veřejné služby, zároveň má povinnost využívat AI tak, aby užití bylo ve shodě se základními etickými hodnotami a lidskými právy a byla minimalizována rizika, která s použitím AI souvisejí. I při použití AI nese konečnou odpovědnost za veškerý obsah a služby, které posluchačům a uživatelům poskytuje, ČRo.

   - **Informovanost a sdílení zkušeností z využití AI napříč organizací** – Zaměstnanci by měli být dobře informováni o používání nástrojů AI před jejich užíváním a měli by absolvovat všechna povinná školení.

   - **Úcta k lidské práci** – Vždy je třeba upřednostňovat talent a oceňovat autentické lidské vyprávění moderátorů, respondentů a tvůrců před výstupy AI. Žádná technologie nemůže plně nahradit lidskou kreativitu a lidskou práci. ČRo je povinen při používání AI vždy brát ohled na práva tvůrců a držitelů práv. Pro tvorbu autorských děl v Českém rozhlase upřednostňujeme spolupráci s autory na úkor obsahu generovaného umělou inteligencí, ať již pro ad hoc příležitost nebo tvořeného systémově.

   - **Bezpečnost** – Modely AI mohou pracovat s osobními údaji, proto je nutné dbát zejména na ochranu soukromí a dodržování příslušných zákonů, zejména nařízení Evropského parlamentu a Rady (EU) 2016/679 o ochraně osobních údajů. Do nástrojů AI, které slouží k dalšímu trénování AI a mají do nich přístup třetí strany, se nesmí vkládat citlivá, tajná nebo obchodní data ČRo apod. Všechny informace, které se sdílejí v nástrojích AI, musí být považovány za veřejné.

   - **Obecný soulad s pravidly a hodnotami ČRo** – Zaměstnanci ČRo jsou povinni využívat AI tak, aby užití bylo ve shodě se všemi pravidly uplatňovanými v ČRo a hodnotami a zásadami ČRo jako média veřejné služby, zejména Kodexem Českého rozhlasu a Žurnalistickými zásadami (zde) a dalšími předpisy. Na obsah vytvořený AI se použijí stejná pravidla jako na veškerý jiný obsah vytvářený v ČRo.

   - **Autenticita a důvěryhodnost** – Je nutné dbát na to, aby užití umělé inteligence nevedlo k podkopávání důvěryhodnosti ČRo a autenticity jeho obsahu. Užití umělé inteligence může ušetřit finanční náklady spojené s výrobou obsahu, nadměrné či neuvážené užití umělé inteligence však může mít za následek snížení důvěryhodnosti v očích veřejnosti. Je nezbytné vždy dobře zvážit, zda benefity z užití umělé inteligence

převažují nad riziky s tím spojenými, zda je takové užití vhodné a není v hrubém nepoměru s naplňováním veřejné služby.

Čl. 5 Hlavní rizika spojená s využitím generativní AI a související opatření

Výčet rizik, která souvisejí s využitím generativní AI, není vyčerpávající, neboť s ohledem na samotnou podstatu a rychlý vývoj generativní AI nelze všechna do detailu popsat a předvídat. ČRo si je těchto rizik vědom a zavazuje se věnovat úsilí jejich minimalizaci a předcházení jim, i za pomocí transparentní komunikace

1. **Etická rizika**

Etická rizika související s využitím generativní AI zahrnují zejména:

a) Riziko zkreslení a zaujatosti z důvodu algoritmických předsudků modelů generativní AI.

b) Předsudky, které typicky vedou k diskriminaci určitých osob nebo skupin osob, se mohou v inteligentním modelu projevit buď z důvodu samotného designu tohoto modelu, který odráží určité chápání světa samotným vývojovým pracovníkem (přitom se může jednat o nevědomý předsudek), nebo z důvodu zakotvení a následné automatické identifikace předsudků z tréninkových dat jako vzorců chování do budoucna.

c) Častá netransparentnost ve využití tréninkových dat, která v kombinaci s jejich masovým používáním může vést k negativním důsledkům pro společnost. U inteligentních systémů nelze zaručit jejich nezávislost a nezaujatost, proto nelze jejich závěry např. považovat za spravedlivé. Modely mohou také produkovat nekvalitní nebo nepřesné výstupy.

