12. 12. 2019

Vezměte, prosíme, na vědomí, že text článku odpovídá platné právní úpravě ke dni publikace.

Umělá inteligence ve světle ochrany soukromí – 1. díl

Právní limity pro vývoj a používání umělé inteligence v oblasti marketingu: princip minimalizace a účelového omezení zpracování

V poslední době se hojně diskutuje o vlivu GDPR na vývoj umělé inteligence, přičemž zaznívají názory, dle kterých GDPR zpomaluje vývoj a pokrok v oblasti umělé inteligence v Evropské unii v porovnání s ostatními regiony.

Na první pohled se může zdát, že GDPR a umělá inteligence jsou postaveny na principech, které se zcela vylučují. GDPR vychází z principu minimalizace dat, striktního účelového omezení, a principu transparentnosti. Machine learning algoritmy naopak potřebují co největší množství dat, aby byly spolehlivé, jejich cílem je generovat co nejvíce nových dat a jejich fungování je komplexní a pro běžného uživatele obtížně pochopitelné.

V tomto článku se zaměřujeme na to, jak je možné tyto protiklady překlenout. V prvním díle se věnujeme principu minimalizace, účelovému omezení a zákonnosti zpracování, a to jak z hlediska vývoje algoritmů umělé inteligence, tak z hlediska dovozování derivovaných dat o subjektech údajů.

V druhém díle se zaměříme na požadavky týkající se automatizovaného rozhodování a výkonu práv subjektů údajů.

PIERSTONE

Princip minimalizace – účel světí prostředek?

Princip minimalizace a účel zpracování

Princip minimalizace zpracování osobních údajů je spolu se zásadami zákonnosti a transparentnosti jednou ze základních zásad GDPR. Je zakotvený v článku 5 odst. 1 písm. c) GDPR a vyžaduje, aby byly zpracovávány jen ty osobní údaje, které jsou pro daný účel *„přiměřené, relevantní a omezené na nezbytný rozsah*“. Pokud tedy správce osobní údaje pro daný účel nezbytně nepotřebuje, neměl by je vůbec uchovávat či jinak zpracovávat, a to ani na základě souhlasu subjektů údajů. Pokud je např. účelem zpracování prodej služeb, neměl by správce pro tento účel zpracovávat údaje např. ze sociální sítě použité k uživatelské registraci, pokud nejsou pro danou službu nezbytné.

Účel vs. prostředek zpracování

Použití algoritmu umělé inteligence však nebude zpravidla účelem zpracovaní, ale pouze tzv. prostředkem zpracování[1]. Rozsah údajů, které může správce v souladu se zásadou minimalizace údajů zpracovávat se tedy nebude odvíjet od potřeb algoritmu umělé inteligence jakožto prostředku

zpracování, ale od potřeb vyplývajících z účelu zpracování, tedy např. poskytování konkrétní služby. V praxi tento rozsah údajů bude často podobný – pokud je služba poskytována s využitím algoritmu strojového učení, bude zpravidla nutné, aby tento algoritmus zpracovával dostatečný rozsah údajů nutný k jeho spolehlivému fungování.

Často však budou existovat údaje, které by pro spolehlivost algoritmu strojového učení byly užitečné, ale nelze již obhájit jejich nezbytnost pro daný účel (např. poskytování služby). Zpravidla to budou údaje týkající se jiného okruhu osob, než je zákaznická báze správce. Příkladem může být zpracování prováděné pojišťovnou za účelem stanovení rizikového profilu konkrétního pojištěnce pro účely stanovení odpovídající výše pojistného. Účelem zpracování je zde poskytování služby, kterou je pojištění. Prostředkem zpracování je algoritmus umělé inteligence vyhodnocující rizikový profil, který umožňuje optimalizovat cenu pojištění tak, aby byla pro zákazníky s nízkým rizikovým profilem výhodná. Algoritmus by fungoval lépe, kdyby mohl zpracovávat i data týkající se jiných osob než zákazníků dané pojišťovny, ale to již nebude pravděpodobně obhajitelné z hlediska účelu zpracování. Takové zpracování by porušovalo princip minimalizace zpracování osobních údajů zakotvený v článku 5 odst. 1 písm. c) GDPR.

