ŘÍZENÍ RIZIK AI V PRAXI

FRANTIŠEK NONNEMANN

ČESKÉ PRÁVO A INFORMAČNÍ TECHNOLOGIE, 12. ZÁŘÍ 2024

- Řízení rizik
- Rizika a AI Act
- Dva účely systému řízení rizik
- Jak řídit AI rizika pro organizaci?
- Case study

ŘÍZENÍ RIZIK

- Co je to riziko?
- Možnost, že se stane něco špatného…
- S určitou pravděpodobností (ne zcela hypoteticky) nastane událost, která bude mít negativní dopad na chráněný zájem (důvěrné informace, kybernetickou bezpečnosti, provozní odolnost, veřejné zdraví, rodinný rozpočet)

PŘÍKLAD Z KYBERNETICKÉ BEZPEČNOSTI

- Regulace kybernetické bezpečnost

Informační aktivum → zranitelnost → hrozba → riziko

- Např. elektronická zdravotnická dokumentace → exponována do internetu → hrozbou je DDoS útok → pokud se realizuje, zdravotničtí pracovníci nemají přístup k dokumentaci, nemohou poskytovat zdravotnické služby

AI ACT

- Vyhlášen v Úředním věstníku dne 12. července 2024
- Účinnost postupně:
  - 2. února 2025: Obecná ustanovení zákaz vysoce rizikových systémů
  - 2. srpna 2025: Zahájení činnosti dozorových úřadů, dozor na úrovni EU a pravidla pro obecné modely AI
  - 2. srpna 2027: Zbytek, včetně pravidel pro vysoce rizikové systémy

AI ACT A RIZIKA

- Čtyři kategorie rizikovosti AI pro dotčené osoby – vyhodnocení rizik provedl normotvůrce
- Nepřijatelné (čl. 5), např.:
  - Podprahové či manipulativní techniky
  - Zneužívání zranitelnosti skupiny osob
  - Sociální scoring
- Vysoké riziko (čl. 6 a příloha III), např:
  - Biometrická identifikace na dálku, rozpoznávání emocí
  - Kritická infrastruktura (směrnice CER a nový zákon)
  - Rozhodování ve vzdělávacím procesu
  - Rozhodování v HR oblasti
  - Přístup k dávkám
  - Hodnocení úvěruschopnosti
  - Vymáhání práva
- Omezené riziko (zvukový, obrazový nebo video obsah podobný realitě)
- Minimální riziko - ostatní

VYSOCE RIZIKOVÉ SYSTÉMY AI

- Poskytovatelé a provozovatelé vysoce rizikových AI systémů musejí plnit řadu povinnosti, mj. zavést systém řízení rizik
- Čl. 9/2/a AI Actu: Systém řízení rizik je chápán jako nepřetržitý opakující se proces plánovaný a prováděný v rámci celého životního cyklu vysoce rizikového systému AI, který vyžaduje pravidelný systematický přezkum a aktualizaci. Zahrnuje následující kroky: … identifikaci a analýzu známých a rozumně předvídatelných rizik, která může vysoce rizikový systém AI používaný v souladu se zamýšleným účelem představovat pro zdraví, bezpečnost nebo základní práva.

ROZDÍL PROTI ŘÍZENÍ BEZPEČNOSTNÍCH ČI OBCHODNÍCH RIZIK?

DVA PŘÍSTUPY K ŘÍZENÍ RIZIK

- V rámci ISMS, projektového řízení, kritické infrastruktury, řízení obchodních/finančních rizik identifikuje a snižujeme **rizika pro organizaci.**

- AI Act řeší **rizika pro dotčené osoby** (a související veřejné zájmy)
  - Ze které další regulace to známe?

