

1.1 Úvod a definovanie pojmu Big Data

Množstvo údajov vytvorených ľuďmi sa každým rokom rýchlo zvyšuje v dôsledku zavádzania nových technológií, rôznych elektronických zariadení a komunikačných kanálov, ako sú napríklad sociálne siete. Veľké údaje sú skupinou obrovských súborov údajov, ktoré nemožno spracovať pomocou typických počítačových metód.

Big Data je masívna zbierka údajov, ktorá sa v priebehu času dramaticky zvyšuje. Ide o súbor údajov, ktorý je taký obrovský a komplikovaný, že ho žiadne typické technológie na správu údajov nedokážu efektívne uložiť ani spracovať. Veľké údaje sú podobné bežným údajom, ibaže sú oveľa väčšie. Analýza veľkých dát je použitie pokročilých analytických techník pre veľmi veľké, heterogénne súbory údajov, ktoré môžu obsahovať štruktúrované, pološtruktúrované a neštruktúrované údaje, ako aj údaje z mnohých zdrojov a veľkostí od terabajtov po zettabajty.

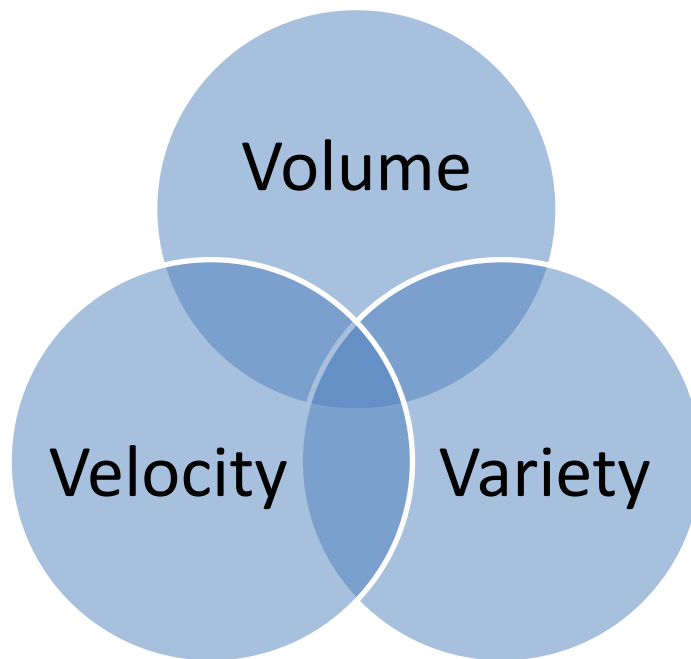


Obrázok 1 Príklady štruktúrovaných, polo-štruktúrovaných a neštruktúrovaných súborov

Pre lepšiu predstavu, Facebook ako sociálna sieť produkuje každý deň 4+ petabyty údajov, čo je približne milión gigabytov. Ak pridáme ešte podrobnejšie štatistiky ohľadom Facebooku, tak každú sekundu je vytvorených päť nových profilov a aktívnych je približne 32 miliárd užívateľov denne.

1.1.1 Charakteristické vlastnosti Big Data

Big data môžu byť opísané tromi základnými charakteristikami – množstvo (Volume), rýchlosť ich tvorby (Velocity) a rôznorodosť (Variety). Tieto charakteristiky zvyčajne znázorňuje nasledovný obrázok:



Obrázok 2 Charakteristické vlastnosti Big Data

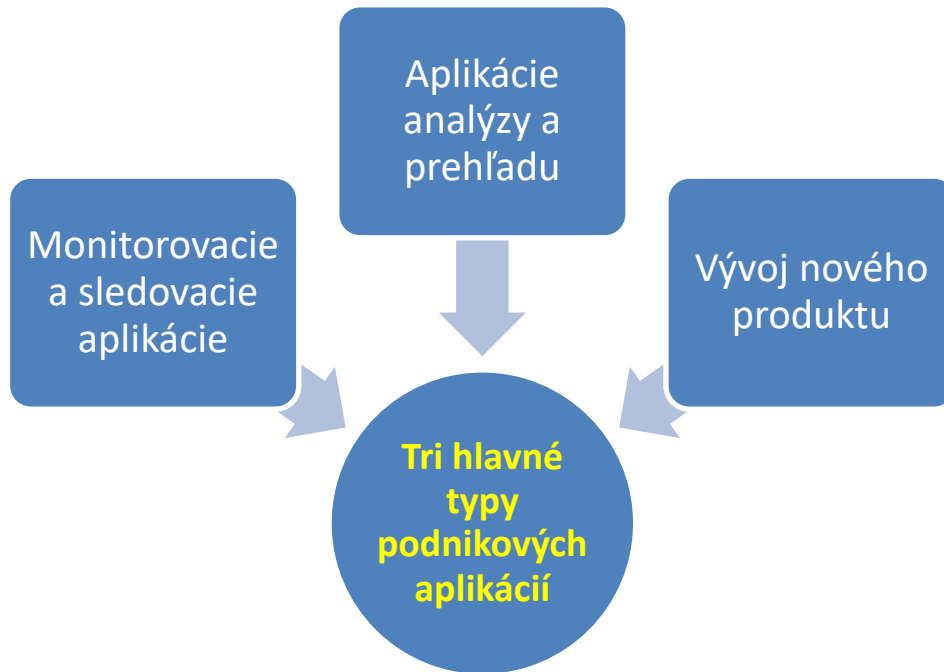
- ✓ **Volume – Množstvo** - veľké množstvo dát. Veľkosť údajov hrá rozhodujúcu úlohu pri určovaní ich hodnoty.
- ✓ **Velocity – Rýchlosť tvorby** - dáta prichádzajú vysokou rýchlosťou zo strojov, sietí, sociálnych médií, mobilných telefónov a iných zdrojov rýchlosťou veľkých dát. Existuje veľký a neustály prílev údajov. To má vplyv na potenciál údajov alebo na to, ako rýchlo sa údaje vytvárajú a spracúvajú na uspokojenie potrieb.
- ✓ **Variety – Rôznorodosť** – sem patrí rozdelenie dát na štruktúrované, pološtruktúrované a neštruktúrované. Štruktúrované údaje sú údaje, ktoré sú usporiadané. Pološtruktúrované údaje sú údaje, ktoré síce majú netradičnú štruktúru. Neštruktúrované údaje sú údaje, ktoré nie sú usporiadané.

Výhody, ktoré plynú zo spracovania veľkých dát sú nasledovné:

- ✓ Podniky môžu pri prijímaní rozhodnutí využívať externé informácie.
- ✓ Prístup k sociálnym údajom z vyhľadávačov a stránok ako Facebook, Twitter umožňujú organizáciám doladiť ich obchodné stratégie.
- ✓ Vylepšený zákaznícky servis.
- ✓ Včasná identifikácia rizika pre produkt/službu, ak nejaké existuje.
- ✓ Lepšia prevádzková efektívnosť.

1.2 Aplikácie súvisiace s Big Data

Všetky údaje musia byť zaznamenané a spracované, čo si vyžaduje veľa odborných znalostí, zdrojov a času. Dáta môžu byť kreatívne a zmysluplne použité na poskytovanie výhod pre podniky. Existujú tri druhy podnikových aplikácií, z ktorých každá má rôznu stupeň revolučného potenciálu.



Obrázok 3 Aplikácie súvisiace s Big Data

Monitorovacie a sledovacie aplikácie:

- ✓ monitorovanie verejného zdravia,
- ✓ sledovanie aktív,
- ✓ monitorovanie dodávateľského reťazca,
- ✓ preventívna údržba strojov a zariadení.

Aplikácie analýzy a prehľadu:

- ✓ prediktívna polícia – kontrola kriminality,
- ✓ víťazstvo v politických voľbách,
- ✓ osobné zdravie.