Vývoj algoritmu a odpovídající účel zpracování

Ačkoli nasazení umělé inteligence v praxi bude zpravidla představovat pouze prostředek zpracování, nikoli účel, v případě vývoje umělé inteligence je situace odlišná. Vývoj a zejména testování a vylepšování algoritmu umělé inteligence vyžaduje, aby měl vývojář k dispozici dostatečné množství testovacích dat. Pokud testovací data zahrnují osobní údaje, musí je vývojář nejen shromáždit, ale zároveň i v souladu s článkem 6 GDPR odůvodnit jejich zpracování pro účely vývoje algoritmu umělé inteligence.

Správce osobních údajů, který si pro vývoj a testování algoritmu přeje využít osobní údaje, kterými disponuje, může v zásadě volit mezi třemi alternativami odůvodnění takového zpracování osobních údajů.

Poskytování služby

V první řadě může dovodit, že vývoj umělé inteligence představuje součást služeb poskytovaných zákazníkovi a zpracování je tak nezbytné za účelem plnění smlouvy mezi správcem a subjektem údajů v souladu s článkem 6 odst. 1 písm. b) GDPR. V důsledku může docházet k výraznému rozšiřování rozsahu zpracovatelských operací, které budou považovány za součást služby a

potenciálně i rozsahu zpracovávaných údajů. Nabízí se však otázka, v jaké míře by takové zpracování bylo pro účel poskytování služby skutečně nezbytné, jak vyžaduje princip minimalizace.

Oprávněný zájem

Pokud nebude možné vývoj umělé inteligence podřadit pod poskytování služeb, mohl by správce osobní údaje zpracovávat na základě svého oprávněného zájmu dle článku 6 odst. 1 písm. f) GDPR. Takové zpracování však vyžaduje provedení tzv. balančního testu, tj. posouzení, zda zájmy subjektů údajů nepřevažují nad oprávněným zájem správce. Pokud by zájmy subjektu údajů převážily, nebylo by možné údaje bez souhlasu zpracovávat.

Test slučitelnosti účelů

Případný oprávněný zájem však zpravidla bude představovat nový účel, odlišný od původního účelu, pro který správce osobní údaje shromáždil (např. pokud správce zpracovává historické údaje svých zákazníků, které byly shromážděny pro účely poskytování služby).[2] Aby mohl pokračovat ve zpracování údajů za tímto novým, sekundárním účelem, musí správce provést test slučitelnosti účelů zakotvený v článku 6 odst. 4 GDPR. Dle tohoto ustanovení musí správce posoudit, zda je sekundární účel slučitelný s primárním účelem, zejména zda mezi účely existují vazby, jaké důsledky sekundárního zpracování pro subjekt údajů plynou a zda jsou zajištěny vhodné záruky práv subjektu (např. šifrování či pseudonymizace).

Jedním z potenciálních oprávněných zájmů správce může být sběrná kategorie zlepšování služeb poskytovaných správcem, u které zpravidla bude možné dovodit vazby na účel primární a lze tedy předpokládat, že test slučitelnosti bude splněn.

Druhým možným oprávněným zájmem je vývoj nových služeb či technologií. V tomto případě však bude slučitelnost zpravidla účelů obtížněji odvoditelná a lze předpokládat spíše negativní výsledek balančního testu. Pro využití údajů tak pravděpodobně bude nutný souhlas subjektů údajů.

Nedostatek dat jako překážka ve vstupu na trh a možnosti regulace

Pro nové hráče na trhu umělé inteligence, kteří vstupními datasety nedisponují, bude situace komplikovanější, neboť zpravidla nebudou mít jinou možnost než opatřit datasety externě, zpravidla na základě souhlasu subjektu údajů. Ačkoli oprávněný zájem s sebou nese dodatečné komplikace v podobě výše uvedeného balančního testu, oproti souhlasu se stále jedná o výhodnější režim, zejména s ohledem na volnou odvolatelnost souhlasu.

Nedostatek „vlastních dat“ tak může představovat jakousi překážku ve vstupu na trh vývoje umělé inteligence nebo na konkrétní relevantní trh konkrétních služeb, na kterém ostatní hráči již algoritmy umělé inteligence využívají. [3]

Za účelem překonání těchto bariér vstupu na trh a snížení množství dat, které nově vyvíjený algoritmus skutečně potřebuje, se mnozí odborníci z oblasti vývoje softwaru a umělé inteligence snaží definovat koncepty a směry, kterými by se umělá inteligence měla ubírat, aby byl zajištěn maximální možný soulad se zásadou minimalizace zpracování. Jedním z těchto směrů je např. vývoj umělé inteligence na principu lidských neuronů, která pro fungování svého algoritmu nepotřebuje takové množství reálných vstupních dat, ale místo toho si dokáže různé vstupní varianty „představit“ (tzv. „thinking human AI“). Další diskutovanou možností je přiblížit proces machine learning, tedy učení umělé inteligence, do roviny počítače uživatele tak, aby umělá inteligence měla možnost se vstupními osobními údaji pracovat, ale do sítě se následně dostaly pouze výsledky tohoto učení umělé inteligence v podobě anonymizovaných dat.

V oblasti regulace se uvažuje o tzv. regulatorních sandboxech, tj. zkušebním či testovacím prostředí, které by vývojářům umělé inteligence umožnilo provádět vývoj a výzkum pod dohledem dozorového orgánu s omezeným či dočasně zcela vyloučeným uplatněním norem regulace. Regulatorní sandboxy zmiňuje jako jeden z krátkodobých cílů (do roku 2021) pro podporu vývoje umělé inteligence také Národní strategie umělé inteligence v České republice.[4]

Co s údaji, které o vás umělá inteligence „vydedukuje“?

Jedním z hlavních přínosů algoritmů umělé inteligence je jejich schopnost vyvodit na základě vstupních dat nové údaje, které správce původně neměl k dispozici. Tato data mohou pomoci zefektivnit procesy, získat lepší analytické nástroje nebo zvýšit prodej výrobků a služeb.

Příkladem může být využití algoritmu machine learning pro optimalizaci prodejů v e-shopu prostřednictvím cílených nabídek. Tyto nabídky budou cíleny nejen na základě údajů získaných z poskytování služeb, jako např. zákaznická registrace, historie nákupů, a prohlížené produkty, ale také na základě údajů, které je provozovatel schopen z těchto údajů dovodit. Může se jednat o např. vzorce chování, zájmy, pravděpodobnost dalšího nákupu, ale také údaje o zdravotním stavu, politickém přesvědčení či jiné zvláštní kategorie osobních údajů. Na základě těchto dovozených údajů je správce schopen individualizovat nabídku svých výrobků či služeb a zvýšit tak prodej.

Účelové omezení zpracování derivovaných dat

Při zpracování derivovaných údajů je otázkou, za jakým účelem je správce tyto údaje oprávněn zpracovávat. V úvahu přicházejí v zásadě dvě alternativy – správce bude oprávněn derivovaná data zpracovávat, jestliže je to nezbytné za účelem poskytování služby, nebo jestliže se jedná o oprávněný zájem správce. V obou případech je správce limitován – v prvním případě se jedná o test nezbytnosti zpracování pro poskytování služby dle čl. 6 odst. 1 písm. b) GDPR, zatímco v druhém případě musí správce provést balanční test mezi svým oprávněným zájmem a oprávněnými zájmy subjektu údajů v souladu s článkem 6 odst. 1 písm. f) GDPR.

Test slučitelnosti při použití historických dat

Pokud správce zamýšlí zpracovávat historická data, která původně shromáždil za jiným účelem, musí kromě výše uvedených testů také posoudit, zda jsou tyto sekundární účely zpracování slučitelné s účelem primárním v souladu s článkem 6 odst. 4 GDPR. Pokud zůstaneme u příkladu e-shopu, aby mohl správce své nabídky individualizovat, bude nucen nejdříve zpracovat historická data o

nákupech svých zákazníků, avšak nyní za zcela jiným účelem (např. cílená reklama, individualizované nabídky). Pokud by účel nebyl slučitelný s účelem původním (např. provedení nákupu, poskytování služby), nebude zpravidla možné založit zpracování na původním právním základu, ať už se jedná o plnění smlouvy nebo oprávněný zájem správce a pro účely tzv. personalizovaných reklam tak bude nutné získat souhlas subjektu údajů.

Derivované zvláštní kategorie osobních údajů

V případě dovozování zvláštních kategorií údajů dle článku 9 odst. 1 GDPR, tj. např. zdravotního stavu nebo rasové či náboženské příslušnosti, to bude ještě složitější. I kdyby zpracování splnilo test slučitelnosti účelů, je kromě právního základu dle článku 6 třeba dovodit také specifický právní základ pro zpracování zvláštních kategorií osobních údajů.

Taxativní výčet těchto právních základů je uveden v článku 9 odst. 2 a za předpokladu, že správce nedisponuje souhlasem subjektů, lze v dané situaci uvažovat pouze o jednom dalším právním základu, kterým je písm. j), zejména zpracování nezbytné pro vědecké účely. Členské státy mohou v souladu s článkem 9 odst. 2 písm. j) a článkem 89 GDPR přijmout právní úpravu stanovující odchylky pro zpracování zvláštních kategorií osobních údajů. Zároveň dle článku 9 odst. 4 GDPR mohou členské státy legislativně odlišně upravit podmínky zpracování genetických údajů, biometrických údajů či údajů o zdravotním stavu.

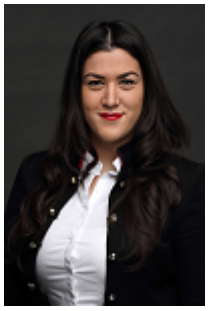
Zajímavá v tomto ohledu může být výše uvedená Národní strategie umělé inteligence vydaná Ministerstvem průmyslu a obchodu České republiky, dle které patří mezi krátkodobé cíle České republiky odstranění administrativní zátěže, usnadnění sdílení a zpřístupnění osobních i neosobních údajů pro účely umělé inteligence ve světle aplikačních zásad GDPR či vytvoření výše zmíněných regulatorních sandboxů. Je možné, že mezi budoucími legislativními změnami bude i nová úprava umožňující zpracovávání derivovaných zvláštních kategorií osobních údajů v souladu s článkem 9 odst. 4, případně odst. 2 písm. j) ve spojení s článkem 89 GDPR.

Přístup dozorových orgánů

Zatím není jasné, jak přísně budou dozorové orgány dodržování principu minimalizace v souvislosti s umělou inteligencí posuzovat. Jedním z faktorů, které by však měly být zohledněny, je potřeba vyvíjet a zajistit spolehlivé algoritmy, které pravděpodobně nebude možné vyvinout a otestovat bez zpracování, které do jisté míry není striktně „nezbytné“ pro konkrétní účel zpracování. Lze však předpokládat, že dozorové orgány pověřené ochranou osobních údajů budou do budoucna hrát významnou roli i v dozoru nad vývojem a nasazováním algoritmů umělé inteligence.

Pokud vás problematika umělé inteligence zajímá, sledujte novinky na www.pierstone.com a https://cz.linkedin.com/company/pierstone, kontaktujte nás na jana.pattynova@pierstone.com nebo teodora.draskovic@pierstone.com nebo se připojte k činnosti Komise pro inovativní sektor v rámci Spolku pro ochranu osobních údajů.

M**gr. Jana Pattynová, LL.M.**



Mgr. Teodora Drašković

PIERSTONE s.r.o., advokátní kancelář

Perlová 371/5
110 00 Praha 1

Tel.: +420 224 234 958

---

[1] Rozlišení mezi účelem a prostředky zpracování vyplývá např. z definice správce dle čl. 4 odst. 7 GDPR.

[2] Mitrou, L., Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'? (2018); Information Commissioner's Office, Big data, artificial intelligence, machine learning and data protection (2017).

[3] Watney C., Duan Ch., R Sheet on Competition in Artificial Intelligence (2018); Kossuth J., Seamans R., The Competitive Landscape of AI (2019).

[4] K dispozici >>> zde.

Další články:

- Dokazování negativních skutečností ve sporném řízení
- Neoprávněný odběr elektřiny – překvapení vlastníka?
- Rodič u dítěte v nemocnici: právo na přítomnost neznamená bez dalšího právo na přespání na jip/jirp
- Pokuta za švarcsystém kurýrů Rohlíku potvrzena Ústavním soudem
- Metropolitní plán schválen. Je Váš projekt v bezpečí?
- Posouzení shody dle AI Act - zkušenosti z praxe
- Začínají soudy zohledňovat náklady podnikatelů při plnění právních povinností v oblasti e-commerce?
- Byznys a paragrafy, díl 35: Ručení za dluhy z podnikání u OSVČ a s.r.o.

- Bezpilotní systémy vlastní konstrukce v kategorii Specific: regulatorní požadavky a praktické aspekty
- Nefungující rozsah péče o dítě. Cesta přes využití terapie a dalších opatření podle ustanovení § 503 zákona o zvláštních řízeních soudních
- De iure traktor, de facto nákladní vozidlo, už ne tolik výhodná dualita