

Modul 3

Ochrana soukromí a
osobních údajů v
návrhu systémů
umělé inteligence



Co-funded by
the European Union



Ethical Engineer

www.ethicalengineer.eu

OBSAH

- 01 Vymezení rizik pro soukromí v systému
- 02 Zahrnutí ochrany soukromí do architektury
- 03 Řízení, hodnocení a poskytování práv
- 04 Ochrana soukromí: Jak se správně rozhodnout?



Co-funded by
the European Union

Tento projekt byl financován s podporou Evropské komise. Za tuto publikaci (sdělení) nese výlučnou odpovědnost autor a Komise nenese žádnou odpovědnost za jakékoli použití informací v ní obsažených. V souladu s novým rámcem GDPR upozorňujeme, že partnerství bude zpracovávat vaše osobní údaje výhradně v zájmu a za účelem projektu a bez jakéhokoli dotčení vašich práv.



Kurz „Etický inženýr“ v kostce

Hlavní obsah a výstupy pro každý modul

Modul 01	Modul 02	Modul 03	Modul 04	Modul 05	Modul 06
Etika a transparentnost	Zaujatost, spravedlnost a inkluze	Soukromí a ochrana osobních údajů	Odpovědnost a bezpečnost	Návrh založený na hodnotách	Udržitelnost a dopad
Etické zásady a hodnoty EU Transparentnost a vysvětlitelnost Analýza případů z oblasti umělé inteligence a robotiky	Zaujatost v průběhu celého životního cyklu umělé inteligence Cíle spravedlnosti a důkazy Příklady inkluzivního inženýrství	Ochrana soukromí již od návrhu a GDPR Správa dat a PET Hodnocení a zpochybnitelnost	Role a odpovědnost Selhání, odolnost a předání Protokolování a poučení z incidentů	Požadavky na etický design Volba dat, modelů a rozhraní Kompromisy a agilní kontrolní body	M Energie, voda a hardware Dopady na společnost a životní cyklus Metriky, správa a přepracování

Představení modulu

Tento modul převádí zásady ochrany soukromí do architektury, podložených důkazů a zodpovědných inženýrských rozhodnutí.

Účel

Vybavit studenty technických oborů schopností identifikovat rizika v oblasti ochrany soukromí a navrhovat přiměřená technická a organizační opatření.

Cílová skupina

Studenti vysokých škol v oborech inženýrství, informatiky, robotiky a příbuzných oborech.

Navazuje na

Moduly 1 a 2: etické zásady, transparentnost, analýza zaujatosti a záznamy o inženýrských rozhodnutích.

[Pokyny pro lektora]

Časová náročnost: 3 minuty. Před zahájením výuky vysvětlíte posloupnost modulů a očekávané výstupy studentů. Přístupnost: předem poskytněte výstupy učení a pracovní list.

Představení modulu

Tento modul převádí zásady ochrany soukromí do architektury, podložených důkazů a zodpovědných inženýrských rozhodnutí.

Průběh výuky

Definujte rizika v oblasti ochrany soukromí, navrhnete kontrolní mechanismy, posudíte povinnosti a obhájte inženýrské rozhodnutí týkající se ochrany soukromí.

Výstupy účastníků

Mapa životního cyklu dat, stručný popis architektury ochrany soukromí, screening DPIA a záznam o inženýrském rozhodnutí.

[Pokyny pro lektora]

Časová náročnost: 3 minuty. Před zahájením výuky vysvětlíte posloupnost modulu a výstupy účastníků.
Přístupnost: předem poskytněte výstupy učení a pracovní list.

Výstupy učení

Na konci modulu budou studenti schopni:

1. Zmapovat osobní údaje, aktéry, účely a rizika pro soukromí v rámci celého životního cyklu systému umělé inteligence.
2. Aplikovat principy „ochrany soukromí již od návrhu“, „ochrany soukromí jako výchozího nastavení“ a GDPR na technické požadavky.
3. Vybrat technologie zvyšující bezpečnost a ochranu soukromí pro danou hrozbu a kontext použití.
4. Posoudit projekt z hlediska posouzení dopadu na ochranu osobních údajů a navrhnout jeho zpochybnitelnost.
5. Zaznamenejte kontrolní opatření, zbytkové riziko, monitorování a technické rozhodnutí s vyvozenou odpovědností

Ochrana soukromí se stává technickým problémem, když změny v návrhu systému zvyšují míru vystavení rizikům

Systémy umělé inteligence mohou pozorovat chování, propojovat záznamy, odvozovat citlivé charakteristiky a ovlivňovat významná rozhodnutí. Riziko pro soukromí proto závisí na architektuře, účelu a použití, stejně jako na důvěrnosti.

Technická otázka je konkrétní: Která data nebo závěry jsou nezbytné pro danou funkci, kdo k nim má přístup nebo s nimi může nakládat, jaká škoda z toho může vyplynout a jaké důkazy podporují daný návrh?

Odůvodněná odpověď propojuje dotčené osoby s požadavky, toky dat, technickými kontrolními mechanismy, právy, mazáním, reakcí na incidenty a zodpovědným schválením.

Rámce propojují povinnosti v oblasti ochrany soukromí s provozními důkazy

Evropské právní předpisy a pokyny definují, proč je nutné chránit soukromí. Provozní rámce pomáhají týmům proměnit tyto povinnosti v opakovatelné technické důkazy.

Požadavky EU:

Evropské etické pokyny EU identifikují ochranu soukromí a správu dat jako podmínku pro důvěryhodnou AI. GDPR stanovuje zásady, práva a odpovědnost. Zákon o umělé inteligenci doplňuje pro příslušné systémy umělé inteligence povinnosti v oblasti správy dat, dokumentace a dohledu

Provozní podpora:

NIST AI RMF organizuje práci v oblasti rizik prostřednictvím fází „Govern“ (řízení), „Map“ (mapování), „Measure“ (měření) a „Manage“ (správa). ISO/IEC 27701:2025 specifikuje systém řízení informací o ochraně soukromí pro správce a zpracovatele osobních údajů (PII).

Vysvětlíte rozdíl mezi požadavky a podporou při implementaci. Evropské právní předpisy a pokyny stanovují povinnosti a očekávání. NIST AI RMF a ISO/IEC 27701 mohou týmům pomoci organizovat podklady pro řízení rizik a ochrany osobních údajů. Zeptejte se účastníků, který artefakt v modulu demonstruje jednotlivé požadavky.

[Zdroje]

<https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
<https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng>
<https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>
<https://www.nist.gov/itl/ai-risk-management-framework>
<https://www.iso.org/standard/27701>

Rámce propojují povinnosti v oblasti ochrany soukromí s provozními důkazy

Evropské právní předpisy a pokyny definují, proč je nutné chránit soukromí. Provozní rámce pomáhají týmům proměnit tyto povinnosti v opakovatelné technické důkazy.

Důkazy v tomto modulu:

Datové mapy, modely hrozeb, architektonická rozhodnutí, screening DPIA, postupy pro uplatnění práv, testy, vlastníci a spouštěče přezkumu.

Otevřeno:

- Etické pokyny EU
- GDPR
- Zákon EU o umělé inteligenci
- NIST AI RMF
- ISO/IEC 27701:2025

Vysvětlíte rozdíl mezi požadavky a podporou při implementaci. Evropské právní předpisy a pokyny stanovují povinnosti a očekávání. NIST AI RMF a ISO/IEC 27701 mohou týmům pomoci organizovat podklady pro řízení rizik a ochrany osobních údajů. Zeptejte se účastníků, který artefakt v modulu demonstruje jednotlivé požadavky.

[Zdroje]

<https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
<https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng>
<https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>
<https://www.nist.gov/itl/ai-risk-management-framework>
<https://www.iso.org/standard/27701>

01

Definujte rizika související s ochranou soukromí v rámci systému

Zmapujte účel, osobní údaje, aktéry,
závěry, sdílení a uchovávání informací a
dotčené osoby.