Vývoj nového produktu:

- ✓ flexibilné poistenie auta – údaje získané z GPS je možné použiť k predvídaní rizika nehôd,
- ✓ maloobchodná propagácia podľa lokality,
- ✓ odporúčanie služieb a tovarov.

1.3 Nástroje používané pre spracovanie Big Data

Apache Hadoop



Umožňuje distribuované spracovanie rozsiahlych súborov údajov naprieč klastrami počítačov. Je to jedna z najvýkonnejších veľkých dátových technológií so schopnosťou rásť z jedného servera na tisíce počítačov.

HPCC Systems



HPCC je nástroj pre veľké dáta vyvinutý spoločnosťou LexisNexis Risk Solution. Prináša spracovanie údajov na jedinej platforme, jedinej architektúre a jedinom programovacom jazyku.

Apache STORM



Storm je bezplatný výpočtový systém s otvoreným zdrojom veľkých dát. Je to jeden z najlepších nástrojov pre veľké dáta, ktorý ponúka distribuovaný systém spracovania v reálnom čase odolný voči chybám. S výpočtovými schopnosťami v reálnom čase.

Qubole



Qubole Data je autonómna platforma na správu veľkých dát. Ide o nástroj s otvoreným zdrojovým kódom pre veľké dáta, ktorý sa sám spravuje, sám sa optimalizuje a umožňuje dátovému tímu zamerať sa na obchodné výsledky.

Statwing



Statwing je jednoducho použiteľný štatistický nástroj. Bol vytvorený analytikmi veľkých dát a pre nich. Jeho moderné rozhranie vyberá štatistické testy automaticky.

Vyššie spomenuté aplikácie sú len výberom z množstva ďalších dostupných aplikácie pre spracovanie Big Data.

1.4 Výzvy pri spracovaní Big Data

Nedostatok správneho pochopenia Big Data

Spoločnosti zlyhávajú vo svojich iniciatívach týkajúcich sa veľkých dát pre nedostatočné pochopenie. Zamestnanci nemusia vedieť, čo sú údaje, ich uloženie, spracovanie, dôležitosť a zdroje. Dátoví profesionáli môžu vedieť, čo sa deje, ale iní nemusia mať jasný obraz. Ak napríklad zamestnanci nerozumejú dôležitosti ukladania údajov, nemusia si uchovávať zálohu citlivých údajov. Možno nepoužívajú správne databázy na ukladanie. V dôsledku toho, keď sú potrebné tieto dôležité údaje, nie je možné ich ľahko získať.

Problémy s rastom údajov

Jednou z najnaliehavejších výziev veľkých dát je správne ukladanie všetkých týchto obrovských súborov dát. Množstvo dát uložených v dátových centrách a databázach spoločností rýchlo narastá. Keďže tieto súbory údajov exponenciálne rastú s časom, je mimoriadne ťažké ich zvládnuť. Väčšina údajov je neštruktúrovaná a pochádza z dokumentov, videí, zvuku, textových súborov a iných zdrojov. To znamená, že ich nemôžete nájsť v databázach.

Zmätok pri výbere nástroja na spracovanie Big Data

Spoločnosti sú často zmätené pri výbere najlepšieho nástroja na analýzu a ukladanie veľkých dát. Je HBase alebo Cassandra najlepšia technológia na ukladanie dát? Je Hadoop MapReduce dosť dobrý alebo bude Spark lepšou voľbou pre analýzu a ukladanie údajov? Tieto otázky firmy trápia a niekedy nevedia nájsť odpoveď. Nakoniec robia zlé rozhodnutia a vyberajú nevhodnú technológiu. Výsledkom je plytvanie peniazmi, časom, úsilím a pracovnými hodinami.