2. **Rizika porušení práv duševního vlastnictví, neoprávněné využívání dat**

a) Rizika při využívání nástrojů generativní AI z pohledu práv duševního vlastnictví lze rozdělit do dvou hlavních kategorií:

b) do dvou hlavních kategorií:

- **Riziko porušení autorských práv:** Generativní AI lze použít k vytvoření obsahu, který je podobný stávajícím autorským dílům, rovněž obsah s kterými AI při vytváření výstupů pracuje, může být nejasný. To může vést k porušení autorských práv, a je potřeba proto vždy pečlivě individuálně posuzovat podmínky užití generativní AI i z hlediska vloženého obsahu a podmínky užití výstupů a autorství výstupů.

- **Riziko porušení jiných práv duševního vlastnictví:** Generativní AI lze také použít k vytvoření obsahu, který obsahuje prvky, které jsou chráněny jinými právy duševního vlastnictví, například ochrannými známkami, průmyslovými vzory nebo patenty. To může vést k porušení těchto práv, a je potřeba proto vždy pečlivě individuálně posuzovat podmínky užití generativní AI i z hlediska vloženého obsahu a podmínky užití výstupů.

c) Generativní AI je v současné době předmětem mnoha sporů v oblasti zneužití práv duševního vlastnictví, a to kvůli užití děl chráněných autorským právem bez řádného svolení autorů. Tato skutečnost může znamenat riziko budoucích soudních sporů z důvodů užití chráněných děl ve výstupech vytvořených těmito technologiemi.

d) Stávající autorské právo neposkytuje právní jistotu ohledně autorství k výstupům generativní AI. Autorem může být v souladu s naším právním řádem pouze fyzická osoba. Je třeba vzít na vědomí, že všechny výstupy generativní AI nejsou chráněny autorským právem. Vyvstává tedy otázka, kdy může být systém AI považován za pouhý nástroj k vytvoření díla a kdo je v takovém případě autorem dle svého jedinečného tvůrčího vkladu do výstupu generativní AI, nebo kdy na konkrétním výstupu generativní AI není dostatečný

jedinečný tvůrčí vklad ani uživatele AI ani tvůrce AI a výstup generativní AI není autorským dílem. Toto je nutné vždy individuálně posoudit.

e) Tréninková data, včetně mediálního obsahu, pocházejí z různých zdrojů na internetu bez souhlasu nebo dokonce přes výslovný zákaz oprávněných osob. Samotný tréninkový proces často pracuje s daty, u kterých autor upravil rozsah jejich volného užívání a jen zřídkakdy před zařazením těchto dat do procesu došlo k vypořádání takových práv nebo např. k informování vlastníků dat, kterými mohou být i veřejnoprávní instituce. Výsledkem jsou tak výstupy, u kterých můžeme jen velmi obtížně ověřit legálnost a zdroj. Tyto výstupy mohou být rovněž nekvalitní, zastaralé nebo zkreslené.

f) I vzhledem k dosud vyvíjející se právní úpravě užití generativní AI a jejich výstupů a právní nejistotě ohledně autorství výstupů generativní AI musí zaměstnanci ČRo používat nástroje AI způsobem, který minimalizuje rizika v oblasti porušení autorských práv a jiných práv duševního vlastnictví.

3. **Rizika v oblasti ochrany osobních údajů**

a) Rizika v oblasti ochrany osobních údajů při využívání nástrojů generativní AI lze rozdělit do dvou hlavních kategorií:

- **Riziko úniku osobních údajů ze strany ČRo**: Toto riziko úzce souvisí s rizikem v oblasti informační bezpečnosti a nastupuje v okamžiku, kdy prostřednictvím nástrojů AI sdílíme informace (zejm. prostřednictvím promptů) obsahující osobní údaje, které se tímto způsobem mohou dostat do nepovolaných rukou.

- **Riziko porušování pravidel ochrany osobních údajů ze strany poskytovatelů nástrojů AI**: Toto riziko souvisí s využíváním tréninkových dat, které obsahují osobní údaje, k čemuž příslušné fyzické osoby nemusely dát souhlas. Nástroje generativní AI byly a stále jsou předmětem šetření lokálních dozorových úřadů, neboť ohledně režimu zajištění dostatečné ochrany osobních údajů panují pochybnosti.

b) Vývoj a zejména testování a vylepšování algoritmu umělé inteligence vyžaduje, aby měl její vývojový pracovník k dispozici dostatečné množství testovacích dat. Pokud testovací data zahrnují osobní údaje, musí je vývojovým pracovníkem nejen shromáždit, ale zároveň i v souladu s platnými právními předpisy odůvodnit jejich zpracování pro účely vývoje algoritmu umělé inteligence. Zároveň je nutné dostát všem dalším povinnostem v oblasti ochrany osobních údajů, což v současné době je pro některé společnosti vyvíjející nástroje umělé inteligence problém.

c) Zaměstnanci ČRo musí používat nástroje AI způsobem, který respektuje soukromí třetích osob, platné právní předpisy a povinnosti Českého rozhlasu jako správce osobních údajů, který je zodpovědný za jejich ochranu.