Riziko narušení soukromí je vlastností systému

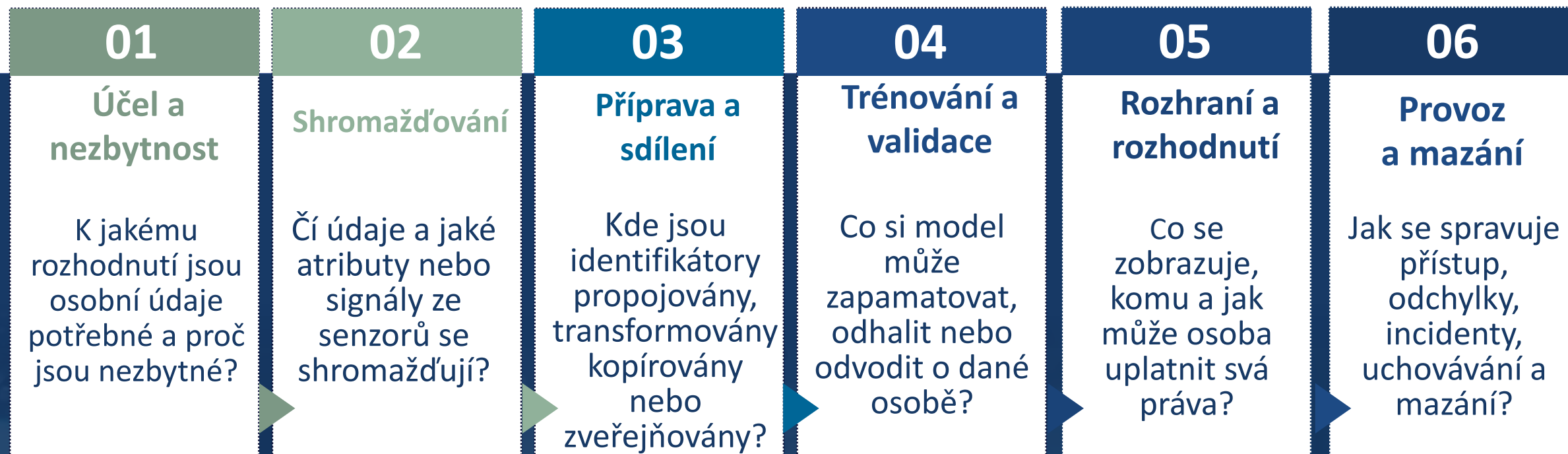
Ochrana soukromí se týká schopnosti osoby porozumět a ovlivnit to, jak jsou o ní shromažďovány, propojovány, odvozovány, využívány a uchovávány informace. Systém může být zabezpečen proti neoprávněnému přístupu a přesto může vést k rušivému sledování, neočekávanému opětovnému použití nebo nespravedlivé ztrátě autonomie.

Riziko vzniká skrze účel, výběr senzorů, identifikátory, propojování dat, chování modelů, rozhraní, pravidla přístupu, vztahy s dodavateli a uchovávání dat.

Inženýři proto zkoumají celý sociotechnický systém, jeho datové toky a rozhodnutí, která podporuje.

Zmapujte ochranu soukromí v průběhu celého životního cyklu systému umělé inteligence

Vyberte jedno inženýrské rozhodnutí. V každé fázi identifikujte data, dotčené osoby, riziko pro soukromí, důkazy a kontrolní mechanismy.



CAPSTONE: Mapa životního cyklu dat propojující každou fázi s účelem, dotčených osob, důkazů, rizika pro soukromí a praktické kontroly.

Analýza ochrany soukromí: Caresecure – detekce pádu

Případová studie:

Ethical Engineer : CareSecure
– detekce pádů.

Rozhodnutí:

Schválit domácí monitorování staršího obyvatele. V každé fázi se zaznamenávají data, dotčené osoby, riziko pro soukromí, důkazy a kontrola.

01

Účel a nezbytnost

- **Údaje:** pohyb potřebný k detekci pádu. Dotčené osoby: obyvatelé a návštěvníci.
- **Riziko/kontrola:** zakázat sledování chování; zdokumentovat bezpečnostní účel a uživatelská nastavení.

02

Shromažďování

- **Údaje:** pohyb na úrovni pokoje a krátká nouzová rezerva. Dotčené osoby: obyvatelé, pečovatelé a návštěvníci.
- **Důkazy/kontrola:** inventární senzory; lokální zpracování a zamezení nepřetržitému ukládání videozáznamů.

[Pokyny pro lektora] Časová náročnost: 10 minut. Odhalujte jednotlivé fáze postupně a požádejte účastníky, aby předpověděli další kontrolní opatření. Zdůrazněte, že uvedený příklad představuje odůvodněný návrh, nikoli univerzální právní odpověď.

Dokončený postup ochrany soukromí: Caresecure – detekce pádu

Případová studie:

Etický inženýr CareSecure –
detekce pádů.

Rozhodnutí:

Schválit domácí monitorování pro staršího obyvatele. V každé fázi se zaznamenávají data, dotčené osoby, riziko pro soukromí, důkazy a kontrola.

03

○ Příprava a sdílení

- **Údaje:** ID zařízení, čas výstrahy a nouzový záznam. Dotčené osoby: obyvatel a schválení pečovatelé.
- **Důkazy/kontrolní opatření:** oddělená identita; šifrování přenosu a vynucení přístupu na základě rolí.

04

○ Školení a ověření

- **Údaje:** souhlasené, simulované a reprezentativní vzorce pádů. Riziko: zapamatování si nebo rutinní dedukce.
- **Důkazy/kontrola:** zabránit exportu surových dat; testovat úniky, falešné poplachy a provozní limity.

[Pokyny pro lektora] Časová náročnost: 10 minut. Odhalujte jednotlivé fáze postupně a požádejte účastníky, aby předpověděli další kontrolní opatření. Zdůrazněte, že uvedený příklad představuje odůvodněný návrh, nikoli univerzální právní odpověď.

Dokončený postup ochrany soukromí:

Caresecure – detekce pádu

Případová studie:

Ethical Engineer CareSecure
– detekce pádů.

Rozhodnutí:

Schválit domácí monitorování pro staršího obyvatele. V každé fázi se zaznamenávají data, dotčené osoby, riziko pro soukromí, důkazy a kontrola.

05

Rozhraní a rozhodnutí

- **Data:** výstraha a spolehlivost. Riziko: výstraha může odhalit zdravotní stav nebo přítomnost osoby.
- **Důkazy/kontrolní opatření:** minimalizujte počet zpráv; před eskalací ověřte situaci a zajistěte přístupné ovládací prvky.

06

Provoz a mazání

- **Data:** auditní protokol a přezkoumaný záznam o incidentu. Riziko: postupné prodlužování doby uchovávání, přístup dodavatele nebo narušení bezpečnosti.
- **Důkazy/kontrola:** plán automatického mazání, kontrola přístupu, postup při incidentech a spouštěč pozastavení.

[Pokyny pro lektora] Časový rámec: 10 minut. Odhalujte jednotlivé fáze postupně a požádejte účastníky, aby předpověděli další kontrolní opatření. Zdůrazněte, že uvedený příklad představuje odůvodněný návrh, nikoli univerzální právní odpověď.

Dokončená cesta k ochraně soukromí: Caresecure – detekce pádů

Případová studie:

Ethical Engineer – detekce
pádů CareSecure.

Rozhodnutí

Schválit omezený pilotní projekt pouze po provedení posouzení dopadu na ochranu osobních údajů (DPIA), stanovení jasných provozních hranic, zajištění lokálního zpracování, přístupu na základě rolí, otestovaných výstrah, definovaného mazání a postupu pro pozastavení nebo ukončení.

[Pokyny pro lektora] Časový rámec: 10 minut. Odhalujte jednotlivé fáze postupně a požádejte účastníky, aby předpověděli další kontrolní opatření. Zdůrazněte, že uvedený příklad představuje odůvodněný návrh, nikoli univerzální právní odpověď.

Osobní údaje zahrnují identifikátory, signály a přiměřené závěry

- Osobní údaje jsou informace týkající se identifikované nebo identifikovatelné osoby. Jasnými příklady jsou jména a čísla účtů. K identifikaci nebo vymezení osoby mohou sloužit také stopy polohy, identifikátory zařízení, hlas, obrázky, pracovní vzorce a kombinace dat ze senzorů.
- Mezi zvláštní kategorie podle GDPR patří údaje o zdraví a biometrické údaje používané k jednoznačné identifikaci osoby. Umělá inteligence může vytvářet citlivé závěry, i když se vstupní data jeví jako běžná.

Každé pole a každý závěr klasifikujte, zaznamenejte, proč je potřebný, a identifikovatelnost považujte za riziko závislé na kontextu.

Účel, minimalizace a zákonný základ utvářejí technický návrh

- Omezení účelu vyžaduje konkrétní a legitimní použití. Minimalizace údajů omezuje shromažďování a zpracování na to, co je přiměřené, relevantní a nezbytné pro dané použití. Před zahájením zpracování musí být stanoven zákonný základ. Souhlas je jedním z možných základů, ale může být nevhodný v případě, že osoba nemá skutečnou možnost volby.
- Převádějte tyto zásady do požadavků: odstraňte nepoužívaná pole, oddělte volitelné funkce, omezte sekundární použití, omezte rozhraní a stanovte pravidla pro uchovávání a mazání.

