

Univerzita Pardubice
Fakulta ekonomicko-správní

Bezpečnost dat ve firemním prostředí v kontextu nasazení umělé inteligence
Diplomová práce

Univerzita Pardubice
Fakulta ekonomicko-správní
Akademický rok: 2024/2025

ZADÁNÍ DIPLOMOVÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení: **Bc. Kateřina Nekorancová**
Osobní číslo: **E23799**
Studijní program: **N0688A140007 Informatika a systémové inženýrství**
Specializace: **Informační a bezpečnostní systémy**
Téma práce: **Bezpečnost dat ve firemním prostředí v kontextu nasazení umělé inteligence**
Zadávající katedra: **Ústav systémového inženýrství a informatiky**

Zásady pro vypracování

Cílem práce je navrhnout sadu ukázkových případů využití nástrojů AI při analytické práci ve zvolené organizaci se zaměřením na bezpečnost a ochranu dat organizace.

Osnova:

- Bezpečnost a ochrana dat
- Bezpečnostní rizika spojená s používáním nástrojů založených na AI ve firemním prostředí
- Využití AI v oblasti datové analýzy
- Tvorba potřebných promptů a návrh sady ukázkových příkladů

Rozsah pracovní zprávy: **50 stran**
Rozsah grafických prací:
Forma zpracování diplomové práce: **tištěná/elektronická**

Seznam doporučené literatury:

BARTOŠ, Pavel, 2003. *A.I. umělá inteligence, úvod do problematiky a současné trendy: (Jak zprovoznit vlastní A.I.)*, Evropská akademie vzdělávání SE 2023, ISBN 978-80-11-04221-9
MIKELSTEN, Daniel, SKALFIST Peter a TEIGENS Vasil *Umělá inteligence: Čtvrtá průmyslová revoluce*, Cambridge Stanford Books 2020 *Artificial Intelligence Safety*, ISBN 9781005168490
GUERRERO, José María, *Mind Mapping and Artificial Intelligence*, Academic Press; 1st edition (26 October 2022) SBN-10 : 0128201193, ISBN-13 : 978-0128201190
SEDLÁK, Petr, KONEČNÝ Martin, *Kybernetická (ne)bezpečnost: problematika bezpečnosti v kyberprostoru*, Zákon o kybernetické bezpečnosti (zákon č. 181/2014 Sb.) ISBN 978-80-7623-068-2
Usnesení Evropského parlamentu ze dne 16. 2. 2017 obsahující doporučení Komise o občanskoprávních pravidlech pro robotiku (2015/2103(INL)) (OJ C, C/252, 18.07.2018, p. 239, CELEX: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52017IP0051>)

Vedoucí diplomové práce: **prof. Ing. Jan Čapek, CSc.**
Ústav systémového inženýrství a informatiky

Datum zadání diplomové práce: **1. září 2024**
Termín odevzdání diplomové práce: **30. dubna 2025**

prof. Ing. Jan Stejskal, Ph.D. v.r.
děkan

L.S.

prof. Ing. Jitka Komárková, Ph.D. v.r.
garant studijního programu

V Pardubicích dne 1. září 2024

Prohlašuji:

Práci s názvem Bezpečnost dat ve firemním prostředí v kontextu nasazení umělé inteligence jsem vypracovala samostatně. Veškeré literární prameny a informace, které jsem v práci využila, jsou uvedeny v seznamu použité literatury.

Byla jsem seznámena s tím, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon), ve znění pozdějších předpisů, zejména se skutečností, že Univerzita Pardubice má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle § 60 odst. 1 autorského zákona, a s tím, že pokud dojde k užití této práce mnou nebo bude poskytnuta licence o užití jinému subjektu, je Univerzita Pardubice oprávněna ode mne požadovat přiměřený příspěvek na úhradu nákladů, které na vytvoření díla vynaložila, a to podle okolností až do jejich skutečné výše.

Beru na vědomí, že v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších předpisů, a směrnicí Univerzity Pardubice č. 7/2019 Pravidla pro odevzdávání, zveřejňování a formální úpravu závěrečných prací, ve znění pozdějších dodatků, bude práce zveřejněna prostřednictvím Digitální knihovny Univerzity Pardubice.

V Pardubicích dne 24. 4. 2025

Kateřina Nekorancová v.r.

PODĚKOVÁNÍ

Na tomto místě bych chtěla poděkovat prof. Ing. Janovi Čapkovi, CSc. za odborné vedení mé práce, za cenné poznámky a připomínky, za podporu a trpělivost při jejím vytváření.

ANOTACE

Diplomová práce se zaměřuje na problematiku ochrany dat ve firemním prostředí v kontextu nasazení umělé inteligence. V teoretické části je popsán vývoj a klasifikace systémů AI, bezpečnostní a etické výzvy spojené s jejich používáním a právní rámec upravující ochranu osobních údajů. Praktická část se věnuje analýze rizik a návrhům konkrétních opatření na zajištění datové bezpečnosti při využívání AI v podnikovém prostředí. Cílem práce je nabídnout firmám přehled možných hrozeb a doporučení pro bezpečné a odpovědné nasazení umělé inteligence.

KLÍČOVÁ SLOVA

Umělá inteligence, bezpečnost dat, etika, GDPR

TITLE

Data Security in Corporate Environments in the Context of Artificial Intelligence Deployment

ANNOTATION

This diploma thesis focuses on the issue of data protection in corporate environments in the context of artificial intelligence deployment. The theoretical part describes the development and classification of AI systems, as well as the security and ethical challenges associated with their use, including the legal framework regulating personal data protection. The practical part analyses risks and proposes specific measures to ensure data security when implementing AI in business settings. The aim of the thesis is to provide companies with an overview of potential threats and recommendations for the safe and responsible use of artificial intelligence.

KEYWORDS

Artificial intelligence, data security, ethics, GDPR

OBSAH

SEZNAM ILUSTRACÍ A TABULEK.....	9
SEZNAM ZKRATEK	10
ÚVOD.....	11
1 TEORETICKÁ ČÁST	12
1.1 Umělá inteligence	12
1.1.1 Historie umělé inteligence	13
1.1.2 Klasifikace AI systémů.....	14
1.1.3 Aplikace metod umělé inteligence.....	15
1.1.3 Systémy využívající umělou inteligenci	20
1.1.4 Large Language Models	21
1.2 Bezpečnost dat	22
1.2.1 Bezpečnostní rizika.....	22
1.2.2 Zneužití dat	23
1.2.3 Chybné nebo záměrně pozměněné analýzy	23
1.2.4 Únik dat.....	24
1.2.5 Black-box efekt a nedostatečná transparentnost.....	25
1.2.6 Zranitelnost AI systémů.....	26
1.2.7 Etické otázky.....	26
1.2.8 Zneužití AI.....	31
1.2.9 Nástroje a technologie pro zajištění bezpečnosti dat.....	32
1.3 Právní rámec	33
1.4 Využití AI v oblasti datové analýzy	34
1.5 Prompt engineering.....	35
1.5.1 Techniky promptování.....	36
1.5.2 Kategorizace a klasifikace dat	37
1.5.3 Čištění dat	37
1.5.4 Shrnutí výsledků	38
1.5.5 Vyhledávání anomálií	39
1.5.6 Generování kódu.....	40
1.5.7 Predikce	40
1.5.8 Statistická analýza.....	41
1.5.9 Vizualizace a reporting	43

2 PRAKTICKÁ ČÁST	44
2.1 Charakteristika společnosti	44
2.2 Rizika identifikovaná při nasazování AI	45
2.2.1 Možnost úniku dat	45
2.2.2 Možnost ztráty kontroly nad daty v trénovacím režimu	46
2.2.3 Riziko rekonstrukce identity uživatele	46
2.2.4 Potenciální bias při predikcích.....	46
2.2.5 Nedostatečné vysvětlení rozhodnutí modelu (black-box efekt)	46
2.3 Návrh opatření pro ochranu dat	47
2.3.1 Anonymizace vstupních dat.....	47
2.3.2 Šifrování komunikace a datového přenosu	48
2.3.3 Nakládání s osobními údaji.....	48
2.3.4 Interní pravidla pro prompt engineering.....	49
2.3.5 Auditní logování interakcí s AI systémy	49
2.3.6 Pravidelné školení zaměstnanců v oblasti datové etiky a práce s AI.....	50
2.3.7 Zpracování AI rizik v rámci interní GDPR agendy.....	50
2.4 Další doporučení	51
2.4.1 Interní testing adversarial scénářů	51
2.4.2 Zajištění dokumentace rozhodovací logiky modelu	51
2.4.3 Použití open-source modelů na vlastním serveru	51
2.5 Využití umělé inteligence v analýze dat	52
2.5.1 Scriptování	52
2.5.2 Reporting	56
ZÁVĚR	61
POUŽITÁ LITERATURA	62

SEZNAM ILUSTRACÍ A TABULEK

Obrázek 1 Marketingové výdaje podle kanálů	43
Tabulka 1 Tržby v regionech	38
Tabulka 2 Týdenní spotřeba elektřiny	39
Tabulka 3 Vygenerovaný SQL dotaz.....	40
Tabulka 4 Měsíční návštěvnost.....	41
Tabulka 5 Měsíční prodeje	42
Tabulka 6 Marketingové výdaje	43
Tabulka 7 Anonymizace vstupních dat.....	48

SEZNAM ZKRATEK

AI	Artificial Intelligence (Umělá inteligence)
AGI	Artificial General Intelligence (Obecná umělá inteligence)
GDPR	General Data Protection Regulation (Obecné nařízení o ochraně osobních údajů)
XAI	Explainable Artificial Intelligence (Vysvětlitelná umělá inteligence)
NLP	Natural Language Processing (Zpracování přirozeného jazyka)
ML	Machine Learning (Strojové učení)
EU	Evropská unie
DPIA	Data Protection Impact Assessment (Posouzení vlivu na ochranu osobních údajů)
AI Act	Návrh nařízení EU o umělé inteligenci
DPA	Data Protection Authority (Úřad pro ochranu osobních údajů)

ÚVOD

V posledních letech jsme svědky bezprecedentního rozvoje umělé inteligence, která zásadním způsobem mění způsob, jakým lidé pracují s informacemi, rozhodují se a interagují s technologiemi. AI již dávno není výsadou akademického výzkumu či technologických gigantů – stává se běžnou součástí firemních procesů napříč sektory, od automatizace zákaznické podpory, přes analýzu chování klientů až po prediktivní modelování ve financích, lidských zdrojích či marketingu. Tento dynamický rozvoj s sebou přináší nejen nové příležitosti, ale také zcela nové výzvy, na které musí organizace reagovat.

S rostoucím nasazením těchto technologií vyvstává zásadní otázka: Jak zajistit, aby využívání umělé inteligence probíhalo bezpečně, eticky a v souladu s ochranou dat? Vývoj AI totiž probíhá rychleji, než se stačí přizpůsobovat legislativa a vnitřní procesy organizací. Firmy tak často vstupují na neprobádané území, kde chybí jasné standardy nebo osvědčené postupy.

Data se mezitím stala jedním z nejcennějších firemních aktiv. Jejich ochrana je klíčová nejen pro zachování důvěry zákazníků, ale také z hlediska regulatorních požadavků, reputace a konkurenční výhody. Umělá inteligence na jedné straně nabízí nové možnosti efektivní analýzy, predikce a automatizovaného rozhodování, na straně druhé ale přináší specifická rizika – od neúmyslného úniku citlivých informací, přes zkreslení výsledků na základě nevyvážených trénovacích dat, až po obtížně vysvětlitelná rozhodnutí tzv. „černých skříněk“ (black-box systems).

Firmy tak čelí strategické výzvě, jak propojit technologický pokrok s odpovědným a bezpečným přístupem k datům, který bude respektovat nejen zákonné normy (např. GDPR), ale i etické principy důvěry, transparentnosti a odpovědnosti.

Tato diplomová práce se zabývá právě tímto průsečíkem: analyzuje bezpečnostní aspekty využívání AI ve firemním prostředí, identifikuje hlavní rizika a navrhuje konkrétní opatření, která mohou organizace implementovat pro zajištění datové bezpečnosti. Zvláštní pozornost je věnována nejen technickým, ale také právním a etickým dimenzím tohoto tématu, přičemž důraz je kladen na praktické využití poznatků v reálném podnikatelském kontextu.

1 TEORETICKÁ ČÁST

Tato kapitola představuje základní teoretický rámec problematiky umělé inteligence. Zabývá se jejím vymezením, historickými souvislostmi, klíčovými milníky vývoje a klasifikací jednotlivých typů systémů umělé inteligence. Cílem je poskytnout ucelený přehled o vývoji a principech fungování umělé inteligence, které tvoří východisko pro analýzu bezpečnostních aspektů jejího nasazení ve firemní praxi.

1.1 Umělá inteligence

Umělá inteligence (AI) představuje interdisciplinární oblast, která se zabývá vývojem systémů schopných napodobovat nebo simulovat lidské kognitivní schopnosti, jako je učení, rozhodování či řešení problémů. Termín *artificial intelligence* poprvé použil John McCarthy v roce 1955 v návrhu výzkumného projektu, který společně s Marvinem Minským, Claudem Shannonem a Nathanielem Rochesterem předložil pro letní školu na Dartmouth College. V tomto dokumentu autoři vyjádřili základní předpoklad, že „každý aspekt učení nebo jakýkoli jiný rys inteligence může být popsán natolik přesně, že lze sestrojít stroj, který ho bude simulovat“ (McCarthy et al., 1955). Tato myšlenka položila základy umělé inteligence jako samostatného vědeckého oboru, který se od té doby dynamicky rozvíjí a nalézá uplatnění v mnoha oblastech lidské činnosti.

Jiná definice pak tvrdí, že “Umělá inteligence je obor, který zkoumá, jak přimět počítače vykonávat činnosti, které lidé v současnosti zvládají lépe” (Rich, Kinght, 1991). Tento pohled se zaměřuje na praktickou stránku umělé inteligence, tedy na schopnost strojů vykonávat úkoly, které tradičně vyžadují lidskou inteligenci. AI zde není definována jako přesná simulace lidského myšlení, ale jako nástroj pro automatizaci úkolů, které lidé běžně vykonávají. Definice naznačuje, že Hranice mezi tím, co je považováno za “inteligentní” chování AI, a tím, co je běžná výpočetní úloha, se neustále posouvá. To, co bylo dříve považováno za projev inteligence (např. rozpoznávání řeči), může být dnes běžnou funkcí software. Definice nenaznačuje, že AI musí člověka nutně nahradit, ale spíše, že může zefektivnit nebo automatizovat úkoly, které lidé vykonávají. To odpovídá současnému trendu, kdy AI slouží jako nástroj pro podporu rozhodování a zpracování velkých objemů dat.

Moderní AI systémy lze rozdělit do několika kategorií, jako je strojové učení, neuronové sítě, zpracování přirozeného jazyka a autonomní agenti. Významným aspektem AI je její schopnost analyzovat a predikovat vzory v datech, což se uplatňuje v oblastech, jako je zdravotnictví, průmysl, finance nebo kybernetická bezpečnost. Rychlý rozvoj AI však přináší

i výzvy spojené s etikou, transparentností a bezpečností, neboť rozhodovací procesy některých modelů mohou být obtížně interpretovatelné (Mařík a spol., 1993).

1.1.1 Historie umělé inteligence

Umělá inteligence je jednou z nejrychleji se vyvíjejících vědeckých a technických disciplín v historii. Počátky AI sahají do 1. poloviny 20. století, kdy byly položeny první teoretické základy a vytvořeny rané modely strojového učení. Mezi první průkopníky AI patřil Alan Turing, který v roce 1935 popsal fungování univerzálního počítačového stroje, který by měl nekonečnou paměť, byl by schopen tuto paměť procházet po jednotlivých znacích, číst a vyhodnocovat tyto znaky a vytvářet další. Tento stroj, pro nějž se vžilo označení univerzální Turingův stroj, je první teoreticky přesný koncept počítače (Šulc, 2024).