4. **Redakční rizika**

Hlavní redakční rizika v médiích veřejné služby při použití nástrojů AI lze rozdělit do následujících kategorií:

a) **Riziko zkreslení a zaujatosti:** Generativní AI je trénována na datech, která mohou obsahovat předsudky a stereotypy. Pokud nejsou tato data pečlivě vybrána a vyčištěna, může to vést k tomu, že generované výstupy budou také obsahovat tyto předsudky. To může mít negativní dopad na důvěryhodnost a objektivitu média veřejné služby, pokud by výstup nebyl podroben řádné kontrole.

- Umělá inteligence může výrazně urychlit a zefektivnit procesy probíhající při redakční činnosti. Uplatňuje se při přepisování, shrnutí delšího textu a klasifikaci obsahu. Je třeba

si však neustále uvědomovat, že se jedná jen o pracovní nástroje, které používají lidé s určitou odborností a jsou za výstupy AI zodpovědní, jako by výstup vytvořili sami.

- Největší riziko, které v redakční oblasti s využitím AI souvisí, spočívá v tom, že AI produkuje chybné či zavádějící informace, které je s ohledem na nedostatky v oblasti uvádění zdrojů obtížné odhalit. Zdroj informace také často není vůbec nikde uveden, protože se při procesu AI nedá přesně detekovat.

- U chatbotů založených na AI existuje tendence k tzv. halucinacím, kdy výsledky, které jsou prezentovány jako pravdivé, jsou zcela vymyšlené nebo věcně nesprávné. To je způsobeno procesem učení generativní AI: modely rozumějí jazyku, ale nerozumějí tomu, co se v něm skrývá. Výstupy AI mohou také trpět nepřesnostmi a zaujatostmi: od kulturních a geografických slepých míst kvůli tréninkovým datovým souborům, které jsou většinou severoamerické, až po historické zkreslení, ale mohou trpět i politickou zaujatostí, demografickou zaujatostí, sexismem, rasismem apod. V případě interakce systému AI s fyzickými osobami, musí být fyzické osoby vyrozuměni, že komunikují se sytémem AI.

b) **Riziko snížení důvěryhodnosti a autenticity:** Nadměrné či neuvážené užití umělé inteligence může mít vážné dopady na vnímání Českého rozhlasu ze strany veřejnosti jako důvěryhodného média. Kupříkladu využití syntetického hlasu namísto výpovědi skutečné osoby může snižovat autenticitu obsahu její výpovědi a ve svém důsledku mít vliv i na celkovou důvěryhodnost Českého rozhlasu.

- Užití nástrojů generativní AI v žurnalistice s sebou nese ještě další specifické etické otázky týkající se důvěryhodnosti zpravodajských příspěvků doprovázených ilustracemi vytvořenými pomocí AI a vlivu AI na kvalitu novinářských výstupů.

- Zaměstnanci ČRo musí používat nástroje AI způsobem, který minimalizuje rizika zaujatosti a zkreslení, nepřenáší na AI zodpovědnost za publikovaný obsah a nepodrývá důvěryhodnost ČRo.

5. **Riziko informační a kybernetické bezpečnosti**

a) Hlavní rizika spojená s užíváním nástrojů AI z pohledu informační bezpečnosti lze rozdělit do následujících kategorií:

- **Riziko kybernetických útoků**: Nástroje AI mohou být cílem kybernetických útoků, které mohou vést k úniku dat, poškození systémů nebo dokonce k zastavení provozu.