I užitečné datové pole stále vyžaduje legitimní účel, platný právní základ, kontrolovaný přístup a důkaz nezbytnosti.

Porušení soukromí zahrnuje odhalení, vyvozování závěrů, sledování a ztrátu svobodné vůle

- K porušení může dojít v důsledku úniku osobních údajů, k poškození soukromí však může dojít i bez porušení. Trvalé sledování může změnit chování. Propojení dat může odhalit totožnost. Modely mohou odvodit zdravotní stav, přesvědčení, polohu nebo zranitelnost. Automatizovaná rozhodnutí mohou osobu zviditelnit pro instituce, zatímco systém zůstává pro tuto osobu neprůhledný.
- Posuďte, kdo se může dozvědět co, jak jistý je daný závěr, jaké kroky z toho mohou vyplynout a zda daná osoba může tento závěr pochopit, opravit nebo zpochybnit.

Zabezpečení chrání data. Inženýrství soukromí také omezuje sledování, odvozování, účel, moc a uchovávání.

Účast zlepšuje požadavky na ochranu soukromí

- Dotčené osoby mohou odhalit očekávání ohledně soukromí, která technický tým nemusí vidět. Obyvatelé, pracovníci, pacienti, návštěvníci, provozovatelé, údržbáři a náhodní kolemjdoucí mohou vnímat stejný senzor nebo model odlišně.
- Využijte rozhovory, společný návrh, reprezentativní testování nebo poradní panel k prozkoumání přijatelných účelů, soukromých prostor, viditelných indikátorů, ovládacích prvků, přístupných oznámení, možnosti odstoupení, opravy a nápravy. Účast musí mít jasně definovanou cestu k požadavkům a schválení.

Zaznamenejte, či názory ovlivnily návrh a jaké nevyřešené obavy přetrvávají.

Praktický příklad ze strojírenství: Exoskelet s ohledem na ochranu soukromí

Továrna navrhuje průmyslový exoskeleton vybavený senzory, který má zlepšit přizpůsobení, detekovat nebezpečné zatížení a plánovat údržbu. Návrh ochrany soukromí určuje, zda systém podporuje pracovníky, nebo se stane nástrojem pro monitorování.

[Pokyny pro lektora]

Čas: 12 minut. Použijte tento příklad k upevnění tématu ochrany soukromí v oboru strojírenství. Zeptejte se, který požadavek by měl být zahrnut do specifikace zařízení a do smlouvy o dodávce.

01

Počáteční zadání

Shromažďujte údaje o úhlech kloubů, silách, stavu zařízení a srdeční frekvenci, aby inženýři mohli vylepšit přizpůsobení a předvídat selhání komponent.

02

Riziko narušení soukromí

Biomechanika může odhalit zdravotní stav, postižení, únavu a tempo. Lze identifikovat zaměstnance, dodavatele a návštěvníky.

03

Upravovaný účel

Podpora bezpečnosti zařízení a individuálního přizpůsobení. Vyloučit hodnocení produktivity, vynucování docházky a disciplinární použití.

Praktický příklad ze strojírenství: Exoskelet s ohledem na ochranu soukromí

Továrna navrhuje průmyslový exoskeleton vybavený senzory, který má zlepšit přizpůsobení, detekovat nebezpečné zatížení a plánovat údržbu. Návrh s ohledem na ochranu soukromí určuje, zda systém podporuje pracovníky, nebo se stane nástrojem pro monitorování.

[Pokyny pro lektory]

Čas: 12 minut. Použijte tento příklad k upevnění tématu ochrany soukromí v oboru strojírenství. Zeptejte se, který požadavek by měl být zahrnut do specifikace zařízení a do smlouvy o dodávce.

04

Architektura

Vyčleňte bezpečnostní funkce zařízení, oddělte identitu pracovníka, přepište surová data z vyrovnávacích pamětí senzorů a zašifrujte události údržby.

05

Přístup a práva

Poskytněte pracovníkům jejich vlastní záznamy. Inženýrům poskytněte agregované trendy. Omezte výklad individuálních zdravotních údajů pouze na oprávněné zaměstnance.

06

Rozhodování a přezkum

Pilotní provoz provádějte pouze na základě zdokumentovaného zákonného základu, za účasti zástupců zaměstnanců, s provedením posouzení dopadu na ochranu osobních údajů (DPIA), plánem mazání údajů, postupem pro podávání stížností a spouštěčem pro zastavení.

Praktický příklad ze strojírenství: Exoskelet s ohledem na ochranu soukromí

Továrna navrhuje průmyslový exoskeleton vybavený senzory, který má zlepšit přizpůsobení, detekovat nebezpečné zatížení a plánovat údržbu. Návrh s ohledem na ochranu soukromí určuje, zda systém podporuje pracovníky, nebo se stane nástrojem pro monitorování.

Ochrana soukromí změnila návrh

Úkol monitorování stavu se proměnil v ohraničený bezpečnostní systém s lokálním zpracováním dat, právy pracovníků a specifikací zakázaného použití.

[Pokyny pro lektora]

Čas: 12 minut. Použijte tento příklad k upevnění tématu ochrany soukromí v oboru strojírenství. Zeptejte se, který požadavek by měl být zahrnut do specifikace zařízení a do smlouvy o dodávce.

02

Zpracujte ochranu soukromí do architektury

Aplikujte zásadu ochrany soukromí již při návrhu, zabezpečení po celou dobu životního cyklu a vhodné technologie zvyšující ochranu soukromí.

Ochrana soukromí již od návrhu je inženýrská metoda

Zásada „soukromí již od návrhu“ zakotvuje ochranu údajů do systému a jeho provozního modelu již v nejranější fázi návrhu. Zásada „soukromí jako výchozí nastavení“ zajišťuje, že ochranná konfigurace platí, aniž by bylo nutné, aby uživatel vyhledával a měnil nastavení.

Sedm zásad „soukromí již od návrhu“ poskytuje užitečné vodítko. Inženýrská práce je proměňuje v požadavky, rozhraní, testy, záznamy a kontrolní mechanismy v rámci celého životního cyklu.

[Pokyny pro lektory]

Doba trvání: 8 minut. Požádejte skupiny, aby každou ze sedmi zásad „ochrany soukromí již od návrhu“ spojily s jedním inženýrským artefaktem.

Ochrana soukromí již od návrhu je inženýrská metoda

01

Účel

Definujte funkci, dotčené osoby, nezbytné údaje a zakázaná použití.
Zpochybňujte jakýkoli sběr údajů, který nemá konkrétní opodstatnění.

02

Architektura

Do systému začleňte minimalizaci, oddělení, lokální zpracování, řízení přístupu, bezpečné výchozí nastavení, uchovávání a mazání.

03

Odpovědnost

Ověřujte kontrolní mechanismy, dokumentujte rozhodnutí, poskytněte práva a nápravná opatření, sledujte využití a reagujte na incidenty a změny.

Ochrana soukromí již od návrhu je kontinuální inženýrská metoda. Proaktivní prevence, ochrana ve výchozím nastavení, zabudované kontrolní mechanismy, návrh s pozitivním součtem, bezpečnost v průběhu celého životního cyklu, transparentnost a respekt k osobě musí být patrné v dokumentaci i v praxi.

Ochrana soukromí ve výchozím nastavení stanovuje bezpečnou základní úroveň

- Výchozí nastavení ovlivňuje skutečné výsledky, protože mnoho lidí je nikdy nemění. Výchozí nastavení chránící soukromí shromažďuje minimální množství dat, vystavuje nejmenší okruh uživatelů, umožňuje nejméně rušivé funkce a uchovává informace po nejkratší odůvodnitelné dobu.
- Volitelná analytika, sdílení nebo sekundární využití by měly vyžadovat jasnou a samostatnou volbu, přičemž souhlas je vhodným základem. Vypnutí funkce by mělo zastavit sběr dat a spustit stanovené pravidlo uchovávání.

Otestujte první zkušenost s používáním, resetování chování, aktualizaci softwaru a konfiguraci správcem. Výchozí nastavení musí obstát při běžném provozu.