Ve své práci *Computing Machinery and Intelligence* (1950) Turing položil základy pro testování strojové inteligence prostřednictvím tzv. Turingova testu. Cílem tohoto testu je zjistit, zda je stroj schopen napodobit lidské myšlení natolik přesvědčivě, že jej lidský pozorovatel nedokáže odlišit od člověka. Turingův test se stal velmi vlivným nástrojem a má řadu praktických využití, zejména při odbourávání nejrozumnějších robotů v digitální komunikaci (Šulc, 2024). Turing však ve své práci nepředstavil pouze test samotný, ale také systematicky zformuloval a následně vyvracel nejčastější námitky proti možnosti existence inteligentních strojů. Mezi tyto námitky patřily například argumenty o nedostatku vědomí, emocí, originality nebo o neschopnosti strojů dělat chyby podobné lidským. Tyto tzv. Turingovy námitky se staly základem pro filozofickou debatu o povaze umělé inteligence a dodnes jsou citovány jako jeden z pilířů diskuse o strojovém myšlení (Mařík a spol., 1993).

V roce 1956 byl na Darmountské konferenci formálně představen pojem “umělá inteligence”. Výzkumníci jako John McCarthy, Marvin Minsky či Allen Newell vyvíjeli první programy simulující inteligentní chování (McCarthy, 1956).

V 70. a 80. letech probíhalo období zvané “zima AI”, kdy došlo ke snížení zájmu a financování kvůli neuspokojivým výsledkům prvních AI systémů (Crevier, 1993). Přesto vznikly první tzv. expertní systémy využívající znalostí převzatých od špičkových odborníků a simulující rozhodovací činnost člověka s cílem dosahovat obdobné kvality rozhodování, jaké dosahuje expert (Mařík, 2024).

V 90. letech a počátkem 21. stol. zaznamenala AI pokrok díky nárůstu výpočetního výkonu. V roce 1997 počítač Deep blue porazil mistra světa v šachu Garryho Kasparova, což bylo považováno za zásadní milník (Campbell, 2002).

V roce 2022 byla společností OpenAI poprvé veřejně zpřístupněna aplikace ChatGPT, založená na jazykovém modelu GPT-3.5. Díky svému intuitivnímu rozhraní a přirozeným odpovědím si rychle získala miliony uživatelů po celém světě. Již během prvních pěti dnů od spuštění aplikace ji využilo již milion uživatelů, v řádu pouhých měsíců se jejich počet vyšplhal na 100 milionů, čímž se stal jednou z nejrychleji rostoucích spotřebitelských aplikací v historii. Hlavní osobností, která stála za tímto úspěchem, byl CEO Open AI Sam Altman. Jeho cílem bylo zpřístupnit AI co nejširšímu spektru uživatelů a zároveň klást důraz na bezpečné, odpovědné a etické využívání těchto technologií. Společnost OpenAI si však uvědomuje i jistá rizika a několikrát varovala, že modely AI, pokud s nimi nebude nakládáno s dostatečnou mírou zabezpečení, mohou způsobit katastrofu. I proto se společnost rozhodla od druhé verze GPT opustit myšlenku otevřeného kódu (Šulc, 2024).

1.1.2 Klasifikace AI systémů

V rámci odborné literatury se umělá inteligence často dělí podle různých hledisek – např. podle způsobu učení (supervizované, nesupervizované), typu dat, s nimiž pracují, nebo podle oblasti nasazení. Jedním z nejčastějších a zároveň nejzásadnějších způsobů klasifikace je však dělení podle úrovně autonomie a kognitivních schopností, které daný systém vykazuje. V tomto kontextu se rozlišují tři základní kategorie:

Úzká AI

Úzká (nebo slabá) umělá inteligence označuje systémy, které jsou navrženy k plnění jednoho konkrétního úkolu, případně omezeného spektra úloh. Nejde o skutečnou „inteligence“ v lidském smyslu, ale o efektivní automatizaci rozhodovacích procesů na základě dat, často s využitím strojového učení. Mezi příklady patří rozpoznávání obrazu, hlasu, překlad jazyků, detekce podvodů, doporučovací algoritmy nebo chatboty. Tyto systémy nejsou schopny přenést naučené poznatky z jedné domény do jiné – např. systém, který hraje šachy, nebude umět hrát dámu. Přesto má úzká AI v současnosti největší praktické uplatnění, a to napříč obory od zdravotnictví po logistiku.

Obecná AI

Obecná umělá inteligence, také *AGI* – silná umělá inteligence (artificial general/strong intelligence) má co do schopnosti analyzovat, myslet, rozhodovat se a tvořit být podobná lidem nebo je překonat (Koukolík, 2024). V současnosti však AGI zatím zůstává pouze teoretickým konceptem, a ačkoli se k ní některé modely částečně blíží, stále chybí univerzální porozumění a autonomní generalizace lidského typu.

Superintelligence

Superintelligence je hypotetická forma AI, která by zásadně překonávala veškeré schopnosti lidského intelektu – včetně logického uvažování, kreativního myšlení, sociální inteligence i emocionálního porozumění. Takový systém by byl schopen řešit problémy, které jsou pro člověka neřešitelné, a vytvářet nové vědecké, technologické či společenské koncepty s minimálním časovým zpožděním.

Koncepce superintelligence je úzce spojena s představou technologické singularity – bodu, kdy by se AI mohla začít sama zdokonalovat bez lidského zásahu. I když jde o čistě teoretický scénář, mnozí odborníci varují před nutností připravit se na etické a bezpečnostní důsledky takového vývoje.

1.1.3 Aplikace metod umělé inteligence

Umělá inteligence nachází uplatnění v řadě odlišných oblastí lidské činnosti. Její schopnost analyzovat rozsáhlé datové sady, učit se z dat a optimalizovat rozhodovací procesy z ní činí klíčovou technologií 21. století. Níže jsou uvedeny vybrané oblasti, kde AI hraje významnou roli.

Business Intelligence

V oblasti Business Intelligence je AI standardem a firmy pomocí počítačových algoritmů běžně zjišťují trendy a poznatky z rozsáhlých databází. Umožňuje jim rychleji provádět kvalifikovaná rozhodnutí v reálném čase a udržovat konkurenceschopnost (Del Monte, 2019). Moderní podniky dnes využívají umělou inteligenci nejen pro základní reporting, ale především pro pokročilé analýzy trendů, predikci vývoje a odhalování skrytých souvislostí, které by byly pro analytika obtížně postřehnutelné. Pomocí strojového učení, zpracování přirozeného jazyka nebo algoritmů klasifikace a shlukování dokáží firmy získat praktické poznatky z různorodých zdrojů dat – od zákaznických interakcí přes finanční výsledky až po chování uživatelů na webu.

Průmysl

V oblasti průmyslu přispívá umělá inteligence ke zvyšování efektivity a flexibility výrobních procesů, zejména prostřednictvím prediktivní údržby strojů, automatické kontroly kvality a optimalizace logistických a dodavatelských řetězců. Díky strojovému učení lze například předvídat poruchy zařízení na základě historických dat ze senzorů, což umožňuje včasný zásah bez nutnosti přerušení výroby. V oblasti vizuální inspekce zase AI dokáže s vysokou přesností detekovat vady výrobků, které by lidské oko přehlédlo nebo posoudilo subjektivně.

Integrace AI do rámce Průmyslu 4.0 navíc otevírá prostor pro plně propojené a adaptivní výrobní prostředí, kde spolu jednotlivé komponenty komunikují v reálném čase. Chytré senzory, autonomní systémy a cloudová infrastruktura umožňují decentralizované řízení procesů, které se na základě aktuálních podmínek a predikcí optimalizují bez nutnosti lidského zásahu (Schwab, 2017).

Důležitým aspektem rozšíření AI v průmyslu není pouze samotná automatizace, ale také změna přístupu k práci a výrobě jako takové. Díky AI se výrobní podniky mohou transformovat z rigidních struktur na dynamické a udržitelné systémy schopné rychlé adaptace na měnící se tržní podmínky. Autor zároveň upozorňuje, že obavy z AI jako hrozby pro pracovní místa je třeba nahradit snahou o kvalifikované zapojení lidí do nových typů rolí, které kombinují technologii a lidskou expertizu (Minařík, 2024).

Chytrá města

V chytrých městech se umělá inteligence stává klíčovým nástrojem pro zvyšování efektivity, udržitelnosti a kvality života obyvatel. AI technologie jsou využívány například pro inteligentní řízení dopravy, kde algoritmy v reálném čase analyzují dopravní situaci a upravují světelné signály podle aktuálního zatížení silnic. Tím přispívají ke snižování dopravních kolon, emisí i spotřeby paliva. AI rovněž hraje důležitou roli při optimalizaci veřejné dopravy, kde predikuje poptávku v různých denních intervalech a pomáhá efektivněji plánovat trasy a intervaly spojů.

V oblasti energetiky slouží AI k monitorování a řízení spotřeby energie v městských budovách, veřejném osvětlení nebo chytrých domácnostech. Díky predikčním modelům je možné efektivněji využívat obnovitelné zdroje energie a předejít přetížení sítě. Podobné principy se uplatňují i při údržbě veřejné infrastruktury, kde AI identifikuje opotřebení nebo poruchy (např. výtluky na silnicích, poruchy veřejného osvětlení) a umožňuje cílené zásahy ještě před vznikem vážnějších problémů.

Významnou roli hraje AI také v oblasti bezpečnosti a monitoringu, kde podporuje rozpoznávání rizikových situací pomocí kamerových systémů, sledování kvality ovzduší nebo identifikaci nebezpečných událostí (např. požáry, úniky plynu). Koncept chytrého města není pouze o technologiích, ale o propojení dat, infrastruktury a lidí, které umožňuje reagovat pružněji a cíleněji na potřeby obyvatel (Batty, 2012).

Zdravotnictví

V medicínském prostředí hraje AI významnou roli již delší dobu, lze se s ní setkat při diagnostice nemocí, při vývoji nových léků, analýze obrazových dat, personalizaci léčby, prediktivní analýze vývoje zdravotního stavu pacienta nebo také při trénování robotů pro chirurgické úkony. AI systémy jsou schopny dosahovat vysoké úspěšnosti např. v rozpoznávání nádorů na radiologických snímcích (Topol, 2019).

Prakticky se AI využívá např. při analýze vyšetření na mamografu. Česká společnost Fomei nabízí přístroje a analytický nástroj pracující s umělou inteligencí Transpara. Podle jejích údajů dokážou až 28 % karcinomů identifikovat přesněji a rychleji, až 35 % intervalových karcinomů identifikovat dříve a až 70 % mamografických vyšetření spolehlivě označit jako negativní (Šulc, 2024).

Archeologie a kulturní dědictví

Umělá inteligence nachází čím dál širší uplatnění i v oblasti humanitních věd, kde přispívá k inovativním formám výzkumu, interpretace a dokumentace kulturní minulosti. V archeologii se AI používá například k rekonstrukci fragmentovaných artefaktů, kdy pomocí strojového učení dochází k automatickému doplňování chybějících částí keramiky, mozaik či nástěnných maleb. Systémy dokážou na základě geometrických a vizuálních charakteristik předpovědět původní podobu předmětu nebo spojit nalezené úlomky.

Dále se AI uplatňuje při rozpoznávání a čtení historických rukopisů a textů, včetně starověkých nebo středověkých písem, která jsou pro běžné OCR systémy obtížně interpretovatelná. Díky neuronovým sítím lze digitalizovat i značně poškozené či neúplné texty, čímž se zpřístupňují nové historické prameny pro badatelskou činnost.

Významná je také schopnost AI modelů predikovat výskyt archeologických nalezišť na základě geografických, geologických a historických dat – například pomocí analýzy leteckých snímků, satelitních dat nebo topografických modelů. Tyto predikce pomáhají zefektivnit terénní průzkumy a minimalizovat náklady spojené s výkopovými pracemi.

AI rovněž hraje stále větší roli v oblasti digitální dokumentace a ochrany kulturního dědictví. Například při tvorbě 3D modelů historických staveb, lokalit a předmětů, které mohou být dále využity pro výzkum, vzdělávání nebo virtuální rekonstrukce zaniklých památek. Tyto technologie pomáhají nejen při zachování historických objektů, ale také při jejich ochraně před poškozením, klimatickými vlivy či válečnými konflikty (Vincent et al., 2022).

Finance

Sektor financí sestává do velké míry z rutinních činností, nabízí tak široký prostor pro využití umělé inteligence. Po finanční krizi v roce 2008 došlo k velkému vzestupu start-upů, které díky AI dokázaly konkurovat zavedeným finančním institucím. Většina bankovních služeb se transformovala do mobilních zařízení a pobočky se staly spíše místem pro konzultace komplexních problémů. Umělá inteligence zde přebírá celou řadu procesů, např. v oblasti zákaznického servisu, kde již řada firem využívá tzv. Chatboty, kteří dokážou v reálném čase komunikovat se zákazníky pomocí textu nebo hlasu.

Dalším procesem využívajícím umělou inteligenci je vytváření rizikového profilu zákazníka. AI systémy dokážou vyhodnotit nejen tradiční finanční ukazatele, ale i alternativní data, jako jsou chování klienta na internetu nebo historie plateb, čímž vytvářejí komplexní obraz o žadateli o úvěr. Společnosti Oculus, Enova International nebo Zest AI nabízejí software, který umožňuje finančním institucím rychle a přesně vyhodnotit rizikovost klienta, a to i v případech, kdy má klient minimální nebo žádnou úvěrovou minulost. Zároveň tyto systémy napomáhají automatizované detekci podvodných žádostí, odhalování neobvyklého chování v transakcích a také při predikci tržních trendů, což je klíčové pro strategické rozhodování v dynamickém finančním prostředí (Šulc, 2024).

Vzdělávání

Umělá inteligence nachází široké uplatnění ve vzdělávání napříč různými úrovněmi školství i formami výuky. Mezi hlavní oblasti využití patří například personalizované vzdělávání, kde AI systémy přizpůsobí obsah, tempo a formu výuky individuálním potřebám studentů. Dále se využívá při automatizovaném hodnocení písemných prací, generování testových otázek, predikci studijního úspěchu nebo detekci studentů, který mohou potřebovat oporu. Významný posun přináší i inteligentní návodné systémy a chatboti, kteří studentům poskytují nepřetržitou zpětnou vazbu nebo odpovědi na časté dotazy.

AI rovněž usnadňuje práci učitelům – od administrativních úkolů až po návrh učebních plánů - a umožňuje analyzovat velké množství dat o studijním chování, což pomáhá při strategickém rozhodování na úrovni škol i vzdělávacích systémů. Vysoký potenciál má také využití generativní AI jako interaktivního nástroje pro podporu kreativity, porozumění textu nebo výkladu obtížných témat.

Zemědělství

Umělá inteligence v zemědělství se stává klíčovým nástrojem pro modernizaci a efektivní řízení agrárních procesů. AI pomáhá farmářům optimalizovat pěstování plodin, chov zvířat i správu půdy díky schopnosti analyzovat data z čidel, satelitních snímků a dronů. Pomocí algoritmů strojového učení lze například předpovídat výnosy, identifikovat škůdce a choroby rostlin, sledovat vývoj počasí či navrhovat přesné dávky hnojiv a zavlažování – tzv. precizní zemědělství.

AI se využívá také v automatizaci zemědělských strojů, například v autonomních traktorech nebo robotických sklízecích systémech, a rovněž při monitoringu zdraví hospodářských zvířat pomocí kamer a bio senzorů. AI může přispět k vyšší udržitelnosti, snížení nákladů a minimalizaci ekologických dopadů zemědělské produkce. Výzvou však zůstává přístup k technologiím pro malé a střední farmáře, a také otázky související s vlastnictvím dat a etickým využitím algoritmů (Mařík, 2024).

Právo

Umělá inteligence v právu přináší nové možnosti pro zefektivnění právních procesů, automatizaci rutinních úkolů a zpřístupnění právní pomoci širší veřejnosti. AI nástroje se dnes používají například pro automatizovanou analýzu smluv, předvídání soudních rozhodnutí, vyhledávání relevantních precedentů nebo třídění dokumentů. Velké jazykové modely nacházejí uplatnění při tvorbě právních shrnutí, generování návrhů právních textů nebo interpretaci paragrafového znění.

AI pomáhá i veřejné správě – například při třídění podnětů, odpovídání na žádosti občanů nebo digitalizaci soudních procesů. Významným směrem vývoje je také prediktivní justice, kde algoritmy analyzují historická rozhodnutí a pomáhají předvídat možné výsledky soudních sporů. Ve Spojených státech se již používá systém COMPAS, který předvídá, kdo pravděpodobně spáchá trestný čin (Coeckelbergh, 2020). Zároveň ale využívání AI v právu vyvolává řadu etických, právních a společenských otázek, především pokud jde o transparentnost rozhodování, odpovědnost za chyby a ochranu práv jednotlivce.