Nedostatok dátových profesionálov

Na prevádzkovanie týchto moderných technológií a nástrojov pre veľké dáta potrebujú spoločnosti skúsených dátových profesionálov. Títo odborníci budú zahŕňať dátových vedcov, dátových analytikov a dátových inžinierov, ktorí majú skúsenosti s prácou s nástrojmi a určovaním významu obrovských dátových súborov. Spoločnosti čelia

problému nedostatku odborníkov na veľké dáta. Je to preto, že nástroje na spracovanie údajov sa rýchlo vyvíjali, ale profesionáli vo väčšine prípadov nie.

Zabezpečenie údajov

Zabezpečenie týchto obrovských súborov údajov je jednou zo skľučujúcich výziev veľkých údajov. Spoločnosti sú často tak zaneprázdnené porozumením, ukladaním a analýzou svojich súborov údajov, že presadzujú bezpečnosť údajov v neskorších fázach. Nie je to však šikovný krok, pretože nechránené dátové úložiská sa môžu stať živnou pôdou pre hackerov.

Integrácia údajov z rôznych zdrojov

Údaje v organizácii pochádzajú z rôznych zdrojov, ako sú stránky sociálnych médií, aplikácie ERP, denníky zákazníkov, finančné správy, e-maily, prezentácie a správy vytvorené zamestnancami. Skombinovať všetky tieto údaje na prípravu správ je náročná úloha. Toto je oblasť, ktorú firmy často zanedbávajú. Integrácia údajov je však kľúčová pre analýzu, zostavovanie a obchodné spravodajstvo, takže musí byť dokonalá.

1.5 Ukladanie Big Data

Prvým ukladacím mechanizmom používaným počítačmi na ukladanie údajov boli dierne štítky. Každá skupina súvisiacich diernych štítkov (Dierne štítky súvisiace s rovnakým programom) bývala uložená v kartotéke a súbory boli uložené v skladoch. Je to veľmi podobné tomu, čo v súčasnosti robíme pri archivácii dokumentov vo vládných inštitúciách, ktoré stále používajú papierovanie denne. Odtiaľ pochádza slovo „Systém súborov“ (File System). Počítačové systémy sa vyvíjali, ale koncept zostáva rovnaký.

Namiesto ukladania informácií na dierne štítky, teraz môžeme ukladať informácie/dáta v digitálnom formáte na digitálnom pamäťovom zariadení, ako je pevný disk, flash disk atď. Súvisiace údaje sú stále kategorizované ako súbory, súvisiace skupiny súborov sú uložené v priečinkoch. Každý súbor má názov, príponu a ikonu. Názov súboru označuje obsah, ktorý má, zatiaľ čo prípona súboru označuje typ informácií uložených v tomto súbore. Napríklad prípona exe označuje spustiteľné súbory, txt označuje textové súbory atď. Systém správy súborov používa operačný systém na prístup k súborom a priečinkom uloženým v počítači alebo akomkoľvek externom úložnom zariadení.

V Big Data sa často stretávame s viacerými klastrami (počítačmi). Jednou z hlavných výhod Big Data je, že presahuje možnosti jedného supervýkonného servera s extrémne

vysokým výpočtovým výkonom. Celá myšlienka Big Data spočíva v distribúcii údajov do viacerých klastrov a vo využití výpočtového výkonu každého klastra (uzla) na spracovanie informácií. Distribuovaný súborový systém je systém, ktorý dokáže spracovať prístup k údajom cez viacero klastrov (uzlov). V ďalšej časti sa dozvieme viac o tom, ako to funguje. Distribuovaný súborový systém pracuje nasledovne:

- ✓ **Distribúcia** - distribúcia blokov množín údajov medzi viacero uzlov. Každý uzol má svoj výpočtový výkon, čo dáva schopnosť systému paralelne spracovávať dátové bloky.
- ✓ **Replikácia** - distribuované súborové systémy budú tiež replikovať dátové bloky na rôznych klastroch kopírovaním rovnakých častí informácií do viacerých klastrov v rôznych stojanoch.

Údaje sú merané v bitoch a Bytoch (vyslovuje sa ako bajt). Jeden bit môže mať hodnotu 0 alebo 1. Osem bitov tvorí Byte. Potom máme Kilobytes (1 000 Bytes), Megabytes (1 000² Bytes), Gigabytes (1 000³ Bytes), Terabytes (1 000⁴ Bytes), Petabytes (1 000⁵ Bytes), Exabytes (1 000⁶ Bytes) a Zettabytes (1 000⁷ Bytes).

Kilobajt	1 000 bajtov
Megabajt	1 000 000 bajtov
Gigabajt	1 000 000 000 bajtov
Terabajt	1 000 000 000 000 bajtov
Petabajt	1 000 000 000 000 000 bajtov
Exabajt	1 000 000 000 000 000 000 bajtov
Zetabajt	1 000 000 000 000 000 000 000 bajtov

1.6 Škálovateľná výpočtová technika cez internet

S cloudovým hostingom je ľahké rásť a zmenšovať počet a veľkosť serverov podľa potreby. To sa dosiahne buď zvýšením alebo znížením zdrojov v cloude. Táto schopnosť meniť plány v dôsledku kolísania veľkosti podniku a potrieb je vynikajúcou výhodou cloud computingu, najmä keď dochádza k náhlemu nárastu dopytu. To sa nazýva aj ako škálovateľnosť. Inak povedané je to schopnosť procesu, siete, softvéru alebo organizácie rásť a riadiť zvýšený dopyt. K tomuto procesu patria nasledovné atribúty:

- ✓ **Vek internetovej výpočtovej techniky**

Miliardy ľudí používajú internet každý deň. V dôsledku toho musia super počítačové lokality a veľké dátové centrá poskytovať vysokovýkonné výpočtové služby veľkému počtu používateľov internetu súčasne. Musíme aktualizovať dátové centrá pomocou rýchlych serverov, úložných systémov a širokopásmových sietí. Účelom je napredovať v sieťových výpočtových a webových službách s novými technológiami.



✓ **Vysokovýkonná výpočtová technika (High Performance Computing)**

Po mnoho rokov systémy HPC zdôrazňujú výkon v surovej rýchlosti. Rýchlosť systémov HPC sa zvýšila z Gflops na začiatku 90. rokov na súčasný Pflops. Toto zlepšenie bolo poháňané najmä požiadavkami vedeckých, inžinierskych a výrobných komúnít. Počet používateľov superpočítačov je však obmedzený na menej ako 10 % všetkých používateľov počítačov. V súčasnosti väčšina používateľov počítačov používa stolové počítače alebo veľké servery pri vyhľadávaní na internete a pri počítačových úlohách riadených trhom.

✓ **Tri nové počítačové paradigmy**

Pokroky vo virtualizácii umožňujú vidieť rast internetových cloudov ako novú výpočtovú paradigmu. Vyspelosť rádiových frekvencií identifikácie (RFID), Globálneho polohovacieho systému (GPS) a senzorových technológií podnietila rozvoj Internetu vecí (IoT).

✓ **Rozdiely výpočtovej paradigmy**

Komunita špičkových technológií sa už mnoho rokov dohaduje o presných definíciách centralizovaných výpočtov, paralelných výpočtov, distribuovaných výpočtov a cloudových výpočtov. Distribuované výpočty sú opakom centralizovaných výpočtov. Oblasť paralelných výpočtov sa do značnej miery prekrýva s distribuovanými výpočtami a cloud computing sa prekrýva s distribuovanými, centralizovanými a paralelnými výpočtami. Toto je výpočtová paradigma, podľa ktorej sú všetky počítačové zdroje centralizované v jednom fyzickom systéme. Všetky zdroje (procesory, pamäť a úložisko) sú plne zdieľané a úzko prepojené v rámci jedného integrovaného operačného systému. Mnohé dátové centrá a superpočítače sú centralizované systémy, ale používajú sa v paralelných, distribuovaných a cloudových aplikáciách.

1.7 Nové dimenzie údajov a úložiská údajov

Veľké dáta často začínajú diskusiu o nových dimenziách definovaných pre dáta. Je potrebné s nimi zaobchádzať iným spôsobom, ako len s veľkými údajmi. Tieto nové výzvy sú:

- ✓ **Údaje v reálnom čase** - tieto údaje sa líšia od tradičnej formy údajov, ktoré ukladáme na našich serveroch. Nezáleží na tom, či to spadá pod typ štruktúrovaných alebo neštruktúrovaných údajov. Kľúčovým aspektom je, že ide o „aktuálne údaje“, nie o staré údaje. Umožňujú situačné uvedomenie si toho, čo sa práve deje. Údaje v reálnom čase vyvolávajú problém rýchlo sa kaziacich a osirotených údajov, ktoré už nemajú platné použitie, no napriek tomu sa naďalej používajú.
- ✓ **Zdieľané údaje** - ide o informácie, ktoré sa zdieľajú v rámci organizácie. To zahŕňa zdieľanie informácií medzi rôznymi aplikáciami a zdrojmi údajov. Na efektívne zdieľanie informácií musia podniky zabezpečiť, aby boli údaje konzistentné, použiteľné a rozšíriteľné. Dôležitým aspektom je, že zdieľanie informácií komplikuje úlohu určenia autority informácií.
- ✓ **Prepojené údaje** - tieto pochádzajú z rôznych zdrojov údajov, ktoré majú vzájomné vzťahy a zachovávajú kontext, aby bol užitočný pre ľudí a počítače. Keď používateľ prepojí údaje, vzťah v týchto údajoch od tohto bodu pretrváva.
- ✓ **Vysoko dôverné údaje** - tieto údaje zachovávajú kontext, podrobnosti, vzťahy a identity dôležitých obchodných informácií. Vo veľkej miere sa to deje prostredníctvom vložených metadát. Vysoko dôverné údaje umožňujú pridať nový význam bez zničenia predchádzajúceho významu údajov.

Dáta sa ukladajú do tzv. Big Data úložísk. Táto technológia bola prvýkrát vyvinutá na začiatku roku 2000, keď spoločnosti čelili ukladaniu obrovského množstva údajov, ktoré nemohli uchovávať na svojich serveroch. Problém bol v tom, že tradičné spôsoby ukladania nezvládli ukladanie všetkých týchto údajov, a tak spoločnosti museli hľadať nové spôsoby, ako ich uchovať. Vtedy vzniklo Big Data úložisko. Pre spoločnosti je to spôsob, ako ukladať veľké množstvo dát bez obáv z nedostatku miesta.

Skladové úložisko a cloudové úložisko sú dve z najpopulárnejších možností na ukladanie veľkých dát. Skladové úložisko sa zvyčajne nachádza na mieste, zatiaľ čo cloudové úložisko zahŕňa ukladanie údajov mimo pracoviska na bezpečnom mieste:

- ✓ **Skladové úložisko** - je jedným z najbežnejších spôsobov ukladania veľkého množstva údajov, má však svoje nevýhody. Ak napríklad potrebujeme okamžitý prístup k svojim údajom a chceme sa vyhnúť oneskoreniam alebo problémom s prístupom k nim cez internet, môžu existovať lepšie možnosti. Skladovanie môže byť tiež drahé, ak hľadáme dlhodobé zmluvy alebo potrebujeme ďalších pracovníkov na správu skladových priestorov.

- ✓ **Cloudové úložisko** - je čoraz populárnejšou možnosťou, pretože je jednoduchšie ako kedykoľvek predtým použiť túto metódu vďaka pokrokom v technológii, ako sú napríklad Amazon Web Services (AWS).

1.7.1 Kde sú uložené všetky dáta z internetu?

Väčšina digitálnych informácií je uložená v troch typoch miest. Prvým je globálna zbierka takzvaných koncových bodov (End Points), ktoré zahŕňajú všetky zariadenia internetu vecí, počítače, smartfóny a všetky ostatné zariadenia na ukladanie informácií. Druhým je tzv. edge, ktorý zahŕňa infraštruktúru, ako sú mobilné veže, inštitucionálne servery a kancelárie, ako sú univerzity, vládne úrady, banky a továrne. Tretie je jadro (Core), väčšina údajov je uložená v tom, čo je známe ako tradičné dátové servery a cloudové dátové centrá.

Na svete je okolo 600 hyperškálových dátových centier – s viac ako 5 000 servermi. Približne 39 % z nich je v USA, zatiaľ čo Čína, Japonsko, Spojené kráľovstvo, Nemecko a Austrália predstavujú približne 30 % z celkového počtu.

Najväčšie dátové servery na svete sú China Telecom Data Centre v Hohhote v Číne, ktoré zaberá 10,7 milióna štvorcových stôp a The Citadel v Tahoe Reno, Nevada, ktoré zaberá 7,2 milióna štvorcových stôp a využíva 815 megawattov energie.



Obrázok 4 Dátové úložiská v Číne a v USA

1.8 Reálne príklady Big Data

Big Data v marketingu a reklame

- ✓ **Netflix** - Netflix má viac ako 150 miliónov predplatiteľov a zhromažďuje údaje o všetkých z nich. Sledujú, čo ľudia pozerajú, kedy to pozerajú, aké zariadenie sa používa, či je relácia pozastavená a ako rýchlo používateľ skončí sledovanie seriálu. Dokonca si robia screenshoty scén, ktoré ľudia pozerajú dvakrát. Prečo? Pretože zavedením všetkých týchto informácií do svojich algoritmov môže Netflix vytvoriť vlastné užívateľské profily. Tie im umožňujú prispôbiť si zážitok odporúčaním filmov a televíznych relácií s pôsobivou presnosťou.

- ✓ **Amazon** - Amazon zhromažďuje obrovské množstvo údajov o svojich používateľoch. Sledujú, čo používatelia kupujú, ako často (a ako dlho) zostávajú online a dokonca aj veci, ako sú recenzie produktov. Amazon dokáže dokonca odhadnúť príjem ľudí na základe ich fakturačnej adresy. Zhromaždením všetkých týchto údajov medzi miliónmi používateľov môže Amazon vytvoriť vysoko špecializované segmentované používateľské profily. Pomocou prediktívnej analýzy potom môžu zacieliť svoj marketing na základe zvykov používateľov pri prehliadaní. Používa sa na navrhovanie toho, čo by ste mohli chcieť kúpiť ďalej, ale aj na veci, ako je zoskupovanie produktov, aby ste zjednodušili nakupovanie.

Big Data v zdravotníctve

- ✓ **Elektronické zdravotné záznamy** - naše lekárske záznamy zahŕňajú všetko od našich osobných údajov až po rodinnú anamnézu, alergie a ďalšie. Po desaťročia boli tieto informácie v papierovom formáte, čo obmedzovalo ich užitočnosť. Zdravotnícke systémy na celom svete však v súčasnosti tieto údaje digitalizujú a vytvárajú značný súbor elektronických zdravotných záznamov.
- ✓ **Big Data a zariadenia na každodenné nosenie** – inteligentné hodinky a náramky dokážu zbierať informácie o našich aktivitách ale napríklad aj o tepovej frekvencii alebo úrovni stresu.

Big Data v cestovaní a logistike

- ✓ **Logistika** - logistické spoločnosti, ktoré sledujú stav skladových zásob, dopravné správy, objednávky produktov a ďalšie, využívajú veľké dáta na zefektívnenie svojich operácií. Dobrým príkladom je UPS. Sledovaním údajov o počasí a údajoch zo snímačov nákladných vozidiel sa UPS naučila najrýchlejšie trasy pre svojich vodičov.
- ✓ **Mestská mobilita** - Big data sú veľkým biznisom v mestskej mobilite, od autopožičovní až po prenájom e-bicyklov a e-kolobežiek. Uber je vynikajúcim príkladom spoločnosti, ktorá naplno využila potenciál veľkých dát. Po prvé, pretože majú veľkú databázu vodičov, dokážu v priebehu niekoľkých sekúnd priradiť používateľov k najbližšiemu vodičovi. Ale tam to nekončí. Uber tiež ukladá údaje o každej uskutočnenej ceste. To im umožňuje predpovedať, kedy bude služba najvyťaženejšia, čo im umožňuje primerane nastaviť svoje cestovné.

Big Data v poľnohospodárstve

- ✓ **Precízne poľnohospodárstvo** - poľnohospodári využívajú veľké údaje na prijímanie informovanejších rozhodnutí o svojich plodinách. Senzory umiestnené na poliach merajú úroveň vlhkosti, teplotu a pôdne podmienky, ako aj na traktoroch a iných poľnohospodárskych strojoch. Keď už hovoríme o

poľnohospodárskych strojoch, tu je nezvyčajný príklad: drony. Vybavenie dronov kamerami môže poskytnúť podrobné letecké pohľady na plodiny, čo pomáha pri zisťovaní chorôb alebo škodcov.

1.9 Big Data a zbieranie informácií o súkromí

Súkromie online sa vzťahuje na vašu schopnosť a právo ponechať si súkromné informácie pre seba. Zatiaľ čo polovica rovnice súvisí s tým, čo uverejňujete online a ako služby, ktoré používate, zdieľajú vaše informácie, zvyšok tvoria zákony a zásady, ktoré sú určené na ochranu spotrebiteľov. Bohužiaľ, len málo regiónov má zavedené prísne pravidlá týkajúce sa toho, ako technologické spoločnosti, inzerenti a online služby chránia vaše súkromie online. A aj keď existujú zákony na ochranu súkromia online, nie všetky spoločnosti ich dodržiavajú.

Jedným z dôvodom, čo robí online súkromie takým komplikovaným je, že aj keď robíte všetko pre to, aby ste ochránili svoje osobné údaje online, spoločnosti môžu byť napadnuté hackermi, zbierať viac údajov, ako je potrebné, alebo to, čo o vás vedia, môžu použiť na pochybné alebo dokonca podvodné účely.

V nasledujúcich bodoch je niekoľko možností ako chrániť súkromie a súkromné informácie online:

1. **Zdieľať menej informácií s aplikáciami a službami** - všetky platformy a aplikácie sociálnych médií zhromažďujú údaje o tom, kto sme, aké máme záujmy a čo robíme online. Všetky tieto zdieľania a dátové body tvoria našu online stopu.
2. **Používať silné a jedinečné heslá s dvojfaktorovým overením** - silné heslá sú najdôležitejšou a niekedy jedinou ochranou proti krádeži identity a hackerom.
3. **Sprísniť nastavenia ochrany osobných údajov na svojich účtoch sociálnych médií** - na zlepšenie súkromia online nemusíme odstraňovať svoje účty na sociálnych médiách. Namiesto toho môže stačiť jednoducho skontrolovať nastavenia ochrany osobných údajov na online účtoch, ktoré pravidelne používame.
4. **Odstrániť nepoužívané mobilné aplikácie a rozšírenia prehliadača** - aplikácie a rozšírenia prehliadača môžu kedykoľvek zmeniť svoje zásady zabezpečenia a ochrany osobných údajov. Ak nástroj aktívne nepoužívame, je najlepšie ho odstrániť.
5. **Neignorovať aktualizácie softvéru alebo operačného systému** – kedykoľvek nás softvér alebo operačný systém vyzýva k aktualizácii, znamená to, že prišlo k prelomeniu niektorej časti a je nutné ju opraviť.

Všeobecne je len ťažko možné zabrániť zbieraniu údajov v online priestore. Treba však dávať pozor, čo všetko zdieľame online.