Mapa toku dat odhaluje hranice důvěry a skryté kopie

- Začněte s fyzickým a digitálním systémem. Zaznamenejte každý senzor, uživatelský vstup, odvozenou vlastnost, model, protokol, mezipaměť, řídicí panel, export, rozhraní dodavatele a zálohu. U každého toku identifikujte odesílatele, příjemce, účel, identifikátor, umístění, pravidlo přístupu a dobu uchovávání.
- Označte hranice důvěry, kde se data přesouvají mezi zařízeními a cloudem, mezi jednou organizací a druhou nebo mezi operačním systémem a výzkumným prostředím. Zahrňte cesty selhání, jako jsou offline fronty, ladicí protokoly a přístup k podpoře.

Tato mapa se stává podkladem pro minimalizaci, návrh zabezpečení, posouzení dopadu na ochranu osobních údajů (DPIA), smlouvy, testování mazání a reakci na incidenty.

Pseudonymizace snižuje riziko propojení, ale neodstraňuje povinnosti vyplývající z GDPR

- Pseudonymizace nahrazuje přímé identifikátory a zajišťuje oddělenou ochranu doplňujících informací. Data zůstávají osobními údaji, pokud je jejich opětovné propojení přiměřeně možné. Může snížit riziko odhalení a podpořit oddělení povinností.
- Cílem anonymizace je zabránit identifikaci prostředky, jejichž použití je přiměřeně pravděpodobné. Pouhé odstranění jmen je nedostatečné, pokud lze osobu identifikovat na základě jedinečných trajektorií, vzácných podmínek nebo propojených datových sad.

Zaznamenejte útočníka, pomocná data, cestu propojení a test opětovné identifikace. Proveďte přehodnocení, jakmile budou k dispozici nové datové sady nebo možnosti.

Zabezpečení v průběhu životního cyklu podporuje důvěrnost, integritu a kontrolované použití

- Šifrování chrání data při přenosu i v klidu, zatímco správa klíčů určuje, zda je tato ochrana skutečně účinná. Přístup založený na rolích, princip nejmenších oprávnění, vícefaktorová autentizace, bezpečné aktualizace, protokolování a kontrola omezují neoprávněný přístup. Kontroly integrity a původ dat chrání rozhodnutí před pozměněnými nebo neúplnými daty.
- Aplikujte kontrolní opatření na trénovací data, artefakty modelů, výzvy, výstupy, protokoly, zálohy a fronty pro mazání. Testujte přístup dodavatelů a správců i externí útoky.

Kompletní kontrolní mechanismus definuje hrozbu, komponentu, konfiguraci, důkazy, vlastníka, reakci na selhání a spouštěč kontroly.

Plochy pro útoky na soukromí přesahují rámec uložených dat

Umělá inteligence může způsobit únik osobních údajů prostřednictvím trénovacích záznamů, chování modelů, aktualizací, rozhraní a metadat. Šifrování chrání důležité cesty, zatímco jiné cesty vyžadují specifická opatření pro daný model.

Inverze nebo rekonstrukce modelu:

Dotazy, výstupy, gradienty nebo parametry mohou odhalit atributy nebo přibližné trénovací příklady.

Odvození členství:

Útočník testuje, zda konkrétní osoba nebo záznam tvořily součást trénovacích dat.

Únik gradientu:

Sdílené aktualizace v distribuovaném učení mohou odhalit informace o lokálních příkladech.

[Zdroje]

- <https://csrc.nist.gov/pubs/ai/100/2/e2025/final>
- <https://www.enisa.europa.eu/sites/default/files/publications/ENISA%20Report%20-%20Securing%20Machine%20Learning%20Algorithms.pdf>

Plochy pro útoky na soukromí přesahují rámec uložených dat

Umělá inteligence může způsobit únik osobních údajů prostřednictvím trénovacích záznamů, chování modelů, aktualizací, rozhraní a metadat. Šifrování chrání důležité cesty, zatímco jiné cesty vyžadují specifická opatření pro daný model.

Scraping a postranní kanály:

Odhalené koncové body, protokoly, velikost paketů, časování nebo jiná metadata mohou odhalit data, chování modelu nebo citlivá témata.

Technická reakce:

Definujte přístup útočníků, omezte výstupy a dotazy, zabezpečte aktualizace, testujte úniky, sledujte zneužití a určete, kdy je třeba nasazení zastavit.

[Zdroje]

- <https://csrc.nist.gov/pubs/ai/100/2/e2025/final>
- <https://www.enisa.europa.eu/sites/default/files/publications/ENISA%20Report%20-%20Securing%20Machine%20Learning%20Algorithms.pdf>

Federované učení uchovává surová trénovací data lokálně, ale vyžaduje dodatečná bezpečnostní opatření

- Federované učení umožňuje zařízením nebo organizacím přispívat k aktualizacím modelů, aniž by sdružovaly své surové datové sady. Může omezit centrální sběr dat a podporovat lokalizaci dat.
- Aktualizace modelů však mohou i nadále odhalovat informace prostřednictvím úniku gradientů, odvození členství nebo zlovolného účastníka. Centrální koordinace, ověřování aktualizací a kompromitace zařízení zůstávají součástí modelu hrozeb.

V odůvodněných případech používejte bezpečnou agregaci, ořezávání aktualizací, diferenciální soukromí, ověřování účastníků a testování úniků. Zaznamenávejte, co opouští každé zařízení a co může koordinátor pozorovat.

Bezpečné více stranové výpočty podporují společnou analýzu bez sdílení vstupů

- Bezpečné víceúčastnické výpočty umožňují stranám vypočítat dohodnutou funkci, přičemž jejich vstupy zůstávají před ostatními skryty. Sdílení tajemství nebo kryptografické protokoly mohou podporovat meziorganizační statistiky, hodnocení modelů nebo porovnávání.
- Protokol chrání vstupy za stanovených předpokladů. Nečiní však nebezpečný výstup neškodným, neopravuje nekvalitní data, nezabraňuje koluzi přesahující prahovou hodnotu ani nerozhoduje o legitimitě výpočtu. Použití může omezovat také výkon a provozní složitost.

Specifikujte účastníky, model hrozeb, předpoklad koluze, odhalený výstup, řešení selhání a nezávislé ověření.

Diferenciální soukromí poskytuje měřitelnou hranici individuálního příspěvku

- Diferenciální soukromí přidává kalibrovanou náhodnost tak, aby přítomnost či absence dat jedné osoby měla omezený vliv na zveřejněný výsledek. Sílu záruky určuje parametr soukromí a složení opakovaných zveřejnění.
- Šum ovlivňuje užitečnost, zejména u malých skupin nebo vzácných událostí. Prohlášení o záruce je neúplné bez uvedení mechanismu, jednotky ochrany, rozpočtu na ochranu soukromí, pravidla ořezávání, složení a důkazů o implementaci.

Použijte diferenciální soukromí pro definovaný proces zveřejnění nebo učení a poté otestujte jak únik soukromí, tak technický výkon.

Vyberte technologie zvyšující ochranu soukromí pro danou hrozbu

- Technologie zvyšující ochranu soukromí je návrhovým opatřením, nikoli obecným označením. Vyberte ji až po definování toku dat, útočníka, jeho schopností, chráněných informací a požadovaného výstupu.
- Federované učení omezuje centralizaci surových dat. Bezpečné víceúčastnické výpočty skrývají vstupy během společného výpočtu. Diferenciální soukromí omezuje informace odhalené výstupy nebo aktualizacemi. Vedle nich může být vyžadováno šifrování, řízení přístupu, zabezpečený hardware a lokální zpracování.

Porovnejte záruku ochrany soukromí, předpoklady, výkon, režimy selhání, provozní zátěž a ověření. Kombinujte kontrolní opatření pouze tehdy, pokud každé z nich řeší zdokumentované riziko.

Edge, IoT a generativní AI mění model ohrožení soukromí

01

Expozice na okraji sítě a v IoT:

Lokální zpracování může omezit přenos dat do cloudu. Terénní zařízení mohou čelit fyzickému přístupu, slabé identifikaci, omezeným zdrojům, nepravidelným aktualizacím, distribuovaným protokolům a obtížnému mazání dat

02

Opatření pro okrajové zařízení a IoT:

Používejte bezpečné spouštění, identifikaci zařízení, princip minimálního množství dat, krátké lokální vyrovnávací paměti, podepsané aktualizace, segmentaci sítě a kompletní inventář po celou dobu životního cyklu.

03

Ovládací prvky generativní AI:

Definujte schválené datové cesty, redigování, oddělení přístupu, uchovávání, podmínky poskytovatele, kontrolu výstupů a testy úniku dat.