Je zásadní, aby systémy využívající AI byly v souladu s principy antidiskriminačního práva a ochrany osobních údajů podle GDPR, a aby bylo vždy možné zpětně ověřit, na základě jakých dat a kritérií bylo rozhodnutí učiněno (Lobotka, 2023).

Politika

Umělá inteligence přináší do politiky zásadní změny, které se mohou výrazně dotknout fungování demokratických systémů. Její vliv se projevuje v celé škále oblastí – od automatizace administrativních agend a predikce volebních trendů, přes monitorování veřejného mínění, až po využití AI při šíření informací (či dezinformací) prostřednictvím sociálních sítí. Zatímco někteří vnímají AI jako nástroj efektivnějšího vládnutí a transparentnější komunikace se společností, jiní upozorňují na rizika spojená s manipulací veřejností, ztrátou soukromí či oslabení lidského úsudku v rozhodovacích procesech.

Ačkoli jsou tyto dopady často vnímány ambivalentně, současný vývoj naznačuje, že AI může být rovněž nástrojem pozitivní transformace společnosti. Stávající vlna umělé inteligence může v důsledcích přinést pozitivní změnu. Postupně totiž i díky AI výrazně více lidí než doposud pochopí skutečnou podstatu demokracie, politiky, vztahů politiky a běžného života. A to v důsledcích má přinést růst zapojení občanů do celkového fungování, tedy onen kýžený lék na aktuální vlnu anomie neboli nedůvěry v systém (Šulc, 2024).

1.1.3 Systémy využívající umělou inteligenci

Umělá inteligence se v posledních letech stala klíčovým nástrojem pro automatizaci, kreativitu i podporu rozhodování napříč různými obory. Na trhu se objevuje stále více specializovaných AI nástrojů, které využívají pokročilé jazykové a multimodální modely k tomu, aby usnadnily tvorbu obsahu, překlad, analýzu dat, vývoj software nebo návrh virtuálních prostředí. Ačkoliv mají různé zaměření, spojuje je společný cíl – zefektivnit lidskou činnost prostřednictvím generativní umělé inteligence.

Konverzační a jazykové AI systémy

Jednou z globálních společností, které hrají a nejspíš i dál budou hrát klíčovou roli v oblasti umělé inteligence, je IBM. Systém Watson je komplexní řešení AI, které IBM nabízí klientům v segmentu B2B, tedy dalším firmám. V České republice využívá tento systém řada významných společností v oblasti e-commerce, ve finančnictví nebo ve zdravotnictví (Šulc, 2024).

Mezi další jazykové modely patří například Claude od společnosti Anthropic, který je navržen s důrazem na bezpečnost a transparentnost, nebo Grok, vyvíjený v rámci technologického portfolia X (dříve Twitter) pod vedením Elona Muska. Pro oblast překladů a vícejazyčné komunikace je hojně využíván DeepL, který vyniká vysokou kvalitou překladů a intuitivním rozhraním.

Tvorba a zpracování textu

Pro technické i kreativní úkoly existují nástroje, které využívají AI ke generování textů, psaní kódu či rozvíjení příběhů. Např. Lore Machine dokáže převádět textové zadání do scénářů, storyboardů a narativních struktur. Platforma AlphaCode.ai (vyvinutá DeepMindem) je pak zaměřena na generování programátorského kódu a řešení algoritmických úloh, a může sloužit jako asistent pro vývojáře.

Generativní vizuální a multimediální nástroje

Zvláštní pozornost si zaslouží také generativní nástroje, které pracují nejen s textem, ale i s obrazem, videem a zvukem. Mezi ně patří např. Sora, nástroj pro generování realistických videí na základě textového popisu, Adobe AI Assistant, integrované AI funkce v rámci Adobe Creative Cloud pro úpravu grafiky a multimédií, Canva.ai, rozšíření platformy Canva o generativní design a vizuální návrhy, Synthesia, platforma pro tvorbu AI videí s generovanými moderátory a dabingem, nebo Descript, nástroj pro editaci videa a podcastů s funkcí přepisu, syntézy hlasu a inteligentních úprav. Tyto nástroje jsou hojně využívány v marketingu, vzdělávání nebo v médiích, kde umožňují rychlou a nákladově efektivní produkci kvalitního obsahu.

3D a virtuální světy

Většina dosud popsaných AI modelů je určena převážně pro jednotlivce, Promethean AI je ale přímo nadesignovaný pro práci v týmech. Důvodem je cílová skupina uživatelů, kterou tvoří vývojáři počítačových her, architekti nebo designeři, a také jeho kompatibilita s 3D editory. Zejména s nástupem náhlavních zařízení typu Meta Quest či Vision Pro je očekáván nárůst počtu vývojářů herních nebo edukačních aplikací, pro něž bude Promethean AI významným pomocníkem při vytváření virtuálního obsahu (Šulc, 2024).

1.1.4 Large Language Models

Velké jazykové modely (Large Language Models – LLM) představují jádro dnešních pokročilých generativních AI systémů. Tyto modely, typicky postavené na architektuře transformer, jsou trénovány na obrovském množství textových dat z internetu, knih, článků, konverzací a dalších zdrojů. Díky tomuto rozsáhlému tréninku jsou schopny rozumět lidskému jazyku a generovat konkrétní výstupy v přirozené formě, ať už jde o odpovědi na otázky, shrnutí textu, překlad, návrh kódu, kreativní psaní či analýzu dat.

LLM fungují na principu predikce dalšího slova na základě předchozího kontextu, přičemž zohledňují nejen gramatiku, ale i sémantiku, styl a intenci zadání. Tím vzniká iluze

„porozumění“ a interakce, která umožňuje využití těchto modelů v celé řadě oblastí – od vzdělávání přes marketing až po vývoj software.

Mezi nejznámější zástupce LLM patří GPT (Generative Pre-trained Transformer) od společnosti OpenAI, PaLM od Googlu, LLaMA od Mety nebo Claude od společnosti Anthropic. Všechny tyto modely pracují s miliardami parametrů a dokážou zpracovávat nejen text, ale i další formy vstupů (např. obrázky nebo tabulky), čímž se posouvají směrem ke multimodální AI.

LLM však nejsou bez rizik – vyvstávají otázky ohledně biasu (zkreslení), pravdivosti generovaných informací, ochrany osobních údajů a vysvětlitelnosti rozhodnutí. I přesto zůstávají jednou z nejdynamičtěji se rozvíjejících oblastí současné umělé inteligence.

1.2 Bezpečnost dat

V dnešní době je bezpečnost dat klíčovým tématem ve firemním prostředí. Data představují nejen cenný zdroj pro rozhodování a optimalizaci procesů, ale také potenciální riziko při jejich špatném zabezpečení. Firemní data jsou často terčem kybernetických útoků, jejichž počet každoročně roste. Podle českého Národního úřadu pro kybernetickou a informační bezpečnost (NÚKIB) v letech 2022 a 2023 z hlediska klasifikace incidentů nejvíce převažovaly incidenty cílící na dostupnost (např. webových stránek či služeb). Dalšími kategoriemi pak byl průnik (např. kompromitace aplikace nebo uživatelského účtu), zásah do informační bezpečnosti (např. neautorizovaný přístup k datům), šíření škodlivého kódu a phishing. (NÚKIB, 2023). Ochrana dat je proto nejen technickou výzvou, ale také právní a etickou povinností organizací. V českém prostředí je důraz kladen zejména na plnění požadavků nařízení GDPR, které definuje pravidla pro zpracování a ochranu osobních údajů (Šmídová, 2020). S rostoucím využitím umělé inteligence (AI) v oblasti analýzy a práce s daty však vyvstávají nové výzvy. AI nástroje mohou na jedné straně přispět k efektivnějšímu zpracování dat, na druhé straně mohou představovat riziko spojené například s neadekvátním nakládáním s osobními údaji (Kučera, 2021). Tato práce se proto zaměřuje na identifikaci bezpečnostních rizik a návrh praktických řešení pro implementaci AI v souladu s principy ochrany dat.

1.2.1 Bezpečnostní rizika

Nasazení umělé inteligence do firemního prostředí přináší kromě řady výhod také nová bezpečnostní rizika, která se odlišují od tradičních hrozeb v oblasti informační bezpečnosti. AI systémy pracují s rozsáhlými datovými sadami, autonomně rozhodují a učí se ze vzorců

chování, což s sebou nese specifické výzvy. Mezi hlavní rizikové faktory patří možnost zneužití dat, úniky citlivých informací, manipulace s výstupy analýz, nedostatečná transparentnost rozhodování (tzv. black-box efekt) nebo zranitelnost vůči specifickým útokům, jako jsou adversarial attacks. Tyto hrozby mohou vést nejen ke ztrátě důvěry zákazníků, ale i k právním postihům a finančním ztrátám v důsledku porušení pravidel ochrany osobních údajů.

S ohledem na rostoucí využití AI v citlivých oblastech – jako jsou zdravotnictví, finance, personalistika nebo veřejná správa – je klíčové, aby organizace rizika včas identifikovaly a implementovaly preventivní i reaktivní bezpečnostní opatření. Následující podkapitoly se věnují nejvýznamnějším hrozbám, které nasazení AI ve firemní praxi přináší, a ilustrují, proč je ochrana dat v kontextu umělé inteligence složitější a strategicky důležitější než kdy dříve.

1.2.2 Zneužití dat

Zneužití dat představuje jedno z nejzávažnějších rizik v oblasti informační bezpečnosti, zejména v kontextu umělé inteligence a automatizovaného zpracování informací. Ke zneužití dochází, když jsou citlivé nebo osobní údaje použity neoprávněným způsobem, např. ke sledování uživatelů, podvodům nebo kybernetickým útokům (Tvrdíková, 2008). Firmy často pracují s velkými objemy dat a neimplementují dostatečně silná opatření k ochraně dat, což umožňuje jejich zneužití nejen externími útočníky, ale i zaměstnanci či obchodními partnery.

V kontextu umělé inteligence hrozí například zneužití trénovacích dat, kdy systémy AI mohou neúmyslně odhalit citlivé informace obsažené v datech, na nichž byly trénovány (Brown et al., 2020). Dalším nebezpečím jsou tzv. data poisoning attacks, při nichž útočníci vkládají škodlivé nebo zavádějící informace do trénovacích dat, čímž ovlivňují výsledky analýzy AI a mohou manipulovat s rozhodováním modelu (Goodfellow et al., 2015). Proto je zásadní, aby organizace zaváděly striktní bezpečnostní politiky, šifrování dat, přísnou kontrolu přístupu a pravidelné audity, které minimalizují riziko neoprávněného přístupu a manipulace s informacemi.

1.2.3 Chybné nebo záměrně pozměněné analýzy

Umělá inteligence je stále častěji využívána k analýze rozsáhlých datových sad a podpoře rozhodovacích procesů ve firmách i veřejném sektoru. Nesprávná interpretace dat nebo manipulace s analytickými výstupy však může vést k chybným rozhodnutím, která mohou mít finanční či právní důsledky pro organizaci.

AI modely jsou závislé na kvalitě a reprezentativnosti trénovacích dat. Pokud model pracuje s nekompletními, zastaralými nebo zkreslenými daty, může dojít ke špatné interpretaci výsledků. Například v oblasti finančních analýz může AI špatně predikovat tržní trendy kvůli nedostatečnému zohlednění makroekonomických faktorů (O'Neil, 2016).

Dalším problémem je overfitting, kdy model "přetrénuje" na specifických vzorcích a ztrácí schopnost zobecnění na nová data. V důsledku toho mohou být výstupy modelu nepřesné nebo nerelevantní pro reálné obchodní prostředí (Goodfellow et al., 2016).

V některých případech mohou být analytické výstupy záměrně pozměněny nebo zkresleny, aby odpovídaly specifickým zájmům stakeholderů. Např.:

- Firemní management může využít AI modely k podpoře rozhodnutí, která jsou výhodná pouze pro určitou skupinu investorů či vedení společnosti, ačkoliv mohou být v dlouhodobém horizontu nevýhodná pro firmu jako celek.
- Marketingové analýzy mohou být upraveny tak, aby uměle zveličovaly úspěšnost určité strategie a ospravedlnily další investice do ní.
- Politická rozhodnutí nebo veřejné politiky mohou být ovlivněny manipulovanými datovými analýzami, které podporují konkrétní narativ nebo ideologii (Bostrom, 2014).

Chybné nebo záměrně zkreslené analýzy mohou vést k finančním ztrátám, právním problémům i narušení důvěry v AI systémy. Z těchto důvodů je klíčové pravidelně auditovat AI modely a validovat jejich výstupy na různých datech (Samek et al., 2019), zajistit transparentnost v rozhodovacích procesech využívajících AI, např. pomocí Explainable AI (XAI) technik, a používat etické principy při vývoji AI a zavádět mechanismy pro detekci manipulativních praktik. Díky těmto opatřením lze minimalizovat rizika spojená s chybnými nebo záměrně zkreslenými analýzami a zajistit, že AI bude sloužit objektivnímu rozhodování na základě reálných dat.

1.2.4 Únik dat

Jedním z klíčových bezpečnostních rizik při využívání umělé inteligence je únik dat, který může nastat jak neúmyslně, tak v důsledku cíleného útoku. AI systémy, zejména ty založené na hlubokém učení a rozsáhlých jazykových modelech, jsou trénovány na obrovských množstvích dat, která mohou obsahovat citlivé informace. Pokud nejsou implementována

vhodná omezení přístupu k trénovacím datům, může AI neúmyslně generovat výstupy obsahující osobní nebo důvěrné údaje (Brown et al., 2020).

Jazykové modely, jako je například GPT, se učí na základě vzorů v trénovacích datech a mohou tak v odpovědích rekonstruovat části těchto dat. Pokud trénovací dataset obsahuje osobní údaje, finanční informace nebo interní firemní komunikaci, model může tyto údaje nedopatřením vygenerovat při interakci s uživateli. Například studie Brown et al. (2020) ukázala, že při dostatečném množství dotazů lze z jazykových modelů extrahovat informace, které byly přítomny v jejich trénovací sadě, včetně jmen, adres či dokonce hesel.

AI systémy mohou být rovněž cílem útoků zaměřených na extrakci informací. Techniky, jako je membership inference attack nebo model inversion attack, umožňují útočnickům získat citlivé údaje z trénovacího datasetu tím, že analyzují chování modelu při zpracování vstupních dotazů (Shokri et al., 2017). Tyto útoky jsou obzvláště nebezpečné v prostředích, kde AI modely pracují s citlivými zdravotnickými, finančními nebo biometrickými údaji, protože by mohly útočnickům poskytnout přístup k osobním informacím uživatelů.

Pro minimalizaci rizika úniku dat je nezbytné implementovat robustní bezpečnostní opatření, jako je diferenciální soukromí, které omezuje schopnost modelu zapamatovat si konkrétní detaily z trénovacích dat (Dwork et al., 2006). Důležitou roli také hraje dodržování legislativních předpisů, jako je GDPR, které stanovuje pravidla pro zpracování osobních údajů a ukládá povinnost zajistit dostatečnou ochranu dat před jejich neoprávněným zpracováním (Evropská komise, 2016).

Vzhledem k tomu, že AI systémy se stále více integrují do firemního i veřejného sektoru, je prevence úniku dat klíčová pro ochranu soukromí uživatelů, firemních tajemství a důvěrných informací. Správná implementace bezpečnostních opatření a etické využívání AI tak hrají zásadní roli v budoucím vývoji těchto technologií.

1.2.5 Black-box efekt a nedostatečná transparentnost

AI modely, zejména hluboké neuronové sítě, jsou známé svou "black-box" povahou, což znamená, že jejich rozhodovací procesy nejsou dostatečně srozumitelné v tom, jak dospěly k určitému výsledku. Jejich rozhodování nemusí být zcela srozumitelné ani pro jejich tvůrce. Modely s vysokou komplexitou (např. GPT nebo DALL-E) jsou obtížně auditovatelné a mohou obsahovat zkreslení, která vedou k neetickým rozhodnutím (Samek et al., 2019). Tento nedostatek transparentnosti přináší problémy s odpovědností, protože pokud AI učiní chybný závěr, není vždy jasné, kdo za něj nese odpovědnost – vývojáři, organizace nebo

uživatelé? (Bostrom, 2014). Navíc nedostatečná vysvětlitelnost modelů komplikuje jejich regulaci, protože chybí efektivní kontrolní mechanismy a právní rámec pro hodnocení jejich rozhodování (NÚKIB, 2023). Z těchto důvodů se zvyšuje důležitost vysvětlitelné umělé inteligence (Explainable AI, XAI), která se snaží zajistit, aby rozhodování modelů bylo srozumitelné a auditovatelné, čímž by se předešlo potenciálním etickým a právním problémům (Samek et al., 2019).