- **Riziko bezpečnostní zranitelnosti**: Nástroje AI mohou obsahovat bezpečnostní díry, které mohou být zneužity k útokům.

b) Generativní AI využívá poskytnutá data uživatelů k tréninku a získávání nových zkušeností, což s sebou nese značná rizika v oblasti informační bezpečnosti. Interní informace, které např. chatbot využívá ke svému tréninku, mohou být zpřístupněny skrze dobře mířené prompty i dalším uživatelům.

c) Stejně jako každý jiný software, může i AI obsahovat bezpečnostní chyby, které mají potenciál být zneužity k úniku dat nebo neoprávněnému přístupu k záznamům konverzací, uživatelským informacím nebo jiným citlivým údajům, což může v konečném důsledku vést ke krádeži identity, zneužití dat či ohrožení soukromí.

d) Kybernetická rizika se mohou objevit i v případě, že bude generativní AI využita pro tvorbu zdrojového kódu k programům/aplikacím použitého v intranetu ČRo, eventuálně s přesahem do internetu (webové servery a služby). I v takovém případě nese za

spolehlivost a bezpečnost kódu odpovědnost pracovník, resp. oddělení, které jej s použitím generativní AI vytvořilo, a to v souladu s principem odpovědnosti dle této směrnice AI.

e) Z pohledu kybernetické bezpečnosti jsou proto zaměstnanci ČRo povinni zabezpečovat své účty na nástrojích AI pomocí unikátních silných hesel, dvoufaktorového ověřování a pravidelně je aktualizovat o bezpečnostní záplaty.

6. **Celospolečenská rizika**

a) Hlavní celospolečenská rizika spojená s užíváním nástrojů AI lze rozdělit do následujících kategorií:

- **Riziko ztráty pracovních míst**: AI může být použita k automatizaci úloh, které jsou v současné době vykonávány lidmi. To může vést k ztrátě pracovních míst a k sociální nestabilitě.

- **Riziko nerovnosti**: AI může být použita k posílení existujících nerovností, například mezi bohatými a chudými, nebo mezi lidmi s různým vzděláním nebo sociálním postavením.

- **Riziko ztráty kontroly**: AI může být použita k tomu, aby lidé byli řízeni nebo ovládáni bez jejich vědomí. To může vést k ztrátě svobody a demokracie.

- **Riziko závislosti na AI**: Generativní AI může být velmi užitečným nástrojem, ale může také vést k závislosti na něm. To může být problematické, protože AI může být chybná nebo může být použita k vytvoření obsahu, který je škodlivý nebo nebezpečný.

b) Nástroje AI se hodí k automatizaci konkrétních úkolů, rozhodnutí nebo pracovních postupů prováděných zaměstnanci, což pravděpodobně v budoucnu ovlivní dostupnost a povahu pracovních míst a způsob, jakým je vnímána lidská práce v oblasti médií i mimo ně.

c) S ohledem na povahu generativní AI, která má potenciál významně zefektivnit lidskou práci, je nutné vnímat i další celospolečenská rizika, která s jejím využitím souvisejí. Generativní AI snižuje náklady pro generování obsahu a produkci zpráv, což umožňuje vstup nových poskytovatelů obsahu na trh. Tito noví konkurenti však mohou mít diametrálně odlišné cíle a úroveň vnímání etiky.

d) ČRo by měl používat umělou inteligenci způsobem, který je prospěšný společnosti a který pomáhá řešit celospolečenská rizika. Měl by poskytovat informace o rizicích spojených s nástroji AI, podporovat výzkum a vývoj bezpečných nástrojů a spolupracovat s dalšími organizacemi, které se zabývají snížením celospolečenských rizik.

7. **Rizika týkající se dětí – zvláštní ohled na děti**

a) Při generování obsahu určeného dětem jsme povinni brát ohled na to, že tyto skupiny obyvatel jsou zvláště zranitelné z důvodu jejich věku, čímž jsou více náchylné k manipulacím a dochází u nich snáze ke zkreslení skutečnosti a vyvolání zmatených představ.

b) Český rozhlas by měl používat umělou inteligenci způsobem, který je bezpečný a který chrání děti.

Čl. 6 Zakázané postupy v oblasti AI

V souladu s evropským nařízením o umělé inteligenci jsou v ČRo zakázané následující systémy umělé inteligence:

a) systémy AI, které používají podprahové techniky a manipulují s chováním osob;

b) systémy AI, které zneužívají zranitelnosti osob (věk, zdravotní stav, ekonomická situace) k manipulaci;

c) systémy AI sociálního bodování, hodnotící sociální chování osob, které vedou k nepřiměřenému zacházení nebo diskriminaci;

d) systémy predikující riziko spáchání trestného činu na základě osobnostních znaků bez objektivních důkazů;

e) systémy AI, které vytvářejí nebo rozšiřují databáze rozpoznávání obličeje prostřednictvím necíleného získávání zobrazení obličeje z internetu nebo kamerových záznamů;

f) systémy, které hodnotí emoce osob na pracovištích a ve školách, s výjimkou lékařských nebo bezpečnostních důvodů;

g) systémy biometrické kategorizace, které kategorizují osoby na základě citlivých biometrických údajů (např. rasa, náboženství, politické názory) za účelem predikce jejich chování;

h) systémy biometrické identifikace na dálku v reálném čase ve veřejně přístupných prostorech pro účely vymáhání práva, bez nezbytných záruk a podmínek.