ODLIŠNÉ HODNOCENÍ STEJNÉHO INCIDENTU

- Hackerský útok na velké zdravotnické zařízení
- Útok je brzy zastaven, systémy a provoz ochráněny
- Dojde k úniku informací o velmi omezeném okruhu osob
  - Identifikační údaje, údaje o zdravotním stavu, kontaktní údaje
  - Následná expozice na internetu
- Z pohledu organizace (NIS, CER) se jedná o málo rizikový incident – systémy i služby fungovaly a fungují
- Z pohledu byť jediného dotčeného pacienta (GDPR) se jedná o vysoce rizikový incident

AI RIZIKA PRO ORGANIZACI

- Řízení rizik podle AI Actu organizaci nezajistí, že pro ni využití AI nebude mít negativní dopady
- Co s tím?
- Řídit rizika AI i z pohledu organizace

ISO 42001

- ISO/IEC 42001:2023: **Systém řízení umělé inteligence (Artificial Intelligence Management System, AIMS).**
- Standardní systém pro řízení rizik pro organizaci
- Shrnutí (dosavadní) best practise
- Lze integrovat s dalšími systémy pro řízení rizik – ISMS, CMS

SYSTÉM ŘÍZENÍ UMĚLÉ INTELIGENCE (AIMS)

1. Porozumění **specifikům a potřebám** organizace
2. Jasný přístup a **závazek vedení** k etickému a transparentnímu využívání AI
3. Přijetí **interní strategie (policy)** pro využití umělé inteligence
4. Určení **rolí a odpovědností**
5. Proces pro **identifikaci a hodnocení rizik** plynoucích z využití umělé inteligence
6. Nastavení **opatření a kontrol ke snížení nepřijatelně vysokých rizik**
7. Zajištění **kvalifikace, odborné přípravy a dostatečných zdrojů** pro zaměstnance podílející se na využití AI
8. Proces pro **pravidelné hodnocení AI rizik**, dostatečnosti kontrol a celého AIMS, průběžné vylepšování
9. Dostupná **dokumentace** celého systému a všech jeho důležitých kroků

PŘÍKLAD Z PRAXE

- Organizace chce zavést AI nástroj pomáhající při IT vývoji (na základě znalostí kódu našeptává vývojářům)
- Co je nutné pro efektivní řízení rizik?
- Jaká typická rizika organizaci plynou a jak je omezit?

CO JE NUTNÉ PRO ŘÍZENÍ RIZIK?

- Vědět, že se rizika mají řídit (shadow IT)
- Znát technologii a pravidla použití
- Mít proces nebo využít obdobné či zavedené metodiky (např. bankovní outsourcing, řízení IT rizik, ISMS)
- Identifikovat rizika, zavést mitiganty
- Vše dokumentovat a po čase přezkoumat

TYPICKÁ RIZIKA

- **Únik důvěrných informací**
  - Osobní údaje, bankovní tajemství, obchodní tajemství (know-how, kód, ceny, plánované produkty)
- **Ohrožení provozu a dostupnosti**
  - Zavlečení slabiny, chyba při testingu, monitoringu, provozu
- **Oslabení kybernetické bezpečnosti a provozní odolnosti**
  - Bezpečnostní chyba, selhání bezpečnostního nástroje, nezachycení/špatné vyhodnocení alertu
- **Právní rizika**
  - Ochrana údajů a informační bezpečnosti
  - Autorské právo
  - Diskriminace
  - Sektorová regulace (např. dokumentovatelnost a rekonstruovatelnost postupů)

JAK RIZIKA SNÍŽIT?

- **Důvěrné informace**
  - Interní klasifikace informací, výběr vhodné licence (nesdílet data s ostatnímu uživateli), pseudonymizace, šifrování dat
- **Provoz a dostupnost**
  - Kontrolovat výstupy (manuálně, automaticky?), ověřit spolehlivost v omezeném rozsahu (pilot), výběr spolehlivého dodavatele, dokumentace
- **Kybernetická bezpečnost**
  - Kontrolovat výstupy, postupné nasazení od méně rizikových procesů, pilot, výběr spolehlivého dodavatele, dokumentace, hodnocení slabin a hrozeb z pohledu využití AI, dostupné náhradní řešení
- **Právní rizika**
  - Znát podmínky (vč. sledování změn), výběr vhodné licence, dokumentace, zohlednění sektorové regulace

VYUŽITÍ AI JE RIZIKOVÉ…

… TAK JI **VYUŽÍVEJME** ODPOVĚDNĚ!

DÍKY ZA POZORNOST

FRANTISEK.NONNEMANN@PARTNERSBANKA.CZ