1.2.6 Zranitelnost AI systémů

AI systémy jsou zranitelné vůči specifickým typům útoků, které mohou manipulovat jejich výsledky. Metoda Adversarial attacks spočívá ve vkládání drobných, pro lidi nepostřehnutelných změn do vstupních dat, které vedou AI model k chybnému rozhodnutí. Například změněný pixel u obrázku může způsobit, že model nesprávně identifikuje objekt (Goodfellow et al., 2015). Útoky mohou ovlivnit rozhodování AI systémů, např. podvádění systémů pro analýzu transakcí za účelem praní špinavých peněz. Zranitelné AI systémy mohou být zneužity např. pro DDoS útoky, pokud nejsou implementovány odpovídající ochranné mechanismy (NÚKIB, 2023).

1.2.7 Etické otázky

Rozvoj umělé inteligence přináší nejen nové technologické možnosti, ale také celou řadu etických dilemat. Mezi takovéto otázky patří např.:

Měla by být v autonomních vozidlech zabudována etická omezení, a jestliže ano, jaká a jak by měla být stanovena (Coeckelbergh, 2020)?

Jak zabránit zneužití AI k manipulaci s veřejným míněním nebo k masovému sledování?

Měla by AI mít přístup k citlivým údajům, jako jsou zdravotní záznamy nebo biometrická data?

Měly by se povolovat autonomní smrtící zbraně (Coeckelbergh, 2020)?

Kdo nese odpovědnost za rozhodnutí umělé inteligence?

Jak zajistit, aby algoritmy nebyly diskriminační nebo předpojaté?

Využití umělé inteligence otevírá nové technologické možnosti, ale zároveň přináší zásadní etické výzvy, které se týkají jak vývojářů a provozovatelů systémů, tak koncových uživatelů a společnosti jako celku. Vývoj AI by měl probíhat v souladu s hodnotami jako jsou lidská důstojnost, rovnost, svoboda a odpovědnost, přičemž důraz by měl být kladen na prevenci

škod a maximalizaci prospěchu pro jednotlivce i společnost. V této souvislosti se často zmiňuje několik klíčových principů:

Transparentnost

Transparentnost je jedním ze základních principů etického nasazení umělé inteligence. Znamená, že AI systémy by měly být srozumitelné, předvídatelné a vysvětlitelné nejen pro vývojáře, ale i pro koncové uživatele a regulátory. Uživatelé mají právo vědět, jakým způsobem AI dochází ke svým závěrům, a zda jsou tyto závěry založené na ověřitelných a spravedlivých datech. Transparentnost se úzce pojí s pojmem *vysvětlitelná AI* (explainable AI, XAI), který označuje přístup k návrhu systému, jejichž rozhodnutí lze zpětně objasnit a obhájit. Tento aspekt je zvláště důležitý v oblastech s přímým dopadem na lidské životy, jako je zdravotnictví, finance, soudnictví nebo nábor zaměstnanců. Např. pokud AI rozhodne o zamítnutí žádosti o úvěr nebo o výběru uchazeče o práci, je nezbytné, aby bylo možné vysvětlit, proč k takovému rozhodnutí došlo a na základě jakých kritérií (Coeckelbergh, 2020).

Nedostatek transparentnosti – typický např. pro hluboké neuronové sítě fungující jako tzv. černé skříňky – představuje zásadní riziko nejen pro důvěru veřejnosti, ale i pro možnost zpětné kontroly, právní odpovědnost a ochranu proti diskriminaci. Transparentní systémy navíc usnadňují odhalování chyb, zajišťují vyšší míru bezpečnosti a podporují odpovědný vývoj technologií.

Transparentnost by proto měla být zajištěna nejen technickými prostředky (např. logováním, dokumentací modelů, vizualizací rozhodovacích procesů), ale také jasnou komunikací směrem k uživatelům a veřejnosti, včetně srozumitelných pravidel pro použití AI a vysvětlení jejich omezení.

Odpovědnost

Princip odpovědnosti (*accountability*) představuje zásadní etický i právní požadavek při vývoji a nasazení systémů umělé inteligence. Pokud AI systém způsobí škodu – např. učiní nespravedlivé rozhodnutí při náborovém řízení, chybně doporučí léčbu ve zdravotnictví nebo provede diskriminační klasifikaci – musí být možné určit, kdo nese za danou situaci odpovědnost: vývojář algoritmu, jeho provozovatel, organizace, která jej nasadila, nebo jiný subjekt.

Tato otázka je obzvláště komplikovaná u autonomních systémů, které se rozhodují samostatně na základě natrénovaných modelů a datových vstupů. U klasického softwaru je odpovědnost

obvykle přiřaditelná vývojáři nebo provozovateli, avšak u komplexních AI systémů, zejména těch, které se průběžně učí (*self-learning*), může být kauzální vazba mezi rozhodnutím a lidským činitelem nejasná.

V právní praxi se tak stále častěji hovoří o potřebě nového rámce právní odpovědnosti pro systémy AI. Evropská unie např. v návrhu AI Act (2021) i v připravované Směrnici o odpovědnosti za škodu způsobenou AI (AI Liability Directive) řeší otázky přenesené odpovědnosti, transparentnosti rozhodnutí a možnosti právního napadení výstupu AI. Právo na vysvětlení rozhodnutí automatizovaných systémů je zakotveno již v článku 22 nařízení GDPR, který dává jednotlivci právo nebýt předmětem rozhodnutí založeného výhradně na automatizovaném zpracování, včetně profilování.

Zajištění odpovědnosti v kontextu AI tak zahrnuje nejen technická a organizační opatření (např. logování rozhodovacích procesů, auditovatelnost systémů), ale také jasné smluvní a právní vymezení rolí a odpovědností jednotlivých aktérů.

Spravedlnost

Jedním z klíčových etických principů při vývoji a využívání systémů umělé inteligence je spravedlnost (*fairness*), která vyžaduje, aby rozhodování AI bylo nediskriminační, nestranné a rovné vůči všem uživatelům. AI by neměla preferovat ani znevýhodňovat žádnou skupinu na základě věku, pohlaví, rasy, etnického původu, zdravotního stavu, náboženského vyznání, sexuální orientace nebo socioekonomického postavení.

K porušení tohoto principu dochází často nepřímo – prostřednictvím tzv. algoritmické zaujatosti (*bias*), která může být přítomna ve vstupních datech, v samotném modelu nebo v jeho implementaci. Pokud jsou například trénovací data zatížena historickými předsudky (např. genderově nevyvážené životopisy nebo rasově podbarvené záznamy z trestní databáze), může model tato zkreslení zautomatizovat a rozšiřovat, aniž by si toho uživatelé byli vědomi. Výsledkem mohou být rozhodnutí, která systematicky znevýhodňují určité skupiny – typicky v náborových systémech, přidělování úvěrů nebo prediktivní justici.

Zajištění spravedlnosti vyžaduje důslednou kontrolu kvality a reprezentativnosti trénovacích dat, pravidelné testování výstupů AI na známky diskriminace a přijetí opatření, která vedou k vyváženému rozhodování (např. tzv. *fairness-aware machine learning*). Dále je třeba zajistit možnost lidského dohledu a zásahu v případě podezření na nespravedlnost, včetně transparentního vysvětlení, jak rozhodnutí vzniklo.

Spravedlnost je také jedním z principů formulovaných v rámci etických doporučení Evropské komise (Ethics Guidelines for Trustworthy AI, 2019), která požaduje, aby vývojáři a provozovatelé AI zohledňovali různé formy zranitelnosti uživatelů a dbali na rovnost přístupu a rovné zacházení.

Soukromí

Ochrana soukromí a osobních údajů patří k nejcitlivějším a zároveň nejvíce diskutovaným tématům v souvislosti s využíváním umělé inteligence. AI systémy často pracují s velkým množstvím citlivých dat – od biometrických údajů a zdravotních záznamů, přes údaje o chování uživatelů na internetu až po lokalizační informace. Zpracování těchto dat v sobě nese riziko zneužití, neautorizovaného sdílení, úniku nebo neúmyslného odhalení informací, což může mít závažné dopady na jednotlivce i organizace.

V rámci Evropské unie je ochrana soukromí regulována zejména Obecným nařízením o ochraně osobních údajů (GDPR), které stanovuje konkrétní práva a povinnosti týkající se sběru, uchovávání a zpracování dat. Mezi základní principy patří:

- právo na informovanost – uživatelé musí být jasně informováni o tom, jaká data jsou shromažďována a za jakým účelem,
- právo na přístup k údajům – každý má právo zjistit, jaké údaje o něm systém eviduje,
- právo na výmaz („právo být zapomenut“) – možnost požádat o odstranění osobních údajů ze systému,
- právo vznést námitku proti automatizovanému rozhodování, zejména pokud má významný právní nebo obdobný dopad.

AI systémy musí být navrženy a provozovány tak, aby tyto principy plně respektovaly. To zahrnuje např. minimalizaci dat (shromažďování pouze nezbytných údajů), anonymizaci nebo pseudonymizaci dat, šifrování, nebo možnost dát či odvolat souhlas se zpracováním dat. Důležitým nástrojem je také tzv. privacy by design, tedy začlenění principů ochrany soukromí již ve fázi návrhu systému.

Nezodpovědné nakládání s daty může vést nejen k narušení důvěry uživatelů, ale i k vážným právním následkům, včetně vysokých pokut za porušení pravidel GDPR. Eticky navržené AI systémy by proto měly klást důraz na transparentní a bezpečné zpracování dat, které je v souladu s legislativními i morálními požadavky.

Udržitelnost

V diskusi o etickém využívání umělé inteligence začíná stále významnější roli hrát také princip udržitelnosti. Vývoj a provoz rozsáhlých AI modelů – zejména v oblasti hlubokého učení (deep learning) – je spojen s vysokou energetickou náročností, která má přímý dopad na životní prostředí. Trénování moderních jazykových modelů nebo neuronových sítí často vyžaduje obrovské výpočetní kapacity, jejichž spotřeba energie vede ke značné uhlíkové stopě.

Eticky odpovědné využívání AI by proto mělo zahrnovat i snahu o minimalizaci environmentálních dopadů. To může zahrnovat:

- optimalizaci algoritmů a výpočetních procesů,
- preferenci energeticky efektivních datových center,
- využívání obnovitelných zdrojů energie,
- nebo lokální výpočty (tzv. edge computing), které snižují potřebu rozsáhlých cloudových infrastruktur a datových přenosů.

Rovněž je důležité zohlednit poměr mezi náklady a přínosy konkrétního modelu – tedy zda výpočetní náročnost odpovídá jeho praktickému využití a společenské hodnotě. Udržitelnost se tak stává součástí širšího konceptu „odpovědné AI“ (responsible AI), která usiluje o vyváženost mezi technologickým pokrokem, ochranou práv a svobod uživatelů a ochranou přírodních zdrojů.

Prospěšnost

Umělá inteligence by měla být vyvíjena a používána tak, aby sloužila člověku a společnosti jako celku – tedy posilovala lidské schopnosti, podporovala vzdělávání, zdravotní péči, ochranu životního prostředí a kvalitu života. V souladu s tímto principem by se mělo zabránit vývoji AI, která by mohla být použita k represím, manipulaci, diskriminaci nebo zneužívání moci.

Singularita

Technologická singularita označuje hypotetický bod ve vývoji umělé inteligence, kdy by AI dosáhla nebo překonala úroveň lidské obecné inteligence (AGI) a následně by se začala autonomně zlepšovat exponenciálním tempem bez lidského zásahu. Tento proces by mohl znamenat zásadní zlom ve vývoji civilizace, ale zároveň přináší hluboké etické a bezpečnostní obavy.

Podle filozofa a odborníka na rizika z AI Nicka Bostroma může být singularita buď největším přínosem nebo největším nebezpečím v dějinách lidstva – v závislosti na tom, zda bude taková superinteligence řízena v souladu s lidskými hodnotami (Bostrom, 2014).

Z tohoto důvodu se stále častěji hovoří o nutnosti výzkumu tzv. *AI alignmentu* – tedy metod, jak zajistit, že budoucí inteligentní systémy budou jednat v souladu s lidskými zájmy a hodnotami. Singularita tak není pouze technologickou otázkou, ale hluboce filozofickou a morální výzvou, která vyžaduje interdisciplinární přístup zahrnující vědu, etiku, právo i politiku.

1.2.8 Zneužití AI

Umělá inteligence přináší řadu výhod, jako je zefektivnění procesů, analýza rozsáhlých dat nebo automatizace opakujících se úloh. Zároveň však otevírá prostor pro její zneužití, a to jak ze strany jednotlivců, tak i organizací či státních aktérů. Jedním z nejviditelnějších rizik je manipulace s veřejným míněním, např. prostřednictvím deepfake videí, která umožňují realisticky zfalšovat výpovědi veřejně činných osob, nebo pomocí automatizovaných definformačních kampaní šířených na sociálních sítích, jejichž cílem může být ovlivnění voleb či polarizace společnosti (Chesney & Citron, 2019).

Dalším závažným rizikem je využití AI pro sofistikované kybernetické útoky. Technologie strojového učení může být zneužita k automatickému generování škodlivého kódu, obcházení bezpečnostních mechanismů nebo vytváření falešných digitálních identit, které napodobují chování skutečných uživatelů a je obtížné je odhalit tradičními bezpečnostními nástroji (Brundage, 2018). Tyto nástroje mohou být využity k útokům na kritickou infrastrukturu, firemní systémy nebo k cílenému sledování osob.

Znepokojující je také možnost diskriminace na základě zkreslených nebo nevyvážených trénovacích dat, což může vést k nespravedlivému zacházení v oblastech jako je zaměstnávání, poskytování půjček nebo pojištění. AI modely, které nejsou správně testovány na přítomnost tzv. *bias* (zkreslení), mohou neúmyslně znevýhodňovat určité skupiny obyvatel (např. na základě pohlaví, rasy či věku), čímž vzniká eticky i právně problematické prostředí (O’Neil, 2016).

Zneužití AI ve zpravodajských službách představuje také významné riziko, a to jak pro národní bezpečnost, tak pro ochranu soukromí jednotlivců. AI umožňuje analýzu obrovského množství dat v reálném čase, což může vést k neoprávněnému sledování obyvatel a porušování lidských práv (Zuboff, 2019). Státní i nestátní aktéři mohou využívat AI k

vytváření dezinformačních kampaní a ovlivňování politických procesů. Zpravodajské agentury mohou využívat AI k provádění sofistikovaných kybernetických operací, vč. narušování infrastruktury jiných států nebo krádeže citlivých informací (Brundage, 2018). Pokročilé IA systémy mohou být nasazeny k provádění cílů bez lidského dohledu, což zvyšuje riziko neetického nebo nekontrolovatelného nasazení vojenské síly. OSN definuje systémy smrtících autonomních zbraní (AWS) jako zbraňové systémy, které nacházejí, vybírají a likvidují lidské cíle bez lidské intervence. AWS se důvodně popisují jako „třetí revoluce ve vedení války” – po střelném prachu a jaderných zbraních (Russell, 2021).

1.2.9 Nástroje a technologie pro zajištění bezpečnosti dat

Vzhledem k rostoucímu množství dat, která společnosti zpracovávají, vč. citlivých osobních údajů, obchodních informací či interních strategií, je klíčové zajistit jejich dostatečnou ochranu proti útokům, zneužití a neoprávněnému přístupu.

Šifrování dat.

Šifrování dat je základní technika ochrany dat při jejich přenosu a ukládání. Umožňuje, aby informace byly přístupné pouze oprávněným osobám. Používají se různé typy algoritmů – symetrické (např. AES), asymetrické (např. RSA) nebo hybridní. Od použitého algoritmu (a délky klíče) se odvíjí síla daného algoritmu. Za bezpečný algoritmus se obecně považuje takový algoritmus, jeho popis byl zveřejněn a podroben zkoumání bezpečnostních expertů a není známá jeho slabina, které by se dalo zneužít (Šulc, 2018).