[Zdroje]

- <https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-generative-artificial-intelligence>
- <https://nvlpubs.nist.gov/nistpubs/specialpublications/NIST.sp.1800-36.pdf>

Edge, IoT a generativní AI mění model ohrožení soukromí

04

Rizika spojená s generativní AI:

Podněty, vyhledané dokumenty, data pro doladění, výstupy a protokoly mohou obsahovat osobní nebo důvěrné informace. Modely si mohou zapamatovat, odvodit nebo vytvářet nepřesné výroky o lidech.

05

Kontrola rozhodnutí:

Zaznamenejte, co zůstává lokálně, co odchází, co lze odvodit, jak se ověřují kontroly a kdo může pozastavit používání.

[Zdroje]

- <https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-generative-artificial-intelligence>
- <https://nvlpubs.nist.gov/nistpubs/specialpublications/NIST.sp.1800-36.pdf>

Případová studie: Faceboard – rozpoznávání obličejů na letišti

Otevřít případovou studii Ethical Engineer FaceBoard a určete, která architektura dokáže splnit účel odbavení s co nejmenším zásahem do soukromí.

- Případová studie „FaceBoard“ z kurzu „The Ethical Engineer“ navrhuje využití rozpoznávání obličeje při nástupu na palubu na letišti. Systém může snížit fronty a omezit podvody, přičemž zpracovává biometrická data v prostředí, kde odmítnutí může mít dopad na cestování. Případová studie vyvolává otázky týkající se účelu, nezbytnosti, zaujatosti, porušení, uchovávání dat a alternativních řešení.
- **Technické zaměření:** vytváření šablon v zařízení, centrální úložiště, ověřování typu „jeden k jednomu“, vymazání po cestě, přístup dodavatele, detekce útoků typu „presentation attack“, lidská kontrola a nebiometrická alternativa.

[Pokyny pro lektora] Časový rámec: 12 minut. Studenti by si měli před hodinou přečíst případovou studii v odkazu nebo využít shrnutí na snímku. Poskytněte textový přepis a před diskusí dejte studentům prostor pro individuální reflexi.

Praktický úkol: Napište rozhodnutí o architektuře ochrany soukromí

[Pokyny pro lektora]

Časový limit: 25 minut. Posuďte srozumitelnost účelu, přesnost toku dat, vhodnost opatření pro kontrolu hrozeb, testovatelnost, práva a odůvodnění zbytkového rizika.

- Použijte FaceBoard nebo příklad exoskeletu. Vytvořte návrhový brief, který by mohl jiný inženýrský tým implementovat a otestovat.
- Definujte funkci, dotčené osoby, osobní údaje, účel, otázku zákonného základu, hranice důvěry a model hrozeb. Porovnejte dvě architektury. Specifikujte minimalizaci, oddělení, přístup, šifrování, uchovávání, mazání, práva, monitorování a kontroly incidentů.

Přidejte akceptační testy, odpovědného vlastníka, zbytkové riziko a podmínku, která vyžaduje REVIZE nebo ZASTAVENÍ.

Výstup

Jednostránkový návrh architektury a dvouminutové odůvodnění založené na důkazech.

03

Řízení, posuzování a poskytování práv

Propojte správu dat podle GDPR, DPIA, zásady napadnutelnosti a zákona o umělé inteligenci s technickými důkazy.

Zásady GDPR se stávají technickými požadavky

- Článek 5 GDPR stanoví zásady zákonného, spravedlivého a transparentního zpracování; omezení účelu; minimalizace údajů; přesnost; omezení uchovávání; integrita a důvěrnost; a odpovědnost. Tyto zásady platí společně.
- Převádějte je do datové specifikace, zásad povoleného použití, testu kvality dat, modelu přístupu, harmonogramu uchovávání, testu mazání, oznámení, postupu pro uplatnění práv, postupu při incidentech a záznamu o rozhodnutí. Definujte odpovědnosti správce, zpracovatele a dodavatele.

Dodržování předpisů vyžaduje organizační úsudek. Inženýři poskytují důkazy o tom, jak se architektura chová v praxi.

Zákonný základ, souhlas a transparentnost vyžadují soudržný postup

- Správce musí pro každý účel zpracování určit vhodný právní základ. Souhlas vyžaduje skutečnou volbu, konkrétní informace a účinný mechanismus pro jeho odvolání. Nerovnováha sil může způsobit, že souhlas není vhodný, a to včetně některých situací v oblasti zaměstnání a základních služeb. Údaje zvláštní kategorie vyžadují dodatečnou podmínku.
- Transparentnost by měla vysvětlovat, jaké údaje se shromažďují, proč, kdo je přijímá, jak dlouho se uchovávají, jaká práva se na ně vztahují a jak je uplatnit. Rozhraní a skutečné toky dat se musí shodovat s oznámením.

Ověřte možnost odvolání, opravy a výmazu v celém procesu, včetně odvozených údajů, dodavatelů, záloh a modelových procesů.

Automatizovaná rozhodnutí vyžadují smysluplné přezkoumání a možnost napadnutí

- Článek 22 GDPR se týká určitých rozhodnutí založených výhradně na automatizovaném zpracování, která mají právní nebo podobně významné účinky, s výhradou stanovených podmínek a výjimek. Související ustanovení o transparentnosti a přístupu mohou v příslušných případech vyžadovat smysluplné informace o logice a důsledcích.
- Proces přezkumu vyžaduje oprávněnou a kompetentní osobu, přístup k relevantním důkazům, čas na změnu výsledku, svobodu nesouhlasit se systémem a zaznamenaný důvod. Tlačítko, které nemůže rozhodnutí změnit, nepředstavuje účinný lidský zásah.

Navrhňte oznámení, vysvětlení, opravu, eskalaci a nápravu jako systémové funkce a poté je otestujte s dotčenými uživateli.

Posouzení dopadu na ochranu osobních údajů (DPIA) propojuje zpracování s vysokým rizikem s rozhodnutími o návrhu

- A Posouzení dopadu na ochranu osobních údajů je vyžadováno v případech, kdy zpracování pravděpodobně povede k vysokému riziku pro práva a svobody osob. Při posuzování by se mělo zohlednit systematické monitorování, inovativní technologie, zranitelné osoby, rozsáhlé zpracování údajů zvláštních kategorií, porovnávání údajů a významná automatizovaná rozhodnutí.
- DPIA popisuje zpracování a jeho účel, posuzuje nezbytnost a přiměřenost, identifikuje rizika, zaznamenává opatření a zbytkové riziko a v případě potřeby zapojuje příslušné odborníky a dotčené osoby.

Začněte v dostatečném předstihu, abyste mohli změnit požadavky a architekturu. Posouzení přehodnoťte, pokud dojde ke změně účelu, údajů, modelu, rozsahu, dodavatele nebo kontextu.

Dokončená obrazovka DPIA:

CareSecure – detekce pádů v domácnosti



**Zákon EU o umělé
inteligenci, správa dat a
povinnosti v oblasti
ochrany osobních údajů
podle GDPR se
vzájemně prolínají**

- U příslušných vysoce rizikových systémů umělé inteligence zákon o AI vyžaduje řízení rizik, správu dat, technickou dokumentaci, protokolování, lidský dohled, přesnost, robustnost a kyberbezpečnost. Datové sady pro trénování, validaci a testování vyžadují odpovídající správu z hlediska relevance, reprezentativnosti, chyb a mezer.
- GDPR a zákon o AI mají odlišnou působnost. Kvalita dat nevyklučuje omezení účelu ani minimalizaci. Posouzení dopadu na ochranu osobních údajů (DPIA) a posouzení dopadu na základní práva mají odlišné právní podněty a obsah, i když důkazy a kontroly mohou být koordinovány.

Udržujte sledovatelný soubor kontrolních opatření, který zaznamenává každý požadavek, důkazní prvek, vlastníka, zbytkové riziko a bod přezkumu.

Důkazy o ochraně soukromí v oblasti zdravotnických technologií jsou zároveň důkazy o bezpečnosti pacientů

Pokud má software umělé inteligence lékařský účel, rozhodnutí týkající se ochrany soukromí souvisejí s nařízením MDR, zákonem o umělé inteligenci a nařízením GDPR. Účel použití a riziko určují příslušné povinnosti týkající se zdravotnických prostředků.

01

Integrita dat

Sledujte zdroj, zákonný přístup, označení, vyloučení, transformace, chybějící údaje a verzi. Poškozená nebo nereprezentativní klinická data mohou změnit diagnózu, třídění pacientů nebo léčbu.

02

Řízení kvality a rizik

Propojte kontroly dat se systémem řízení kvality, dokumentací o řízení rizik, klinickým hodnocením, kyberbezpečností, řízením změn a sledováním po uvedení na trh.

03

Ochrana soukromí pacientů

Omezte přístup k údajům zvláštní kategorie, oddělte identity, ověřujte přístup a mazání a chraňte protokoly, modelové artefakty a rozhraní dodavatelů.

Důkazy o ochraně soukromí v oblasti zdravotnických technologií jsou zároveň důkazy o bezpečnosti pacientů

Pokud má software umělé inteligence lékařský účel, rozhodnutí týkající se ochrany soukromí souvisejí s nařízením MDR, zákonem o umělé inteligenci a nařízením GDPR. Účel použití a riziko určují příslušné povinnosti týkající se zdravotnických prostředků.

04

Důkazy před uvedením na trh

Klinická výkonnost podle příslušné skupiny, původ dat, testování hrozeb, lidská kontrola, postup při incidentech, zbytkové riziko a schvalovací orgán.

05

Technický závěr

Selhání v oblasti ochrany osobních údajů, které narušuje integritu, dostupnost nebo vhodné použití dat, se může stát selháním z hlediska bezpečnosti zdravotnického prostředku.

[Zdroje]

- <https://eur-lex.europa.eu/eli/reg/2017/745/oj/eng>
- https://health.ec.europa.eu/document/download/b78a17d7-e3cd-4943-851d-e02a2f22bbb4_en
- <https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng>
- <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>

Dilema: Bezpečnostní vyšetřování nebo funkční posun?

- Exoskeleton zaznamená téměř nehodu, při které pracovník zvedá břemeno mimo schválený rozsah. Bezpečnostní tým požaduje údaje o individuální biomechanice za poslední tři měsíce, aby mohl prošetřit výkon zařízení. Vedoucí linky žádá o stejná data, aby mohl identifikovat pracovníky, kteří jsou pomalí nebo nedodržují předpisy. Původní oznámení se týká přizpůsobení, bezpečnosti a údržby.
- U každé žádosti se rozhodněte, zda ji POVOLIT, UPRAVIT nebo ZAMÍTNOUT. Definujte účel, nezbytnost, oprávněný přístup, časové období, zapojení pracovníka, důkazy, výmaz a přezkum.

Vysvětlete, které pravidlo technického oddělení a správy zabraňuje tomu, aby se legitimní bezpečnostní funkce stala sledováním výkonu.

04

**Přijměte a obhajte
rozhodnutí týkající
se ochrany soukromí**

Porovnejte projektové případy,
provedte revizi návrhu a obhajte
rozhodnutí na základě důkazů.

Projektový případ A: Umělá inteligence ve zdravotnictví a domácí monitorování

- Společnost CareSecure využívá lokální senzory k detekci pádu a upozornění pečovatele. Přínos pro bezpečnost je značný, zatímco nepřetržité sledování uvnitř domácnosti může narušit soukromí, autonomii a důvěru. Může dojít také k zachycení návštěvníků a pečovatelů.
- Definujte minimální požadovaný signál, soukromé prostory, lokální zpracování, zveřejnění údajů v nouzových situacích, souhlas nebo jiné otázky týkající se zákonného základu, přístupné ovládací prvky, kontrolu falešných poplachů, mazání údajů a reakci na incidenty.

Otevřete případovou studii Ethical Engineer v oblasti zdravotnictví. Rozhodněte, jaké důkazy jsou nutné před zahájením omezeného pilotního projektu a jaké způsoby použití musí být zakázány.

[Pokyny pro lektora]
předvedení.

Časový rámec: 12 minut. Přidělte role pacienta, pečovatele, inženýra a správce dat. Příkladové role osobních údajů. Přístupnost: povzte písemné popisy rolí namísto živého předvedení.

Projektová případová studie B: AI Witness a citlivá výpověď

- AI Witness zaznamenává výpovědi svědků prostřednictvím konverzačního systému. Tato služba může zachovat detaily a zkrátit zpoždění při shromažďování vysoce citlivých výpovědí, které mohou ovlivnit bezpečnost a spravedlnost. Průběh výslechu může také ovlivnit to, co osoba v tísní prozradí.
- Definujte účel, kategorie údajů, bezpečné zaznamenávání, jazykovou dostupnost, lidskou kontrolu, opravy, zákonné zveřejnění, uchovávání, dopad porušení bezpečnosti, přístup poskytovatele modelu a neautomatizovanou alternativu.

Otevřete případovou studii Ethical Engineer o svědeckých výpovědích. Rozhodněte se, zda by návrh měl automatizované výslechy podporovat, omezovat nebo odmítnout.

[Pokyny pro lektora] Časový rámec: 12 minut. Použijte přístup zohledňující traumatické zážitky. Studenti mohou analyzovat návrh systému, aniž by museli hrát roli rozrušeného svědka.

Porovnejte případy před výběrem opatření na ochranu soukromí

- CareSecure se týká trvalého sledování v soukromém domě za účelem zajištění bezprostřední bezpečnosti. AIWitness se týká časově omezené, ale vysoce citlivé interakce, jejíž záznam může být použit v soudním řízení. Oba případy vyžadují bezpečnost, liší se však jejich účely, mocenské vztahy, doba uchovávání, práva a důsledky.
- U každého případu uveďte subjekt údajů, přihlížející osoby, správce, účel, indikátory vysokého rizika, hranice důvěry, nejhorší uvěřitelnou újmu na soukromí a požadovanou lidskou autoritu.

Užitečný závěr může vést k volbě odlišných architektur, otázek týkajících se právního základu, prahových hodnot pro důkazy a podmínek pro ukončení. Vysvětlete, proč je každé opatření vhodné pro daný kontext.

Týmová aktivita: Proved'te přezkum návrhu ochrany soukromí

Vyberte si CareSecure, FaceBoard, AIWitness nebo exoskeleton. Přidělte role v oblasti systémů, dat, bezpečnosti, práv, zainteresovaných stran a zapisovatele.

1. Definujte funkci, účel a dotčené osoby.
2. Vyplňte mapu životního cyklu dat v šesti fázích.
3. Nakreslete toky dat a hranice důvěry.
4. Zkontrolujte, zda je k dispozici posouzení dopadu na ochranu osobních údajů (DPIA), a vyberte přiměřená kontrolní opatření.
5. Uved'te práva, důkazy, zbytkové riziko a monitorování.
6. Rozhodněte se: POKRAČOVAT, UPRAVIT nebo ZASTAVIT.

Výstup:

Záznam rozhodnutí a dvouminutová prezentace o ochraně soukromí.

[Pokyny pro lektora]

Časový rámec: 30 minut. Vytvořte skupiny po čtyřech až šesti členech. Požádejte každou roli, aby přispěla jedním důkazem nebo jednou nevyřešenou obavou.

Co obsahuje silný návrh ochrany soukromí

- Silný návrh vychází z konkrétní funkce a dokládá, proč je každý prvek osobních údajů nezbytný. Rozlišuje mezi osobními údaji, citlivými závěry a neosobními systémovými údaji. Mapuje aktéry, kopie, hranice důvěry, účely a dobu uchovávání.
- Architektura kombinuje minimalizaci, oddělení, bezpečnost, technologie posilující ochranu soukromí, práva k užívání, lidskou kontrolu, mazání, monitorování a reakci na incidenty. Každé kontrolní opatření má přiřazenou hrozbu, konfiguraci, akceptační test a odpovědnou osobu.

Záznam o rozhodnutí obsahuje otázky týkající se zákonného základu a posouzení dopadu na ochranu osobních údajů (DPIA), připomínky zainteresovaných stran, alternativy, zbytkové riziko, schvalovací orgán, spouštěč přezkumu a zakázaná použití

Pracovní list: Záznam o rozhodnutí v oblasti inženýrství ochrany soukromí

System, funkce a dotčené osoby:

.....

Konkrétní účel a zakázaná použití:

.....

Osobní údaje, citlivé závěry a nezbytnost:

.....

Toky údajů, příjemci a hranice důvěry:

.....

Právní základ a kontrolní otázky k posouzení dopadu na ochranu osobních údajů (DPIA):

.....

Technická kontrola a akceptační test:

.....

[Pokyny pro lektora] Časový rámec: 20 minut. Poskytněte pracovní list jako upravitelný dokument. Akceptujte vyplnění na počítači, ručně, zvukově nebo nahrané.

Pokračování na dalším snímku



Práva, lidská kontrola, výmaz a náprava:

.....

Odpovědné role a vlastník důkazů:

.....

Rozhodnutí: POKRAČOVAT UPRAVIT ZASTAVIT

.....

Zbytkové riziko, monitorování a spouštěč přezkumu:

.....

Nesouhlas nebo alternativní architektura:

.....

[Pokyny pro lektory] Časový rámec: 20 minut. Poskytněte pracovní list jako upravitelný dokument. Akceptujte vyplnění na počítači, ručně, zvukově nebo nahrané.

Role Play: DPIA & Procurement Review Panel

- Univerzita plánuje pořídit kamerový systém s umělou inteligencí, který zaznamenává činnost v laboratoři, detekuje nebezpečné chování a generuje individuální skóre dodržování předpisů. Dodavatel nabízí zpracování v cloudu a uchovává záznamy pro vylepšování modelu. Návrh uvádí celkovou přesnost, ale neobsahuje žádné informace o návštěvnicích, zdravotním postižení, ochranném oděvu, mazání záznamů ani přístupu zaměstnanců.
- Role: laboratorní technik, zástupce studentů, vedoucí bezpečnosti, pověřenec pro ochranu osobních údajů, dodavatel, poradce pro přístupnost a vedoucí zadávacího řízení.

Toky dat, otázky k posouzení dopadu na ochranu osobních údajů (DPIA), rozsah pilotního projektu, smluvní podmínky a důkazy o přijetí. Uveďte jednu podmínku pro UPRAVIT a jednu pro ZASTAVIT. Zaznamenejte, kdo může pozastavit používání a jak může osoba získat přístup k záznamu, opravit jej nebo proti němu vznést námitku.

[Pokyny pro lektora] Časový rámec: 25 minut. Komise musí uvést podmínky, které musí dodavatel splnit, aby mohl poskytnout službu. Komise musí také určit, kdo může pozastavit používání a jak může osoba získat přístup k záznamu, opravit jej nebo proti němu vznést námitku.

Individuální reflexe & Přenos

Napište krátké technické zdůvodnění

1. Který tok dat nebo který závěr vyvolal největší obavy ohledně ochrany soukromí?
2. Která volba návrhu nejúčinněji snížila riziko?
3. Která otázka týkající se zákonného základu nebo posouzení dopadu na ochranu osobních údajů (DPIA) zůstává k posouzení odborníkem?
4. Který důkaz by změnil vaše rozhodnutí?
5. Které použití by mělo být zakázáno, i kdyby bylo technicky možné?
6. Jak by se analýza lišila v jiné technické disciplíně?

Napište:

Předložte písemné, zvukové nebo nahrané důkazy. Propojte dotčené osoby, inženýrské opatření, důkazy a odůvodněné rozhodnutí.

[Pokyny pro lektora] Časový rámec: 10 minut. Studenti mohou předložit písemné, zvukové nebo videozáznamy. Před sdílením poskytněte studentům čas na přípravu v klidu.

Sedm principů „Privacy-By-Design“ jako inženýrské kontroly

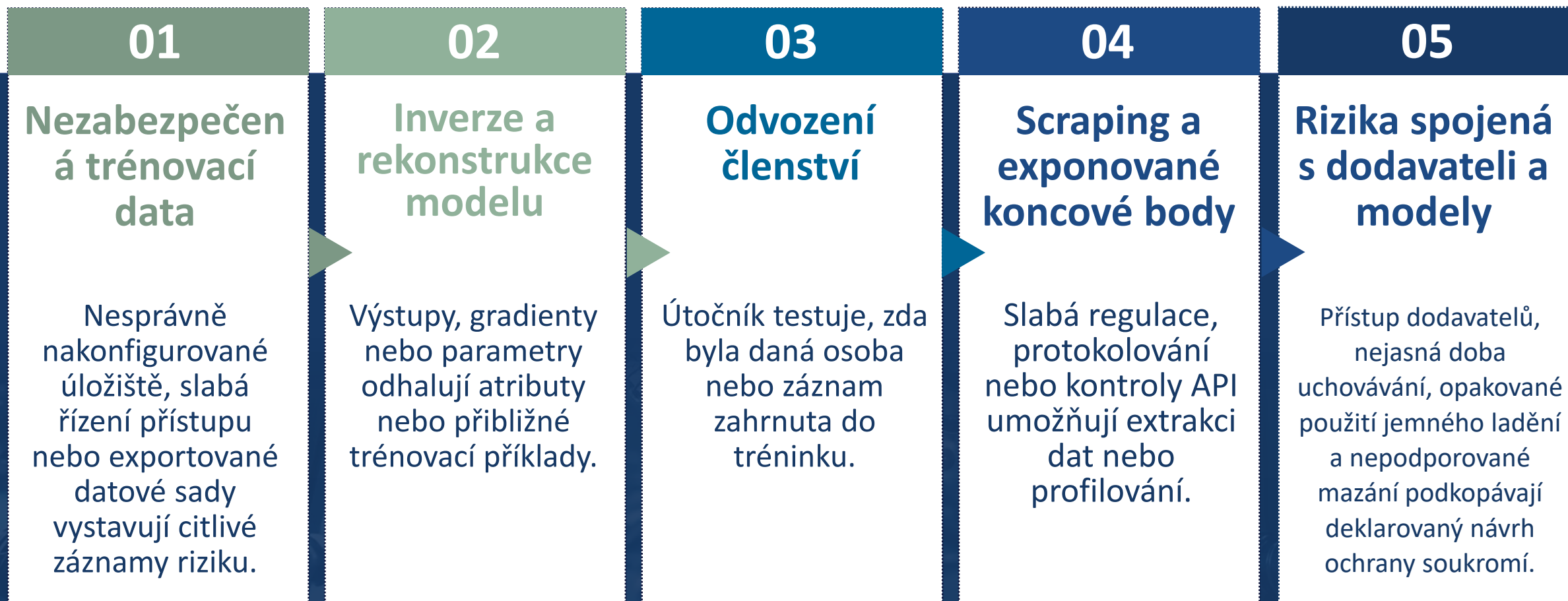
Použijte těchto sedm principů jako rychlou diagnostiku před schválením návrhu AI.

- **Proaktivní prevence:** identifikujte pravděpodobná selhání ochrany soukromí ještě před vývojem nebo nasazením.
- **Výchozí ochrana:** ve výchozím nastavení shromažďujte, zveřejňujte a uchovávejte co nejméně dat.
- **Integrovaná ochrana soukromí:** začlenit do architektury minimalizaci, oddělení, přístup, mazání a PET.
- **Návrh s pozitivním součtem:** splnit technické požadavky, aniž by bylo soukromí považováno za kompromis, který je třeba odstranit.
- **Bezpečnost od začátku do konce:** chraňte data, modelové artefakty, výzvy, protokoly, zálohy a fronty pro mazání.
- **Viditelnost, transparentnost a orientace na uživatele:** dokumentujte kontroly, vysvětlujte dopady a poskytněte použitelná oprávnění.

Úniky prostřednictvím postranních kanálů a metadat v systémech AI

- Šifrování chrání obsah při přenosu i v klidu, ale systémy umělé inteligence mohou stále únik informací prostřednictvím metadat, časování, chování dotazů nebo výstupních vzorců.
- **Příklad rizika:** konverzace v modelu LLM chráněná protokolem TLS může i tak odhalit vzorce citlivých témat prostřednictvím analýzy velikosti paketů nebo časování.
- **Technické řešení:** definujte metadata jako součást plochy útoku na soukromí, omezte exponované protokoly, omezte chování dotazů a sledujte neobvyklý přístup.
- V případě vysokého rizika testujte inferenci na základě výstupů, podnětů, gradientů, vzorců paketů a provozní telemetrie.
- Důkazy o přijatelnosti by měly prokázat, že testování úniků, monitorování a reakce na incidenty pokrývají útokové cesty specifické pro daný model.

Zranitelnosti v oblasti ochrany soukromí v provozní AI, které je třeba otestovat



Trojice prvků správy a provozní standardy

Vyspělý proces ochrany soukromí v oblasti umělé inteligence propojuje právní, soukromí a kyberbezpečnostní odbornost od samého počátku návrhu systému.

01

Právní

Právní základ, práva, smluvní odpovědnost, soulad s předpisy a dokumentační povinnosti.

02

Ochrana osobních údajů

Minimalizace údajů, posouzení dopadu na ochranu osobních údajů (DPIA), transparentnost, postupy pro uplatnění práv, uchovávání údajů a doklady o správě ochrany osobních údajů.

