

VELKÁ DATA





PŘEDSTAVENÍ 4. PRŮMYSLOVÉ REVOLUCE

Následující materiál vznikl v rámci projektu „Průmysl 4.0 - INTRO 4.0“, který je financován ze zdrojů Evropské unie. Jeho cílem je představit Průmyslu 4.0. v evropském kontextu.

Obsah výukových materiálů slouží jako zdroj nejdůležitější informací o Průmyslu 4.0 určený učitelům (odborného vzdělání a přípravy a vysokých škol), vzdělávacím zařízením, zaměstnavatelům, zaměstnancům, široké veřejnosti a osobám podílejících se na nových inovačních přístupech.

Uvedené informace vycházejí ze zprávy „Současného stavu Průmyslu 4.0“ a ze „Souhrnné zprávy odborných rozhovorů/dotazníků a specifického výzkumu ve zpracovatelském průmyslu“, na jejichž tvorbě se podíleli všichni uvedení partneři projektu.

OBSAH

<p>2 Obsah & cíle studia</p> <p>3 Úvod</p> <p>4-5 Co jsou velká data?</p> <p>6-12 K čemu slouží velká data?</p> <p>13-14 Osvědčené postupy</p>	<p>15-17 Přínos pro společnosti</p> <p>18-23 Budoucí vývoj</p> <p>24-30 Pokročilé informace</p> <p>31-32 Vzdělávání</p> <p>33 Literatura a osobní přínos</p>
--	--



OBSAH PŘÍNOSNÝ PRO
SPOLEČNOSTI



OBSAH PŘÍNOSNÝ PRO
ŠIROKOU VEŘEJNOST



CÍLE STUDIA

- ❖ Porozumět pojmu velká data.
- ❖ Identifikovat příležitosti v oblasti velkých dat.
- ❖ Rozpoznat a zlepšit nejdůležitější dovednosti.
- ❖ Sledovat organizační výzvy a osvědčené postupy.
- ❖ Pozvednou firemní zdroje na vyšší úroveň.
- ❖ Provést užitečné strategie.
- ❖ Zjistit praktická využití.
- ❖ Poskytnou užitečné informace o kurzech a certifikacích.











ÚVOD

Pojem **velká data (big data)** označuje velké množství velmi rychle vytvářených dat velkým množstvím různých zdrojů.



Cíle studia

-  Porozumět pojmu velká data
-  Identifikovat příležitosti v oblasti velkých dat
-  Rozpoznat a zlepšit nejdůležitější dovednosti
-  Sledovat organizační výzvy a osvědčené postupy
-  Pozvednou firemní zdroje na vyšší úroveň
-  Provést užitečné strategie
-  Zajistit praktická využití
-  Poskytnout užitečné informace o kurzech a certifikacích

Snadnější rozhodování

Identifikace rizik ohrožující produkty a služby



K čemu velká data slouží?

Služby poskytované zákazníkům

Provozní efektivita



6 klíčových bodů v oblasti Big Data

Jsou navržena tak, aby ekonomicky extrahovala velký objem širokého spektra dat vysokorychlostním snímáním, jejich objevováním nebo analýzou.



BUDOUCÍ VÝVOJ



CO JSOU VELKÁ DATA?



Napříč celou společností, od zdravotnictví po zemědělství a dopravu, od energetiky po změnu klimatu a bezpečnost, odborníci v každém vědním oboru oceňují potenciál obrovského množství údajů, které jsou každý den generovány. Úkolem velkých dat je zachytit, spravovat a zpracovat informace a na jejich základě získat smysluplné výsledky. Data mohou být vytvářena jak lidmi, tak stroji, jako jsou senzory shromažďující informace o klimatu, satelitní snímky, digitální obrázky a videa, záznamy transakcí, GPS signály atd. Pokrývají mnoho odvětví, od zdravotní péče po dopravu a energii, komunikaci a maloobchod.



Obr. č. 1. Využití velkých dat
Zdroj: www.edureka.com

Vytváření určité hodnoty v různých fázích hodnotového řetězce dat bude ústředním bodem pro budoucí ekonomiky. Správné použití údajů může také přinést příležitosti tradičním odvětvím, jako je doprava, zdravotnictví nebo výroba.



CO JSOU BIG DATA?



- Transformují odvětví služeb generováním široké škály inovativních informačních produktů.
- Zvyšují produktivitu všech sektorů ekonomiky pomocí vylepšené obchodní inteligence.
- Pomáhají řešit problémy, kterým společnosti čelí.
- Zlepšují výzkum a urychlují inovace.
- Dosahují snižování nákladů prostřednictvím individualizovaných služeb.
- Zvyšují efektivitu ve veřejném sektoru.
- Digitalizují evropský průmysl.



Obr. č. 2. Prediktivní analýza
Zdroj: www.dreamstime.com



XV-XIX

- Vznik statistiky.
- První zaznamenaný experiment ve statistické analýze dat.
- Využití Hollerith Tabulate Machine snižující práci z 10 let na 3 měsíce.



XX

- Počátky moderního ukládání dat.
- Počátky obchodní inteligence.
- Vznik velkých datových center.
- Vznik internetu.



XXI

- Počátky nápadu velkých dat.
- Internet věcí "Internet of Things"
- Web 2.0 zvyšuje objem dat.
- Objevuje se dnešní použití termínu 'Big Data'



K ČEMU SLOUŽÍ VELKÁ DATA?

Velká data představují jedinečné příležitosti, které nám pomáhají vyvíjet nové kreativní produkty a služby, např. aplikace na mobilních telefonech nebo produkty pro inteligentní obchodování společností. Mohou podpořit růst a zaměstnanost v Evropě a také zlepšit kvalitu života.

Zdravotní péče

Velká data se podílejí na diagnostice a léčbě nemocí při zachování soukromí pacienta. Nabízejí účinnější zpracovávání informací o zdravotní péči pacienta, což je hodnotné jak pro společnosti, tak pro veřejný sektor i samotné občany. Analýza rozsáhlých klinických datových souborů může vést k optimalizaci a efektivitě z hlediska nákladů na nové léky nebo na léčebné postupy, pacienti tak mohou těžit z včasější a vhodněji zvolené léčby. Interoperabilita dat je nanejvýš důležitá, protože data pocházejí z různých a heterogenních zdrojů, jako je biologie, genomika, zdravotní záznamy a klinické laboratorní testy. Cílem technologií je poskytovat pacientům, zdravotnickým pracovníkům a klinickým výzkumným pracovníkům přístup ke zdravotním údajům jednotným způsobem a v anonymizované a agregované formě, který zároveň chrání jejich soukromí.