Systémy pro detekci a prevenci narušení (IDS/IPS).

IDS (Intrusion Detection Systém) a IPS (Intrusion Prevention Systém) monitorují síťový provoz a hledají podezřelé aktivity, které by mohly signalizovat útok. V prostředí s AI modely mohou tyto systémy detekovat neobvyklé chování modelu, např. při pokusu o adversarial attack nebo neoprávněný přístup k API.

Mechanismy pro správu oprávnění a přístupu

RBAC (Role-Base Access Control) je metoda řízení přístupu, která přiděluje oprávnění na základě role uživatele v systému. V kontextu AI systémů se využívá pro řízení přístupu k datům, modelům i výpočetním zdrojům. Moderní systémy často kombinují RBAC s kontextově citlivými politikami, jako jsou ABAC nebo Policy-as-Code.

Specifické nástroje pro bezpečnost AI (např. metody detekce útoků na modely strojového učení)

S rostoucím využitím AI narůstá i potřeba specializovaných nástrojů pro ochranu modelů a dat, na kterých AI pracuje. Patří sem:

- Detekce adversarial attacks (např. pomocí metod jako Feature Squeezing, adversarial training)
- Ochrana proti modelu stealing
- Auditovatelnost a interpretovatelnost modelů (např. LIME, SHAP)
- Monitoring neobvyklých výstupů modelu v reálném čase

1.3 Právní rámec

S rostoucím významem a rozšířením umělé inteligence vyvstává potřeba její regulace, která má zajistit bezpečné, etické a transparentní využívání těchto technologií. Na úrovni Evropské unie vzniklo Nařízení o umělé inteligenci (AI Act), které je pravděpodobně prvním komplexním právním rámcem pro AI. Act AI zavádí rizikově založený přístup, kdy jsou systémy rozděleny podle úrovně rizika – od minimálního až po nepřijatelné, přičemž ty s vysokým rizikem (např. ve zdravotnictví, justici nebo vzdělávání) podléhají přísnějším pravidlům, včetně požadavků na transparentnost, auditovatelnost a lidský dohled.

Důležitým úkolem nové evropské legislativy v oblasti Evropské unie je ochrana práv a svobod občanů, podobně jako je tomu například v Obecném nařízení o ochraně osobních údajů (GDPR), zejména v souvislosti s rostoucím nasazením AI ve veřejném i soukromém sektoru. Podobně jako GDPR chrání soukromí jednotlivců v digitálním prostoru, má i AI Act za cíl zajistit, aby AI systémy nebyly zneužívány k diskriminaci, neoprávněnému sledování či omezování svobody projevu. Umělá inteligence má potenciál výrazně ovlivňovat rovné zacházení, a proto je nutné na ni již od počátku aplikovat principy prevence diskriminace“. Tento pohled podporuje vznik právního rámce, který nejen technicky reguluje AI, ale zároveň chrání jednotlivce před automatizovanými rozhodnutími s potenciálně negativními dopady na jejich důstojnost, rovnost a soukromí (Lobotka, 2024).

V České republice se AI řeší v rámci Národní strategie umělé inteligence, která byla schválena v roce 2019. Jejím cílem je vytvořit prostředí pro etický rozvoj a využití AI technologií, podpořit výzkum, vzdělávání a inovace a zároveň připravit veřejnou správu i soukromý sektor na dopady automatizace a digitalizace. Hlavními cíli AI Act jsou zejména:

1. řešení rizik specificky spojených s aplikacemi AI
2. zákaz praktik AI, které představují nepřijatelná rizika
3. vytvoření seznamu vysoce rizikových aplikací AI
4. stanovení požadavků na systémy AI pro vysoce rizikové aplikace
5. definování konkrétních povinností poskytovatelů vysoce rizikových aplikací AI
6. vyžadování posouzení shody před uvedením daného systému AI do provozu nebo uvedením na trh
7. zavedení vymáhání povinností týkajících se AI po uvedení daného systému na trh
8. vytvoření struktury řízení v oblasti AI na evropské úrovni

V únoru loňského roku (2024) byl, jako součást Evropské komise, zřízen Evropský úřad pro umělou inteligenci (The European AI Office). Ten by měl dohlížet na implementaci Nařízení o umělé inteligenci v jednotlivých členských zemích Evropské unie a na jeho řádné vymáhání (Štědroň, 2024). Podle oficiálních stránek Evropské unie má být jeho cílem vytvoření prostředí, kde technologie umělé inteligence budou respektovat lidskou důstojnost, lidská práva a důvěru lidí (European Commission, 2024).

Evropská komise také loni spustila Inovační balíček pro umělou inteligenci (AI Innovation Package) na podporu malých a středních podniků při vývoji a zavádění důvěryhodné a etické umělé inteligence, která je v souladu s hodnotami a pravidly EU. Tento balíček zahrnuje finanční, legislativní i infrastrukturní nástroje, které mají pomoci firmám využívat výhody generativní AI, aniž by přitom ohrožovaly bezpečnost nebo soukromí občanů. Součástí tohoto balíčku je iniciativa GenAI4EU, která podporuje praktické zavádění generativní AI ve strategických oblastech, jako jsou zdravotnictví, vzdělávání, energetika nebo veřejná správa. Cílem této iniciativy je propojit výzkumné organizace, startupy a průmyslové partnery a podpořit tak inovace v rámci evropského ekosystému AI.

1.4 Využití AI v oblasti datové analýzy

Umělá inteligence hraje v datové analýze stále významnější roli – od základního zpracování dat až po jejich automatickou interpretaci a vizualizaci. Mezi základní typy analytických úloh patří deskriptivní (popisné) a prediktivní (předpovědní) modelování.

Deskriptivní modely přispívají k porozumění základním procesům, dominantním strukturám nebo vazbám v modelovaných datech. Cílem je popsat analyzovanou doménu, popř. proces, který vedl k vygenerování analyzovaných dat. Doplnkovou třídou úloh se pak stává hledání struktur a pravidel (označované také jako hledání nugetů), které se zaměřuje na identifikaci a

popis příkladů specifických v dané doméně a neklade si nárok na úplnost popisu (Hand, 2001). Při řešení deskriptivních úloh postupujeme podobně jako u úloh prediktivních. Používáme sice často odlišné algoritmy pro tvorbu modelů, ale podobně jako u úloh prediktivních řešíme otázky hodnotící funkce modelu, optimalizace parametrů modelu, práce s daty a nezávislosti odhadu generalizační schopnosti modelu. Častým rysem deskriptivních úloh je obtížnější návrh hodnotící funkce, protože se nenabízí možnost přímého porovnání předpovědi a skutečnosti (Mařík, 2003).

S rozvojem AI se tyto klasické metody rozšiřují o neurální sítě, autoenkodéry a jazykové modely, které dokážou najít komplexní vzory i v nestrukturovaných datech (texty, obrázky, zvuk). Moderní přístup zahrnuje např.:

- automatické vyhledávání klíčových proměnných pomocí výběrových algoritmů a feature engineeringu,
- detekci anomálií s využitím hlubokého učení,
- nebo interpretaci výsledků prostřednictvím generativních modelů (např. shrnutí nálezů z tabulky pomocí jazykové AI).

Datová analýza s podporou AI se tak stává nástrojem, který nejen popisuje minulost, ale stále více i vysvětluje a navrhuje budoucí kroky, a to i v prostředí, kde je objem dat pro tradiční manuální analýzu neudržitelný.

1.5 Prompt engineering

Prompt engineering je specifický proces navrhování vstupních instrukcí (tzv. *promptů*) pro generativní modely umělé inteligence, zejména pro velké jazykové modely (LLM), jako je GPT-4. Cílem tohoto procesu je vytvořit takové zadání, které povede k co nejpřesnějšímu, nejrelevantnějšímu a nejužitečnějšímu výstupu. Prompt je typicky formulován jako text v přirozeném jazyce, který může mít podobu otázky, příkazu, úkolu či popisu situace, kterou má AI zpracovat.

Prompt engineering představuje praktickou disciplínu na pomezí lingvistiky, psychologie a výpočetní techniky. Je takovou formou dovednosti, která spojuje jazykové schopnosti s porozuměním fungování AI. Jeho cílem je najít ideální rovnováhu mezi jednoduchostí a přesností, aby model správně pochopil kontext i očekávání uživatele. (Mizrahi, 2024)

Správně formulovaný prompt má zásadní vliv na kvalitu, srozumitelnost i přesnost výstupu modelu. Vzhledem k tomu, že jazykové modely generují odpovědi prediktivním způsobem

(na základě pravděpodobnosti dalšího slova v kontextu předchozích slov), má podoba vstupu přímý dopad na výsledné chování modelu. Prompt engineering proto vyžaduje nejen jazykovou zdatnost, ale také porozumění způsobu, jakým AI model pracuje s kontextem, instrukcí a strukturou textu (Reynolds & McDonell, 2021).

Příklad jednoduchého promptu:

Prompt:

Jaký je rok?

Odpověď:

Dnes je rok 2025.

Tento ukázkový prompt ilustruje základní princip; model rozpozná dotaz a na jeho základě vygeneruje odpověď, která odpovídá aktuálnímu (nebo v případě offline modelu trénovanému) kontextu.

1.5.1 Techniky promptování

Pro dosažení kvalitního a konzistentního výstupu generativních modelů je zásadní nejen správně formulovaný dotaz (prompt), ale také použití vhodné techniky promptování, která ovlivňuje způsob, jakým model úlohu chápe a zpracovává. V praxi se využívá několik přístupů, které se liší mírou vedení modelu a složitostí úkolu:

- *Zero-shot learning* – Model dostává pouze samotné zadání bez jakéhokoliv příkladu nebo kontextu. Tento přístup je rychlý a uživatelsky přívětivý, avšak nemusí vést k nejpřesnějším výsledkům u složitějších úloh.
- *Few-shot learning* – Prompt obsahuje jeden nebo více vzorových příkladů, které modelu pomáhají pochopit, jaký typ odpovědi se očekává. Tento přístup je vhodný pro jemné doladění chování modelu bez nutnosti trénování.
- *Chain-of-thought prompting* – Model je veden k tomu, aby své myšlení rozdělil do kroků a vysvětloval, jak k výsledku dospěl. To pomáhá zvýšit přesnost u úloh vyžadující logické, vícekrokové uvažování, jako jsou matematické výpočty, deduktivní úlohy či analýza dat.
- *Self-consistency* – Tato technika rozšiřuje chain-of-thought prompting tím, že model vygeneruje více různých uvažování ke stejnému problému, a pak se vybere nejčastější

nebo nejlogičtější výsledek. Tento přístup zvyšuje stabilitu a kvalitu odpovědí, zejména u komplexních rozhodovacích úloh.

- *Tree-of-thought* – Tento pokročilejší přístup simuluje stromové uvažování. Model generuje alternativní větve možného řešení a průběžně hodnotí, která z nich vede k nejlepšímu výsledku. Tento přístup lze využít například při hledání strategie, plánování nebo řešení úloh s otevřeným koncem. Umožňuje modelu reflektovat více alternativ a dosáhnout promyšlenějšího závěru.

1.5.2 Kategorizace a klasifikace dat

V analytických úlohách pomáhá prompt engineering s celou řadou úkonů. Například ho můžeme využít při kategorizaci či klasifikaci dat. Pomocí správně navrženého promptu může model analyzovat např. e-mailové zpráva a určit, zda se jedná o spam či legitimní komunikaci. Příkladem takového promptu může být:

Prompt:

Urči, zda následující e-mailové zprávy vypadají jako spam, nebo ne:

1) Předmět: Nejlepší ceny na internetu

Odesílatel: info@notifikace.allegro.cz

2) Předmět: Dárkový poukaz Ryanair v hodnotě 12 500 Kč

Odesílatel: contactkfpquhyw@landmarks.beastinme.com

3) Předmět: Bud' vidět. A k tomu ušetři.

Odesílatel: adidas@cz-news.adidas.com

Odpověď:

Ne – nevypadá jako spam

Ano – vypadá jako spam

Ne – nevypadá jako spam

1.5.3 Čištění dat

Další oblastí využití je čištění dat, tedy detekce a odstranění nekonzistentních, neúplných nebo irelevantních informací z datových sad. Pomocí promptů lze AI navést k tomu, aby

identifikovala gramatické chyby, duplikáty, nevhodné hodnoty nebo chybějící údaje a navrhla jejich opravu.

Prompt:

Převeď následující seznam telefonních čísel do jednotného formátu +420 XXX XXX XXX:
604 123 456
+420-723-456-789
00420771987654

Odpověď:

+420 604 123 456
+420 723 456 789
+420 771 987 654

1.5.4 Shrnutí výsledků

Jedním z běžných využití generativní umělé inteligence v praxi je schopnost automaticky shrnout a interpretovat datové výstupy, např. obchodní výsledky, finanční přehledy nebo marketingové reporty. Pomocí správně formulovaného promptu může AI rychle identifikovat klíčové trendy, porovnat vývoj mezi obdobími a upozornit na odchylky nebo neobvyklé hodnoty, což značně usnadňuje a zrychluje rozhodovací procesy.

Prompt:

Vytvoř souhrn následující tabulky tržeb za poslední čtvrtletí podle jednotlivých regionů a uveď hlavní trendy.

Tabulka 1 Tržby v regionech

<i>Region</i>	<i>Leden (Kč)</i>	<i>Únor (Kč)</i>	<i>Březen (Kč)</i>
<i>Praha</i>	<i>1 200 000</i>	<i>1 350 000</i>	<i>1 500 000</i>
<i>Brno</i>	<i>800 000</i>	<i>780 000</i>	<i>820 000</i>
<i>Ostrava</i>	<i>300 000</i>	<i>320 000</i>	<i>340 000</i>

Zdroj: Vlastní zpracování

Odpověď:

Ve sledovaném období došlo k postupnému růstu tržeb ve většině regionů. Nejvýraznější nárůst zaznamenala Praha, kde se tržby zvýšily o 25%. Stabilní výkonnost vykazovalo Brno, které si udrželo podobnou úroveň s mírným růstem v březnu. Ostrava zaznamenala postupný růst každý měsíc, ačkoli zůstává nejslabším regionem podle absolutní výše tržeb. Celkově lze pozorovat pozitivní trend v prodeích, zejména v hlavním městě.

1.5.5 Vyhledávání anomálií

Dalším využitím AI v datové analýze je detekce anomálií, které mohou naznačovat chybu, výjimečnou událost nebo bezpečnostní hrozbu. Identifikace těchto nepravidelností je zásadní např. při monitoringu spotřeby, kontrole kvality, detekci podvodů nebo predikci výpadků. Pomocí vhodného promptu lze AI model navést k tomu, aby automaticky vyhodnotil datovou sadu a označil hodnoty, které vybočují z běžného rozsahu. Takové prompty lze snadno aplikovat i v každodenní praxi, například:

Prompt:

Vyhodnot' následující týdenní spotřebu elektřiny v kWh pro jednotlivé budovy a sleduj, zda některý z údajů nevykazuje podezřelou odchylku.

Tabulka 2 Týdenní spotřeba elektřiny

<i>Budova</i>	<i>Spotřeba (kWh)</i>
<i>Budova A</i>	<i>12 340</i>
<i>Budova B</i>	<i>11 950</i>
<i>Budova C</i>	<i>94 800</i>
<i>Budova D</i>	<i>12 410</i>

Zdroj: Vlastní zpracování

Odpověď:

Budova C vykazuje výrazně vyšší spotřebu ve srovnání s ostatními budovami, jejich spotřeba se pohybuje okolo 12 000 kWh. Tento údaj je pravděpodobně anomální a může být způsoben měřicí chybou, chybným zadáním nebo výpadkem systému.

1.5.6 Generování kódu

Generativní AI se stále více uplatňuje i v oblasti kódování a datového skriptování, kde slouží jako inteligentní asistent pro analytiku, datové vědce i začínající uživatele bez hlubších znalostí programování. Jednou z praktických funkcí je schopnost generovat kód (např. v SQL, Pythonu nebo R) na základě přirozeně formulovaného zadání. Pomocí jednoduchého promptu může uživatel zadat, co chce z dat zjistit, a model mu vygeneruje potřebný dotaz – vč. správné syntaxe, názvů sloupců, funkcí i agregací. Například:

Prompt:

Vygeneruj mi SQL dotaz, který spočítá průměrnou hodnotu objednávek pro jednotlivé zákaznické segmenty ze sloupce segment v tabulce orders, kde sloupec order_value obsahuje hodnotu objednávky.