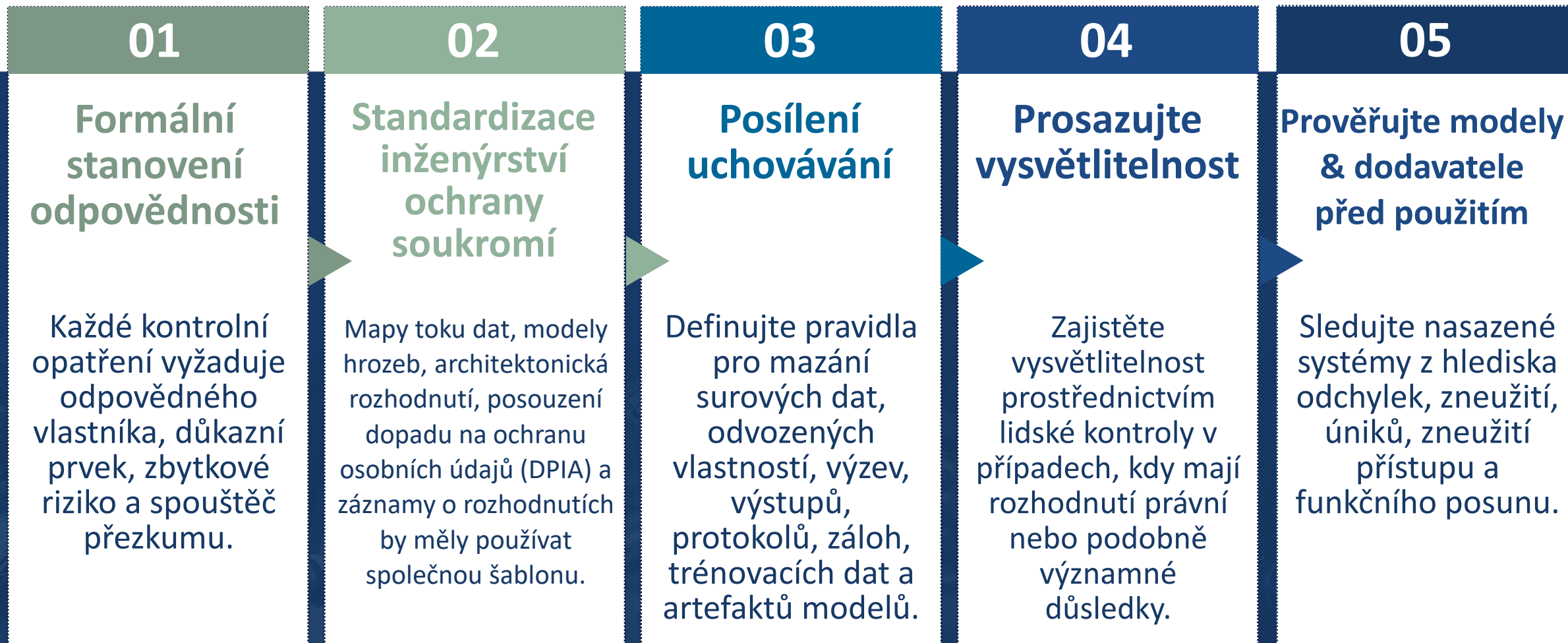
03

Kyberbezpečnost

Důvěrnost, integrita, dostupnost, testování za předpokladu útoku, zajištění dodavatelů, monitorování a reakce na incidenty.

Využijte rámec NIST AI RMF k organizaci práce v oblasti rizik a normu ISO/IEC 27701 ke strukturování dokladů o správě ochrany soukromí pro správce a zpracovatele.

Kontrolní seznam pro strategickou implementaci



Volitelná výuková aktivita připravená k použití

01

Teachable Machine:

Prozkoumejte otázky soukromí a souhlasu v souvislosti s biometrickými údaji pomocí klasifikace obrázků nebo hlasu.

02

Nástroj „What-If“:

Prozkoumejte efekty podskupin, prahové hodnoty a kontrafaktuální scénáře týkající se citlivých atributů.

03

Fasety:

Prověřte datové sady z hlediska nevyváženosti, anomálií, chybějících hodnot a zbytečných polí.

Volitelná výuková aktivita připravená k okamžitému použití

04

ARX:

Demonstrace kompromisů mezi pseudonymizací, k-anonymitou a reidentifikací.

05

Větvené scénáře H5P:

Simulujte rozhodnutí týkající se souhlasu, sledování a využívání dat.

06

DEDA nebo Ethical OS:

Strukturovejte dialog se zainteresovanými stranami a dlouhodobé zvažování rizik v oblasti ochrany soukromí.

Kvíz

1. Proč může být systém bezpečný a přesto ohrožovat soukromí?
2. Jaký je rozdíl mezi pseudonymizovanými a anonymizovanými údaji?
3. Proč musí být PET propojen s modelem hrozeb?
4. Co by měla obsahovat mapa toku dat?
5. Uveďte dva ukazatele, které mohou podpořit rozhodnutí o provedení posouzení dopadu na ochranu osobních údajů (DPIA).
6. Proč je lidská kontrola důležitá pro automatizované rozhodování?
7. Co patří do záznamu o rozhodnutí v oblasti ochrany soukromí?

[Pokyny pro lektora]. Čas: 8 minut. Nejprve nechte studenty hledat informace samostatně, poté diskutujte ve dvojicích. Následující snímek obsahuje stručné odpovědi a odůvodnění.

Odpovědi na kvíz a pedagogické zdůvodnění

1. Bezpečné zpracování může i nadále umožňovat rušivé sledování, vyvozování závěrů, opětovné použití nebo ztrátu svobodné vůle.
2. Pseudonymizovaná data lze i nadále propojit s dalšími informacemi; anonymizovaná data by měla odolávat identifikaci pomocí přiměřeně pravděpodobných prostředků.
3. Každá technika ochrany soukromí (PET) chrání konkrétní informace za stanovených předpokladů a má odlišné způsoby selhání.
4. Zdroje, příjemci, účely, identifikátory, transformace, kopie, hranice důvěry, přístup a uchovávání.
5. Systematické monitorování, zranitelné osoby, inovativní technologie, citlivá data, porovnávání nebo významná automatizovaná rozhodnutí.
6. Oprávnění, kompetence, důkazy, čas, svoboda nesouhlasit a pravomoc změnit výsledek.
7. Účel, data, rizika, kontrolní mechanismy, důkazy, role, práva, zbytkové riziko, monitorování, spouštěče rozhodnutí a přezkumu.

Kontrola výstupů učení

Než skončíte, ujistěte se, že umíte:

1. Zmapujte osobní údaje, aktéry, účely, závěry a rizika v rámci celého životního cyklu umělé inteligence.
2. Aplikujte principy „ochrany soukromí již od návrhu“, „ochrany soukromí jako výchozího nastavení“ a zásady GDPR na dané požadavky.
3. Vyberte a odůvodněte technologii zvyšující ochranu soukromí pro danou hrozbu.
4. Provádějte posouzení nutnosti provedení DPIA a navrhujte smysluplná práva a možnost napadnutí rozhodnutí.
5. Dokumentujte kontroly, důkazy, zbytková rizika, monitorování a odpovědné rozhodování.

Důkazy:

Mapa životního cyklu dat, stručný popis architektury, posouzení DPIA a záznam o technickém rozhodnutí.

[Pokyny pro lektora]

Čas: 5 minut. Požádejte studenty, aby ukázali na artefakt, který demonstruje každý výsledek. Zaznamenejte všechny výsledky, které vyžadují další procvičení.

Příspěvek do kurzu

Přidejte následující podklady z modulu 3 do svého etický inženýrský spis:

1. Vyplněná mapa životního cyklu dat a toku dat.
2. Specifikace účelu, nezbytnosti a zakázaného použití.
3. Architektura ochrany soukromí a výběr PET.
4. Záznam o posouzení dopadu na ochranu osobních údajů (DPIA) a nevyřešené odborné otázky.
5. Práva, mazání, incidenty a kontrolní mechanismy monitorování.
6. Záznam o rozhodnutích v oblasti inženýrství ochrany soukromí a reflexe.

Tyto podklady navazují na témata odpovědnosti, designu založeného na hodnotách a udržitelnosti v následujících modulech.



Oznámení o financování a použití ve vzdělávacích účelech

Tento otevřený vzdělávací zdroj slouží k podpoře výuky a diskuse. Nepředstavuje právní, regulační, bezpečnostní ani odborné inženýrské poradenství. Při aplikaci jeho obsahu se řiďte platnými zákony, normami, institucionálními zásadami a oborově specifickými pokyny.

**Spolufinancováno programem Erasmus+
Evropské unie.**





Ethical Engineer

www.ethicalengineer.eu

Děkujeme

Dokončili jste modul 3



Co-funded by
the European Union