Obchod s daty

Informační technologie měly v posledních desetiletích, přímo nebo nepřímo, vliv na evropský hospodářský růst. Přístup k tvorbě hodnotných údajů v nové ekonomice se stal rozhodující, přičemž Evropa se i nadále nachází v dobré výchozí pozici, aby mohla pokračovat v dalším rozvoji.

Doprava: méně nehod a kongescí

Odvětví dopravy může těžit z velkého množství dat shromážděných prostřednictvím senzorů, GPS dat a zejména prostřednictvím sociálních médií. Inteligentní využití velkých dat podporuje optimalizaci multimodální dopravy a řízení dopravních toků, díky čemuž se naše města stávají chytřejšími. Občané a společnosti mohou např. ušetřit čas pomocí systémů podpory plánování tras.



K ČEMU SLOUŽÍ VELKÁ DATA?

Životní prostředí: snížená spotřeba energie

Revoluce v oblasti velkých dat přináší nové způsoby porozumění a řešení environmentálních problémů. Lepší využití celosvětově dostupných národních a místních datových souborů pomáhá vědcům při výzkumu a umožňuje tvůrcům politik činit na důkazech založená rozhodnutí, bojovat proti přírodním pohromám, změně klimatu a snižovat náklady. Inteligentní města poskytují zázemí datovým centrům, která přizpůsobují spotřebu energie veřejných budov dostupnosti obnovitelné energie. Současně se naše mobilní zařízení stávají chytřejšími díky integraci analytických nástrojů, které snižují spotřebu energie a šetří peníze.

Otevřená data

Otevřená data se vztahují na informace shromážděné, vytvořené nebo zaplacené veřejnými orgány a volně přístupné k opakovanému použití. Informace veřejného sektoru jsou informace, které jsou výhradně v rukou veřejného sektoru. Směrnice o opakovaném využití informací veřejného sektoru poskytuje společný právní rámec pro celý evropský trh. Tento rámec je postaven na klíčových pilířích vnitřního trhu: volný tok dat, transparentnost a spravedlivá hospodářská soutěž.

Zemědělství: bezpečnější potraviny a zvýšená produktivita

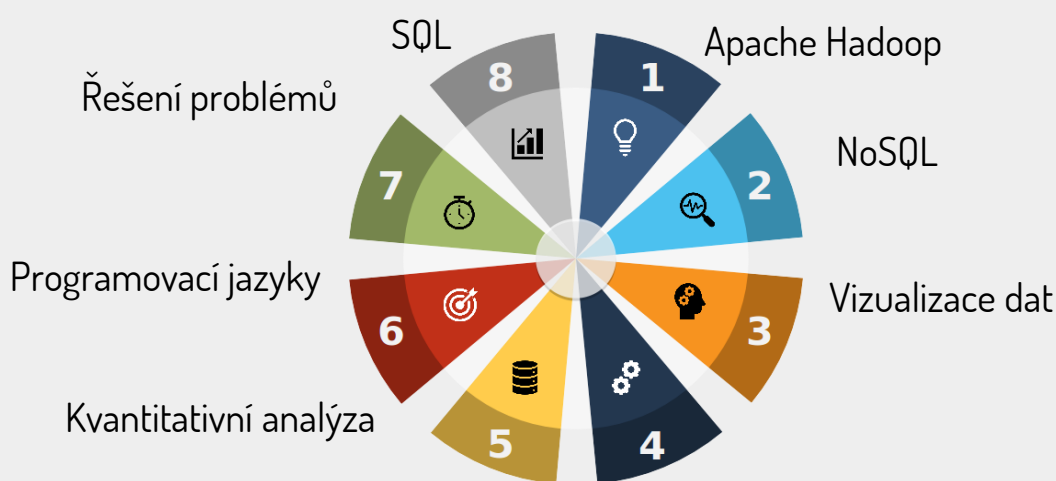
Inteligentní využití velkých dat v zemědělství může současně zvýšit produktivitu, bezpečnost potravin a příjmy zemědělců. Inteligentním a rozšířeným využitím dat ze senzorů a pozorování Země, jako jsou data z programu Copernicus, lze způsoby, kterými dnes hospodaříme, změnit k lepšímu. Může dojít k efektivnějšímu využívání přírodních zdrojů (včetně vody nebo slunečního záření) v zemědělských postupech. Díky pokročilým technologiím mohou mít zemědělci přístup k datům o fungování jejich zemědělských strojů nebo k historickým vzorcům počasí, topografii a úrodnosti.

Dopad na průmysl/Technologie velkých dat/Výzkum

Maximální využití dostupných údajů je pro průmyslovou konkurenceschopnost stále důležitější. Přístup k relevantním údajům se stává stále obtížnějším kvůli velikosti a složitosti datových souborů. Maximální využití dat vyžaduje flexibilní přístup, kdy musí inženýři data prozkoumat způsoby, které současné aplikace nepodporují. Inženýři tráví až 80 % svého času řešením problémů přístupu k datům. Uvolnění času odborníka vede k přidané hodnotě v podobě hlubší analýzy a lepšího rozhodování.



K ČEMU SLOUŽÍ BIG DATA?

TOP 8 BIG DATA PRACOVNÍCH DOVEDNOSTÍ

Obr. č.3. Top 8 Big Data pracovních dovedností
Zdroj: vlastní zpracování

Apache Hadoop: Hadoop, programovací rámec vytvořený v programu Java, je určen pro zpracování rozsáhlých datových sad v distribuovaném výpočetním prostředí. Spouští několik aplikací v distribuovaných systémech s tisíci uzly zahrnující petabajty informací. Má distribuovaný systém souborů, nazývaný Hadoop Distributed File System nebo HDFS, který umožňuje rychlý přenos dat mezi jednotlivými uzly. Implementace Hadoop již nyní zahrnuje celou řadu souvisejících projektů, které poskytují bohatou nabídku velkých datových služeb:

- Apache Spark je distribuovaný procesní nástroj, který provádí vysoce výkonné zpracování velkých datových souborů.
- Apache Hive poskytuje vestavěné funkce pro ukládání dat do systému Hadoop pomocí přístupových metod typu SQL pro ověřování dat a jejich analýzu.
- Apache HBase je škálovatelná, distribuovaná NoSQL sloupcová databáze postavená na vrcholu HDFS.
- Apache Zeppelin je webový víceúčelový notebook, který umožňuje interaktivní zpracování dat včetně funkcí přijímání, průzkumu, vizualizace a spolupráce pro Hadoop a Spark.



K ČEMU SLOUŽÍ BIG DATA?

NoSQL: Databáze NoSQL, včetně Couchbase, MongoDB, nahrazuje tradiční databáze SQL, jako je DB2, Oracle aj. Distribuované databáze NoSQL pomáhají ukládat a zpřístupnit velká data. Dále jsou doplněny odbornými znalostmi společnosti Hadoop. Odborníci se specializací NoSQL mohou v dnešním světě najít uplatnění téměř kdekoli.

Vizualizace dat: Nástroje vizualizace dat, jako je QlikView, Tableau, mohou pomoci porozumět analýze provedené analytickými nástroji. Komplexní pochopení technologií a procesů velkých dat je často velmi obtížné a právě zde se do popředí dostává jejich vizualizace. Profesionál dobře obeznámený s nástroji vizualizace dat může dostat šanci kariérního růstu u mnoha velkých organizací.

Porozumění nástrojům & práce s daty: Manipulace s daty a porozumění nástrojům patří mezi důležitá pole v oblasti velkých dat. Ačkoli pole působení velkých dat je obrovské, tyto dva ukazatele významně ovlivňují celý proces. Odborníci, kteří jsou schopni efektivně využívat přístroje pro provádění prediktivní a předepisovací analýzy, jsou v dnešní době stále ještě vzácní. Tato pole působení mohou přispět ve vývoji doporučujících systémů, klasifikaci a personalizaci. Odborníci, kteří umí efektivně manipulovat s daty patří mezi nejlépe placené skupiny lidí.

Kvantitativní analýza: Kvantitativní a statistická analýza je významnou součástí velkých dat, jelikož se zabývá samotnými čísly. Silný kompetenční základ ve statistice a matematice je velkou výhodou. Znalost nástrojů, jako je SAS, SPSS, R aj., je bezpochyby velmi přínosná při rozšiřování rámce dovedností.

Programovací jazyky: Určité univerzální programovací jazyky mohou pomoci získat nad ostatními důležité konkurenční výhody. Mezi tyto programovací jazyky patří Java, Python, C, Scala atd. I programátoři se zkušenostmi s analýzou dat jsou velmi žádanými.



K ČEMU SLOUŽÍ BIG DATA?

Řešení problémů: I když jste obeznámeni se všemi nástroji a technologiemi v oblasti velkých dat, také schopnost řešit problémy a kreativita může výrazně pomoci plnit úkoly. Implementace technik velkých dat pro efektivní řešení bude u profesionálů vyžadovat obě tyto schopnosti.

SQL: SQL je jazyk zaměřený na data, který funguje jako základ pro éru velkých dat. Znalost Strukturovaného dotazovacího jazyka (SQL) bude pro programátory další výhodou při práci na technologiích velkých dat, jedná se např. o NoSQL. Je také důležitou součástí výbavy Hadoop Scala.



NEJHODNOTNĚJŠÍ DOVEDNOSTI:

1. Týmová spolupráce
2. Obchodní prozíravost
3. Intelekt
4. Schopnost řešit problémy
5. Komunikační dovednosti

Ne všechny aplikace technologií související s velkými daty jsou určeny pro analýzu dat. Některé se používají při rozmístění webů pro sociální média nebo herní aplikace, další se používají pro velké obchody, které poskytují informační přístup k velkému množství dokumentů. Příklady zahrnují:

Analýzy (např. manipulace s daty, vícerozměrná analýza, vizualizace dat)

Operace (např. provoz webu, zpracovávání online objednávek)

Informační přístup (např. přístup k informacím na základě vyhledávání, normalizace, přístup prostřednictvím obsahu a datových zdrojů)



K ČEMU SLOUŽÍ BIG DATA?

Zdokonalením dovedností v oblasti Big Data budete schopni...



Obr. č. 4. Schopnosti při zlepšování dovedností Big Data.
Zdroj: vlastní zpracování



K ČEMU SLOUŽÍ BIG DATA?

Jak rozvíjet dovednosti v oblasti velkých dat?

Vizualizace dat a analytické dovednosti

Nástroje velkých dat provádějí analýzu dat, která je z důležitých informací z velkých datových souborů schopna odvodit důležité poznatky. Znalost obchodní oblasti vám může pomoci pochopit data, pro která je analýza prováděna. Odborníci na práci s daty mají schopnost interpretovat data jejich vizualizací. K pochopení komplexních dat pomocí kreativity a představitivosti je zapotřebí zvláštní vědecká a matematická přednost. Naučit se analytické nástroje vám může pomoci rozvinout vizualizaci dat a analytické dovednosti.

Schopnost programování

Schopnost kódovat a provádět statistické a kvantitativní analýzy je významným požadavkem na trhu velkých dat. Dobrý základ znalostí v matematice k tomu může významně přispět. Získat znalosti účelně orientujících se jazyků a základní znalost datových struktur a algoritmů může být však běh na dlouhou trať. Je důležité mít povědomí o třídících algoritmech, typech dat atd.

Seznámení se s technologiemi

Je důležité, aby profesionál z oblasti velkých dat byl obeznámen s řadou nástrojů a technologií, které jsou v průmyslu používány. Zvyšuje se tak počet nástrojů, se kterými můžete pracovat. Tyto nástroje pomáhají při provádění výzkumné analýzy a stanovování závěrů. Mezi typické nástroje patří SPSS, Excel, SQL, SAS, R, MatLab, Python, Linux, Hadoop, Scala atd. Existuje mnoho technologií s otevřeným zdrojovým kódem napsaným v jiném jazyce, což dává výhodu technickým odborníkům.

Zkušenosti z první ruky a osobní rozvoj

Při vzdělávání je důležité získat praktické zkušenosti s nástroji velkých dat. Jelikož se technologie mění velmi rychle, k rozvoji může přispět účast na některých kurzech. Interakce s databázemi může pomoci lépe porozumět datovým nástrojům. Strojové učení je jedním ze způsobů, jak získat lepší zkušenosti s nástroji velkých dat. Dobrým způsobem je účastnit se online vzdělávacích kurzů, které pomáhají dozvědět se více informací o nových technologiích.



OSVĚDČENÉ POSTUPY



Dnes probíhá téměř veškerá interakce přes internet nebo prostřednictvím spotřeby zboží a služeb. Veškeré informace jsou sledovány, ukládány a cíleným způsobem využívány. To vedlo k představě velkého množství dat, která odrážejí chování a jednání různých lidí. Vědci v oblasti manipulace s daty a různé platformy pro sběr dat jsou nyní schopni výpočetně organizovat petabajty a exabajty dat, takže je podstatně snazší analyzovat a identifikovat vzorce, které by jinak nemusely být zachyceny.



Technologie velkých dat jsou nasazovány na podporu procesů v komerčních, neziskových nebo vládních organizacích. Výzvy a problémy, kterým organizace čelí, nejsou velkými datovými výzvami, ale spíše obchodními nebo organizačními výzvami, které jsou pouze ovlivněny velkými daty. Případy nasazení technologie velkých dat lze nalézt napříč podnikovými procesy, jako například:

- Řízení vztahů se zákazníky (prodej, marketing, zákaznický servis atd.)
- Dodavatelský řetězec a operace
- Administrace (zaměřená na finance a účetnictví, lidské zdroje, právní, atd.)
- Výzkum a vývoj
- Řízení informačních technologií
- Řízení rizik



OSVĚDČENÉ POSTUPY



Historie BBVA je historií mnoha různých lidí, kteří byli součástí více než stovky finančních institucí, které se připojily k podnikovému úsilí od svého vzniku v polovině 19. století. Během hospodářského rozvoje šedesátých let až do současnosti se BBVA rozšířila, získala další banky a vytvořila solventní finanční skupinu. Prestižní finanční publikace ocenily efektivitu integrace BBVA a označily ji za nejlepší banku na světě (Forbes) a ve Španělsku (The Banker), v roce 2000 nejlepší banku v Latinské Americe (Forbes) a v roce 2001 za nejlepší banku v Evropě (Lafferty).

Společenská odpovědnost je jádrem jejího obchodního modelu. BBVA podporuje finanční vzdělávání a začleňování a podporuje vědecký výzkum a kulturu. Funguje s nejvyšší integritou, dlouhodobou vizí a používá nejlepší postupy. Věří, že znalosti získané z finančních údajů mohou transformovat bankovní sektor. Implementuje a využívá nejpokročilejší analytiku a umělou inteligenci, aby nabídli nejlepší digitální interakci se zákazníkem. Výzvou společnosti není jen shromažďování a uchování dat, ale také jak nejlépe těmto datům porozumět - jak je dobře využívat.

Příklady vedoucích společností:





PŘÍNOS PRO SPOLEČNOSTI

Řešení managementu velkých dat:

- ❑ Poskytovat společností možnost přidávat různé údaje ze stovek různých zdrojů. Tímto způsobem může dojít k zvýšení angažovanosti klienta na základě účinnější interakce a lepších marketingových návrhů, které nakonec povedou k pevnějšímu a efektivnějšímu vztahu společností s klientem.
- ❑ Odstranit mezery na trhu, aby společnosti mohly získat jedinečný pohled na zákazníky, který umožní sestavit podrobné záznamy o jejich aktivitách. Sestavené profily poskytují společností možnost porozumět chování svých klientů na globální úrovni.
- ❑ Poskytovat společností kompletní profily zákazníků, což umožňuje více personalizované zkušenosti zákazníků v každém bodě, kde dochází ke kontaktu se společností.

Skutečnost je taková, že s rostoucím množstvím dat, se jejich poskytování společností exponenciálně zvyšuje. To následně společností umožňuje s přesností převádět nezpracovaná data v relevantní údaje, předpovědi a trendy.

Velká data jsou aktuální: 60 % každého pracovního dne tráví pracovníci se znalostmi velkých dat snahou najít je a spravovat je. Velká data mohou okamžitě poskytovat aktuální a nejnovější zprávy.

Velká data jsou přístupná: Polovina vedoucích pracovníků naznačuje, že přístup ke správným údajům je často obtížný.

Velká data jsou celistvá: Pohyb informací probíhá v současné době v mnoha společnostech jen v omezené míře. Například marketingová data lze nalézt ve webové analýze, mobilní analýze, sociální analýze, CRM systémech, testovacích nástrojích A/B, e-mailových marketingových systémech a mnoha dalších webech, každý se zaměřením na své struktury.



PŘÍNOS PRO SPOLEČNOSTI

Velká data jsou spolehlivá: I věci tak jednoduché, jako je zajištění správnosti zákaznických údajů prostřednictvím kontroly více systémů může při nesprávném odeslání znamenat ztrátu tisíce dolarů.

Velká data jsou relevantní: 43 % společností není spokojeno se schopností svých nástrojů filtrovat irelevantní data.

Velká data jsou bezpečná: Porušení zabezpečení dat stojí stovky dolarů na každého zákazníka.

Velká data jsou přesná: Podniky mají potíže s objektivními verzemi skutečnosti na základě poskytnutí rozdílných informací. Díky kombinaci několika spolehlivých zdrojů mohou podniky vytvořit přesný odhad.

Velká data jsou užitečná: Mnoho společností činí špatná rozhodnutí kvůli zastaralým nebo špatným datům. Velká data mohou zajistit, že bezchybná data.



Výhody

- Lepší rozhodování
- Zvýšená produktivita
- Snížení nákladů
- Lepší zákaznický servis

- Zvýšené příjmy
- Zvýšená pohyblivost
- Větší inovativnost
- Rychlejší uvedení na trh
- Detekce podvodů



Nevýhody

- Potřeba talentu
- Problémy s kvalitou dat
- Potřeba kulturní změny
- Soulad s předpisy

- Rizika kybernetické bezpečnosti
- Technologické změny
- Hardwarové potřeby
- Problémy s integrací starších systémů

Obr. č. 5. Výhody a nevýhody Big Data
Zdroj: vlastní zpracování



BUDOUCÍ VÝVOJ



Význam velkých dat se netočí kolem množství dat, kterými společnost disponuje, ale jak společnost shromážděná data využívá. Každá společnost používá data svým vlastním způsobem; čím účinněji používá svá data, tím její potenciál roste. Společnost může při hledání odpovědí brát data a analyzovat je z jakéhokoli zdroje považuje za vhodné.



Obr. č. 8: Přínos velkých dat pro společnosti
Zdroj: vlastní zpracování

Používání velkých dat se stává běžnou součástí boje s konkurencí. Ve většině průmyslových odvětví budou stávající konkurenti i noví účastníci používat strategie vyplývající z analyzovaných údajů a inovací.

Velká data pomáhají společnostem vytvářet nové příležitosti k růstu a zcela nové kategorie společností, které mohou kombinovat a analyzovat průmyslová data. Tyto společnosti mají bohaté informace o produktech a službách, kupujících a dodavatelích nebo preferencích spotřebitelů.



BUDOUCÍ VÝVOJ



Praktické využití velkých dat:

Každé průmyslové odvětví používá velká data jinými způsoby. V našem seznamu jsme zkomplementovali použití velkých dat a jaká průmyslová odvětví je používají.

Sledování polohy: Logistické společnosti již delší dobu pomocí analýzy polohy sledují a zadávají objednávky. Velká data poskytují možnost sledovat stav zboží v tranzitu a odhadovat ztráty. Nyní je možné shromažďovat údaje o dopravě a počasí a definovat trasy pro přepravu. To může pomoci logistickým společnostem zmírnit rizika v dopravě, zvýšit rychlost a spolehlivost dodávek.



Precizní medicína: Velká data mohou nemocnicím pomoci zlepšit úroveň péče o pacienty. Monitorování 24x7 lze zajistit pacientům s intenzivní péčí bez nutnosti přímého dohledu. Kromě toho lze účinnost léků zlepšit analýzou minulých záznamů o pacientech a léčích, které jim byly doposud poskytovány. V případě některých

biofarmaceutik existuje mnoho proměnných, které ovlivňují konečný produkt. Například při výrobě inzulínu je třeba věnovat zvýšenou pozornost zajištění produktu požadované kvality. Analýzou všech faktorů ovlivňujících finální lék může analýza velkých dat poukázat na klíčové faktory, které by mohly zkomplikovat výrobu.



BUDOUCÍ VÝVOJ



Detekce a manipulace s podvody:

Bankovní a finanční sektor používá Big Data k předvídání a prevenci počítačových zločinů, k odhalování podvodů s kartami, k archivaci auditních záznamů atd. Analýzou dat o předchozích útocích mohou banky předpovídat budoucí útok. Velká data pomáhají nejen při předpovídání počítačové kriminality, ale také při řešení problémů souvisejících s chybnými transakcemi a selháním v internetovém bankovníctví. Komise pro cenné papíry (SEC) používá velká data ke sledování finančních trhů pro možné nezákonné obchody a podezřelé činnosti. SEC používá síťovou analytiku a procesory přirozeného jazyka k identifikaci možných podvodů na finančních trzích.



Reklama: Inzerenti jsou jedním z největších hráčů v oblasti velkých dat. Ať už je to Facebook, Google, Twitter nebo jakýkoli jiný online gigant, všechny přesně sledují chování uživatelů. Tito internetoví giganti poskytují velké množství uživatelských údajů inzerentům, kteří je využívají pro jejich kampaně. Vezměte si například Facebook, zde můžete cílit na lidi na základě jejich

nákupního záměru, návštěv na webových stránkách, zájmů, atd. Všechna tato data jsou shromažďována pomocí algoritmů Facebooku pomocí technik analýzy velkých dat. Totéž platí pro Google, když zacílíte na lidi na základě kliknutí, získáte různé výsledky a když vytvoříte kampaň pro potenciální zákazníky, získáte jiné výsledky.

Zábava a média: Důležitou roli hraje zveřejnění správného obsahu ve správný čas. Na základě předešlého online chování se zobrazí různá doporučení, která cílí na konkrétní priority uživatelů. Netflix a Youtube tuto techniku běžně používají ke zvýšení zapojení a zvýšení výnosů. To umožní větší příjem z reklam a poutavější uživatelského prostředí.





BUDOUCÍ VÝVOJ



TRENDY:

❖ Rychle rostoucí síť internetu věcí (IoT)

Dnes je již celkem běžné, že se naše chytré telefony používají k ovládní domácích spotřebičů, a to díky technologii zvané Internet věcí (IoT). S inteligentními zařízeními, jako jsou Google Assistant a Microsoft Cortana, je stále větším trendem domácností automatizovat konkrétní úkoly a investovat do novějších technologií. To povede k více způsobům sběru obrovského množství dat a spolu s tím i k jejich řízení a analýze. Odezvou odvětví je prosazovat více nových zařízení, která jsou schopnější sbírat, analyzovat a zpracovávat data.

Internet věcí (IoT)



Internet věcí, nebo IoT, je systém vzájemně propojených výpočetních zařízení, mechanických a digitálních strojů, objektů, zvířat nebo lidí, které jsou vybaveny jedinečnými identifikátory (UID) a schopností přenášet data přes síť, aniž by vyžadovaly interakci člověka s počítačem.

❖ Přístupná umělá inteligence (AI)

Umělá inteligence je nyní běžněji využívána, aby pomohla velkým i malým společnostem zlepšit jejich procesy v obchodování. Programy umělé inteligence nyní mohou provádět úkoly, díky nimž jsou rychlejší a přesnější než lidé, omezují chyby a zlepšují celkový chod. To umožňuje lidem lépe se soustředit na své úkoly a dále zvyšovat kvalitu služeb. Dobrou zprávou je, že každý může mít přístup k předinstalovaným strojům, které provozují aplikace AI a řeší rostoucí poptávku, která v rámci stejných odvětví vyrovnává podmínky mezi společnostmi. Jednotlivé organizace mohou získat výhodu, pokud najdou nejefektivnější způsob, jak integrovat mechanismy do svého obchodního procesu.



BUDOUCÍ VÝVOJ



❖ Vzestup prediktivní analýzy

Analýza velkých dat byla vždy klíčovou strategií pro podniky, které chtěly dosahovat konkurenčních výhod. Používají potřebné analytické nástroje ke zpracování velkých dat, na základě kterých určují důvody, proč k určitým událostem dochází. Nyní může prediktivní analýza pomocí velkých dat pomoci předpovědět, co se může v budoucnu stát. Není pochyb o tom, že tento druh strategie je vysoce účinný při analýze shromážděných informací a předpovídání chování spotřebitelů. To společně umožňuje určit opatření, která jim pomohou předpovídat kroky zákazníků dříve než je znají oni sami. Analýza může dokonce poskytnout více údajů, které pomou pochopit celkový kontext.

❖ Migrace „tmavých“ dat do claudu

Informace, které se teprve mají transformovat do digitálního formátu, se nazývají „tmavá“ data a v současnosti se jedná o obrovskou zatím nevyužitou zásobárnu. Očekává se, že tyto analogové databáze budou digitalizovány a přesunuty do claudu, takže je lze použít pro prediktivní analýzu, která je velkým přínosem pro podniky.

❖ Větší role pozice Chief Data Officers

Nyní, když se Big Data stále více stává nezbytnou součástí provádění obchodních strategií, má pozice Chief Data Officer (odpovědnost za správu dat) stále významnější roli. Očekává se, že zaujmou aktivnější pozici, která povede společnosti správným směrem. Tento trend otevírá dveře „vývojářům dat“ při svém rozvoji a kariérním růstu.



BUDOUCÍ VÝVOJ



❖ Kvantové výpočty

Získání analýzy a interpretace velkého množství dat může při současných technologiích, které používáme, zabrat spoustu času. Pokud by jsem dokázali během několika minut zpracovat miliardy dat najednou, společnosti by měly šanci učinit včasná rozhodnutí při dosahování požadovaných výsledků. Tento obrovský závazek lze dosáhnout pouze kvantovými výpočty. Přestože jsou stále na počátku, v současné době už se experimenty provádějí na kvantových počítačích s cílem pomoci při praktickém a teoretickém výzkumu napříč různými průmyslovými odvětvími. Brzy velké technologické společnosti, jako jsou Google, IBM a Microsoft, začnou testovat kvantové počítače a integrovat je do svých obchodních procesů.

❖ Kybernetická bezpečnost

Organizace, které se nevyhnuly skandálům týkajících se hackerství a jiného narušení systému, i tak rostou rapidním způsobem. Předěšlé události je však přiměly zaměřit se na posílení důvěrnosti informací. Internet věcí je také důvodem k obavám ze všech shromažďovaných údajů. Pro řešení této trvalé hrozby se společnosti zabývající se velkými daty zasazují o pomoc organizacím používat analytiku dat jako nástroj k předvídání a detekci hrozeb kybernetické bezpečnosti. Velká data lze integrovat do strategie kybernetické bezpečnosti prostřednictvím dat protokolu zabezpečení, která lze použít k poskytování informací o minulých hrozbách. To může společností pomoci předcházet a zmírňovat dopad budoucích hacků a narušení dat.



BUDOUCÍ VÝVOJ



❖ Řešení s otevřeným zdrojovým kódem

K dispozici je mnoho veřejných datových řešení, jako je například software s otevřeným zdrojovým kódem, který významně zrychluje zpracování dat. Z tohoto důvodu se očekává, že budou od roku 2019 velmi žádané. Není pochyb o tom, že software s otevřeným zdrojovým kódem je levnější, protože může pomoci vašemu podnikání snížit provozní náklady.

❖ Edge Computing

Edge computing je způsob, kdy dochází ke zpracovávání dat blíže k místu jejich vytvoření, vyhýbá se tak jejich přesouvání do cloudu. Poskytuje lepší výkon a snižuje náklady, protože v síti v konečném důsledku proudí méně dat. Společnost může také těžit z nepotřebných nákladů na skladování a infrastrukturu. Okrajové výpočty mohou navíc urychlit analýzu dat a poskytnout společností dostatek času reagovat.

❖ Chytřejší Chatbots

Díky inteligentnější AI využívají společnosti chatové služby, prostřednictvím kterých vyřizují dotazy zákazníků, poskytují přizpůsobující se interakci a současně eliminují potřebu skutečného lidského personálu. Velká data mají hodně co do činění s příjemnější zákaznickou zkušeností, protože specializované přístroje jsou schopny poskytnout relevantní odpovědi na základě zadaných klíčových slov při zákaznickém dotazování. Během interakcí jsou také schopni shromažďovat a analyzovat podrobnější informace o zákaznících. Tento proces může obchodníkům pomoci vyvinout efektivnější strategii k dosažení lepších výsledků.



POKROČILÝ OBSAH

Rozsáhlé sbírky informací dostupných na webu a v cloudu by mohly pomoci zabránit další finanční krizi. Klíč spočívá v tom, že všechno (ať už je to osoba, firma nebo produkt), poskytuje jedinečný identifikátor.

Trh s údaji v EU byl v posledních letech analyzován několika studii a zprávami. Navzdory rostoucímu trhu však existují určité překážky:

- Evropa ve srovnání s USA pomalu přijímala datové technologie.
- Mezera v datových dovednostech.
- Komplikovanost existujících norem může zpomalit inovace.
- Ochrana soukromí a údajů. Spolehlivý právní rámec, jako je GDPR, je složitý.
- Oslovení nejrůznějších malých a středních podniků a start-upů - společnosti nejčastěji vznikající z podnikatelských prostředí, jako jsou akcelerátory nebo inkubátory, nejsou obvykle spojeny s iniciativami EU.

Služba Data Market Services se zrodila za účelem překonání bariér malých a středních podniků založených na datech, pokud jde o dovednosti v oblasti dat, podnikatelských příležitostí, právních otázek a standardizace prostřednictvím poskytování bezplatných podpůrných služeb.



POKROČILÉ INFORMACE

Chcete-li se naučit velká data, je důležité získat také praktické zkušenosti kromě teoretických znalostí. Organizace, které hledají analytiku dat a vědce, často preferují talent se speciální certifikací velkých dat v terénu.

Při budování kariéry v oblasti velkých dat existují některé důležité a užitečné certifikáty:

Cloudera Certifications

Postaven na platformě Apache Hadoop, poskytující nástroje k získávání maxima z vašich zákaznických dat.

MongoDB

Databáze s možností škálovatelných a flexibilních dokumentů

Apache Storm

Bezplatný a veřejně distribuovaný výpočetní systém. Storm usnadňuje spolehlivé zpracování neomezených datových toků

Scala

Programovací jazyk používaný pro funkční programování a statické systémy.

Hortonworks Certifications

Pomáhá profesionálům v Big Data Hadoop vytvořit platné a důležité renomé Hadoop.

R software

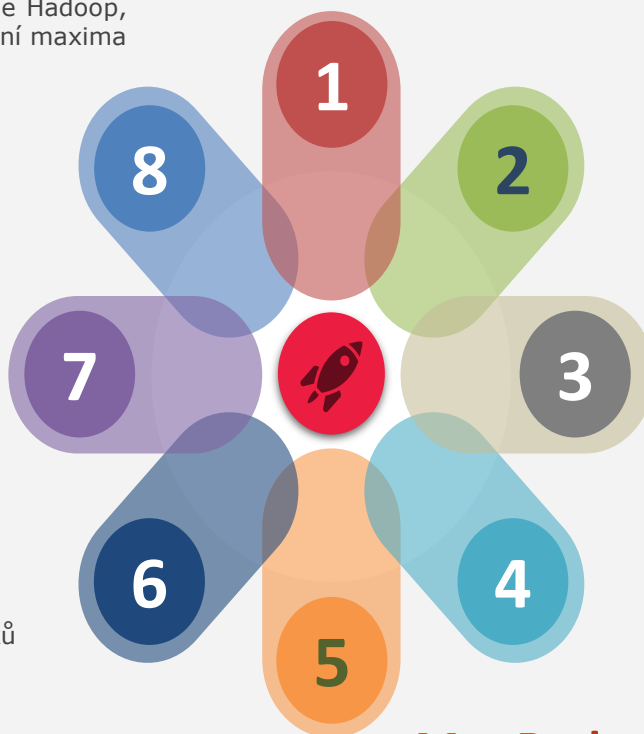
Jazyk a prostředí pro statistické výpočty a grafiku. R je integrovaná sada softwarových zařízení pro manipulaci s daty, výpočty a grafické zobrazování.

Python

Programovací jazyk na vysoké úrovni s integrovanou dynamickou sémantikou primárně určen pro vývoj webových aplikací.

MapReduce

Programový model pro distribuované výpočty založené na programu Java.



Obr. č. 9: Certifikace v budování kariéry v oblasti velkých dat
Zdroj: vlastní zpracování



POKROČILÉ INFORMACE

Vytvoření strategie

1

Identifikace záměru:

Váš konečný cíl má největší dopad na průběh vaší celkové strategie. Musíte se rozhodnout, zda chcete zvýšit efektivitu založenou na stálých zákaznících, zlepšit provozní efektivitu, zvýšit příjmy, poskytnout lepší zákaznické zkušenosti nebo zlepšit marketing. Cíl, který máte, by měl být přesný, jistý a přímý. Jakákoli strategie, jejímž jediným účelem je prozkoumat možnosti, bude pravděpodobně zmatená a nesprávná. Na základě vašeho cíle si můžete vybrat metodologii, najmout zaměstnance a vybrat správné zdroje dat. Vytvořte tedy cíle SMART (Specifické, měřitelné, dosažitelné, relevantní a aktuální) a podle toho plánujte.

2

Využití ověřené strategie v oblasti velkých dat

Existují 4 osvědčené způsoby, jak vytvořit fungující strategii velkých dat. Na základě vašeho konečného cíle a dostupnosti dat si můžete vybrat některou z níže uvedených strategií pro lepší dosažení výsledků:

- A. Řízení výkonu:** Jedná se o použití transakčních dat, jako je historie nákupů zákazníků, obrat a úroveň zásob, rozhodování o správě obchodů a nadřazenosti provozu. Tato data jsou v rámci organizace k dispozici a umožňují nahlédnout do témat týkajících se krátkodobého rozhodování a dlouhodobého plánování. Funguje dobře u společností s velkými historickými databázemi. Může také pomoci s lepší segmentací a cílením na zákazníky.



POKROČILÝ OBSAH

- B. Zkoumání dat:** Tento přístup se snaží získat data k jejich výzkumu a nacházení řešení a korelací, které nelze snadno zjistit pomocí vlastních dat. V současné době je využíván společnostmi zaměřujícími se na robustní marketing, který vytváří přehled o chování potenciálních zákazníků na webových stránkách. Pomáhá identifikovat nové segmenty dat a poskytuje informace o chování a preferencích zákazníka.
- C. Sociální analytika:** Sociální analytika měří netransakční data na různých sociálních sítích a recenzovaných webech, jako je Facebook, Twitter a Google+. Je založen na analýze konverzací a recenzí, které se na platformách objevují. Uvádí tři primární analýzy: povědomí, angažovanost a ústní vyjadřování. Techniky analýzy dat, jako je analýza postojů a mínění, se v těchto případech ukázaly jako velmi efektivní. Poskytuje informace o identitě značky a názorech zákazníků na nové nabídky a služby. Sociální analýza se také osvědčila při předpovídání maximální poptávky po určitých produktech.
- D. Rozhodování:** Způsob rozhodování se týká experimentů a analýzy netransakčních dat, jako je obsah generovaný spotřebitelem, nápady a recenze. Jde spíše o zkoumání možností než měření cílů. Na rozdíl od sociální analýzy, která je založena na analytice angažovanosti, strategie založená na rozhodování zaměřuje na testování hypotéz a myšlenkový proces. Zahrnuje rozsáhlé používání analýzy textů a postojů k pochopení názorů zákazníků na nové služby a schémata.



POKROČILÝ OBSAH

3

Identifikace změn v infrastruktuře

Chcete-li využít velká data, zejména historické databáze, budete muset vytvořit mnoho infrastrukturních změn ve společnosti. Pokud byla stará firemní data uložena v tradičních formátech, může to zkomplikovat funkci algoritmů a analýz. Různá oddělení mohou navíc potřebovat vzájemnou integraci, aby mohla efektivněji pracovat. Integrace mezi různými odděleními je klíčem k zavádění a provádění změn. Pokud vaše stávající infrastruktura není správně propojena, budete se muset připravit na velké změny.

4

Založení Talent Pool:

Lidské zdroje jsou jedním z nejdůležitějších aspektů vytváření strategie velkých dat. Váš tým musí mít statistické údaje, které dávají smysl, obchodní analytiku, kteří sdělují informace nadřízeným osobám, které jsou schopny tým vést. Bez řádně vedeného týmu se mohou diskuse o velkých datech stále točit kolem praktik, které nejsou nikomu jasné. Je třeba vytvořit vhodný jazyk, který usnadní diskuse mezi vedoucími podniků a technickým týmem.

5

Brát v potaz názory zákazníka

Klíčovým využitím velkých dat je vytváření poznatků, které mohou společností pomoci být v lepším kontaktu se svými zákazníky. Marketing orientovaný na zákazníka je nový způsob, jak se prosadit na trhu a dosáhnout větších výnosů. I na konci dne musíte být schopni zákazníkovi sdělit, že jste tady, abyste vyřešili jeho problém a nejen vydělali peníze. Společnosti s velkými daty poskytují takové informace o názorech svých zákazníků, které lze využít při zlepšení a změnách současných marketingových praktik. Další věc, na kterou je dobré zaměřit se, je vytvořit hranici mezi shromažďováním údajů a nezneužíváním soukromí. Vaši zákazníci by se neměli cítit, jako by byli neustále pod dohledem.



POKROČILÝ OBSAH

6

Zajistit použitelnost

Mnohokrát se stává, že poznatky vytvořené statistikem jsou pro zaměstnance nepochopitelné. Data, analýzy a poznatky, které shromažďují analytici, je třeba přesně sdělit implementačnímu týmu. Informace by měly být srozumitelné a reprezentovány tak, aby je mohli pochopit i lidé, kteří nepocházejí ze statistického prostředí. Toho lze dosáhnout pomocí grafické reprezentace a předáním přímých pokynů zúčastněným týmům.

7

Být agilní

Při implementaci technologií by mohlo dojít k mnoha překážkám, o kterých se původně neuvažovalo. Musí dojít k úpravě rozpočtu, lidé a ideologie se musí přizpůsobit okolnostem a současným poznatkům. Nejlepší je vytvořit plán a podle potřeby provádět změny. Možná přijdete s plánem, který se vymyká současné strategii, do budoucna se však vyplatí.



POKROČILÉ INFORMACE

NÁSTRPOJE V OBLASTI VELKÝCH DAT:



FusionCharts Suite XT

Interaktivní grafy pro zprávy, dashboardy, analýzy, monitory a průzkumy.

QlikView

Software pro obchodní inteligenci a vizualizaci dat. Můžete analyzovat data a využít jejich objevy při rozhodování.

Tibco Spotfire

Je velmi výkonný nástroj pro vizualizaci dat, který uživatelům umožňuje přístup a kombinování dat. Poskytuje rychlý přehled o vašich datech.

Watson Analytics

Je inteligentní služba pro analýzu a vizualizaci dat v cloudu, která pomáhá téměř každému rychle objevit vzory a význam svých datech.

Sisense

Je to platforma, která poskytuje analytickým tvůrcům nástroje ke zjednodušení složitých dat.

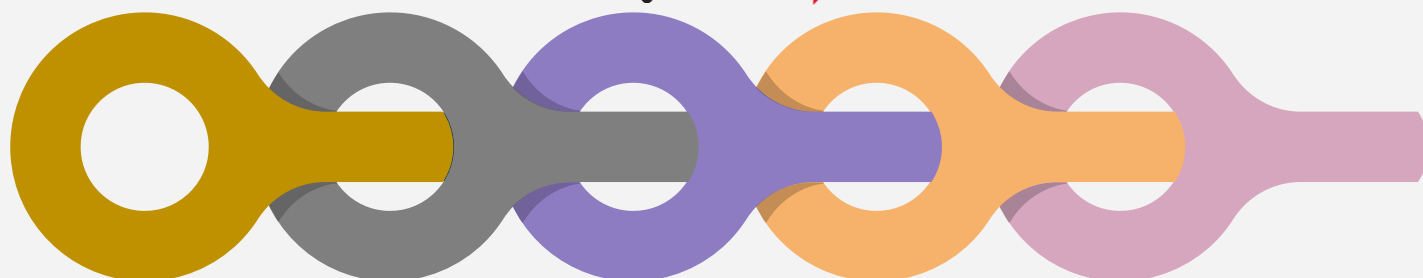


Datwrapper



infogram

plotly



Tableau

Můžete vytvářet interaktivní a flexibilní uživatelské rozhraní využívající vlastní filtry a funkce.

Datwrapper

Je snadno použitelný nástroj pro vytváření vizualizací, jako jsou infografiky, mapy, datové tabulky a grafy.

Microsoft Power BI

Jedná se o analytický obchodní nástroj, který usnadňuje obchodníkům vizuální analýzu dat a tvorbu strategií.

Infogram

Usnadňuje přístup k datům tím, že umožňuje editovat data a připojit se k požadované cloudové službě.

Plot.ly

Grafické uživatelské rozhraní pro import a analýzu dat a použití statistických nástrojů.



VZDĚLÁVÁNÍ



Otestujte si své znalosti nástrojů velkých dat v cloudu pomocí následujícího kvízu:

Nástroj pro osobní hodnocení:

<https://searchcloudcomputing.techtarget.com/quiz/Test-your-knowledge-of-big-data-cloud-services>

Stupně vzdělání

- MSIT: BUSINESS INTELLIGENCE & DATA ANALYTICS (BIDA) - Carnegie Mellon University's Heinz College
- M.S. in Statistics: Data Science - Stanford University
- Big Data, Strategic Decisions: Analysis to Action - Stanford Graduate School of Business
- Master of Science in Data Science – ETH Zurich

Online vzdělávací kurzy

- Big Data Analysis: Hive, Spark SQL, DataFrames and GraphFrames -Coursera
- Big Data Applications: Machine Learning at Scale -Coursera
- Managing Big Data with MySQL - Coursera
- Intro to Machine Learning - Udacity



VZDĚLÁVÁNÍ



Externí příručky s dalšími informacemi

- Data Analytics Made Accessible, by A. Maheshwari
- Lean Analytics: Use Data to Build a Better Startup Faster, by A. Croll and B. Yoskovitz
- Big Data and Hadoop Tutorial - Intellipaat
- Introduction to Machine Learning
- Introduction to Data Science
- Big Data Quick Exploratory Self-Assessment Guide

Certifikace

- Cloudera Certified Professional
- Intellipaat Big Data Hadoop Certification
- Microsoft's MCSE : Data Management and Analytics
- Hortonworks Hadoop Certification



LITERATURA

- ❖ European Commission (2018). *Big Data*. Digital Single Market. Retrieved from <https://ec.europa.eu/digital-single-market/en/big-data>
- ❖ European Commission (2014). *Worldwide Big Data Technology and Services - 2012-2015 Forecast*. Digital Single Market. Retrieved from <https://ec.europa.eu/digital-single-market/en/news/worldwide-big-data-technology-and-services-2012-2015-forecast>
- ❖ European Commission (2013). *FACTSHEET: What is Big Data?*. Retrieved from http://europa.eu/rapid/press-release_MEMO-13-965_en.htm
- ❖ Gaitho, M. (2018). *How Applications of Big Data Drive Industries*. Retrieved from <https://www.simplilearn.com/big-data-applications-in-industries-article>
- ❖ Simplilearn. (2018). *9 Must-have skills you need to become a Data Scientist, updated*. Retrieved from <https://www.kdnuggets.com/2018/05/simplilearn-9-must-have-skills-data-scientist.html>
- ❖ Burtch, L. (2014). *The Must-Have Skills You Need to Become a Data Scientist*. Burtch Works. Retrieved from <https://www.burtchworks.com/2014/11/17/must-have-skills-to-become-a-data-scientist/>
- ❖ SAS. (2013). *Big Data Analytics An assessment of demand for labour and skills, 2012-2017*. E-skills UK. Retrieved from <https://ec.europa.eu/digital-single-market/en/news/big-data-analytics-assessment-demand-labour-and-skills-2012-2017>
- ❖ *5 Practical Uses of Big Data*. (2017). Retrieved from <https://www.newgenapps.com/blog/5-practical-uses-of-big-data>
- ❖ *Top 13 Best Big Data Companies of 2019*. (2019). Retrieved from <https://www.softwaretestinghelp.com/big-data-companies/>
- ❖ Vesset, D., Morris, H.D., Little, G., Borovick, L., Feldman, S., Eastwood, M., ... Yezhkova, N. (2012). *Worldwide Big Data Technology and Services 2012 – 2015 Forecast*. Framingham, USA: IDC. Retrieved from https://www.researchgate.net/profile/Shafagat_Mahmudova/post/How_can_big_data_analytics_and_AI_apply_to_risk_and_contingency_management/attachment/59d6525979197b80779aa96a/AS%3A511969745489920%401499074507185/download/Big_Data_Analytics_as_a_Service_for_Business_Intelligence1.pdf



OSOBNÍ PŘÍNOS



- ★ Mám již jasnou představu o tom, co představují velká data?
- ★ Jaké nástroje mám při práci s velkými daty ve zvyku používat?



- ★ Zním výhody, které mohou velká data přinést mé společnosti?
- ★ Umím rozpoznat výhody a nevýhody implementace velkých dat v mé společnosti?



PŘEDSTAVENÍ 4. PRŮMYSLOVÉ REVOLUCE

Tento projekt byl realizován za finanční podpory Evropské unie. Za obsah publikací odpovídá výlučně autor. Publikace nerepresentují názory Evropské komise a Evropská komise neodpovídá za použití informací, jež jsou jejich obsahem.