Odpověď:

Tabulka 3 Vygenerovaný SQL dotaz

<i>Sql</i>
<i>SELECT</i> <i>Segment,</i> <i>AVG(order_value) AS average_order_value</i> <i>FROM</i> <i>orders</i> <i>GROUP BY</i> <i>segment;</i>

Zdroj: Vlastní zpracování

1.5.7 Predikce

Jedním z významných přínosů generativní AI v oblasti datové analýzy je její schopnost generovat predikce na základě historických dat. Pomocí vhodně formulovaného promptu lze model navést k tomu, aby analyzoval vývoj v čase, identifikoval vzorce a navrhl odhad budoucího vývoje, a to i s ohledem na sezónnost, růstové trendy nebo předchozí výkyvy. Tento přístup je užitečný zejména při plánování rozpočtů, predikci poptávky, nastavování kapacit či vyhodnocování kampaní. Prompt může vypadat například takto:

Prompt:

Na základě historických dat o měsíční návštěvnosti webu predikuj očekávaný počet návštěvníků pro následující čtvrtletí. Zohledni sezónní trendy a dosavadní růst .

Tabulka 4 Měsíční návštěvnost

Měsíc	2023	2024
Duben	85 000	94 000
Květen	89 000	99 500
Červen	91 500	104 000

Zdroj: Vlastní zpracování

Odpověď:

Z dat vyplývá, že návštěvnost webu vykazuje pravidelný meziroční nárůst kolem 10-12% v jarních měsících. Tento nárůst může být ovlivněn sezónními kampaněmi, zvýšenou poptávkou v daném období a optimalizací webového obsahu. Na základě trendu lze predikovat, že v roce 2025 bude návštěvnost činit přibližně:

- Duben: 103 000 až 105 000*
- Květen: 110 000 až 112 000*
- Červen: 115 000 až 118 000 návštěvníků*

Doporučujeme připravit marketingovou strategii pro udržení tempa růstu a zvážit investice do výkonové reklamy a SEO optimalizace během tohoto období.

1.5.8 Statistická analýza

Generativní AI dnes pomáhá i v oblasti statistického zpracování dat, kde dokáže na základě správně zadaného promptu navrhnout a provést vhodné statistické testy, interpretovat jejich výsledky a doporučit další postup. Jedním z případů je využití lineární regrese k určení závislosti mezi dvěma proměnnými – např. prodeji a výdaji na reklamu. Prompt může být formulován tímto způsobem:

Prompt:

Proveď lineární regresní analýzu k předpovědi výše měsíčních prodejů (v tis. Kč) na základě výdajů na reklamu (v tis. Kč). Následující tabulka obsahuje data za 6 měsíců.

Tabulka 5 Měsíční prodeje

Měsíc	Výdaje na reklamu (tis. Kč)	Prodeje (tis. Kč)
Leden	50	210
Únor	60	230
Březen	55	225
Duben	70	250
Květen	80	275
Červen	65	240

Zdroj: Vlastní zpracování

Odpověď:

Provedeme jednoduchou lineární regresi, kde:

- *Nezávislá proměnná (X): výdaje na reklamu*
- *Závislá proměnná (Y): prodeje*

Regresní rovnice má tvar: $Y = a + bX$

Na základě výpočtů (např. metodou nejmenších čtverců) vychází

přibližná regresní rovnice: $Prodeje = 110 + 2.0 \times Výdaje\ na\ reklamu.$

Interpretace:

- *Koeficient 2.0 znamená, že každý tisíc Kč investovaný do reklamy zvýší prodeje průměrně o 2 tis. Kč.*
- *Intercept (110) reprezentuje odhadovaný objem prodejů při nulových výdajích – slouží jako výchozí hodnota.*
- *Korelační koeficient $R \approx 0.98$ značí velmi silnou pozitivní lineární závislost mezi proměnnými.*

Tento model lze použít pro predikci budoucích prodejů – např. při plánované investici 90 tis. Kč do reklamy lze očekávat prodeje kolem 290 tis. Kč.

1.5.9 Vizualizace a reporting

Vizualizace dat je klíčovým nástrojem pro rychlé pochopení informací a efektivní komunikaci analytických závěrů. Generativní AI dokáže na základě jednoduchého promptu navrhnout vhodný typ grafu, vytvořit vizualizaci z tabulkových dat nebo vygenerovat kód pro nástroje jako Python (Matplotlib) či R, čímž výrazně zjednodušuje a zrychluje proces tvorby reportů. Například pro srovnání marketingových nákladů podle kanálů lze použít tento typ promptu:

Prompt:

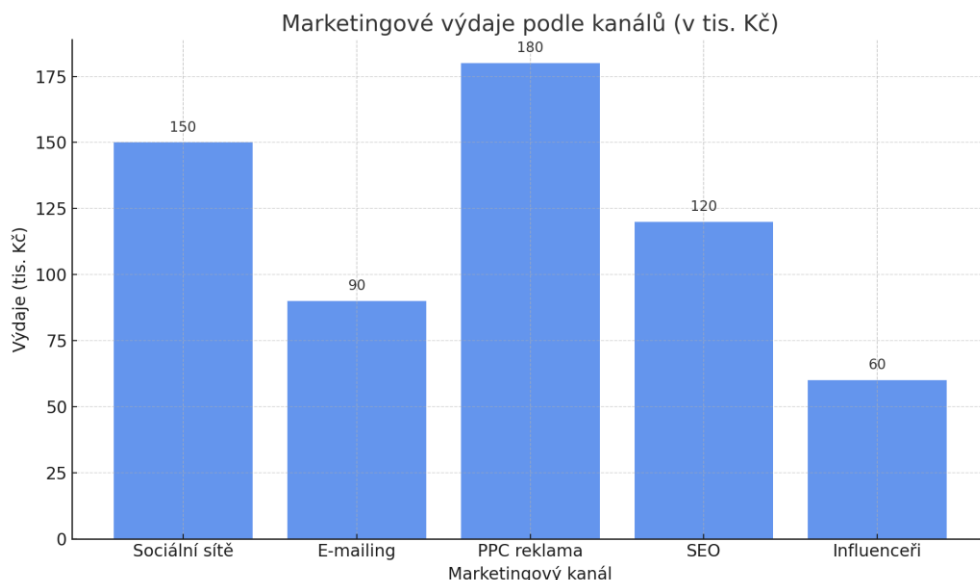
Vytvoř graf, který vizuálně porovná marketingové výdaje podle jednotlivých kanálů.

Tabulka 6 Marketingové výdaje

<i>Marketingový kanál</i>	<i>Výdaje (tis. Kč)</i>
<i>Sociální síť</i>	<i>150</i>
<i>E-mailing</i>	<i>90</i>
<i>PPC reklama</i>	<i>180</i>
<i>SEO</i>	<i>120</i>
<i>Influenceri</i>	<i>60</i>

Zdroj: Vlastní zpracování

Odpověď:



Obrázek 1 Marketingové výdaje podle kanálů (ChatGPT)

2 PRAKTICKÁ ČÁST

Cílem praktické části je přiblížit reálné možnosti využití umělé inteligence ve firemním prostředí se zaměřením na oblast správy dat, analýzy chování uživatelů a zabezpečení přístupu. Na konkrétních příkladech a simulacích bude demonstrováno, jak lze pomocí generativních modelů a analytických nástrojů podpořit efektivní rozhodování, odhalování nestandardních vzorců chování a predikci budoucích trendů. Zvláštní pozornost je věnována tomu, jak umělá inteligence interaguje s osobními a citlivými údaji, jaká rizika z toho plynou a jak je možné těmto rizikům čelit.

Praktická část vychází z prostředí společnosti MultiSport, kde jsou analyzována data o využívání benefičních karet, přístupech uživatelů k jednotlivým službám a s tím spojená bezpečnostní opatření. V rámci jednotlivých kapitol je ukázáno, jak lze pomocí přesně formulovaných promptů využít AI k tvorbě výpočtů, detekci podezřelých událostí či generování prediktivních modelů – a to i bez pokročilé znalosti programování. Cílem není pouze demonstrovat technické možnosti nástrojů, ale i reflektovat jejich dopad na firemní kulturu, přístup k datům a celkové řízení rizik.

2.1 Charakteristika společnosti

Společnost Multisport Benefit s.r.o. je poskytovatelem benefičního programu pro sport a volnočasové aktivity v České republice i dalších evropských zemích. Zaměřuje se na spolupráci s firmami a zaměstnavateli, kterým nabízí možnost motivovat své zaměstnance prostřednictvím přístupu k široké síti sportovišť. V posledních letech společnost výrazně investuje do digitalizace služeb a zefektivnění zákaznické práce, a právě v této oblasti začala využívat vybrané nástroje umělé inteligence. Společnost klade důraz na zodpovědné zacházení s daty, protože její činnost zahrnuje zpracování citlivých osobních, transakčních a smluvních informací.

V rámci digitální transformace nasadila společnost několik AI řešení:

- chatbot pro zákaznickou podporu, který zajišťuje odpovědi na časté dotazy, např. o dostupnosti sportovišť, ztrátě karty nebo registraci
- interní využití nástrojů generativní AI (Copilot)
- systém pro analýzu uživatelských dat, který vyhodnocuje frekvenci využívání karty, docházení na konkrétní sportoviště nebo třeba sezónní trendy

2.2 Rizika identifikovaná při nasazování AI

Nasazování nástrojů umělé inteligence v organizacích, jako je společnost MultiSport, přináší nejen funkční výhody, ale také nová specifická rizika spojená s bezpečností dat. V rámci interní analýzy byly identifikovány následující klíčové oblasti ohrožení, které je třeba při implementaci AI nástrojů systematicky řídit:

2.2.1 Možnost úniku dat

Při využívání generativních jazykových modelů, jako jsou ty nabízené prostřednictvím veřejných API (např. OpenAI, Google Bard nebo Anthropic Claude), existuje reálné riziko neúmyslného odeslání citlivých nebo osobních údajů mimo infrastrukturu firmy. Tato rizika vznikají zejména v případech, kdy jsou do systému zadávány dotazy (tzv. prompty), které obsahují identifikovatelné informace o klientech, zaměstnancích, smlouvách, finančních datech nebo interních procesech.

Pokud nejsou vstupy do modelu dostatečně anonymizovány, šifrovány nebo předem validovány, může dojít k tomu, že citlivé informace budou zpracovány nebo dočasně uloženy na serverech poskytovatele modelu, často mimo území EU, čímž dochází k porušení principů stanovených například nařízením GDPR. U některých platform navíc chybí plná transparentnost ohledně toho, zda a jak jsou uživatelská data využívána pro trénink modelů nebo k jiným účelům.

Toto riziko je zvláště problematické v případě modelů provozovaných třetími stranami bez přímé smluvní garance bezpečnosti a ochrany dat, kdy firma nemá možnost ovlivnit, jak bude s vloženými informacemi nakládáno. V praxi se tak může stát, že i jednorázové zadání s citlivým obsahem způsobí únik dat mimo organizaci, aniž by si toho byl uživatel vědom.

Z pohledu řízení rizik je proto klíčové:

- nastavit interní pravidla pro práci s AI nástroji,
- omezit přístup k citlivým systémům přes veřejná API,
- a případně preferovat lokální nebo privátní nasazení jazykových modelů, kde má firma plnou kontrolu nad datovým tokem.

V případě práce s externími modely je vhodné zajišťovat pravidelný audit datových interakcí, proškolit zaměstnance ohledně rizik promptování a posílit technická opatření pro filtraci a anonymizaci vstupů ještě před jejich odesláním do AI systému.

2.2.2 Možnost ztráty kontroly nad daty v trénovacím režimu

V případech, kdy organizace trénuje vlastní modely nebo využívá externího poskytovatele k jejich tréninku, vzniká riziko, že osobní údaje zůstanou v modelu „zakódovány“ natrvalo, bez možnosti jejich pozdějšího odstranění. Pokud nebyl tréninkový proces správně nastaven – například bez anonymizace, bez právního základu nebo bez odpovídajících smluvních záruk – může takové použití dat porušovat zásady GDPR, zejména právo na výmaz podle článku 17. Modely hlubokého učení navíc mohou neúmyslně generovat výstupy, které citlivé informace znovu odhalí. Proto je zásadní, aby organizace před tréninkem zajistily jak technická opatření, tak právní soulad zpracování s ochranou osobních údajů.

2.2.3 Riziko rekonstrukce identity uživatele

I v případě anonymizovaných dat existuje riziko rekonstrukce identity uživatele na základě jeho chování, vzorců využívání služby nebo kombinací několika zdánlivě neškodných datových atributů – tento jev je označován jako re-identifikace. Umělá inteligence, která má schopnost nacházet i velmi subtilní souvislosti v datech, může tuto schopnost neúmyslně posilovat, a to i bez přímého cílení na identifikaci jednotlivce. Z tohoto důvodu roste význam technik, jako je diferenciální soukromí, které mají za cíl omezit možnost zpětného odhalení identity z výstupů modelu. Stejně tak je nezbytné kontrolovat přístup k výstupům AI systémů, zejména pokud jsou výsledky analýz sdíleny mezi odděleními nebo s externími partnery.

2.2.4 Potenciální bias při predikcích

AI modely mohou ve svých výstupech odrážet zkreslení obsažené ve vstupních datech. Pokud jsou například data historicky nerovnoměrně rozložená (např. více mužských uživatelů než ženských), může dojít k tomu, že prediktivní model začne nespravedlivě favorizovat určité skupiny uživatelů. Takové situace jsou rizikové zejména při doporučování služeb, personalizaci nebo cílení nabídek.

2.2.5 Nedostatečné vysvětlení rozhodnutí modelu (black-box efekt)

Modely na bázi hlubokého učení (deep learning) nebo velké jazykové modely (LLM) fungují jako tzv. black-box systémy – tedy systémy, u nichž není snadné zpětně vysvětlit, jakým způsobem dospěly ke konkrétnímu závěru či doporučení. To představuje problém jak z hlediska uživatelské důvěry, tak i z pohledu legislativních nároků na auditovatelnost rozhodnutí, zejména v oblasti veřejných služeb nebo personalistiky.

2.3 Návrh opatření pro ochranu dat

Na základě analýzy identifikovaných rizik, praktických příkladů využití AI ve společnosti MultiSport a současných právních požadavků byla navržena soubor technických a organizačních opatření, jejichž cílem je minimalizovat bezpečnostní hrozby a zajistit soulad s platnou legislativou, zejména s Obecným nařízením o ochraně osobních údajů (GDPR).

Tato opatření zohledňují nejen právní aspekty, ale i potřebu uchovat důvěru uživatelů, minimalizovat riziko úniku nebo zneužití dat a umožnit auditovatelnost rozhodovacích procesů AI systémů. Vycházejí také z principů prevence, transparentnosti a odpovědnosti za zpracování dat v celém jejich životním cyklu – od sběru, přes analýzu až po uchování a případné výmaz. Níže jsou uvedena klíčová opatření navržená v rámci ochrany dat v kontextu AI:

2.3.1 Anonymizace vstupních dat

Při využívání externích generativních modelů – jako jsou například jazykové modely poskytované prostřednictvím API od společnosti OpenAI, Google nebo Anthropic- je naprosto zásadní, aby veškerá vstupní data procházela anonymizací nebo alespoň pseudonymizací ještě před jejich odesláním do cloudu. Tyto systémy často běží na infrastruktuře mimo přímou kontrolu organizace, a i přes smluvní záruky existuje riziko, že data mohou být uložena, využita k dalšímu tréninku, nebo zneužita při případném narušení bezpečnosti.

Anonymizací se nevratně odstraňují všechny identifikátory – jména, e-mailové adresy, rodná čísla, ale také kombinace atributů, které by mohly vést k identifikaci uživatele. V případě pseudonymizace se naopak využije nahrazení citlivých údajů za zástupné hodnoty (např. ID), přičemž propojení se skutečnou identitou je uchováváno odděleně.