Čl. 7 Proces hodnocení rizik a schvalování využití generativní AI v ČRo

1. Vzhledem k rizikům výše uvedeným, které využití generativní AI může představovat, je užití generativní AI v ČRo regulováno a podléhá v případech dále uvedených povinnému procesu posouzení rizik a dopadů.

2. Proces vyhodnocování rizik se vztahuje na:

   - každé nové systémové využití nástrojů generativní AI zaměstnanci Českého rozhlasu. Systémovým využitím se rozumí organizované a opakované využití generativní AI v pracovních procesech;

   - užití systémů AI, které mohou s ohledem na zamýšlený účel užití představovat vysoké riziko újmy na zdraví a bezpečnosti nebo na základních právech osob s přihlédnutím jak k závažnosti možné újmy, tak k pravděpodobnosti, že nastane (jedná se zejména o systémy biometrické identifikace na dálku, v oblasti zaměstnávání, náboru či řízení zaměstnanců či vzdělávání) (dále jen „vysoce rizikové systémy“);

   - užití generativní AI, jehož výsledkem jsou výstupy, které mohou být zaměnitelné s autorskými díly, tj. jedná se např. o nějaký literární útvar, hudební útvar, vizuální útvar (dále jen jako „autorské užití AI“) a jsou užity ke sdělování veřejnosti.

3. Individuální využití nástrojů generativní AI jednotlivými zaměstnanci nepodléhá procesu hodnocení rizik a schvalování využití dle této směrnice AI, ale musí být vždy v souladu s hlavními zásadami využití nástrojů AI popsanými v této směrnici AI.

4. Systémové využití generativní AI, užití vysoce rizikových systémů a autorské užití generativní AI je s ohledem na uvedená rizika v ČRo možné pouze po řádném provedení posouzení rizik a dopadů tohoto způsobu užití na činnost ČRo (dále jen jako „posouzení rizik a dopadů“).

5. Proces posouzení rizik a dopadů je nutné zahájit před započetím využívání generativní AI konkrétním způsobem. V případě, že využití vyžaduje vynaložení finančních prostředků, je nutné toto vyhodnocení provést před tím, než dojde k závaznému rozhodnutí o vynaložení těchto prostředků daným způsobem. Součástí procesu posouzení rizik a dopadů může být i návrh opatření ke zmírnění těchto rizik a dopadů.

6. Zahájení procesu posouzení rizik a dopadů iniciuje příslušný vedoucí zaměstnanec na prvním stupni řízení, který plánuje využít generativní AI způsobem dle odstavce 2 tohoto článku v rámci činnosti jím řízených útvarů.

7. Posouzení rizik a dopadů je nutné provádět opakovaně, jestliže se změní podmínky, za kterých bylo původní posouzení rizik a dopadů provedeno. Změnou podmínek se rozumí zejména změna parametrů daného využití generativní AI, změna právních okolností (závazné soudní rozhodnutí týkající se daného způsobu užití, změna právních předpisů apod.) nebo změna podmínek užití daného nástroje ze strany poskytovatele služby aj.

8. Příslušný vedoucí zaměstnanec na prvním stupni řízení, který v rámci činnosti jím řízených útvarů využívá generativní AI, je povinen průběžně vyhodnocovat použití generativní AI s ohledem na rizika a dopady, která byla v rámci hodnocení definována a vyhodnotit účinnost přijatých opatření ke zmírnění těchto rizik, pokud byla v rámci procesu posouzení rizik a dopadů navržena.

9. Vyhodnocení rizik a dopadů využití generativní AI v ČRo provádí pracovní skupina Strategického výboru k využití AI v ČRo. Účast příslušného zaměstnance na prvním stupni řízení na jednání pracovní skupiny, kde se posuzují rizika a dopady, je povinná. Výsledkem procesu posouzení rizik a dopadů je vyjádření pracovní skupiny k jednotlivým rizikům daného využití generativní AI, které se provede formou zápisu.

10. Pokud externí dodavatel poskytuje Českému rozhlasu systém AI nebo jakýkoli produkt, který takový systém využívá, je příslušný věcný garant povinen vynaložit maximální úsilí, aby zajistil, že dodavatel bude vázán Pravidly pro externí dodavatele Českého rozhlasu pro využívání umělé inteligence, která tvoří přílohu této směrnice.

11. Pracovní skupina Strategického výboru k využití AI v ČRo také navrhuje vyjádření ČRo k využívání generativní AI určená veřejnosti.

12. V oblasti využití AI v rámci pracovních procesů v ČRo poskytuje metodickou pomoc Pracovní skupina Strategického výboru k využití AI v ČRo. Kontaktní osobou je vedoucí Oddělení strategického rozvoje.

Čl 8. Autorské užití AI

1. K autorskému užití AI může dojít buďto prostřednictvím zaměstnanců ČRo nebo prostřednictvím externistů. V případě, že autorské užití AI bylo realizováno na základě smlouvy s externistou, jsou informace o využití AI a rozsahu tohoto využití případně další důležité okolnosti využití AI součástí smlouvy s tímto externistou. Rozsah údajů potřebných pro řádné vedení evidence vysílání či evidence rozhlasových pořadů vyplývá z aktuálních požadavků útvarů odpovědných za vedení těchto agend.

2. V případě, že dochází k autorskému užití AI, je nutné zajistit splnění všech podmínek v oblasti evidence vysílání a evidence rozhlasových pořadů, které stanoví platné právní předpisy, vnitřní předpisy a smluvní závazky vyplývající ze smluv s kolektivními správci ve stejném rozsahu jako v případě, kdy není k tvorbě zvukového obsahu využita AI.

3. Pro autorské užití AI je zvláště důležité dodržení zásady, která stanoví povinnost úcty k lidské práci, a která upřednostňuje talent, lidskou kreativitu a lidskou práci, která má přednost před vytvářením obsahu pomocí AI. Odchýlení se od tohoto základního pravidla je možné pouze v souladu s touto směrnicí vč. jejích příloh, nebo pokud se v rámci procesu posouzení rizik a dopadů dojde k závěru, že takové užití je v zájmu ČRo.

Čl. 9 Evidence jednotlivých užití AI v ČRo

1. Pracovní skupina Strategického výboru k využití AI vede evidenci jednotlivých způsobů užití generativní AI v ČRo. V rámci této evidence jsou pak specifikovány jednotlivé nástroje AI, rozsah jejich užití a rizika, která byla v rámci procesu posouzení rizik a dopadů definována.

2. Evidence posuzovaných nástrojů AI v ČRo je dostupná všem zaměstnancům ČRo na Intranetu zde.

Čl. 10 Závazná pravidla pro konkrétní způsoby využití AI v ČRo

Závazná pravidla pro využití AI konkrétními způsoby, u kterých již byla posouzena rizika a dopady těchto způsobů využití, jsou uvedena v přílohách č. 1 až 4.

Čl. 11 Pravidla zadávání veřejných zakázek s AI komponentou

Zadávání veřejných zakázek s AI komponentou se řídí příslušnými pravidly dle Směrnice o zadávání veřejných zakázek v Českém rozhlasu a pravidly uvedenými v příloze č. 6 této směrnice Pravidla pro užívání umělé inteligence v ČRo.

Čl. 12 Závěrečná ustanovení

1. Směrnice AI je závazná pro všechny zaměstnance ČRo.

2. Směrnice AI nabyla účinnosti dnem 15. ledna 2024, byla e novelizována 1. novelou s účinností od 1. října 2024, 2. novelou s účinností od 1. července 2025 a 3. novelou s účinností od 1. září 2025.

Mgr. René Zavoral
generální ředitel ČRo

Přílohy
Příloha č. 1 Pravidla využití AI pro syntézu řeči
Příloha č. 2 Pravidla využití AI pro vytváření vizuálního obsahu
Příloha č. 3 Pravidla využití AI pro generování zdrojových kódů k programům a aplikacím
Příloha č. 4 Pravidla pro použití AI pro úpravu audio a audiovizuálního obsahu
Příloha č. 5 Pravidla pro externí dodavatele Českého rozhlasu pro využívání umělé inteligence
Příloha č. 6 Pravidla zadávání veřejných zakázek s AI komponentou