Příkladem částečné anonymizace dat před odesláním do externího AI API může být následující tabulka znázorňující záznamy o návštěvách sportovišť. Identita uživatele je zastoupena pseudonymem CardUserId, který neobsahuje přímo identifikující údaje, ale může být interně zpětně dohledatelný. I když čas návštěvy, identifikátor zařízení a název města zůstávají zachovány, taková forma pseudonymizace již výrazně snižuje riziko úniku citlivých informací. Pro použití v cloudu by však bylo vhodné provést další kroky, jako je agregace dat podle časových intervalů nebo zobecnění lokality, čímž by se minimalizovalo riziko možné re-identifikace uživatele kombinací více atributů.

Tabulka 7 Anonymizace vstupních dat

CardUserId	Anonimization	VisitDate	FacilityID	CityName
52175831	2	31.10.2024 21:51	3573	Praha
52175831	2	30.10.2024 16:20	5069	Plzeň
52158584	2	16.10.2024 14:19	1337	Kolín
52157993	2	24.09.2024 16:48	3709	Český Brod
52158179	2	23.09.2024 14:48	1337	Kolín
52158387	2	23.09.2024 13:35	5194	Luhačovice
52175831	2	22.09.2024 18:07	839	Praha
52157650	2	07.09.2024 17:41	789	Chrudim
52158339	2	07.09.2024 13:20	5579	Praha
52157993	2	07.09.2024 10:49	1337	Kolín

Zdroj: Vlastní zpracování

2.3.2 Šifrování komunikace a datového přenosu

Veškerý přenos dat mezi firemními systémy a externími AI službami – například při využívání cloudových API – by měl probíhat výhradně prostřednictvím šifrovaných komunikačních protokolů, jako je TLS 1.3 (Transport Layer Security). Tento moderní standard zajišťuje důvěrnost a integritu dat při přenosu tím, že znemožňuje jejich odposlech, zachycení nebo neoprávněnou manipulaci ze strany třetích osob.

Zároveň by měly být uplatňovány silné kryptografické algoritmy, například AES-256 (Advanced Encryption Standard), který je v současnosti považován za jeden z nejbezpečnějších šifrovacích mechanismů pro ochranu dat v pohybu. Toto opatření je zásadní zejména při práci s osobními nebo citlivými údaji, protože snižuje riziko jejich úniku v důsledku útoků typu „man-in-the-middle“, přesměrování provozu nebo zranitelností v síťové infrastruktuře.

Z pohledu souladu s GDPR i kybernetickými standardy (např. NIS2) je šifrování dat při přenosu považováno za základní bezpečnostní opatření, které by mělo být uplatňováno systematicky, bez ohledu na typ dat nebo použitou AI službu. Důležité je rovněž pravidelně revidovat nastavení šifrování, udržovat certifikáty aktuální a omezit přístup pouze na ověřené koncové body.

2.3.3 Nakládání s osobními údaji

Organizace by měla mít jasně definovaná interní pravidla a směrnice pro práci s osobními údaji při využívání nástrojů umělé inteligence. Zejména je nezbytné stanovit omezení pro používání veřejně dostupných AI platforem, jako jsou ChatGPT, Bard, Claude nebo Gemini,

kteře jsou provozovány třetími stranami a běží na cloudové infrastruktuře mimo kontrolu organizace.

Zadávání osobních nebo citlivých údajů do těchto systémů – například jména, e-maily, zdravotní informace nebo interní identifikátory – by nemělo být povoleno, pokud k tomu neexistuje výslovný a informovaný souhlas subjektu údajů, nebo jiný relevantní právní základ dle čl. 6 GDPR (např. oprávněný zájem, smluvní plnění, zákonná povinnost). V opačném případě hrozí neoprávněné předání dat do třetích zemí, riziko jejich uložení nebo využití pro trénink modelu, což může vést k porušení zákonných povinností i reputačním škodám.

Je také vhodné zřídít tzv. „zásady bezpečné práce s AI“, které budou součástí školení zaměstnanců a budou definovat, jaké typy dat lze zpracovávat v AI nástrojích, kdo má k těmto nástrojům přístup, a jaká opatření (např. anonymizace) musí být před odesláním dat provedena. Současně je vhodné omezit přístup k veřejným AI službám z firemních zařízení, pokud nejsou provozovány v bezpečném a smluvně ošetřeném režimu (např. prostřednictvím enterprise API s vypnutým logováním dotazů).

2.3.4 Interní pravidla pro prompt engineering

Společnost by měla vytvořit vzorové šablony promptů, které mohou zaměstnanci využívat při interakci s generativními AI nástroji. Tyto šablony slouží jako vodítko pro správné formulování dotazů a zároveň zajišťují, že do promptů nebudou nevědomky vkládány citlivé, osobní nebo obchodně důvěrné informace. Vzory mohou obsahovat doporučenou strukturu dotazu, příklady vhodných formulací i upozornění na nevhodný obsah.

Jejich cílem je snížit riziko lidské chyby, zvýšit konzistenci v komunikaci se systémy AI a zároveň podpořit bezpečné a odpovědné využívání těchto technologií v souladu s interními směrnici a právními předpisy. Pravidelné aktualizace šablon by měly reflektovat vývoj nástrojů i nově identifikovaná rizika. Zaměstnanci by také měli být proškoleni v tom, kdy a jak šablony používat a jak případně konzultovat neobvyklé nebo komplexní dotazy s pověřenou osobou (např. DPO nebo datovým analytikem).

2.3.5 Auditní logování interakcí s AI systémy

Zavedením systému auditního logování je možné zpětně dohledat, jaká data byla použita v interakci s AI nástrojem, kdo je odeslal, kdy a v jakém kontextu. Tento dohledový mechanismus umožňuje efektivní kontrolu nad využíváním AI systémů a je klíčový pro zajištění transparentnosti, odpovědnosti a souladu s interními pravidly i legislativními požadavky.

Logování záznamů může rovněž sloužit jako nástroj pro detekci nestandardního chování, například v případech, kdy by do modelu byly zadány nevhodné nebo rizikové informace. V případě bezpečnostního incidentu poskytují auditní záznamy důkazní základnu pro forenzní šetření a zároveň pomáhají vyhodnotit, zda bylo jednání zaměstnance nebo systému v souladu s přijatými pravidly. Důležité je, aby byl přístup k auditním logům omezen a chráněn, a aby jejich zpracování probíhalo v souladu se zásadami ochrany osobních údajů.

2.3.6 Pravidelné školení zaměstnanců v oblasti datové etiky a práce s AI

Zaměstnanci, kteří přicházejí do styku s AI nástroji, by měli ideálně absolvovat pravidelná školení zaměřená na etickou práci s daty, rizika spojená s AI a konkrétní pravidla a zásady platné v rámci organizace. Taková školení by neměla být jednorázová, ale průběžná – ideálně aktualizovaná v návaznosti na nové technologické trendy, legislativní změny nebo interní bezpečnostní události.

Součástí programu by mělo být také praktické seznámení s příklady nevhodného nebo rizikového použití AI nástrojů, které by mohlo vést k úniku dat, porušení smluvní mlčenlivosti nebo dokonce k porušení Obecného nařízení o ochraně osobních údajů (GDPR). Cílem je zvýšit povědomí o možných dopadech běžné pracovní činnosti a podpořit zodpovědný a informovaný přístup zaměstnanců k technologiím, které mají vysoký potenciál, ale i riziko.

Zvláštní důraz by měl být kladen na školení pracovníků z oblastí jako je analytika, HR, marketing či zákaznická podpora, kde je práce s osobními nebo důvěrnými informacemi běžná. Správně nastavené vzdělávání může výrazně snížit riziko lidské chyby a přispět k vytvoření bezpečné a datově uvědomělé firemní kultury.

2.3.7 Zpracování AI rizik v rámci interní GDPR agendy

V rámci interních dokumentů a procesů GDPR by měla být vytvořena speciální příloha věnovaná rizikům spojeným s využíváním AI. Tato část by měla obsahovat např. klasifikaci dat podle citlivosti, pravidla pro práci s trénovacími daty a seznam AI nástrojů schválených pro firemní použití. Pravidelně by měla být také prováděna DPIA (Data Protection Impact Assessment) u nových AI projektů.

2.4 Další doporučení

2.4.1 Interní testing adversarial scénářů

Jedním z nejčastějších bezpečnostních rizik u AI systémů jsou tzv. adversarial attacks, tedy záměrně vytvořené vstupy, které mají za cíl zmást nebo manipulovat chování modelu. Společnosti využívající AI by měly testovat své systémy proti takovým scénářům, a to buď interně, nebo pomocí externího bezpečnostního auditu. Cílem je zjistit, jak model reaguje na nestandardní, škodlivé nebo neobvyklé vstupy, a zda existuje prostor pro zneužití. Např. u chatbotu může jít o zadání, která testují, zda model neodhalí interní informace nebo nezačne generovat nevhodný obsah.

2.4.2 Zajištění dokumentace rozhodovací logiky modelu

Vzhledem k povaze některých AI nástrojů jako „black-box systémů“ je klíčové zajistit minimální úroveň srozumitelnosti rozhodovacích mechanismů – zejména v oblastech, kde má model dopad na uživatele (např. personalizace, zákaznická doporučení). Doporučuje se vytvořit nebo nasadit nástroje pro explainable AI (XAI), jako je např. SHAP (SHapley Additive exPlanations) nebo LIME (Local Interpretable Model-Agnostic Explanations), které umožní:

- analyzovat, na základě jakých vstupů model rozhoduje,
- interpretovat výstupy i pro netechnické uživatele,
- podpořit auditovatelnost a transparentnost rozhodnutí.

Tato dokumentace je zároveň nezbytná z pohledu souladu s GDPR, pokud AI systém zpracovává osobní údaje a dochází na jeho základě k automatizovanému rozhodování.

2.4.3 Použití open-source modelů na vlastním serveru

Při zpracování citlivých nebo osobních dat (např. v rámci interních HR systémů, zákaznické segmentace apod.) lze doporučit využití open-source jazykových modelů, které je možné nasadit lokálně na firemní infrastruktuře. Tento přístup má několik výhod:

- vyšší kontrola nad daty – nic neopouští vnitřní systém,
- možnost customizace modelu podle konkrétního oboru nebo stylu komunikace,
- nižší právní riziko, protože nedochází k přenosu dat přes API třetích stran.

Modely jako LLaMA, Mistral, Falcon nebo BLOOM umožňují běh na vlastním serveru bez nutnosti napojení na cloudové poskytovatele. I přesto je potřeba zajistit dostatečnou kapacitu výpočetních zdrojů a technické know-how pro jejich správu a zabezpečení.

2.5 Využití umělé inteligence v analýze dat

Umělá inteligence představuje silný nástroj pro analýzu rozsáhlých datových souborů, a to jak ve fázi zpracování, tak i interpretace a predikce. Ve firemním prostředí slouží AI k odhalování vzorců v uživatelském chování, identifikaci trendů, predikci budoucí poptávky nebo k personalizaci služeb. Společnost MultiSport využívá umělou inteligenci zejména při analýze dat o využívání benefitních karet, sledování frekvence návštěv a identifikaci sezónních výkyvů, což pomáhá přizpůsobovat nabídku konkrétním preferencím uživatelů a optimalizovat síť zapojených sportovišť.

2.5.1 Scriptování

Jedním z praktických přínosů generativní umělé inteligence ve firemní analýze dat je její schopnost automaticky generovat kód nebo dotazy na základě zadání zadaného v přirozeném jazyce. Tento přístup výrazně zrychluje práci analytiků, zejména v případech, kdy je třeba vytvářet komplexní SQL dotazy, jejichž ruční sestavení by bylo časově náročné nebo náchylné k chybám.

Níže je uveden příklad, jak byla umělá inteligence využita při vytváření SQL dotazu pro analýzu benefitních karet (plastových a virtuálních), jejich návazností a zjišťování, zda měl uživatel v daném období přístup přes mobilní aplikaci. Jazykový model na základě dobře formulovaného promptu dokáže vytvořit složitý dotaz zahrnující práci s dočasnými tabulkami, spojeními více databázových struktur a podmínkami vyhodnocujícími časovou návaznost záznamů.

Prompt:

Vytvoř SQL script na výběr aktivních plastových a virtuálních karet. Podkladem je tabulka Cards_SubStatus, kde jsou historické stavy plastových karet (CardServiceId, DateFrom, StatusID). Potřebuji zjistit souvislá aktivní období, pro aktivní platí StatusID = 1, a to bez mezer mezi záznamy (mezi DateTo a následujícím DateFrom max 1 den). V tabulce MobileVirtualAccess jsou pak virtuální přístupy (DateFrom, DateTo, EuzAccountId), tabulka je propojena přes CardUser a Card_Service až na CardServiceID. Výstupem má být přehled všech aktivních období plastových karet a

informace, jestli v daném období byla karta virtuální. Výstup: CardServiceID, DateFrom, DateTo, JeVirtualni. Ideálně zkus využít dočasné tabulky a spoj je do jedné výsledné.

Odpověď:

```
IF OBJECT_ID('tempdb..#Aktivni') IS NOT NULL DROP TABLE #Aktivni;
IF OBJECT_ID('tempdb..#Virtualni') IS NOT NULL DROP TABLE #Virtualni;
-- 1) uložení plastové karty do dočasné tabulky
WITH sml AS (
    SELECT
        [CardServiceId],
        [DateFrom],
        [StatusID],
        [ClientAccountId],
        [DateFrom2],
        LEAD([DateFrom], 1) OVER (PARTITION BY [CardServiceId] ORDER BY
DateFrom DESC) AS previous_datefrom,
        MIN(DateFrom) OVER (PARTITION BY [CardServiceId]) AS Card_Start_date,
        CASE
            WHEN [DateFrom] <> ISNULL(LEAD([DateFrom], 1) OVER (PARTITION BY
[CardServiceId] ORDER BY DateFrom DESC), 0)
            THEN 1 ELSE 0
        END AS hledaj
    FROM [dbo].[Cards_SubStatus]
),
sml2 AS (
    SELECT
        sml.Cardserviceid,
        CAST(sml.DateFrom AS DATE) AS DateFrom,
        sml.StatusID,
        CAST(ISNULL(DATEADD(DAY, -1, LAG(DateFrom, 1) OVER (PARTITION BY
[CardServiceId] ORDER BY DateFrom DESC)), DATEFROMPARTS(2999, 12, 31))
AS DATE) AS DateTo
    FROM sml
```

```

),
zdroj AS (
    SELECT * FROM sml2 WHERE StatusID = 1
),
CTE AS (
    SELECT
        Cardserviceid AS CardID,
        DateFrom,
        DateTo,
        LAG(DateTo) OVER (PARTITION BY CardServiceId ORDER BY DateFrom) AS
        PrevDateTo
    FROM zdroj
),
Grouped AS (
    SELECT
        CardID,
        DateFrom,
        DateTo,
        SUM(CASE WHEN PrevDateTo IS NULL OR DATEDIFF(DAY, PrevDateTo,
        DateFrom) > 1 THEN 1 ELSE 0 END)
        OVER (PARTITION BY CardID ORDER BY DateFrom) AS GroupID
    FROM CTE
)
SELECT
    CardID AS CardServiceID,
    MIN(DateFrom) AS DateFrom,
    MAX(DateTo) AS DateTo
INTO #Aktivni
FROM Grouped
GROUP BY CardID, GroupID;
-- 2) uložení virtuální karty do dočasné tabulky
WITH CTE2 AS (
    SELECT

```

```

        cs.CardServiceId AS CardID,
        CAST(mva.[DateFrom] AS DATE) AS DateFrom,
        CAST(mva.[DateTo] AS DATE) AS DateTo,
        LAG(CAST(mva.[DateTo] AS DATE)) OVER (PARTITION BY cs.CardServiceId
ORDER BY CAST(mva.[DateFrom] AS DATE)) AS PrevDateTo
    FROM [dbo].[MobileVirtualAccess] mva
    INNER JOIN dbo.Card_User AS CU ON MVA.EuzAccountId =
CU.EndUserAccountId
    INNER JOIN dbo.Card_Service AS CS ON CU.CardUserId = CS.CardUserId
),
Grouped2 AS (
    SELECT
        CardID,
        DateFrom,
        DateTo,
        SUM(CASE WHEN PrevDateTo IS NULL OR DATEDIFF(DAY, PrevDateTo,
DateFrom) > 1 THEN 1 ELSE 0 END)
        OVER (PARTITION BY CardID ORDER BY DateFrom) AS GroupID
    FROM CTE2
)
SELECT
    CardID,
    MIN(DateFrom) AS DateFrom,
    MAX(DateTo) AS DateTo
INTO #Virtualni
FROM Grouped2
GROUP BY CardID, GroupID;

```

-- 3) Výsledný SELECT s propojením plastových a virtuálních karet

```

SELECT
    p.CardServiceID,
    p.DateFrom,
    p.DateTo,

```

```

CASE
    WHEN v.CardID IS NOT NULL THEN 1
    ELSE 0
END AS JeVirtuální
FROM #Aktivni p
LEFT JOIN #Virtualni v
    ON p.CardServiceID = v.CardID
    AND v.DateFrom <= p.DateTo
    AND v.DateTo >= p.DateFrom
ORDER BY p.CardServiceID, p.DateFrom;

DROP TABLE #Aktivni;
DROP TABLE #Virtualni;

```

Uvedený prompt vygeneroval komplexní SQL script, jehož cílem je identifikovat souvislá aktivní období plastových a virtuálních karet a zjistit, zda se v uvedeném období jednalo vstup přes mobilní aplikaci nebo přes fyzickou plastovou kartu. Použití dočasných tabulek *#Aktivni* a *#Virtualni* umožňuje přehledně pracovat s historickými stavy a propojit záznamy z různých tabulek. Díky tomu lze získat přehled o tom, kdy a v jaké formě měl uživatel přístup do provozovny, což je klíčové např. pro analýzu chování klientů nebo kontrolu oprávněnosti přístupu. Závěrečné příkazy *DROP TABLE* slouží k odstranění dočasných tabulek z databáze pro dokončení výpočtu, aby nezatěžoval systém.

2.5.2 Reporting

Další oblastí, kde může umělá inteligence významně podpořit datovou analýzu, je automatizovaný reporting. AI nástroje umožňují generovat srozumitelné souhrny, vizualizace a přehledy dat na základě jednoduchého textového zadání. Tento přístup šetří čas analytikům a zároveň zvyšuje dostupnost datových výstupů pro širší okruh zaměstnanců, kteří nemusí ovládat analytické nástroje nebo jazyky jako SQL.

Prompt:

Vytvoř počítaný sloupec, který bude počítat životnost plastových karet, tzn. pokud je měsíc a rok ve sloupcích DateFrom a _virtual_from a zároveň DateTo a _virtual_to, vrátí sloupec Zivotnost_mesice.

Odpověď:

```
Zivotnost_plastove_karty =  
VAR _dateFrom = SQL_Cards_Active[DateFrom]  
VAR _dateTo = SQL_Cards_Active[DateTo]  
VAR _virtualFrom = SQL_Cards_Active[_virtual_from]  
VAR _virtualTo = SQL_Cards_Active[_virtual_to]  
RETURN  
IF(  
    ((_virtualFrom) && ISBLANK(_virtualTo))  
    ||  
    (YEAR(_dateFrom) < YEAR(_virtualFrom) || (YEAR(_dateFrom) =  
YEAR(_virtualFrom) && MONTH(_dateFrom) < MONTH(_virtualFrom)))  
    &&  
    (YEAR(_dateTo) < YEAR(_virtualFrom) || (YEAR(_dateTo) =  
YEAR(_virtualFrom) && MONTH(_dateTo) < MONTH(_virtualFrom)))  
),  
SQL_Cards_Active[Zivotnost_mesicu],  
BLANK()  
)
```

Uvedený příklad demonstruje, jak lze s pomocí AI asistenta vytvořit automatizovaný výpočet pro reportovací nástroj, konkrétně v prostředí Power BI. Prompt formuluje požadavek na výpočet životnosti plastové karty ve specifickém časovém intervalu, přičemž odpověď poskytuje kompletní výraz v jazyce DAX, který lze okamžitě použít v modelu. Výpočet zohledňuje různé varianty aktivace plastových a virtuálních karet a zajišťuje, že do výstupu budou zahrnuty pouze ty karty, jejichž doba platnosti se odehrála před spuštěním virtuální verze.

Tento přístup umožňuje rychle a přesně zpracovat logiku výpočtu bez nutnosti ručního psaní kódu, čímž výrazně zvyšuje efektivitu analytické práce. Zároveň poskytuje jasně strukturovaný a opakovatelný výstup, který je srozumitelný i pro uživatele, kteří nejsou experty na programování v DAXu. V praxi to znamená, že i méně technicky orientovaní pracovníci mohou vytvářet reporty s pokročilými metrikami, což přispívá k širšímu využití dat v organizaci a demokratizaci datové analytiky.

2.5.3 Detekce podvodů

Jedním z důležitých využití umělé inteligence v oblasti analýzy dat je odhalování podezřelého nebo podvodného chování. V prostředí využívání benefitních karet, jako je tomu u společnosti MultiSport, může AI sloužit jako nástroj pro automatické rozpoznání nestandardních vzorců používání karet, které by mohly signalizovat pokus o zneužití služby, technickou chybu nebo systémovou slabinu.

Pomocí dobře definovaného dotazu (promptu) lze AI instruovat k vyhledání případů, kdy například došlo k opakovanému selhání registrace karty, neobvyklému přepnutí mezi zařízeními, nebo k pokusům o přístup s neaktivní či zneplatněnou kartou. Takový přístup pomáhá firmě rychle reagovat na rizikové situace a chránit své služby před zneužitím.

Prompt:

Najdi v podkladových datech neobvyklé nakládání s kartami. Např. pokud načtení karty skončilo ve stavu „Registration failed“ a karta byla poté použita v jiném zařízení, tento stav se pro danou kartu objevuje několikrát po sobě, atd.

Odpověď:

Identifikované vzorce podezřelého chování v datech

1. Opakované chyby u stejné karty v krátkém čase

Např. karta č.52582977 byla během několika minut opakovaně skenována v zařízení IRON Gym Tábor a v každém případě došlo k chybě (Registration failed). Tento vzorec může naznačovat:

- *neplatnou kartu, kterou se uživatel snaží opakovaně použít*
- *systémovou chybu, kdy zařízení opakovaně nezpracuje legitimní přístup*
- *možné zneužití nebo testování systému přístupu*

2. Smíšené výsledky pro jednu kartu

U karty 516131951 a 51641952 došlo k úspěšné registraci, zatímco u jiných karet se opakovaně objevuje neúspěšná registrace – to by bylo v pořádku, pokud by šlo o různé pokusy různých uživatelů, ale u opakovaných pokusů jedné osoby na stejné lokaci to může znamenat nestandardní zacházení nebo chybu zařízení.

3. Shodné zařízení, opakovaný výskyt chyb

Všechny chyby se vážou ke stejnému zařízení (IRON Gym Tábor). To může poukazovat například na:

- *lokální technický problém se čtečkou karet,*
- *neaktuální data v zařízení (např. karta je aktivní, ale zařízení nemá aktuální seznam platných přístupů),*
- *potřebu aktualizace firmware nebo audit logiky validace přístupů.*

Uvedené příklady ukazují, jak může umělá inteligence pomoci odhalit nestandardní nebo potenciálně podvodné vzorce chování v reálných datech. Namísto ručního procházení rozsáhlých logů dokáže systém upozornit na situace, které by mohly signalizovat bezpečnostní incident, chybu v infrastruktuře nebo snahu obejít pravidla používání služby. Identifikace opakovaných selhání, neobvyklých přechodů mezi zařízeními nebo výskytu chyb v konkrétní lokalitě umožňuje firmě rychle reagovat, provést kontrolu zařízení, ověřit platnost karet nebo upravit logiku zpracování přístupů.

Zároveň je důležité, aby byly výsledky takové analýzy srozumitelné i pro netechnické uživatele. Díky využití promptu v přirozeném jazyce může i pracovník bez znalosti SQL nebo pokročilé analytiky získat jasný přehled o problému a jeho možných příčinách. To zvyšuje využitelnost AI nástrojů napříč různými odděleními firmy – od zákaznické podpory po IT bezpečnost.

2.5.4 Predikce

Umělá inteligence umožňuje firmám nejen analyzovat minulá data, ale také předpovídat budoucí vývoj na základě rozpoznávaných vzorců a trendů. V oblasti sportovních benefitů je možné využít prediktivní modelování např. pro odhady návštěvnosti konkrétních sportovišť, predikci zatížení sítě provozoven nebo optimalizaci kapacit podle sezónnosti. Takové predikce pomáhají zefektivnit plánování, vyhnout se přetížení některých lokalit a zároveň lépe řídit spolupráci s partnery.

Prompt:

Mám k dispozici data o poměru vstupů do různých sportovních provozoven, potřebuji naplánovat podíl vstupů k celkovému počtu návštěv na období 4-25 až 12-25, jaké různé postupy můžeme využít?

Odpověď:

1. *Spočítej průměr podílů za měsíce duben-prosinec v předešlých letech pro každé zařízení a ten použij jako predikci pro rok 2025.*

=PRŮMĚR(D13:N13)

2. Spočítej průměr pouze z roku 2024, jako nejaktuálnější trend:

=PRŮMĚR(J13:R13)

3. škálování historické hodnoty podle poměru změny v dosud známých měsících letošního roku oproti stejnému období loňského roku

=(SUMA(\$BY\$14:\$CA\$14)/SUMA(\$BM\$14:\$BO\$14))*BP14

4. Použij lineární trend (regresi):

=TREND(historické_hodnoty_y;historické_hodnoty_x;budoucí_x
)

5. Vypočítej sezónní průměr (např. podíl dubna na každý rok), vezmeme jen dubnové hodnoty z několika let a uděláme průměr jen pro duben, totéž pro další měsíce, atd.

=PRŮMĚR(2021.4;2022.4;2023.5;2024.5)

6. Využij trend z celkového vývoje. Pokud návštěvnost konkrétního zařízení klesá nebo stoupá, můžeš spočítat trend (např. pomocí lineární regrese) a tím upravit predikci nahoru nebo dolů.

Uvedené návrhy ukazují, že pro predikci návštěvnosti sportovních provozoven je možné využít celou řadu metod, které se liší svou složitostí, nároky na historická data i schopností zohlednit aktuální vývoj. Nejjednodušší přístupy pracují s průměry – buď z několika předchozích let, nebo z posledního dostupného období – a jsou vhodné zejména pro rychlé odhady. Složitější metody, jako je škálování podle meziročních změn nebo lineární trend (regrese), umožňují zohlednit skutečný vývoj návštěvnosti a přizpůsobit předpověď aktuálním podmínkám. Velmi užitečná je i sezónní analýza, která bere v úvahu opakující se vzorce chování zákazníků v jednotlivých měsících.

Tyto postupy lze realizovat i v běžném nástroji, jako je Microsoft Excel, což z nich činí dostupný nástroj pro každodenní plánování. Klíčové přitom je, aby byly predikce nejen přesné, ale také interpretovatelné – tedy srozumitelné pro uživatele, kteří se nepohybují v oblasti datové analytiky. Právě v tom spočívá síla dobře zvoleného promptu: umožňuje formulovat dotaz přirozeným jazykem a získat návrh konkrétních výpočtů, které lze snadno ověřit a aplikovat.

ZÁVĚR

Umělá inteligence se stává nedílnou součástí digitální transformace firem a její vliv na způsob práce s daty, rozhodovací procesy a zákaznickou zkušenost je dnes již nepopíratelný. Tato diplomová práce si kladla za cíl prozkoumat, jaké bezpečnostní, právní a etické výzvy přináší nasazení AI technologií ve firemním prostředí, a nabídnout doporučení pro jejich odpovědné a efektivní využívání.

V teoretické části byla nejprve představena podstata umělé inteligence, její historický vývoj a hlavní směry aplikace. Pozornost byla dále věnována etickým principům, jako je transparentnost, odpovědnost, spravedlnost, ochrana soukromí a udržitelnost – tedy hodnotám, které by měly být nedílnou součástí každého systému využívajícího AI. Na příkladech z různých odvětví bylo ukázáno, jakým způsobem může AI přinášet prospěch, ale i ohrožovat základní práva jednotlivců či integritu datových systémů.

Praktická část práce pak ukázala, jak lze generativní modely využít v reálném firemním kontextu, a to například při detekci podezřelého chování, predikci využívání služeb nebo tvorbě automatizovaných výpočtů a reportů. Důraz byl kladen nejen na samotnou funkčnost promptů a výsledků, ale také na související rizika – zejména možnost neúmyslného úniku dat, zkreslených interpretací nebo ztráty kontroly nad zpracováním citlivých informací.

Závěrem lze konstatovat, že přínosy AI jsou značné, ale nelze je oddělovat od odpovědného přístupu k datům. Bezpečnost a etika musí jít ruku v ruce s inovacemi, jinak může důvěra ve využívání AI snadno erodovat. Klíčem k úspěšnému nasazení AI ve firemní praxi proto není pouze technologická připravenost, ale především schopnost propojit technické, právní a lidské aspekty tohoto nástroje. Práce ukazuje, že takový přístup je nejen možný, ale i nezbytný pro udržitelný rozvoj digitálních technologií v podnicích budoucnosti.

POUŽITÁ LITERATURA

1. Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., . . . Portugali, Y. (2012). Smart Cities of the Future. *The European Physical Journal Special Topics*.
2. Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press.
3. Brown, T. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems (NeurIPS)*.
4. Brundage, M. (2018). *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation*. Načteno z Cornell University: <https://arxiv.org/abs/1802.07228>
5. Campbell, M., Hoane, A., & Hsu, F. (2002). Deep Blue. *Artificial Intelligence*.
6. Crevier, D. (1993). *AI: The Tumultuous History of the Search for Artificial Intelligence*. New York: BasicBooks .
7. Dwork, C. (2012). *Fairness Through Awareness*. ACM.
8. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. Cambridge: MIT Press.
9. Chesney, R., & Citron, D. (2019). *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*. California Law Review.
10. Koukolík, F. (2024). *Nezapomenu! Navždy vděčná vaše umělá inteligence*. Praha: Galén.
11. Kučera, P. (2021). *Kybernetická bezpečnost a ochrana dat ve firemním prostředí*. Brno: Masarykova univerzita.
12. *Kybernetická bezpečnost v roce 2023*. (2023). Načteno z NÚKIB: <https://www.nukib.cz>
13. Lobotka, A. (2023). *Umělá inteligence z pohledu antidiskriminačního práva a GDPR*. Praha: Wolters Kluwer.
14. Mařík, V., Štěpánková, O., & Lažanský, J. (1993). *Umělá inteligence*. Praha: Academia.
15. Mařík, V., Trčka, M., & Černý, D. (2024). *Proč se nebát umělé inteligence? AI pohledem nejen českých odborníků*. Brno: JOTA.
16. McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (1955). *A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence*. Hanover: Dartmouth College.
17. Mizrahi, A. (2024). *AI Governance: Policy and Ethics in the Age of Automation*. New York: Routledge.

18. *Obecné nařízení o ochraně osobních údajů (GDPR)*. (2016). Načteno z Evropská komise: <https://eur-lex.europa.eu/eli/reg/2016/679/oj>
19. O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown Publishing Group.
20. Reynolds, E., & McDonnell, T. (2021). *AI and Human Rights: Challenges and Recommendations*. Načteno z Amnesty International: <https://www.amnesty.org/en/documents/>
21. Rich, E., & Knight, K. (1991). *Artificial Intelligence*. New York: McGraw-Hill.
22. Russell, S. (2021). *Jako člověk. Umělá inteligence a problém jejího ovládní*. Praha: Dokořán.
23. Samek, W. (2019). *Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models*. Načteno z Cornell University: <https://arxiv.org/abs/1708.08296>
24. Shokri, R. (2017). *Membership Inference Attacks Against Machine Learning Models*. Načteno z Cornell University: <https://arxiv.org/abs/1610.05820>
25. Schwab, K. (2016). *The Fourth Industrial Revolution*. Ženeva: World Economic Forum.
26. Šmídová, L. (2020). *GDPR v praxi: Průvodce ochranou osobních údajů*. Praha: Grada.
27. Šulc, S. (2024). *Umělá inteligence: Vítejte v nové realitě*. Praha: Práh.
28. Topol, E. (2019). *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. New York: Basic Books.
29. Vincent, M. (2022). Artificial Intelligence in Archaeology: Potentials and Challenges. *Journal of Archeological Method and Theory*.
30. Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs.