

# Metodika zpracování rozsáhlých datových souborů v organizacích

## Bakalářská práce

*Studijní program:*

B2646 Informační technologie

*Studijní obor:*

Informační technologie

*Autor práce:*

**Martin Vrátný**

*Vedoucí práce:*

Ing. Věra Pelantová, Ph.D.

Ústav mechatroniky a technické informatiky





## Zadání bakalářské práce

# Metodika zpracování rozsáhlých datových souborů v organizacích

*Jméno a příjmení:* **Martin Vrátný**  
*Osobní číslo:* M16000065  
*Studijní program:* B2646 Informační technologie  
*Studijní obor:* Informační technologie  
*Zadávací katedra:* Ústav mechatroniky a technické informatiky  
*Akademický rok:* 2020/2021

### Zásady pro vypracování:

1. Vytvořte úvod do metod zpracování rozsáhlých datových souborů.
2. Provedte srovnávací analýzu metod zpracování rozsáhlých datových souborů.
3. Provedte průzkum současného stavu zpracování rozsáhlých datových souborů v organizaci.
4. Stanovte vztah zpracování rozsáhlých dat vůči procesnímu přístupu a Průmyslu 4.0.
5. Stanovte doporučení pro organizace.
6. Závěr.

*Rozsah grafických prací:*  
*Rozsah pracovní zprávy:*  
*Forma zpracování práce:*  
*Jazyk práce:*

dle potřeby dokumentace  
30–40  
tištěná/elektronická  
Čeština



### **Seznam odborné literatury:**

- [1] Hruška, L. – Foldynová, I. – Sciskalova, M. – Fujak, R. – Hrušková, A. – Kubáň D. Metodika optimalizace sběru dat. Ostrava: PROCES – Centrum pro rozvoj obcí a regionů, s.r.o, 2016. Výzkum systému sběru dat ve veřejné správě. Reg. Č. TB9500MV006.
- [2] Karafiát, M. Big Data Metody zpracování a analýzy velkých dat. [Bakalářská práce.] Zlín: Univerzita Tomáše Bati ve Zlíně, FAI. 2017. [Online.] Dostupné na: <https://theses.cz/id/g8dbsk/fulltext.txt>
- [3] Hendl, J. – Remr, J. Metody výzkumu a evaluace. Praha: Portál, 2017. ISBN:978-80-262-1192-1.

*Vedoucí práce:*

Ing. Věra Pelantová, Ph.D.  
Ústav mechatroniky a technické informatiky

*Datum zadání práce:*

9. října 2020

*Předpokládaný termín odevzdání:*

17. května 2021

prof. Ing. Zdeněk Plíva, Ph.D.  
děkan

L.S.

doc. Ing. Milan Kolář, CSc.  
vedoucí ústavu

## Prohlášení

Prohlašuji, že svou bakalářskou práci jsem vypracoval samostatně jako původní dílo s použitím uvedené literatury a na základě konzultací s vedoucím mé bakalářské práce a konzultantem.

Jsem si vědom toho, že na mou bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb., o právu autorském, zejména § 60 – školní dílo.

Beru na vědomí, že Technická univerzita v Liberci nezasahuje do mých autorských práv užitím mé bakalářské práce pro vnitřní potřebu Technické univerzity v Liberci.

Užiji-li bakalářskou práci nebo poskytnu-li licenci k jejímu využití, jsem si vědom povinnosti informovat o této skutečnosti Technickou univerzitu v Liberci; v tomto případě má Technická univerzita v Liberci právo ode mne požadovat úhradu nákladů, které vynaložila na vytvoření díla, až do jejich skutečné výše.

Současně čestně prohlašuji, že text elektronické podoby práce vložený do IS/STAG se shoduje s textem tištěné podoby práce.

Beru na vědomí, že má bakalářská práce bude zveřejněna Technickou univerzitou v Liberci v souladu s § 47b zákona č. 111/1998 Sb., o vysokých školách a o změně a doplnění dalších zákonů (zákon o vysokých školách), ve znění pozdějších předpisů.

Jsem si vědom následků, které podle zákona o vysokých školách mohou vyplývat z porušení tohoto prohlášení.

23. dubna 2021

Martin Vrátný

## **Anotace**

Tato práce se věnuje metodám zpracování dat v malých a středních podnicích. Zkoumá aktuální stav zpracování datových souborů ve středně velkém výrobním podniku. Popisuje nedostatky výrobního podniku z pohledu datové politiky a možnosti přechodu podniku k Průmyslu 4.0 a procesnímu přístupu. Na závěr práce jsou uvedena doporučení autora pro malé a střední organizace ohledně zkoumané problematiky.

## **Klíčová slova**

Analýza, Big data, Cloud, FMECA, Procesní přístup, Průmysl 4.0

## **Annotation**

This bachelor thesis describes data processing methods in small and medium enterprises. It examines the current state of data file processing in a medium-sized manufacturing company. It describes the shortcomings of the manufacturing company in terms of data policy and the possibility of the transition of the company to Industry 4.0, and the process approach. At the end of the thesis, the author gives recommendations for small and medium-sized organizations regarding the researched issues.

## **Keywords**

Analysis, Big data, Cloud, FMECA, Industry 4.0, Process approach

# Obsah

1 Úvod k velkým data setům.....	12
1.1 Před Big Daty.....	12
1.2 Big data.....	13
1.3 Problémy Big dat.....	14
1.4 Trendy Big dat.....	15
2 Analýza Big dat.....	17
2.1 Hledání pravidel asociace.....	17
2.2 Genetické algoritmy.....	17
2.3 Umělá inteligence.....	18
2.4 Textová analytika.....	19
3 Typy datových analýz.....	20
3.1 Deskriptivní a diagnostická analýza.....	21
3.2 Prediktivní analýza.....	21
3.3 Preskriptivní analýza.....	23
4 Porovnání metod analýzy pomocí korelační matice QFD metody.....	24
5 Průzkum aktuálního stavu podniku.....	28
5.1 O společnosti.....	28
5.2 Předpokládaný vývoj.....	29
5.3 IT ve společnosti.....	30
5.3.1 Systémy IT.....	30
5.3.2 Problémy z IT strany.....	32
5.4 Popis funkčnosti firmy.....	32
5.5 Výrobní proces a jeho neshody.....	34
5.5.1 Nedostatek odborného personálu.....	39
5.5.2 Absence datové politiky podniku.....	39
5.5.3 Zastaralý ERP systém a Shadow IT.....	40
5.5.4 Nedostačující komunikace v podniku.....	42

5.5.5 Nedostatečná automatizace datových toků.....	42
6 Procesně řízená Big data v Průmyslu 4.0.....	44
6.1 Procesní řízení.....	44
6.1.1 Životní cyklus dat.....	45
6.1.2 Z datového pohledu.....	45
6.2 Průmysl 4.0.....	46
6.2.1 Technologie a problémy.....	48
6.2.2 Význam pro zkoumaný podnik.....	49
7 Cloudová řešení.....	50
8 Doporučení pro podniky.....	54
9 Závěr.....	57



## Seznam obrázků

Obrázek 1 Korelační matice QFD metody, porovnání metod analýzy pro potřeby podniku [Vlastní zpracování podle [49]].....	25
Obrázek 2 Hierarchická organizační struktura společnosti DZD [21].....	29
Obrázek 3 Schéma IT systémů v podniku [Interní zdroj podniku].....	31
Obrázek 4 Legenda FMECA analýzy [Vlastní zpracování podle [47]].....	35
Obrázek 5 Architektura pro nedostatky ERP řešení [Vlastní zpracování].....	41
Obrázek 6 Životní cyklus dat se souvisejícími technologiemi a metodami [Vlastní zpracování podle [48]].....	45

## Seznam tabulek

Tabulka 1 Klasifikace přístupů k datům podle Dai [31].....	20
Tabulka 2 Hlavní metody prediktivní analýzy podle Sheng [25].....	22
Tabulka 3 Typy analýz podle Sheng [25].....	23
Tabulka 4 FMECA analýza neshod v podniku [Vlastní zpracování na základě [46]]...	36

## Seznam zkratk

AI	Artificial Intelligence (Umělá inteligence)
BI	Business intelligence
CSP	Cyber Security Practitioner (kyberfyzický systém)
DML	Data Manipulation Language (Jazyk pro manipulaci s daty)
DZD	Družstevní závody Dražice – strojírna s.r.o.
ERP	Enterprise Resource Planning (Podnikové plánování zdrojů)
ETL	Extract, Transform, Load (Extrakce, transformace, nahrání)
FMECA	Failure Modes, Effects and Criticality Analysis
HW	Hardware
IBM	International Business Machines Corporation
IoT	Internet of Things (Internet věcí)
IT	Informační technologie
MES	Manufacturing execution system (Výrobní informační systém)
ML	Strojové učení
MRP	Manufacturing Resource Planning (Výrobní plánování zdrojů)
MSSQL	Microsoft Structured Query Language
NLP	Zpracování přirozeného jazyka
NoSQL	Not only SQL
ODBC	Open Database Connectivity
QFD	Quality Function Deployment
RDBMS	Relačně databázový systém managementu
SQL	Structured Query Language
SW	Software
WMS	Warehouse Management System (Systém řízení skladu)
WWW	World Wide Web (web)
XML	Extensible Markup Language

# Úvod

Přístup k informacím je dnes pro většinu lidí samozřejmostí. Internet, který umožňuje získat jakékoliv informace pomocí několika kliknutí, je všude přítomný. Tento technologický pokrok je takřka nezastavitelný. Každý z nás generuje obrovské množství dat, ze kterých se může kdokoli dozvědět velmi komplexní věci, má-li k těmto datům přístup a znalosti potřebné k práci s nimi. Adoptovaly tento přístup k informaci i menší a středně velké výrobní podniky? Jak tyto podniky zpracovávají svá data?

Cílem bakalářské práce je seznámení čtenáře s Big daty, jejich problémy a aktuálními trendy. Následuje popsání rozdílných přístupů k datové analýze a porovnání jejich metod. V praktické části práce byl proveden průzkum středně velkého výrobního podniku, který byl zaměřen zejména na výrobní úsek a IT. V neposlední řadě se práce věnuje vztahu mezi procesním přístupem a Průmyslem 4.0 z pohledu datové problematiky.

Výstupem je porovnání analytických metod, které mohou být využity při zpracování Big dat, identifikace neshod ve zkoumaném podniku, jejich soupis a zmínění možných následků. Na tyto neshody jsou pak v práci navržena možná opatření. Byl stanoven vztah procesního přístupu a Průmyslu 4.0 vzhledem k problematice Big dat. V závěru práce jsou zpracována obecná doporučení pro malé a středně velké podniky, jež jsou založena na rešerši a průzkumu vybraného podniku.

# 1 Úvod k velkým data setům

Dnešní svět je digitální. Lidé stále více využívají a spoléhají na internetové služby. Když něco nelze najít online, jako by to neexistovalo. Tento trend je stále více posilován rostoucí výpočetní silou počítačů, levnějším hardware a čím dál tím větší sítíovou propustností. Tato kombinace generuje obrovské množství dat. Správné zacházení s těmito daty dokáže velmi dobře monitorovat a následně i popsat jakékoliv interakce. Mezi nejvýznamnějšími společnostmi světa se dnes nenajde firma, která by svou hodnotu neodvozovala od online přístupu k obrovskému množství dat, definujících preference jejich zákazníků. Další velké světové firmy navyšují investice do sběru dat a jejich analýzy. Firmy považují tvorbu datové infrastruktury a získání kvalifikovaných lidí v oblasti datové analýzy jako jedno z největších rizik, kterým aktuálně čelí. S velkými daty přichází i velká zodpovědnost v oblasti zabezpečení dat. Pro firmy jde o zásadní riziko, kterým je poškození reputace. Největší problém velkých dat je především správná interpretace. Vzhledem k vlastnostem velkých dat v nich lze nalézt téměř cokoliv a je jen na odborném pracovníkovi, aby rozeznal pravdu [15].

## 1.1 Před Big Daty

Vždy tomu tak nebylo. První pohled směřuje do minulosti, před období „Big data“.

V roce 1970 představil E. F. Codd relační databázový model [16]. Tento model změnil pohled na dosavadní databáze. Tvrdí se, že by aplikace měla namísto procházení ukazatelů data vyhledávat. Coddův model byl založen na matematických odvětvích, nazývaných teorie množin a predikátová logika.

Nový model, nazývaný Entitně-relační, navrhl P. Chen v roce 1976 [17]. Tento model umožnil designérům soustředit se na aplikaci dat, namísto na struktury logické tabulky. V této době se také objevuje termín relačně databázový systém managementu nebo také RDBMS.

Po roce 1980 síťový a hierarchický model ukončují vývoj a relační model vítězí. Structured Query Language (SQL) se stává standardem pro práci s daty v relačních databázích [18].

Na začátku 90. let se začínají rozmáhat nástroje, umožňující vlastní vývoj a tvorbu, například MS Excel, MS Access a Open Database Connectivity (ODBC). Velký zlom pro databáze bylo zpřístupnění World Wide Web (WWW) veřejnosti, které nastalo v polovině 90. let minulého století.

V roce 1997 byl představen Extensible Markup Language (XML). Definoval sadu pravidel pro zakódování dokumentů, které byly čitelné jak pro lidi, tak pro stroje. Přestože byl XML vytvořen se zaměřením na dokumenty, rozšířil se i mezi další datové struktury, třeba na web.

Začátkem 21. století se vývoj nových technologií ustálil, avšak nezastavil stálý růst databázových aplikací. Nový trend zavádí do databázové struktury programovací logiku. Začínají se objevovat velké datové sklady a s nimi i otázky, týkající se uložení a zpracování dat. Vznikají nové techniky, například: Data mining, Data warehousing, Data marts.

Objevují se Not only SQL (NoSQL) databáze, které se snaží popisovat nerelační databázové řešení a distribuovaná datová úložiště. NoSQL bývají optimalizovány pro zápis a přidávání záznamů a většinou nenabízí další funkcionality kromě uložení záznamů.

Dnes se databáze potýkají s dalším významným vývojovým krokem. Velká změna přichází v architektuře databázového řešení, kde už není nutnost mít všechna data v jedné databázi nebo v jednom data centru. Rozšiřují se distribuované databáze a cloudová řešení [19].

## 1.2 Big data

Nyní je možno popsat myšlenku Big dat. Různé zdroje uvádí odlišné definice termínu Big data. První zmínka o velkých datech přišla od vědců z NASA. Roku 1997 vydali Cox a Ellsworth publikaci [13] s první definicí, ve které popisují rostoucí nároky na paměť. NASA identifikovala tento problém jako problém velkých dat [2], [4]. V roce 2001 META Group analytik Doug Laney definoval příležitosti a výzvy s růstem dat jako trojrozměrné (volume, velocity and variety) [1], [3], [5]. Roku 2013 byla definice velkých dat rozšířena firmou Gartner, která definovala velká data jako koncept [20]: „*Velkoobjemové, vysokorychlostní a/nebo různorodé informační prostředky, které*

*vyžadují nákladově efektivní inovativní formy zpracování informací pro lepší přehled, rozhodování a optimalizaci procesu“ [1].* Definice od firmy Gartner je i nadále rozšiřována dalšími dimenzemi, a v době psaní této práce se mluví o 5V (volume, velocity, variety, veracity, value)<sup>1</sup>, která jsou obecně uznávána [7]. Další definice pochází od IBM a popisuje velká data jako data, která jsou svojí velikostí nad rámec tradičních relačních databází, takže je nelze ukládat, spravovat a zpracovávat s nízkou latencí, tedy v rozumném čase [8]. Definici IBM podporuje i Microsoft SQL Server team [9]. Google přidává k definici IBM i cenovou efektivitu a zdůrazňuje komplexitu a složitost uložení velkých dat na on-premise úložištích. Upřednostňuje uložení velkých dat v cloudu [10].

### **Hlavní oblasti využití**

Nasazení Big dat ve výrobních podnicích se zabývá monitorováním, predikcí, systémem podpory v rozhodování a plánováním. Tyto čtyři kategorie zastřešují 71 % Big data aplikací ve výrobních podnicích [38].

Bylo identifikováno šest oblastí Big dat ve výrobě. Těmito oblastmi jsou systémová integrace, data, predikce, sdílení zdrojů a síťování, udržitelnost a nízkonákladový HW. Na těchto šest oblastí navazuje devět komponent Big data „ekosystému“, jimiž jsou příjem dat, uložení dat, výpočetní analytika, vizualizace, pracovní a datové postupy, správa dat, a v neposlední řadě výpočetní infrastruktura a model zabezpečení [38].

## **1.3 Problémy Big dat**

Podle Al-Sai [27] je bezpečnost Big dat základní stavební kámen, se kterým bezprostředně souvisí zajištění soukromí a ochrany dat. Další výzvou jsou dostatečně vzdělaní zaměstnanci a náležitá infrastruktura. Autoři také poukazují na rychlý rozvoj technologií, který se Big dat dotýká.

Oussous a spol. se zaměřují na zpracování před samotnou analýzou [28]. Poukazují na důležitost a náročnost předpřípravy dat. Kontrola důvěryhodnosti datových zdrojů a ověření kvality dat je podle nich důležitá součást Big dat. Popisují způsoby agregace nových a stávajících dat a poukazují na důležitost zvolení vhodné technologie, která se pojí s investicemi podniku.

---

<sup>1</sup> Objem, rychlost změn, různorodost, důvěryhodnost a hodnota jsou charakteristickými vlastnostmi Big dat.

Management datové kvality považuje Talha a spol. jako stěžejní prvek datové strategie. Podle autorů napomáhá management datové kvality se zajištěním bezpečnosti dat [29].

Borodo a spol. pojednává o bezpečnosti dat v návaznosti na nároky infrastruktury. Vybízí k využití cloudových platforem za účelem snížení počátečních investic. Podotýká různé distribuované přístupy a MapReduce model [30].

## 1.4 Trendy Big dat

Podle informací od firmy Gartner se za pomoci technik strojového učení (ML), optimalizace a zpracování přirozeného jazyka (NLP) bude nadále zlepšovat dostupnost, rychlost a přesnost umělé inteligence (AI). Zvýší se transparentnost AI modelů, což zaručí lepší propojení člověka a stroje, a nadále bude prohlubovat důvěru v nové technologie v oblasti AI [39], [40].

Změna nastala v reportování z datových zdrojů. Došlo k odstoupení od statických reportů a datových sestav. Objevují se automatizované analýzy, vytvářené přesně pro uživatele a jeho zaměření, které jsou aktuálně ve fázi startupu [39], [40].

Dalším trendem je adopce inteligentního rozhodování. Gartner předpovídá, že do roku 2023 bude více než 33 % korporátů využívat data a datové modely, vytvořené pro podporu v rozhodování [39].

Gartner také spoléhá s rozšířením tzv. X analýzy. X označuje analýzu, kde X je datová proměnná pro řadu různých strukturovaných a nestruturovaných obsahů, jako je textová analýza, video analýza, audio analýza a další. Pomocí X analýzy by se měly řešit ty nejobtížnější problémy, jako například globální oteplování, prevence pandemie nebo ochrana životního prostředí [39].

Přesun ke cloudovým řešením je jeden z nejpopulárnějších trendů. Do roku 2022 bude veřejný cloud stát za 90 % veškerých inovací, které jsou spojené s daty a jejich analýzou [39]. Rozšíření a využití Edge computingu, což je zpracování dat přímo v zařízení, které je pořizuje, je pro některé systémy (například autonomní dopravu) nutností [40].

Dochází k prohloubení dat a datové analýzy. Doposud byly tyto dvě kategorie chápány samostatně, avšak Gartner podotýká, že je nevyhnutelná prolnutí těchto kategorií a předpokládá jejich propojení. Předpokládá se vznik nových rolí v IT a změna vývoje nástrojů pro práci s daty, jejich správu a analýzu [39].

Rozvoj datových trhů a obchodu s daty nabývá na síle. Gartner považuje obchod s daty třetích stran jako důležitý článek datové analýzy a předpovídá monetizaci datových zdrojů a jejich růst [39].

Vzestup DataOps. DataOps je metodologie, která vychází z DevOps. DataOps se zabývá tokem dat napříč organizací a snaží se odstranit překážky, limitující využití dat v organizaci. Dalším úkolem DataOps je integrace dat třetích stran do již existující struktury organizace [40].



## 2 Analýza Big dat

Big data analýza popisuje proces analýzy rozsáhlých datových souborů za účelem nalezení vzorů, neznámých vazeb, trendů, uživatelských preferencí a dalších cenných informací, které by nebylo možné získat tradičními nástroji [12]. Analýza velkých dat souvisí s ukládáním dat do cloudových úložišť a snahou je co nejlépe využít. Uvažována je analýza textů, zvuků, obrazů a kvantitativních strukturovaných údajů. Význam procedur pro analýzu velkých dat vzrostl v poslední době s pokrokem informačních technologií a prostředků pro sběr a ukládání dat. Celosvětová roční míra růstu Big data technologií a služeb byla odhadována na zvýšení o 36 % od roku 2014 do roku 2019 [12]. Výsledkem je osamostatnění celé oblasti metod, které se říká datová analytika. Několik pokročilých analytických technik (strojové učení, dolování dat a umělá inteligence) a také potenciální strategie, jako je paralelizace, „rozděl a panuj“, přírůstkové učení, vzorkování a granulární výpočet, dokážou převést velké problémy na malé, a tím umožňují kvalitnější rozhodování, snížení nákladů a efektivnější zpracování [12]. Tato oblast představuje dnes důležitou část informačních technologií [3].

### 2.1 Hledání pravidel asociace

Jedná se o oblast vyhledávání klasifikačních pravidel, jež zahrnuje analýzu a vyhledávání vztahů mezi daty. Výsledkem je kategorizace dat s určitými společnými vlastnostmi, která se využívá při kategorizaci textů. Například význam webových stránek často závisí na počtu jejich čtenářů. V této oblasti se zjišťují počty uživatelů stránky a význam stránky [3].

### 2.2 Genetické algoritmy

Genetické algoritmy jsou překladem biologického konceptu evoluce do algoritmických receptů. Jsou součástí mnoha inteligentních systémů, často jsou součástí strojového učení a umělé inteligence, jejichž cílem je konstrukce metod, které napodobují či dokonce překonávají lidskou inteligenci [14]. Genetické algoritmy se používají například pro identifikaci nejčastěji sledovaných videí, televizních pořadů a jiných forem médií [3].

## 2.3 Umělá inteligence

Podle Simmonse a Chappella pojem umělá inteligence označuje chování stroje, které je považováno za inteligentní, pokud je shodné s chováním člověka [23].

Umělá inteligence (dále jen AI) je obecná oblast, jež zahrnuje, mimo strojového<sup>2</sup> a hlubokého<sup>3</sup> učení, i řadu dalších přístupů, ve kterých není učení zahrnuto. Odborníci se domnívali, že by umělé inteligence na úrovni člověka mohlo být dosaženo, pokud by programátoři měli soubor explicitních pravidel pro manipulaci s poznatky v dostatečném rozsahu. Tomuto přístupu se říká symbolická AI a nejpopulárnější byla v 80. letech během boomu expertních systémů.

Symbolická AI se projevila jako vhodné řešení pro dobře strukturované problémové úlohy. Zároveň se také ukázalo, že není možné zjistit explicitní pravidla pro řešení složitějších nejasných problémů, jako je například rozpoznání obrazu. Na tento přístup navazuje AI za pomoci strojového učení [22].

### Umělé neurální sítě

Neurální sítě jsou množina klasifikátorů, které jsou uspořádány do vrstev, v nichž je výstup z jedné vrstvy vstupem do vrstvy další. Mezi vstupní a výstupní vrstvou se nacházejí skryté vrstvy a s jejich počtem roste i hloubka, viz hluboké učení. Po roce 2006 byly navrženy další techniky umožňující rozvoj. Pro trénování je potřeba velké množství dat, výpočetní kapacity a času. Vytvořené uzly jsou mezi sebou propojené, vykonávají práci na vstupech a předávají je dalším uzlům [3]. Podle Al-Abasasi [26] se značně využívají při prediktivní analýze.

### Strojové učení

V této oblasti se jedná o kategorizaci a určení pravděpodobných výsledků na základě specifické množiny dat [3]. Ve strojovém učení je kombinována statistika a paradigmatu učení, jež jsou využívána umělou inteligencí. Výzkum strojového učení je v posledním desetiletí velmi úspěšný. Je možné jej rozdělit na dvě hlavní kategorie, kterými jsou učení pod dozorem a učení bez dozoru [3], [14]. U učení pod dozorem je predikována známá proměnná a cílem je nalezení funkce, která nejlépe predikuje cílovou proměnnou.

---

<sup>2</sup> Strojové učení je součástí umělé inteligence, která zkoumá možnosti nebo principy učení strojů

<sup>3</sup> Hluboké učení je specifickým rozšířením strojového učení.

Při učení bez dozoru není k dispozici cílová proměnná a cílem je porozumět přirozeně daným strukturám [3].

Na rozdíl od statistiky má strojové učení schopnost vypořádat se s rozsáhlými a složitými datovými množinami (např. miliony obrázků), pro které by klasická statistická analýza, např. Bayesovská, byla nepraktická [22].

### **Hluboké učení**

Hluboké učení je specifickou podskupinou strojového učení. Využívá hluboké neuronové sítě k mapování vstupů na cíle prostřednictvím hluboké posloupnosti jednoduchých transformací dat (vrstev). V tomto kontextu učení znamená zjišťování množiny hodnot pro váhy všech vrstev v síti tak, aby síť správně mapovala příklady vstupů na přidružené cíle. Základním prvkem hlubokého učení je ztrátová funkce, která vyjadřuje přesnost celé sítě. Podle tohoto zpětnovazebního signálu se upravují hodnoty v síti tak, aby ztrátová funkce vracela co nejmenší hodnotu. Celý trénovací cyklus se mnohokrát opakuje. Síť, pro kterou jsou výstupy tak blízko cíle, jak jen mohou být, je označována termínem natrénovaná síť [22].

## **2.4 Textová analytika**

Zpracování textových dat, nebo také textová analytika, je činnost, při níž se počítač snaží porozumět psanému textu a získat cenné informace. Zpracování textu se dělí na dvě základní oblasti, jimiž jsou příprava dat a extrakce informace. Tradičně využívanými nástroji jsou algoritmy z oblasti zpracování přirozeného jazyka, známé také jako počítačová lingvistika [24].

Zpracování přirozeného jazyka je netriviální problém. Většinu problémů je nutné řešit za pomoci sofistikovaných technik, například s využitím hlubokého učení. Zpracování přirozeného jazyka také komplikuje velké množství jazyků s různorodou syntaxí<sup>4</sup>, sémantikou<sup>5</sup> a pragmatikou<sup>6</sup> [24].

---

<sup>4</sup> Syntaxe (též skladba) je jazykovědná disciplína, která se zabývá tím, z jakých jednotek se skládají věty, jaké jsou vztahy mezi těmito jednotkami a jaké je jejich uspořádání ve větě.

<sup>5</sup> Sémantika je nauka o významu výrazu slovních spojení a vět.

<sup>6</sup> Pragmatika se zabývá sledováním jazykových prostředků, které jsou užívány za určitým účelem v rámci komunikace, též ji lze označit jako studium významu v kontextu komunikačních znalostí.

### 3 Typy datových analýz

Pochopení a správné využití vhodné analýzy přímo napomáhá získání informace a jejímu porozumění. Každá analýza se specializuje na konkrétní otázku nebo problém strukturovaných či nestrukturovaných dat. Často se analýzy kombinují za účelem hlubšího pochopení a získání přesnějších výsledků.

Deskriptivní a diagnostické analýzy jsou reaktivní přístup k řešení problému, zatímco prediktivní a preskriptivní analýzy jsou proaktivní přístup. Odlišné přístupy popisuje Tabulka 1 [31].

Tabulka 1 Klasifikace přístupů k datům podle Dai [31]

Otázky	Přístupy	Použití
<b>Deskriptivní analýza</b> Co se stalo?	<ul style="list-style-type: none"> <li>Asociační analýza</li> <li>Klastrování, sekvenční dolování vzorů</li> <li>Dotazování, statistické reporty</li> <li>Datová vizualizace</li> </ul>	Zachycení vzoru nevhodného chování Kontrola stavu produktu Srovnávací analýza
<b>Diagnostická analýza</b> Proč se to stalo?	<ul style="list-style-type: none"> <li>Diagnostické uvažování</li> <li>Bayesovská analýza</li> </ul>	Diagnostika poruch Analýza příčin selhání Detekce anomálií
<b>Prediktivní analýza</b> Co se může stát?	<ul style="list-style-type: none"> <li>Klasifikace, regrese</li> <li>Strojové učení</li> <li>Hluboké učení</li> </ul>	Predikce trajektorie Predikce chování zákazníka Predikce údržby zařízení
<b>Preskriptivní analýza</b> Co by se mělo udělat?	<ul style="list-style-type: none"> <li>Simulace</li> <li>Optimalizace</li> <li>Zpětnovazební učení</li> <li>Rozhodování, Saatyho metoda (AHP), Technika pro pořadí preferencí podle podobnosti s ideálním řešením (TOPSIS)</li> </ul>	Odolnost systému Spolehlivost systému Optimalizace systému

## 3.1 Deskriptivní a diagnostická analýza

Tyto dvě analýzy prezentují data ve srozumitelném formátu a zjišťují vztahy mezi příčinami a následky. Deskriptivní analýza shrnuje data z minulosti za účelem poskytnutí přehledu potenciálních vzorů a trendů. Diagnostická analýza poskytuje historický pohled na existující činnosti, ze kterých je možné identifikovat problémy a příležitosti [25].

Tyto algoritmy obsahují statistické ukazatele jako jsou: průměr, medián nebo rozptyl. Sofistikovanější metody, například: korelace, shlukování nebo generativní modely, dosahují přesnějších a konzistentních výsledků. Na druhou stranu jsou tyto metody komplikované a finančně náročné [26].

### Korelační metody

Korelační metody jsou statistickým tématem, které se většinou zaměřuje na atributy neustále se měnících Big dat. Příkladem metod může být Chi-kvadrát pro kategorická data a Pearsonův korelační koeficient pro numerická data [26].

### Shlukování

Zatímco se korelace zabývá změnou hodnot, metody shlukování se snaží nalézt podobné skupiny atributů. Sem spadají: dělicí algoritmy, algoritmy založené na hustotě, modelové algoritmy a algoritmy shlukování přes mřížku [26].

### Generativní modely

Tyto modely generují data podle předdefinovaných pravidel. Reálná data lze porovnat s generovanými za pomoci pravděpodobnostní funkce, která hledá podobnosti. Modely využívající pravděpodobnostní funkce jsou: Naivní Bayes, Skryté Markovovy modely a maticová faktorizace [26].

## 3.2 Prediktivní analýza

Prediktivní analýza se zabývá událostmi, které nastanou v budoucnosti. V prediktivní analýze jsou využívány statistické metody, které analyzují aktuální a historická data za účelem vytvoření předpovědi, týkající se budoucích událostí a/nebo budoucího chování. Je možné ji rozdělit do tří kategorií, kterými jsou statistická inference, strojové učení a metody pro analýzu nestrukturovaných dat, viz Tabulka 2 [25].

Tabulka 2 Hlavní metody prediktivní analýzy podle Sheng [25]

Kategorie	Popisek	Použití	
<b>Statistická inference</b>	Využití statistického přístupu k analýze velkého množství dat a multi-dimenzionální datové sady.	Alokování zdrojů; hodnocení rizik; předpověď odchodu zákazníka.	
<b>Strojové učení</b>	Regrese, neurální sítě, klasifikace, klastrování, asociace, hluboké učení	Pomocí algoritmů strojového učení lze vytvářet složité modely na velkých datových souborech a předpovídat tak budoucí výsledky.	Plánování kapacit; profilování rizik; segmentace zákazníků; předpověď poptávky; předpověď prodeje; detekce poruch.
<b>Metody pro analýzu nestructurovaných dat</b>	Dolování z textu, analýza obrazu, analýza sentimentu, web mining, analýza sociálních sítí, vizualizace	Kvantifikace nestructurovaných dat a snaha nalézt a vysvětlit vzorce v lidském chování.	Predikce chování uživatele; cílená reklama; monitorování produktu či služby.

### Metody regrese

Regresní metody mohou být využívány v deskriptivních i v prediktivních modelech. Mezi regresní algoritmy patří lineární regrese, logistická regrese a další [26].

### Rozhodovací stromy

Modely rozhodovacích stromů se skládají ze stromové struktury, kde jsou datové záznamy rozdělovány do větví. Tyto stromy mohou být obrovské. Za pomoci *prořezávání* se redukuje větve, které mají slabý příspěvek ke klasifikaci, čímž se zvyšuje přesnost a rychlost klasifikace [26].

### Bayesovská statistika

Jedná se o statistické metody založené na Bayesově větě. Tyto metody dokážou predikovat s vysokou přesností, je-li dodržena podmínka nezávislosti [26].

### 3.3 Preskriptivní analýza

Preskriptivní řešení určují akce a hodnotí jejich dopad na podnikové cíle, požadavky a limity, a tím podporují rozhodování podniku. Sheng dělí metody preskriptivní analýzy do čtyř kategorií: matematické programování, simulace, evoluční výpočty a modely založené na logice, viz Tabulka 3 [25].

Tabulka 3 Typy analýz podle Sheng [25]

Kategorie	Popis	Použití	
<b>Optimalizace</b>	Lineární programování a jeho rozšíření	Technika optimalizace lineární funkce. Rozšířením je celočíselné programování (proměnné jsou celá čísla) a binární programování (proměnné jsou binární).	Plánování kapacit; marketing; nastavení ceny.
	Nelineární programování	Technika optimalizace, kde jsou funkce či její omezení nelineární.	Cenová strategie; správa majetku.
	Stochastická optimalizace	Kolekce optimalizačních metod, kde výsledná funkce závisí na náhodných elementech.	Lidské zdroje; management; řízení zásob.
	Bayesovská optimalizace	Technika optimalizace funkce, která je komplexní a/nebo náročná na výpočet.	Inkasování dluhů; veřejné zdraví.
	Evoluční algoritmy	Algoritmy inspirované biologickou evolucí pro globální optimalizaci.	Alokace a plánování zdrojů; návrhy projektů.
<b>Simulace</b>	Technika modelování skutečných či hypotetických operací nebo systémů za účelem predikce chování a objasnění rozhodnutí.	Chytrá výroba; zdravotnictví; management; vývoj strategických map.	
<b>Logické modely</b>	Hypotetická reprezentace teorie změny a jejího následku.	Plánování údržby; robotika; chytrá města.	

## 4 Porovnání metod analýzy pomocí korelační matice QFD metody

Metody zmíněné v kapitole 2 a 3 byly porovnány korelační maticí metody QFD (Obrázek 1). Touto metodou se vytváří systém plánování a komunikace, s pomocí kterého mají být koordinovány všechny zdroje podniku, čímž má dojít ke zvýšení konkurenceschopnosti.

Autor práce vybral deset oblastí pro zhodnocení analytických metod viz Obrázek 1. První oblastí je záběr využití, který označuje možnost využití napříč podnikem. Druhá oblast je náročnost aplikace metody. Tato oblast popisuje, jak složité je danou metodu správně využít. Třetí oblast se věnuje výpočetní náročnosti nebo také hardwarovým požadavkům. Náročnost na hardware lze rozdělit na několik kategorií, například zda je potřeba specializovaný hardware, jako je dedikovaný výpočetní server, nebo zda stačí kancelářský notebook. Čtvrtá oblast se zabývá finanční náročností dané metody. Jak finančně náročné bude nasazení metody? Bude potřeba koupit licence? Jedná se o volně dostupný software? Pátá oblast je věnována přípravě dat, nad kterými bude metoda nasazena. Musí se data speciálně transformovat, než budou použita v určité metodě? Jaká bude dostupnost dat? Oblast šestá se zabývá časovou stopou metody. Popisuje, jak dlouho trvá obdržení prvních výsledků. Sedmá oblast se věnuje přesnosti výsledku a popisuje, jak moc je možné se na výsledek spolehnout. Osmá oblast mluví o datové náročnosti. Jak moc metodu ovlivní množství dat, které bude metoda využívat? Zpřesní se výsledek, pokud bude využito více dat? Devátá oblast se zaměřuje na otázku personálu. Bude potřeba specialista nebo metodu zvládne použít i řadový zaměstnanec se zájmem o datovou analytiku, či zaměstnanec po proškolení? Poslední oblastí je infrastruktura společnosti. Jak náročné bude pro podnik a jeho datové toky využití dané metody?

Těchto deset oblastí bylo vybráno za účelem zodpovědět nejčastější otázky vedoucích pracovníků ohledně implementace analytických metod. Metoda QFD byla vybrána z podstaty metody. QFD převádí požadavky zákazníka na vlastnosti produktu. Pomocí této matice se snaží odpovídat na neurčité otázky managementu.



č.	Problém	Priorita	Ideal	Hledání pravidel asociace	Genetické algoritmy	Umlá inteligence	Textová analytika	Korelace	Shlukování	Generativní modely	Regrese	Rozhodovací stromy	Bayesovská statistika	Matematické programování	Simulace	Logické modely	Statistická inference	Strojové učení	Analýza nestrukturovaných dat	Datová vizualizace	Umlé neurální sítě	Hluboké učení
1	Záběr využití	1	9	1	1	9	1	9	9	9	9	6	3	1	3	1	9	9	3	9	9	9
2	Nízká náročnost aplikace	1	9	3	3	1	1	9	9	1	9	9	3	3	3	3	3	3	1	9	3	3
3	Nízká výpočetní náročnost	1	9	3	3	3	1	9	9	3	9	3	3	3	3	9	9	3	3	9	1	3
4	Nízká cena realizace	1	9	3	3	1	1	9	9	3	9	3	9	9	3	3	9	3	1	9	1	3
5	Potřeba upravit data před využitím	1	9	1	3	1	3	3	3	9	9	3	3	3	1	3	3	1	9	3	1	1
6	Nízká časová náročnost	1	9	3	3	3	3	9	9	3	9	3	9	9	1	3	3	3	1	9	1	3
7	Vysoká přesnost výsledku	1	9	3	3	9	9	9	9	3	3	9	9	9	9	3	3	3	3	9	9	3
8	Vysoká datová náročnost	1	9	3	9	9	3	3	3	1	3	1	9	1	3	3	9	9	9	1	9	9
9	Odborný personál	1	9	3	3	1	3	3	3	1	9	9	3	3	1	3	1	1	1	9	3	1
10	Infrastruktura	1	9	1	3	1	3	9	9	1	9	9	3	9	3	3	3	1	3	3	1	1
	<b>Celkový výsledek</b>	<1; 9>	90.00	24.00	34.00	38.00	28.00	72.00	72.00	34.00	78.00	58.00	54.00	50.00	30.00	34.00	52.00	36.00	34.00	70.00	38.00	36.00
		100%	26.67	37.78	42.22	31.11	80.00	80.00	80.00	37.78	86.67	64.44	60.00	55.56	33.33	37.78	57.78	40.00	37.78	77.78	42.22	40.00

Obrázek 1 Korelační matice QFD metody, porovnání metod analýzy pro potřeby podniku [Vlastní zpracování podle [49]]

Možnost využití metod zpracování dat z pohledu malého až středně velkého výrobního podniku autor práce vidí zejména v datové vizualizaci a metodách, které nevyžadují tým specialistů, nebo nejsou finančně náročné. Takovými metodami jsou, mimo již zmíněné vizualizace, například: regrese, korelace, shlukování, rozhodovací stromy a Bayesovská statistika. Maximálního potenciálu těchto metod dokáže využít i zaměstnanec se zájmem o problematiku a odpovídajícím školením. Pomocí při využití vybraných metod jsou dnes i předpřipravené funkce a metody v různých programovacích jazycích a frameworkcích. Tyto nástroje zpřístupňují metody analýzy a rozšiřují jejich aplikaci. Za další výhodu by se dalo považovat jejich využití ve složitějších modelech a komplexnějších přístupech, jakými jsou: strojové učení, umělá inteligence a hluboké učení.

Pro výrobní podniky si autor myslí, že analýza nestrukturovaných dat a textová analytika nemají vysoký potenciál pro využití přímo ve výrobě, a to zejména z důvodu formátu dat. Autor takto usuzuje na základě pozorování v podniku. Jejich data z výroby jsou strukturovaná a v naprosté většině případů číselného typu.

U generativních modelů a simulací vidí autor využití zejména v oddělení vývoje, v oblasti údržby a z pohledu kvality. Tyto metody již vyžadují pokročilé znalosti a také náležité výpočetní prostředky. Pokud je ovšem vyřešena otázka infrastruktury a potřebných datových toků, je možné experimentovat a vyvíjet pomocí digitálního dvojčete. Digitální dvojče eliminuje nutnost testovat nové nápady přímo ve výrobě. Předchází tak rizikům, které s sebou nese vývoj a testování v ostrém provozu.

Podle autora textová analytika do výroby nezapadá, protože textová data se vyskytují ve výrobě velmi zřídka a jsou-li přítomna, tak nejsou předmětem analýzy. Může se jednat například o poznámky údržby k opravám.

Hledání pravidel asociace je nelehký problém zejména pro infrastrukturu a datovou politiku podniku. Z podstaty hledání asociace je potřeba přístup k různým datům z různých datových zdrojů. Pokud na to není podnik připraven, nelze využít algoritmů a metod pro hledání asociace v plném rozsahu. Autor práce si myslí, že jedním z hlavních problémů pro menší a střední výrobní podniky je stav IT a datové strategie podniků. Tento problém je spojený s potřebnou změnou myšlení podniku a případným upravením strategie firmy, aby mohla být upravena nebo vybudována datová

infrastruktura, na které bude možné realizovat hledání pravidel asociace či jiné datově náročné metody analýzy.

Většina metod vyžaduje přístup k datům napříč organizací, čehož není možné dosáhnout bez náležité infrastruktury a architektury příslušných systémů. S tímto netriviálním problémem přichází i další požadavky na předpřípravu dat, které jsou odlišné pro různé metody. Podnik může mít nastavenou datovou politiku a mít normalizovaná data, ale to nemusí být dostačující pro využití například strojového učení. Předpříprava dat je nutností pro korektní použití metody a dosažení správných výsledků.

Jako vhodný začátek by autor pro výrobní podniky volil datovou vizualizaci, která umožní lepší přehled informací ve výrobě. Jedná se o metodu, která přináší podniku informace, které již většinou má, ale efektivně nevyužívá. Autor se setkal se situacemi, kdy podnik sbíral a ukládal data, ale dále s nimi nepracoval. Předání dat správným zaměstnancům v reálném čase může zvýšit produktivitu.

U menších a středních podniků autor doporučuje zvážit využití infrastruktury třetí strany, jako je například Microsoft. Tento přístup snižuje počáteční investice a zbavuje podnik dalších povinností, jako jsou například nutná bezpečnostní opatření při správě vlastního hardware. Microsoft navíc nabízí širokou škálu služeb a uživatelsky přívětivých analytických řešení v prostředí Azure.

## 5 Průzkum aktuálního stavu podniku

V této kapitole se autor zaměří na zkoumání současného stavu zpracování Big dat ve výrobním podniku. Popíše provázanost Big dat, procesního přístupu a Průmyslu 4.0 a navrhne doporučení pro práci s Big daty, zejména ve výrobních podnicích.

### 5.1 O společnosti

Družstevní závody Dražice – strojírna je společnost s ručením omezeným (DZD) se sídlem v Benátkách nad Jizerou. Je to dceřiná společnost skupiny NIBE. Vizí společnosti je vyrábět nejlepší ohřívač vody a její misí je dodávat teplou vodu. Předmětem podnikání společnosti jsou následující činnosti:

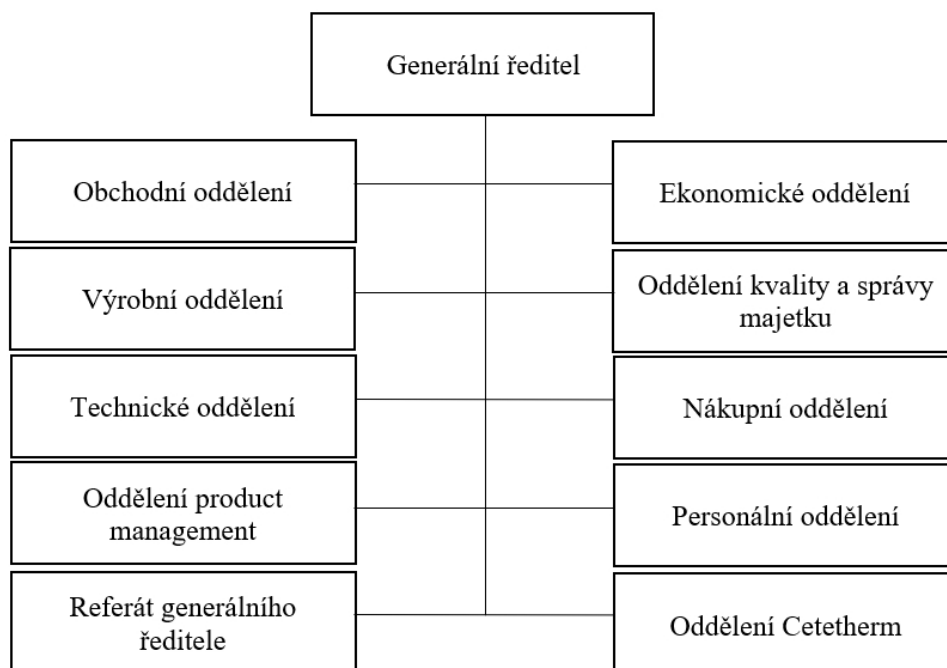
- zámečnictví, nástrojářství,
- výroba, instalace, opravy elektrických strojů a přístrojů, elektronických a telekomunikačních zařízení,
- výroba, obchod a služby neuvedené v přílohách 1 až 3 Živnostenského zákona,
- montáž, opravy a rekonstrukce chladících zařízení a tepelných čerpadel.

Společnost vyrábí sériově tyto druhy produktů:

- ohřívače a zásobníky teplé vody,
- akumulární nádrže,
- klimatizační jednotky,
- fotovoltaická řešení (ohřev vody),
- příslušenství.

V majetkové struktuře společnosti ke konci roku 2019 představují nejvýraznější část aktiv hmotné movité věci a jejich soubory a stavby. Na straně pasiv vlastní kapitál výrazně převyšuje cizí zdroje. [21]

Obrázek 2 popisuje funkční hierarchickou organizační strukturu společnosti. Oddělení jsou si mezi sebou rovna a nad všemi stojí generální ředitel. Praktická část práce byla vypracovávána za pomoci technického a IT oddělení.



Obrázek 2 Hierarchická organizační struktura společnosti DZD [21]

### **Certifikace**

Podnik disponuje řadou certifikací. Certifikát ČSN EN ISO 9001:2016 firma drží již od roku 2004 a 3. 7. 2020 jeho držení prodloužila. ČSN EN ISO 14001:2016 a ČSN ISO 45001:2018 firma obdržela ve stejný den jako certifikát řady 9001. Tyto certifikáty byly podniku uděleny organizací IQNet. Podnik také disponuje zlatým certifikátem pro integrovaný systém managementu v souladu s požadavky QMS, EMS a SM BOZP.

## **5.2 Předpokládaný vývoj**

Firma se momentálně potýká s následky koronavirové krize, takže je velmi obtížné predikovat budoucí vývoj. Plán pro rok 2020, který byl schválen před vypuknutím pandemie, počítá s dalším růstem tržeb i přes očekávané ukončení některých dotačních programů, jako například kotlíkové dotace.

Společnost DZD se i nadále bude zaměřovat na efektivní řízení všech procesů a odhalování rezerv, a to jak v oblasti plánování, tak v oblasti nákupu, výroby, expedice atd. Cílem společnosti je soustředit se na posílení kontrolních procesů ve všech oblastech své činnosti.

## **Výzkum a vývoj**

V roce 2019 se firmě podařilo prosadit nové vnitřní systémové jednotky pro tepelná čerpadla. Tím také DZD vstoupilo výrobně do segmentu tepelných čerpadel. Stěžejním vývojovým projektem pro oddělení vývoje za rok 2019 byl projekt nové řady ohřívačů v objemech 5 a 10 litrů.

Firma stále podporuje intenzivní vývoj v oblasti systému pro topení, ohřev a skladování teplé vody, například vývojem aplikací pro tepelná čerpadla, vývojem elektronického termostatu 4. generace a dalšími aplikacemi pro chytrou domácnost a IoT.

## **5.3 IT ve společnosti**

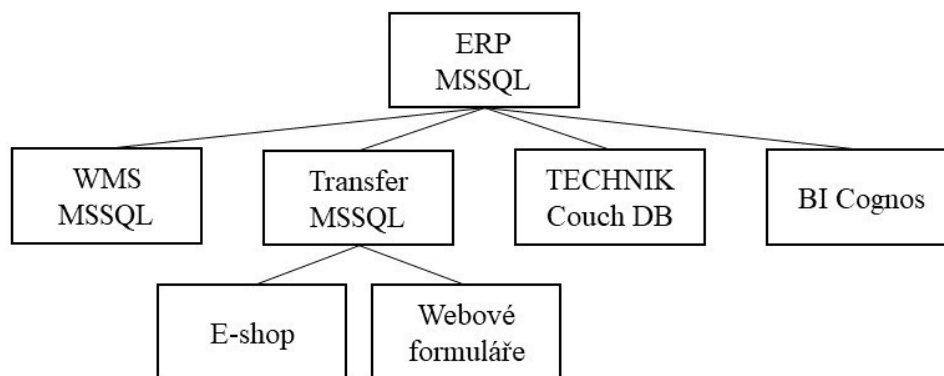
IT záležitosti spadají pod finanční oddělení, protože má podnik pouze dva zaměstnance, dedikované pro IT. Hlavní náplní práce pro tyto dva pracovníky je:

- podpora uživatelů (SW i HW),
- správa SW i HW,
- správa active directory,
- rozvoj IT podniku.

Do chodu a pravidel IT také přímo zasahuje skupina NIBE, kterou je DZD součástí. Jedním z příkladů IT řízení ze švédské strany je například omezení komunikačního software pouze na MS Teams.

### **5.3.1 Systémy IT**

Pod svojí správou mají 5 databázových serverů ve dvou serverovnách. Všechny 5 serverů je pravidelně zálohováno. Celkový obsah těchto serverů se momentálně pohybuje kolem 500 GB. Každý server představuje samostatný systém a není provázán s ostatními. Obrázek 3 mapuje vybrané systémy v podniku.



Obrázek 3 Schéma IT systémů v podniku [Interní zdroj podniku]

Ve zkoumaném podniku je nasazen ERP systém, který je poskytován firmou ORCZ a firmou Ortex. Licence ERP systému je placená ročně na použité moduly nebo na přihlášené uživatele. Každých 5 minut je vytvořen snapshot ERP systému, který je považován za hlavní systém společnosti. Podle správce je systém neintuitivní, se zastaralým vizuálním rozhraním a pro efektivní práci s ním je nutné zaškolení. Problémem by také mohl být přístup firmy poskytující ERP. Podle konzultanta má firma nedostatečný počet kvalitních vývojářů, což se odráží v nedostatečných aktualizacích součástí ERP. Na tento systém je napojen BI systém Cognos od firmy IBM. Cognos slouží jako nástroj pro management. V systému Cognos mají uživatelé přístup k vybraným datům z ERP a k datům docházky zaměstnanců. Každý den ERP systém vytváří datovou kostku pro jednotlivá oddělení podle předdefinovaných pravidel.

Systém TECHNIK slouží servisním technikům, kteří pomocí mobilní aplikace (popřípadě na tabletu) schvalují instalace tepelných čerpadel. Tato aplikace kompletně nahradila papírovou verzi schvalování. TECHNIK je jako jediný postaven na Couch DB a funguje jako pomocná databáze ERP systému. Jedná se o mezičlánek ERP a mobilní aplikace. Do budoucna je v plánu rozšíření na další servisní úkony, například uznání záruky nebo záznam servisní činnosti.

Transfer zajišťuje komunikaci mezi ERP a aplikačními systémy. Jedná se o e-shop firmy a o webové formuláře. Je postaven na MSSQL Express.

Warehouse Management System (WMS) zařizuje řízení skladu. V roce 2020 byl otevřen nový skladový areál, který je kompletně digitalizovaný. S otevřením této haly začala firma paletizovat. Každá paleta je opatřena jedinečným identifikátorem a sledována v ERP.

Paralelně se záznamy ve WMS vzniká plán nakládek v Excelu, který si zaměstnanci tvoří sami. Tento soubor je uložený kompletně mimo WMS systém a je dostupný pouze zaměstnancům, kteří se podílejí na tvorbě jeho obsahu. Tento postup je problematický, protože jsou informace nedostupné pro další zaměstnance a management. Tato praxe je ve firmě známa a aktivně se řeší. Vedení zatím tento postup schvaluje, jelikož se potýká s nedostatečnou kvalifikací zaměstnanců, kteří nezvládnou pracovat se sofistikovaným systémem.

Podnik využívá SW nástrojů od Microsoft, konkrétně balíčku Office 365. Zejména se jedná o MS Word, MS Excel, MS Access, MS Teams, Outlook a další.

Začátkem roku 2021 je plánované spuštění intranetu, v němž všichni zaměstnanci naleznou postupy a informace ohledně dění v podniku. Aktuální průběh tohoto plánu není autorovi práce znám.

Ve výrobě je mnoho zařízení s vlastním SW. Společně s postupným pořizováním strojů narůstá i různorodost používaných SW.

### **5.3.2 Problémy z IT strany**

Největším problémem z IT strany je počet zaměstnanců, kteří jsou vedeni jako IT; jedná se pouze o dva zaměstnance. Od počtu zaměstnanců se přímo odvíjí kvalita a možnosti vykonané práce. Není zde konkrétní specialista zodpovědný za konkrétní systém. To znamená, že oba zaměstnanci musí mít obecné znalosti všech systémů, a v praxi potom fungují stylem „všichni pracují na všem“.

Obrázek 3 není vyčerpávající soupis SW a systémů ve firmě. Z pozorování a rozhovorů s vedoucím zaměstnancem za IT se ve firmě najde spousta případů takzvaného „Shadow IT“, což je termín označující IT řešení mimo správu IT oddělení ve firmě. Tato řešení vznikají jako reakce na aktuální problém a obcházejí IT oddělení z důvodu rychlosti spuštění do produkce. Jako příklad lze uvést řešení od oddělení kvality, které si vytvořilo vlastní aplikaci v Accessu, která je zmiňována dále v textu.

## **5.4 Popis funkčnosti firmy**

Podnik vyrábí na zakázku. Podnět k výrobě dává zákazník svojí objednávkou. Podnik zároveň operuje s naskladněnými kusy některých modelů, které nebyly ještě objednány,



ale jsou uloženy na skladu jako rezervní. Podnik například udržuje stálou rezervu 20 kusů svého nejprodávanějšího ohřívače, a je tak schopen expedovat drobné objednávky a doplňovat objednávky, které by se jinak zpozdily z různých důvodů.

ERP MRP systém tvoří plán výroby, prodeje a objednávky materiálu. Výroba dostává materiál podle plánu výroby. Logistika doplňuje materiál systémem „Just in Time“. Funguje zpětná kontrola výdeje materiálu.

Samotnou výrobu lze rozdělit na dvě části podle toho, zda je zaznamenána v ERP a má tedy digitální podobu, nebo zaznamenána není. Část výroby, která není v ERP, se sestává z kompletace polotovaru. Tato část výroby není v ERP zaznamenávána, protože při dokončení polotovaru prochází pecí, která znemožňuje jakýkoliv pokus o označení. Podnik se v minulosti několikrát pokoušel najít způsob, jak označit polotovary, ale bezúspěšně. Bylo konstatováno, že další vývoj či investice do tohoto problému by podniku nepřinesly žádné úspory nebo zisky. Výrobní proces pro ERP systém začíná proto až s kompletovaným polotovarem.

Výrobek je tedy poprvé v ERP systému zaznamenáván jako výrobek až po označení, a není možné dohledat z jakého přesně materiálu (polotovaru) byl kompletován. Část výrobního procesu, která je zaznamenána v ERP systému, se skládá ze čtyř částí. Jedná se o aktivaci, el. test, kontrolu, balení, v tomto pořadí. Podnik má k dispozici dvě výrobní linky.

Pokud se při výrobě vyskytne neshoda, výrobek je odstaven a kontrolován zodpovědnou osobou, která rozhodne o dalším postupu. Tyto neshody jsou zaneseny do systému v Access a jsou kompletně mimo ERP systém. Pokud jsou odstavené výrobky vhodné k předělání, jsou přesunuty na začátek výroby a prochází výrobním procesem znovu a se stejným identifikátorem. ERP systém tedy nerozlišuje opravované výrobky. Pokud je výrobek nevhodný k opravě, je likvidován.

Kvalita výroby je uchovávána v Access aplikaci. Na rozdíl od ERP je v Access aplikaci kompletní proces výroby, tedy od získání materiálu a jeho vstupních testů až po balení hotového výrobku. Access aplikace byla vytvořena na základě nedostatků ERP modulu pro oddělení kvality. Oddělení kvality si tedy vytvořilo vlastní řešení, které odpovídalo jejich požadavkům a podléhá jejich správě. Každý den zaměstnanec zodpovědný za chod aplikace dostává export z ERP systému v podobě Excel tabulky. Jedná se o ruční

práci, kdy se tato tabulka importuje do Access aplikace. Aplikace má také funkcionalitu předpřipravených reportů, které denně rozesílá na email předurčeným zaměstnancům. Připraveno zde je 7 až 10 reportů v závislosti na neshodách.

V podniku se také vzhledem ke kvalitě nachází další systém, ve kterém jsou zaznamenávány systémové neshody, bezpečnost a ochrana zdraví při práci, požáry, environmentální neshody a údržba strojů. Jedná se o PALSTAT.

### **Data ve firmě**

Podnik v tuto chvíli nevlastní žádná Big data. Využívá některé technologie z oblasti Průmyslu 4.0, nicméně aktuální stav IT oddělení neumožňuje růst a adekvátní podporu podniku po IT stránce.

Podnik částečně využívá vizualizace vybraných dat v denních, týdenních a měsíčních reportech. Jedná se o jednoduché součtové reporty, které zobrazují například počet vadných kusů nebo počet aktuálně naskladněného materiálu.

Pro plnohodnotné využití náročnějších praktik, jako je umělá inteligence, digitální dvojče nebo generativní modely, potřebuje podnik nejprve upravit stávající infrastrukturu, datové toky a datovou politiku podniku. Dále tyto metody vyžadují odborně vzdělané zaměstnance, kterými podnik nedisponuje. Nutno podotknout, že podnik nemá jasnou datovou strategii. Podnik nevěnuje této části IT dostatečnou pozornost, čemuž odpovídá aktuální stav infrastruktury. Chybí centrální správa SW a HW ve firmě, kterou by mělo zajišťovat IT oddělení.

## **5.5 Výrobní proces a jeho neshody**

Následující podkapitoly se věnují popisu jednotlivých vybraných neshod, jejich možným i reálným následkům, a autorem navrhovaným řešením. Vzhledem k růstu podniku jsou níže vyjmenované neshody autorem považovány za závažné, protože již dnes nestačí centrální podpora IT a podnik sužuje Shadow IT. Na tyto neshody je pohlíženo z pohledu datové politiky na výrobním procesu.

Neshody jsou vypsány v Tabulce 4. Byla zvolena upravená tabulka FMECA analýzy. Sloupec rozpad v tabulce popisuje další neshody, které vyvstávají z hlavních neshod vypsáných v prvním sloupci. Tabulka má také autorem upravené hodnocení sloupců význam, výskyt a odhalení, které je popsáno v Obrázku 4. Výpočet sloupce RC (úroveň

rizika) zůstal stejný podle výpočtu FMECA. Nejdůležitější případy jsou rozepsány v kapitolách následujících pod Tabulkou 4.

<b>Četnost</b>	<b>Popis</b>
Všudypřítomné	Neshoda se vyskytuje celopodnikově
Časté	Neshoda se vyskytuje v částech podniku
Občasné	Neshoda se vyskytuje u specifických řešení

<b>Význam</b>	<b>Popis</b>
1	Neovlivní rozvoj podniku
2	Velmi malý vliv na rozvoj podniku
3	Umožňuje rozvoj, ale zvyšuje náklady
4	Omezuje možnost rozvoje a zvyšuje náklady
5	Nutné řešení, bez kterého nelze rozvíjet podnik

<b>Výskyt</b>	<b>Pravděpodobnost výskytu</b>
1	Nepravděpodobná
2	Zřídka kdy
3	Přichází v úvahu
4	Častá
5	Jistá

<b>Odhalení</b>	<b>Pravděpodobnost odhalení</b>
1	Jistá
2	Střední
3	Malá
4	Velmi malá
5	Téměř žádná

Obrázek 4 Legenda FMECA analýzy [Vlastní zpracování podle [47]]

Tabulka 4 FMECA analýza neshod v podniku [Vlastní zpracování na základě [46]]

Neshoda	Roypad	Četnost	Následky	Příčina	Kontrola	Význam	Výskyt	Odhalení	
<b>Nedostatek odborného personálu</b>	Počet zaměstnanců IT oddělení	Všudy-přítomné	Obtížný rozvoj podniku, přetížení stávajících pracovníků	Nemožný nábor, nehlásí se korektní lidé	Personální požadavky oddělení	5	5	1	25
	Nutné „ruční“ zásahy pro zpracování chybových hlášení komunikace mezi systémy TRISS a ERP	Občasné	Lidská práce nad automatizovanými procesy	Nedostatečně definovaná pravidla v systému komunikace	Pravidelné hlášení z logů systému	1	2	1	2
<b>Absence datové politiky podniku</b>	Pravidla pro práci s daty	Všudy-přítomné	Nejednotná řešení	Nedefinování vyšším managementem	Rozhovory v podniku	3	4	2	24
	Technologie/protokoly uznané podnikem	Všudy-přítomné	Nejednotná řešení	Nedefinování vyšším managementem	Rozhovory v podniku	3	4	1	12
	Absence datového katalogu	Všudy-přítomné	Absence přehledu držení dat v podniku	Nedefinování vyšším managementem	Rozhovory v podniku	5	5	1	25
	Neefektivní práce s datovými zdroji	Časté	Prodloužení dodání výsledku	Infrastruktura podniku a podpora IT oddělením	Rozhovory v podniku	3	3	4	36
	Nekonzistence dat	Časté	Nekonzistence a duplikace dat, předání neaktuálních informací	Nedostačující nástroje pro komunikaci. Nevzdělání zaměstnanci.	Rozhovory v podniku	4	3	2	24

<b>Neshoda</b>	<b>Rozpad</b>	<b>Četnost</b>	<b>Následky</b>	<b>Příčina</b>	<b>Kontrola</b>	<b>Význam</b>	<b>Výskyt</b>	<b>Odhalení</b>	<b>RC</b>
	Nedefinovaná datová bezpečnost podniku	Všudy- přítomné	Šedá místa ohledně přístupu k datům	Neangažovanost vedení podniku	Rozhovory v podniku	5	5	1	25
	Neprospojené databáze	Občasné	Náročnější zpráva databází	Při tvorbě řešení nebyl důvod k propojení	Při tvorbě řešení	2	5	1	10
<b>Zastaralý ERP systém</b>	Rozvoj nové funkcionality	Všudy- přítomné	Obtížný posun vpřed a stagnace inovace	Nedostačující vývoj poskytovatelem	Rešeře a zkušenosti administrátora	5	5	1	25
	Nedostačující stávající moduly	Občasné	Vznik samostatných řešení pomocí jiných nástrojů	Nedostačující vývoj poskytovatelem	Rešeře a zkušenosti administrátora	5	3	2	30
	Konektivita nových systémů	Občasné	Náročné či nemožné připojení nově vznikajících řešení	Nedostačující vývoj poskytovatelem	Rešeře a zkušenosti administrátora	4	3	1	12
	Neobsahuje kompletní data ohledně výrobního procesu	Všudy- přítomné	Ztráta informace vedoucí k nemožnosti kvantifikování a následné optimalizaci	Velmi obtížné označení výrobu pro uchování digitální stopy	Znamý fakt	5	5	1	25
	Tvorba vlastního řešení pro jednotlivá oddělení	Časté	Ztráta času a finančních prostředků podniku	Nedostatečná infrastruktura a nepřehlednost možností podniku	Rozhovory v podniku	5	4	2	40
<b>Nedostatečná automatizace datových toků</b>		Časté	Ztráta času a finančních prostředků podniku	Nedostatek odborných pracovníků a špatná komunikace v podniku	Při škálování aktuálních řešení nebo rozšíření funkcionality	4	3	2	24

<b>Neshoda</b>	<b>Rozpad</b>	<b>Četnost</b>	<b>Následky</b>	<b>Příčina</b>	<b>Kontrola</b>	<b>Význam</b>	<b>Výskyt</b>	<b>Odhalení</b>	<b>RC</b>
<b>Nedostačující komunikace napříč podnikem</b>		Všudy- přítomné	Neinformovanost zaměstnanců	Neangažovanost vedení podniku	Podání návrhu řešení, které již existuje	3	4	2	24
	Nedostačující komunikace napříč odděleními ohledně realizovaných řešení a aktuálních problémů	Všudy- přítomné	Duplikace řešení, prodloužení času dosažení výsledku, finanční ztráty	Funkční řízení	Rozhovory v podniku	4	3	2	24
<b>Chybí ohlašování úkonů operátorem</b>		Občasné	Chybná data v náležitých systémech	Lidský faktor	Pravidelné kontroly	1	2	1	2

### **5.5.1 Nedostatek odborného personálu**

Nedostatek odborného personálu se vztahuje především na IT oddělení. Vyplývají z něj další neshody, které budou zmíněny. Dva zaměstnanci, kteří aktivně řeší veškeré problémy spojené s IT, již pro podnik nejsou dostačující. Vedení podniku se již několik měsíců snaží o nábor třetího zaměstnance. Hledají osobu zběhlou v IT, která by byla schopná řešit všechny záležitosti týkající se tohoto oddělení.

Tento krok je však pouze řešením dosavadní vyčerpání IT oddělení. Nebude tím odstraněn kořen problému, a pokud bude podnik i nadále prosperovat, tento přístup k IT oddělení a jeho vedení nebude dostačující pro držení kroku s jeho vývojem a potřebami.

#### **Následky této neshody**

Náplň práce IT oddělení již byla popsána. Pokud je přítomen pouze jeden ze zaměstnanců, musí si pečlivě plánovat svou denní pracovní aktivitu a často reagovat i na nepředvídatelné situace. Například, pokud je zaměstnanec přítomen na pravidelné poradě a přijde mu do emailu chybové hlášení týkající se čteček čárových kódů z výroby, musí neprodleně opustit poradu a začít řešit tento problém. Nejhorším možným následkem problémů se čtečkou může být pozastavení výroby nebo také chybějící záznamy v ERP či WMS z důvodu nefunkčnosti čtečky, což by mohlo způsobit další chyby následných procesů.

Kvůli nedostatečné podpoře IT infrastruktury a pravidel navíc nemůže podnik využívat vlastněná data efektivně, a také ztrácí možnost získání informací z těchto dat.

#### **Navrhované řešení**

Podnik se potýká s obtížným náborem nových zaměstnanců. Hlavním řešením tohoto problému by měla být informovanost a vzdělávání stávajících zaměstnanců. Dalším možným řešením by mohlo být přesunutí adekvátních zaměstnanců, kteří spravují řešení na úrovni oddělení, do IT oddělení.

### **5.5.2 Absence datové politiky podniku**

Tento problém vychází přímo od vedení podniku, a měla by jím být řešena za úzké spolupráce s odborníky ze všech oddělení. Aktuálně neexistují jasná pravidla pro práci s daty a datovými zdroji. Nejsou předdefinovány ověřené postupy pro práci s daty, jako jsou protokoly, HW pro práci s daty, a uživatelská rozhraní pro konzumování dat.

### **Následky této neshody**

Strategie podniku pro IT záměry je podporována mnoha způsoby. Datová politika je jedním z hlavních. Bez náležité datové infrastruktury a dat nelze řešit problémy spojené s chytrými systémy. Mezi aktuální problémy podniku patří: neefektivní práce s datovými zdroji; sběr dat pouze pro uložení nebo archivaci; nedostupnost dat mimo oddělení, která je pořizuje; různorodé datové formáty; různorodá bezpečnostní kritéria; nekonzistence dat a duplicita řešení.

### **Navrhované řešení**

Vyjasnění cílů podniku v rámci Průmyslu 4.0, které bude odrážet IT strategie podniku. Je potřeba umožnit IT zaměstnancům věnovat čas této problematice a vývoji. Důležité je také přijetí těchto nově vzniklých pravidel zaměstnanci. Nestačí je pouze vydefinovat a očekávat, že zaměstnanci změní své chování či práci. Takové vynucení může poukázat na nesprávně přijatá pravidla a jejich úpravu. V datové politice podniku by se měla stanovit pravidla datové bezpečnosti, vlastnictví a správy dat, a také dostupnost k uloženým datům.

### **5.5.3 Zastaralý ERP systém a Shadow IT**

ERP systém, který by měl být „datovým centrem“ podniku, je podle slov jeho správce zastaralý. IT oddělení v podniku nemělo možnost zásadně změnit architekturu. Vedoucí IT převzal existující řešení, které se snaží udržovat v chodu a rozšiřovat ho na základě požadavků jiných oddělení. Tento přístup je nevyhovující a může stát za rozšířením Shadow IT, se kterým se momentálně podnik potýká. Výše zmíněné nejasnosti s infrastrukturou a daty v podniku umožnily uživatelům vytvářet vlastní řešení mimo správu a podporu IT. V práci je popsán případ mezi ERP a oddělením kvality, ve kterém moduly ERP nebyly pro oddělení kvality dostačující, a tak si jeho zaměstnanci vytvořili vlastní řešení. Vedení podniku se momentálně touto situací nezabývá.

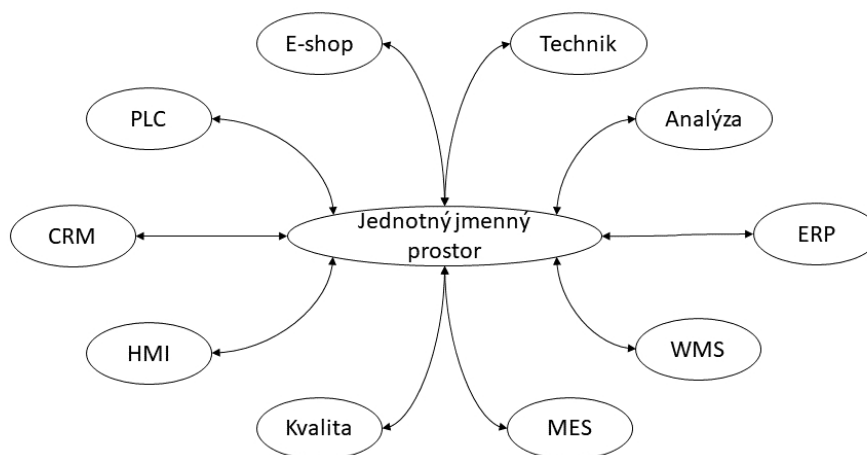
### **Následky této neshody**

Již dnes má podnik nepopiratelné problémy s ERP systémem, který je nedostačující pro potřeby podniku. Nedokáže totiž dostatečně pokrýt požadavky jednotlivých oddělení, a následkem toho vznikají vlastní řešení konkrétních problémů. Tato řešení nejsou připravena pro růst podniku, nemají náležitou podporu IT, a nejsou do budoucna ani udržitelná, pokud se bude podnik chtít rozvíjet ve smyslu Průmyslu 4.0.



### Navrhované řešení neshod v ERP systému

Možným řešením je revidování aktuálního ERP systému. Bylo by potřeba zjistit jeho nedostatky a důvody těchto nedostatků. Po důkladné rešerši je nutno diskutovat výsledky s odborníky oddělení a vedením podniku. Na základě diskuse je pak možno určit, zda umožňuje ERP systém náležitou podporu firmě, nebo zda ji brzdí v expanzi. Autor práce si uvědomuje významnost ERP ve firmě a nevidí jeho změnu jako vhodnou variantu řešení. Na Obrázku 5 předkládá proto návrh architektury, která by rozšířila možnosti podniku při zachování ERP systému.



Obrázek 5 Architektura pro nedostatky ERP řešení [Vlastní zpracování]

Architektura na Obrázku 5 využívá centrální bod pravdy, nazvaný „jednotný jmenný prostor“. Zde lze najít veškerá data, která jsou v podniku přítomna. Nedochozí zde k duplikaci dat a připojení k tomuto prostoru předchází skrytým propojením napříč systémy podniku. Systémy pracující s daty fungují s jedinou změnou. Tato změna spočívá v tom, že tato data získává z jednotného jmenného prostoru a po zpracování je do něj znovu ukládá. Tento přístup by usnadnil práci IT oddělení se správou dat a umožnil přístup k nim. Nad takto formátovanými a dostupnými daty lze postavit komplexní analytická řešení.

### Navrhované řešení neshod se Shadow IT

Shadow IT by mohlo pomoci podniku zjistit nedostatky centrálních systémů. Aktuální postoj vedení k tomu problému podle autora není správný. Podnik by měl přezkoumat a nalézt individuální řešení jednotlivých oddělení, která nejsou pod centrální správou a pokud je to možné, tak se pokusit o integraci.

### **5.5.4 Nedostačující komunikace v podniku**

Komunikace je důležitou součástí Průmyslu 4.0. Dalo by se říct, že cílem Průmyslu 4.0 je předání potřebných informací do správných rukou ve správný čas a ve správném formátu. Zde není myšlena pouze komunikace vedení s podřízenými, ale i opak. Nejvíce informací a zkušeností se nachází přímo ve výrobě v hlavě operátora zařízení. Pokud operátor dostane možnost vyjádřit své zkušenosti za účelem zlepšení procesu, měl by ho vedoucí pracovník vyslechnout.

Jednotlivá oddělení mezi sebou mají omezenou představu, jaká data se na konkrétním oddělení nacházejí. Například neví, kdo data vlastní, v jakém jsou formátu a jak nad nimi pracovat (pokud je to vůbec možné). Podnik by měl věnovat pozornost datové politice, která by jasně vymezovala datové vlastníky a procesy spojené s daty přístupnými podnikem. Data nacházející se pouze na jednom oddělení by měla být zpřístupněna pro podnik jako celek.

#### **Následky této neshody**

Neinformovanost zaměstnanců může mít různé následky, od nepochopení vedení až po ztrátu krokových zlepšení přímo ve výrobě. Správně informovaný zaměstnanec může lépe plnit svoji práci. Často se komunikací mezi zaměstnanci odstraní i palčivé problémy, jako například nepochopení vykonávané práce, a tudíž nemožnost předejít jasně chyby, které se zaměstnanec dopouští svou nevědomostí. Takové situace mohou mít za následek zpoždění dodání produktu, zhoršení psychického i fyzického zdraví zaměstnance nebo nechuť ke spolupráci.

#### **Navrhované řešení**

S tímto problémem by mělo pomoci nasazení procesního přístupu a nových datových toků. Podnik již navrhoval informační kiosky pro zaměstnance, zejména pro operátory výroby, kde by našli informace o aktuálním dění v podniku a odpovědi na často kladené otázky.

### **5.5.5 Nedostatečná automatizace datových toků**

Příkladem může být odhlašování činností výrobního procesu zaměstnanci. Odhlášení znamená, že zaměstnanec dokončil svou práci, a výrobek pokračuje k dalšímu úkonu. Stává se, že zaměstnanec zapomene odhlášení provést, nebo se nenačte správně karta

a v systému jsou následkem toho nekompletní údaje o provedeném úkonu. Automatizace datových toků je důležitá, protože odstraňuje chybu lidského faktoru, zrychluje komunikaci a zkvalitňuje práci. Dalším příkladem mohou být sbíraná data, která pracovník dostane k dispozici až po požádání nebo vlastní aktivitou. Data, týkající se výrobního procesu, která jsou sbírána a jsou pomocná k výkonu práce operátora, by měla být předávána v reálném čase přímo na odpovídající pracoviště.

### **Následky této neshody**

Časové úspory a přesnost a kvalita informace jsou cílem automatizace datových toků. Pokud operátor nemá potřebné informace ohledně svého pracoviště, mohou nastat komplikace a chyby. Výsledkem včas informovaného operátora může být předejití chyb, zkvalitnění jeho práce, a v neposlední řadě také jeho spokojenost.

Z rozhovoru se zaměstnancem podniku byla zjištěna existence databází, které pouze uchovávají data. Nad těmito daty nejenže není prováděna prakticky žádná další práce, ale nejsou ani dostupná pro zaměstnance, kteří by s nimi mohli pracovat.

### **Navrhované řešení**

Většina ručně způsobených přenosů dat může být zautomatizována. Bylo by vhodné informovat zaměstnance, kteří pravidelně spouští datové toky nebo s nimi pracují, o těchto možnostech automatizování, o něž by si pak mohli zažádat. Tento krok by mohlo uspíšit i nahlédnutí do logů hlavních systémů, ze kterých se data čerpají a nabídnutí pravidelných exportů. Příkladem je denní stahování určitých dat za účelem osobního reportu pro ověření nově nasazeného postupu.

## 6 Procesně řízená Big data v Průmyslu 4.0

Pro stanovení vztahu mezi Big daty a jejich zpracováním s návazností na procesní řízení a Průmysl 4.0 je nejprve třeba tyto dva pojmy uvést.

### 6.1 Procesní řízení

Procesní řízení je přístup k řízení organizace založený na zajištění nepřerušovaného toku logicky seřazených činností, které přidávají hodnotu v jednoznačně definovaných a vzájemně propojených podnikových procesech. Procesní řízení zajišťuje dlouhodobou prosperitu a konkurenceschopnost organizace v prostředí, které je charakteristické neustálými změnami a nejistým budoucím vývojem. Jednoznačně definované a kvantitativně vyjádřené strategické cíle organizace jsou naplňovány trvale zlepšovanými procesy [42].

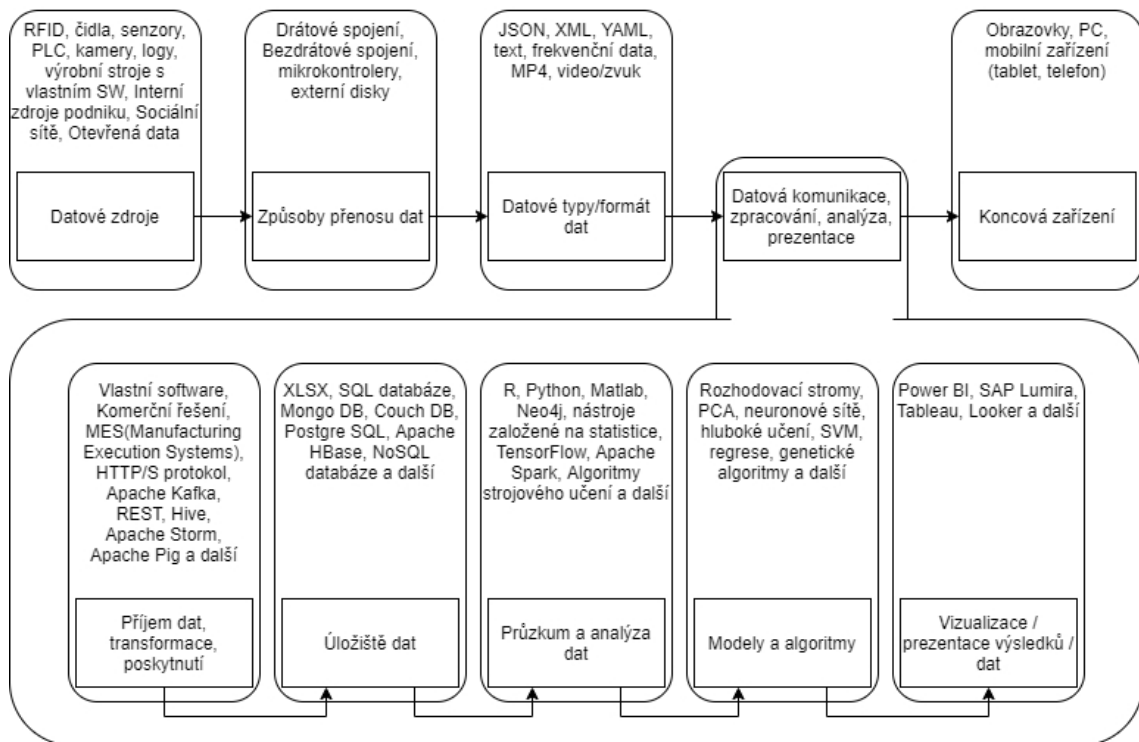
Podnik bývá většinou rozdělen podle organizační struktury, kde jsou jednotlivá oddělení specializovaná na konkrétní problematiku a následně zodpovědná za její zpracování. Toto vymezení hranic napříč podnikem může vést k nepochopení mezi odděleními a k nedostatečné komunikaci uvnitř podniku. Procesní přístup se snaží tyto bariéry překonat a soustředí se na hlavní procesy, procházející celým podnikem.

Procesní přístup by měl být nasazován pro celý podnik, protože to vede na procesy vytvářející hodnotu. Následné nastavení procesu a jeho činností napříč podnikem by mělo probíhat bez ohledu na organizační strukturu. Procesní přístup generuje organizační strukturu.

Procesní přístup se snaží zejména o zvýšení efektivity podniku. Podnik by měl být řízen procesně, aby mohl kompletně převzít myšlenku Průmyslu 4.0. Big data v tomto ohledu potřebují jasně popsané procesy, které budou uznávány napříč celým podnikem.

## 6.1.1 Životní cyklus dat

Životní cyklus dat popisuje všechna stádia, kterými data prochází za dobu své existence. Každé stádium obsahuje své problémy a řešení. Obecně se životní cyklus dat skládá ze stádií, která jsou vyobrazená na Obrázku 6.



Obrázek 6 Životní cyklus dat se souvisejícími technologiemi a metodami [Vlastní zpracování podle [48]]

Datové zdroje, způsob přenosu dat a datové formáty jsou nedílnou součástí Big dat a předchází jakékoliv analytické metodě. Autor práce považuje za důležité zmínit prerekvizity datové analýzy. V Obrázku 6 jsou jen některé z mnoha možností, které obsahuje každá sekce, čímž se snaží ilustrovat nesourodost možných datových toků, které se v podniku mohou vyskytnout.

## 6.1.2 Z datového pohledu

Data prostupují celým podnikem, stejně jako proces vytvářející hodnotu. Dá se říct, že data tento proces dokumentují, podporují a umožňují měřit kvalitu a hodnotu procesu.

Obrázek 6 popisuje obecný životní cyklus dat, který má každý podnik svým způsobem upravený. Celý tento cyklus lze spojit s procesním přístupem. Zdroje dat v podniku jsou různorodé stejně jako jejich formát. Datové zdroje lze najít napříč celým podnikem,

například senzory výrobní linky, operátoři výroby nebo internetové stránky firmy či její sociální síť. Existují i datové zdroje, které jsou mimo správu podniku, například odborná internetová fóra nebo podobně zaměřené podniky. Uvnitř podniku lze přenos dat z takových zdrojů definovat či vytvořit podle schopností podnikového specialisty. Avšak při přenosu dat ze zdrojů, které jsou mimo správu podniku, je nutné pracovat s předdefinovaným rozhraním. Následně je před uložením nutno data kontrolovat, aby nedocházelo ke snížení datové kvality například tvorbou duplikací či vložením špatného formátu. Uložení dat by mělo doprovázet vytvoření záznamu v datovém katalogu. Takto uložená data lze předat odborným pracovníkům k analýze nebo přímo k prezentaci. Tento proces datového života v organizaci lze považovat jako část procesního přístupu.

Data, která jsou náležitě zpracována a odborníci z různých oddělení firmy k nim mají přístup, jsou pro podnik důležitá. Nad stejnými daty může každý odborník podle svého zaměření najít jinou optimalizaci, proto je důležité, aby data procházela napříč celým podnikem. Dále taková data lze využít k získání informací napříč odděleními. Odborní pracovníci nejsou limitováni zdroji oddělení a mohou využít komplexních informací podniku.

Procesní přístup zároveň pomáhá odbourat duplicitu řešení. Pokud je v podniku vyřešen problém, je vhodné mít řešení náležitě popsané a mít přístup k datům, která s řešením souvisí. Nemělo by se potom stávat, že jedno oddělení vyřeší problém s přenosem dat a druhé oddělení stále tápe nad stejným přenosem. Tím se lze dostat k roli IT oddělení v podniku. Řešení datové komunikace spadá pod činnost IT oddělení, které by mělo spravovat veškeré procesy související s Obrázkem 6. IT by mělo určovat oficiálně podporované nástroje podniku, být nápomocné při školení a práci s těmito nástroji, a v neposlední řadě také vyžadovat po uživatelích správné zacházení s daty.

## **6.2 Průmysl 4.0**

V České republice vydala vláda dokument s názvem Iniciativa Průmysl 4.0 v roce 2017 [32]. Cílem tohoto dokumentu je poskytnout klíčové informace související s tématem čtvrté průmyslové revoluce, ukázat možné směry vývoje a nastítnit návrhy opatření, která by mohla nejen podpořit ekonomiku a průmyslovou základnu ČR, ale i pomoci připravit celou společnost na absorbování této technologické změny. Česká iniciativa si vzala za vzor německou strategii nazvanou Industrie 4.0 [33].

Jako další program na podporu čtvrté průmyslové revoluce lze jmenovat například program Advanced Manufacturing Partnership (AMP) z USA nebo Čínskou iniciativu Made in China 2025 a mnohé další [34].

Všechny tyto iniciativy se snaží přiblížit budoucnosti výrobních podniků takzvané Chytré výroby. Chytrá výroba byla definována Národním Institutem standardizace a technologií (NIST) jako plně integrovaný, spolupracující výrobní systém, který v reálném čase reaguje na měnící se požadavky a podmínky v továrně, v napájecí síti a v potřebách zákazníků [35].

Výrobní průmysl využívá rozsáhlý výběr SW a systémů automatizace pro zvýšení efektivity a produktivity – od programovatelných logických automatů (PLC), přes dispečerské řízení a sběr dat (SCADA), Výrobní informační systémy (MES), počítačem podporované technologie (CAX: CAD, CAAD, CAM, CAE), Řízení životního cyklu výrobku (PLM), Environmentální systém managementu (EMS), Provoz a údržbu (O&M), Správu dodavatelského řetězce (SCM), Řízení vztahů se zákazníky (CRM) a mnohé další. Cílem Chytré výroby je propojení těchto systémů. Tento úkol má dva hlavní problémy. Zaprvé, tyto systémy a SW jsou vyvíjeny mnoha prodejci nad různými protokoly přenosu dat. Zadruhé, tito prodejci většinou postrádají informace přímo z výroby, kde jsou jejich systémy nasazeny. Nemají tedy potřebná data k tomu, aby jejich systémy mohly reagovat v reálném čase na problémy z výroby nebo na náhlou změnu požadavků zákazníka.

Jako řešení těchto dvou problémů jsou navrhovány dva koncepty. Digitální vlákno a Digitální dvojče, které prostupují celým životním cyklem produktu a snaží se propojit fyzické nástroje s kyberprostorem výroby [36]. Kyberfyzický systém (CPS) je fyzický systém, který je monitorován, kontrolován, integrován a řízen za pomoci počítačových algoritmů [37]. Internet věcí (IoT) je nedílnou součástí Chytré výroby. Přistupuje k datům kyberprostoru a umožňuje jejich zpracování v reálném čase. Digitální dvojče a Digitální vlákno staví na CPS a IoT. Propojením těchto technologií vzniká obrovské množství dat, které je třeba náležitě zpracovat. Tato data měla procházet procesy sběru dat, uložení, agregace a integrace, analýzy a prezentace informací výrobcí.

Právě Big data by mohla být způsobem, jak využít koncepty Digitálního vlákna a Digitálního dvojčete. Pomocí tohoto spojení má také prudký vývoj v oblastech Cloud computingu, datové vědy a AI.

Princip Chytré výroby za pomoci Big dat a poznatků z výroby dodává informace v reálném čase. S využitím těchto informací lze identifikovat chybné výrobky, realizovat prediktivní údržbu strojů a zlepšit datové podklady pro systém MES, který může lépe plánovat výrobu.

### **6.2.1 Technologie a problémy**

Technologie a nástroje pro práci s Big daty přináší možnosti pro Chytrou výrobu. Mnoho nástrojů je vyvíjeno pod záštitou velkých firem, jako jsou: Google, Microsoft nebo Facebook.

Využití dostupných nástrojů v metodologii Chytré výroby přineslo několik problémů, které jsou specifické pro výrobní podniky. Zaprvé, tyto nástroje nerozlišují mezi Big daty z výroby a internetovými Big daty. Přístroje z výrobního průmyslu musí splňovat určité standardy, které jsou specifické pro výrobu nebo přístroj. Výrobci průmyslových zařízení musí tyto standardy dodržovat, ale už není definováno, jakým způsobem. Každý výrobce má vlastní uživatelské rozhraní. Přístroj také může využívat řadu protokolů, které mohou být prodejcem upraveny pro konkrétní výrobek. Tyto rozdíly oproti internetu, kde jsou data většinou ve formátu přirozené řeči nebo v polostrukturovaných formátech, jako jsou XML nebo JSON, nebyly brány v potaz při vývoji Big data nástrojů. Tento nedostatek vyplývá z prostředí, ve kterém se hlavní vývojářské firmy pohybují, jedná se o SW firmy se zaměřením na internetové služby. Zadruhé, nástroje pro práci s Big daty jsou masivní, různorodé nástroje. Vybrání správného nástroje pro konkrétní použití není triviální problém, a často vyžaduje konkrétní definici užití. Zatřetí, výrobní systémy mají velmi specifické úkoly ohledně výroby. Často nelze použít obecná řešení. Na rozdíl od internetových Big dat se výrobní Big data liší již od začátku jejich životního cyklu. Big data z výroby jsou poskytována senzory a mají jasně danou strukturu. Při zjišťování stavu přístroje je možné mít libovolný počet senzorů, a proto je potřeba vytvořit model stavu přesně na míru tomuto stroji [41].



## 6.2.2 Význam pro zkoumaný podnik

Podnik se snaží o digitalizaci, která však není centrálně podporována. Pro IT oddělení se dvěma zaměstnanci není možné pokrýt nároky celého podniku ve smyslu digitalizace a Průmyslu 4.0, natož stavět Big data řešení. Tento problém by neodstranil ani outsourcing dané problematiky třetí straně, tedy pořízení infrastruktury a řešení například od Amazonu nebo Microsoftu.

Digitalizaci a další následné kroky by v podniku měl řešit vyšší management za podpory svých odborníků nebo analytiků třetí strany, který by měl zvážit strategii podniku v dalších letech. Pokud podnik chce vyrábět pod záštitou čtvrté průmyslové revoluce, musí změnit své dosavadní smýšlení a strukturu. Průmysl 4.0 nemá jasnou a přímočarou definici. Management by měl být dostatečně informován ohledně pokroku a měl by za pomoci svých odborníků formulovat konkrétní strategii a změny v podniku v souladu s Průmyslem 4.0 a procesním řízením. Bez náležitého rámce nedává smysl budování Big dat v podniku jen za účelem toho, aby se dalo říct, že podnik má Big data.

Lze konstatovat, že by se podnik měl zaměřit na aktuální řešení a přezkoumat, zda je možná integrace aktuálních řešení do nového konceptu, který bude v souladu s procesním přístupem a pomůže podniku dospět do souladu s Průmyslem 4.0, a následně umožní naplno využít potenciál Big dat. Takto velké změny jsou pro podnik finančně náročné a mohou podniku zaručit stabilní růst, ale při nesprávné implementaci způsobit i škody. Kladen by měl být důraz na podrobnou rešerši stávajícího stavu uvnitř firmy, a poté využití rámcových doporučení, která se již povedlo implementovat [41]. Malé a střední podniky již nemusí experimentovat a riskovat. Velké firmy již tuto fázi překonaly a vytvořily pro menší a střední podniky řešení, která se osvědčila. Malé a střední podniky mohou tedy těžit z úspěchů i neúspěchů velkých firem.

V neposlední řadě je i sebelepší strategie předurčena k neúspěchu, pokud s ní nejsou zaměstnanci seznámeni a nemají šanci se podílet na její realizaci či pokud nerozumí smyslu strategie. Zaměstnanec by měl být uvědomen o změnách v podniku. Hlavní je porozumění zaměstnanců nastávajícím změnám, jaký je jejich důvod a jaký budou mít dopad na jejich práci. Komunikace v podniku je základní stavební kámen pro úspěšnou digitální transformaci, která je jedním z kroků k přijetí konceptu Průmyslu 4.0.

## 7 Cloudová řešení

Jedním z trendů zmíněných v kapitole 1.4 je přesun ke cloudovým řešením. Nad rámec této práce byl proto vytvořen ukázkový příklad převedení Excel či CSV souboru do cloudové databáze za pomoci Python skriptu.

Tento převod do databáze by zaměstnancům podniku umožnil přístup k datům odkudkoliv za předpokladu, že mají přístupová práva k databázi a internetovému připojení. Soubory nahrané do databáze lze spravovat centrálně, zaručit jejich dostupnost, kompletně je logovat a spravovat přístupy k databázi. Všechny tyto vlastnosti by bylo velmi složité nastavit pro Excel soubory. Nad takto nahranými daty lze pokračovat například metodami uvedenými na začátku práce, tedy některou analýzou nebo těžením informace. Datový steward může nahrávaná data konzultovat přímo s tvůrci a vytvářet propojení nahrávaných dat, které uživateli pomůže vytvořit si obecný obrázek o jeho konkrétní otázce, kterou se snaží podpořit datovým výzkumem. U dat nahraných do cloudové databáze je také možnost neomezeného škálování, jež vyhoví budoucímu růstu podniku a jeho dat.

### Microsoft Azure

Autor práce zvolil využití Microsoft Azure SQL Database [43], protože nabízí bezplatné vyzkoušení svých služeb. Zároveň byl výběr databáze ovlivněn zkušenostmi autora s cloudovými službami od společnosti Microsoft. Jedná se pouze o ukázkový příklad a autor práce nevyzdvihuje toto řešení nad jinými cloudovými službami, jako jsou: AWS, Google nebo Oracle. V tomto příkladu byla využita bezplatná licence, registrovaná na školní účet autora práce.

Popis veškeré funkčnosti a provázanosti produktů MS Azure je mimo rozsah práce. V této práci byl použit portál Azure pro přehled databáze a také kontrolu nahraných dat pomocí Query editoru, který umožňuje dotazovat databázi jazykem SQL.

### Nastavení databáze

Na webovém portálu služeb Azure byla zvolena Azure SQL. Po kliknutí na vytvoření nové služby má uživatel na výběr ze tří možností. První možností je vytvoření nové SQL databáze nebo vytvoření takzvaného Elastic pool, což umožňuje management více SQL databází najednou. Byla zvolena tvorba jedné nové SQL databáze. Byla zvolena

Bezplatná zkušební verze, v níž je zdarma k dispozici kredit v hodnotě 170 Euro po dobu 30 dnů. Tato verze byla vhodná pro zpracování vzorového úkolu. Licenci je možno také kdykoliv rozšířit na jednu z předpřipravených licencí od Microsoftu nebo na licenci vytvořenou přesně na míru podniku [44]. Dále je nutné vyplnit skupinu prostředků. Jedná se o kolekci prostředků, které sdílí životní cyklus, oprávnění a jiná pravidla. Další nutná pole jsou název databáze a server, který musel být vytvořen nově. Byly vyplněny požadované údaje v průvodci vytvořením serveru. Volba Elastic pool byla zaškrtnuta jako „Ne“ jelikož se jedná o školní projekt, který naplno nevyužije ani tu nejskromnější databázi, kterou lze vydefinovat.

Při konfiguraci nové SQL databáze lze vybrat z několika kategorií. Existují tři kategorie založené na Database Transaction Unit (DTU), jimiž jsou: Basic, Standard a Premium. Druhou možností jsou kategorie založené na vCores, neboli virtuálních jádrech. Jedná se o General Purpose, Hyperscale a Business Critical. Pro tento případ byla vybrána možnost, která bude stát nejméně přiděleného volného kreditu. Touto možností byla kategorie General Purpose s možností Serverless, a to na základě výpočtů z kalkulačky, která je přítomna po dobu celého nastavení a zobrazuje výpočet výsledné ceny. Maximální možný počet využitelných jader byl nastaven na minimum, což bylo jedno, a minimální počet nastaven taktéž na minimum, což bylo půl jádra. Pro ušetření nákladů byla zaškrtnuta možnost automatické pauzy, protože v tomto formátu se platí za každou sekundu aktivně využívaného jádra. Automatická pauza umožňuje pozastavení a uvolnění prostředků, pokud je databáze neaktivní po specifikovaný čas jedné hodiny. Poslední volbou byla maximální velikost databáze, která byla nastavena na 5 GB. Konečná částka byla 0,000134 Euro za využití jednoho virtuálního jádra na jednu sekundu.

### **Připojení k databázi**

Azure nabízí generované připojovací řetězce ve formátu ADO.NET, JDBC, ODBC, PHP a Go. Zároveň je pod generovaným řetězcem odkaz pro stažení odpovídajícího ovladače pro SQL server. Pro připojení k této databázi byl využit Open Database Connectivity (ODBC) konektor a Microsoft ODBC Driver 17 for SQL Databases (x64) [45].

Autor práce se k databázi připojuje pomocí skriptu napsaného v jazyce Python. Pro urychlení práce s ODBC konektorem byl využit modul pyodbc. Autor ve svém skriptu využívá řadu knihoven pro urychlení a ulehčení práce.

## Práce v Pythonu

Vytvořený skript obstarává připojení k výše definované databázi. Pomocí příkazového řádku přebírá tři povinné argumenty pro spuštění, kterými jsou jméno a heslo pro přihlášení do databáze a cestu k souboru, který chce uživatel nahrát. Po vyplnění povinných polí a spuštění skriptu je proveden pokus navázání spojení s databází a také pokus o přihlášení do databáze. Pokud jsou přihlašovací údaje v pořádku a uživatel má práva k práci s databází, skript zkontroluje existenci vybraného souboru, a pokud soubor na disku existuje, nahraje ho do databáze. První řádek souboru je databází chápán jako záhlaví sloupců.

Jelikož jde o ukázkový skript, tak nejsou definované přesné požadavky. Skript byl proto vytvořen s myšlenkou obecnosti a jeho hlavním úkolem je přesunutí souboru s formátem .csv; .xls; .xlsx; .xslm; .xlsb; .odf; .ods; .odt. Skript neopravuje chyby vstupního souboru, ani jej jiným způsobem neupravuje. Datové typy nejsou optimalizovány a nejsou hlídána integritní omezení. Smyslem tohoto příkladu je umožnění nahrání souborů s výše zmíněnými příponami do databáze, tak jak jsou.

Zdrojový kód 1 psaný v Pythonu verze 3.8.4 popisuje hlavní logiku přiloženého skriptu pro přenos dat. Funkce *parse\_input* využívá modul *argparse* pro zpracování vstupů z příkazové řádky. Proměnná *engine* uchovává object pro připojení k databázi a přebírá parametry, které odpovídají přihlašovacímu jménu a heslu databáze. Dále je proveden pokus o připojení k databázi, které ověří správnost zadávaných přihlašovacích údajů. Komunikaci s databází pomocí DML obstarává modul *sqlalchemy*. Třetí vstupní údaj je cesta k souboru, který chce uživatel nahrát do databáze. Pokud soubor existuje, skript načte data pomocí modulu *pandas* do datového rámce. Metoda nahrání dat je rozdělena podle přípony souboru na *upload\_csv\_data* a *upload\_excel\_data*. Při konkretizovaném zadání je možné skript snadno rozšířit v plnohodnotný ETL upravením těchto funkcí, kde může uživatel, znalý Python kódu, upravovat svůj soubor dle potřeb. Obě metody přebírají datový rámec, připojení k databázi a název souboru, který je nahráván. Metody kontrolují existenci tabulek stejného názvu. Pokud se v databázi již taková tabulka vyskytuje, uživatelem zadávaná data nebudou nahrána. Metoda *upload\_excel\_data* přidává k názvu tvořené tabulky ještě jména listů daného excel souboru. Tvoří tedy pro každý list samostatnou tabulku, která nese název podle jména souboru a jména listu. Formát názvu je spojen podtržítkem.

```

1  def main():
2      args = parse_input()
3      engine = azure_db_conn_alch(args.uid, args.passwd)
4      try:
5          engine.connect()
6      except exc.DBAPIError as e:
7          if e.code == 'rvf5':
8              sys.exit('Nebylo možné přihlášení do databáze, zkontrolujte přihlašovací údaje.')
9
10     if os.path.exists(args.filepath):
11
12         file_name, file_format = check_input_format(args.filepath)
13
14         if file_format == '.csv':
15             df_data = load_csv(args.filepath)
16             if upload_csv_data(df_data, engine, file_name) == 11:
17                 sys.exit('Tabulka s tímto názvem již existuje!')
18         elif file_format in ['.xls', '.xlsx', '.xlsm', '.xlsb', '.odf', '.ods', '.odt']:
19             df_data = load_excel(args.filepath)
20             if upload_excel_data(df_data, engine, file_name) == 11:
21                 sys.exit('Tabulka s tímto názvem již existuje!')
22         else:
23             sys.exit('Nevalidní formát datového souboru. Využijte -h pro nápovědu.')
24     else:
25         sys.exit('Vámi zadaný soubor neexistuje.')
26     sys.exit('Data byla nahrána.')

```

*Zdrojový kód 1 Ukázka hlavní funkce Python skriptu [Vlastní zpracování]*

## 8 Doporučení pro podniky

Na základě provedené rešerše a poznatků získaných v rámci praktické části autor práce navrhuje následujících pět doporučení pro malé a střední výrobní podniky:

- Vytvoření strategie, případně úprava stávající.
- Vybrání odpovídajících nástrojů a odpovědného personálu.
- Implementování standardů řízení, postupů pro práci s daty a zřízení jejich bezpečnosti.
- Vybrání vhodného případu použití pro demonstraci výsledků.
- Nepřetržité sledování vývoje technologií a s nimi související studie.

### **Vytvoření strategie, případně úprava stávající**

Už z názvu Čtvrtá průmyslová revoluce vyplývá, že se jedná o zásadní změnu. Tato změna musí být řízena z vedení podniku a měla by mít podporu celého podniku. Tento krok musí řešit a prosazovat vrcholný management za pomoci specialistů na danou problematiku. Zásadním problémem tohoto kroku je neznalost a nevzdělanost managementu v tomto odvětví, což se může odrazit na jejich rozhodnutí. Nedělitelnou součástí Průmyslu 4.0 je IT podniku, které má za úkol nastavit pravidla a vytvořit infrastrukturu podniku. Je žádoucí, aby byli všichni zaměstnanci podniku seznámeni s novou strategií. Ta by měla zmiňovat možnosti přístupu k datům, jejich klasifikaci, možnosti využití a bezpečnostní pravidla. Měly by zde být podnikové cíle spojené s digitalizací a informačními technologiemi, které umožní čtenáři pochopit nastavení pravidel a porovnání mezi aktuálními možnostmi podniku a jeho cílem v dohledné době. Strategie by měla rámcově definovat budoucí cíle organizace, kterými mohou být například přesunutí infrastruktury na cloud nebo zřízení firemního IoT.

### **Vybrání odpovídajících nástrojů a odpovědného personálu**

Jak již bylo zmíněno v šesté kapitole, výběr správných nástrojů je netriviální problém, který nelze perfektně vyřešit hned napoprvé, bez využití znalostí podniku. Zde je nutná analýza současného řešení podniku a jasné stanovení cílů. Pro vybrání správných nástrojů je potřeba přesně vědět, čeho chce podnik dosáhnout. Následně má podnik možnost vzdělání svých zaměstnanců, což je pro podnik z pohledu lidských zdrojů lepším přístupem, a druhou možností je nábor nových odborníků. Obě možnosti mají

své výhody a úskalí. Vzdělávání zaměstnanců je časově náročné a vyžaduje zájem samotných pracovníků. Na druhou stranu, tito zaměstnanci mají zkušenosti s prací v podniku a vědí, jak podnik aktuálně funguje. Mají tedy know-how o podniku, které jim umožňuje uplatnění nově nabytých vědomostí. Náborem nových zaměstnanců se podnik vyhýbá financování vzdělávání. Nový zaměstnanec ovšem nemůže své znalosti začít uplatňovat ihned po nástupu do podniku, aby mohl naplno využít své zaměření. Potřebuje k tomu i pochopení chodu podniku, které mu oproti stávajícímu zaměstnanci schází.

### **Implementování standardů řízení, postupů pro práci s daty a zřízení jejich bezpečnosti**

Tento bod je spojen s procesním řízením, které výrobní podniky dobře znají. Stejně jako jsou definovány procesy pro výrobu, tak musí být definovány procesy pro práci s daty během celého jejich životního cyklu. Zde nastupuje IT oddělení a jejich soupis pravidel a procesů od pořízení dat, přes jejich přenos, až po uložení a zřízení přístupu k nim. Mezi taková pravidla by mohlo patřit každodenní uložení všech přírůstkových dat klasifikovaných jako interní data podniku do určeného centrálního úložiště s náležitým popisem dat. Dalším případem mohou být data pořízená senzorikou nasazenou ve výrobě, která však musí být vyhodnocována v reálném čase, a proto pro jejich zpracování bude potřeba použít odlišnou technologii. Tato data budou po vzniku automaticky nahrávána do datového skladu nebo na cloud. Je tedy nutno určit pro každý typ dat odlišné úložiště. Standardizace práce s daty přináší jasný rámec pro všechny zaměstnance a zřizuje řád pro jejich práci s nimi.

Autor práce doporučuje využití otevřené architektury. Velké společnosti si mohou dovolit využití odlišných technologií v různých částech firmy, protože mají odborný personál nebo dostatečný rozpočet pro zaplacení osob třetích stran. Malé a střední podniky mohou tento přístup částečně imitovat využitím otevřených přístupů, což znamená, že podnik bude budovat na otevřených technologiích. To znamená, že podnik nepřijde o možnost adopce nové technologie nebo propojení stávajícího řešení s novými funkcionalitami. V otevřeném přístupu bude připojení nových funkcionalit a technologií možné vyzkoušet v přípravném projektu a zhodnotit přínosy pro podnik, namísto zavrnutí z důvodu nemožnosti připojení ke stávajícímu řešení. S otevřeným přístupem

nenastane situace, v níž by systém neumožňoval adoptovat cizí práci nebo vyžadoval neúnosné finanční náklady na adaptaci.

### **Vybrání vhodného případu použití pro demonstraci výsledků**

Vybrání pilotního projektu pro implementaci je důležité. Nejzkušenější zaměstnanci napříč odděleními, každý s jiným zaměřením, by se měli shodnout na nejlepším využití datové inteligence v podniku. Podnik nelze převést na datově orientovaný ze dne na den. Vhodným případem jsou problémy, které jsou jasně definované. Na takových problémech se odhalí nedostatky nastavených pravidel a šedá místa v infrastruktuře podniku. Při prvních projektech se podnik a odpovědní zaměstnanci teprve učí s novými nástroji zacházet, a proto je důležité jasně a přesně nastavit cíl projektu. Pilotní projekt má ukázat možnosti nově pořízených technologií, a hlavně popsat přínos pro podnik.

### **Nepřetržité sledování vývoje technologií a s nimi související studie**

Byl zmíněn prudký vývoj v oblasti Big dat a příslušných technologiích, který je nutné sledovat. Držet krok s vývojem je velmi důležité v jakékoli oblasti podnikání. Mluví se zde o sledování trendů a metodologií. Vybraná technologie bude používána dlouhodobě, a je proto nutné provést kvalitní průzkum potřeb podniku a možných řešení. Je nelogické nasadit v podniku například řešení od Microsoftu a po pilotním projektu se rozhodnout pro vyzkoušení Amazonu. Sledovat by se měl vývoj kolem vybraných technologií, které budou v podniku nasazeny a možnosti jejich rozšíření či zlepšení. Myšlenka neustálého zlepšování je tu již od třetí průmyslové revoluce a platí stále. Sem také spadá vzdělávání odborného personálu a pravidelná aktualizace nasazených prostředků. Je nutná úprava nastavených pravidel podle dat a zjištění, která přijdou s používáním nových technologií týkajících se Průmyslu 4.0.



## 9 Závěr

Cílem bakalářské práce bylo seznámení se s metodami analýzy dat a provedení jejich srovnání, které bylo uskutečněno pomocí korelační matice QFD metody. V práci jsou zmíněny kategorie analýzy a k nim byly vybrány jednotlivé metody. Tato porovnání byla provedena na základě potřebných prekvizit a na výstupu metod. Celkem bylo vybráno deset kritérií, která představují neurčité otázky managementu ohledně analýzy dat.

Každý podnik má v této době potenciál pro práci s Big daty. Po provedeném průzkumu řešení vybraného podniku však autor práce došel k závěru, že v tomto podniku se jedná pouze o potenciál, nikoliv realitu. Tato práce zkoumala zejména výrobní proces, systémy s ním související a byla zaměřena na data a datové toky podporující výrobní procesy. Autor práce spolupracoval s IT oddělením a oddělením kvality, odkud čerpal veškeré informace o podniku. Byly zjištěny závažné nedostatky připravenosti podniku pro přechod na Průmysl 4.0, které ovlivňují veškerou práci s daty. IT oddělení schází zaměstnanci i podpora ze strany vedení.

V podniku byly nalezeny neshody týkající se počtu zaměstnanců a jejich komunikace, výrobního procesu a datové politiky podniku, na něž byla aplikována FMECA analýza. Autor práce dále rozepisuje hlavní neshody podniku a přikládá k nim možné následky. K hlavním neshodám autor uvádí možné řešení na základě provedené analýzy a svých pracovních zkušeností. Autorem navrhovaná řešení nejzávažnějších neshod jsou uvedena v kapitole 5.5 a obecná doporučení v kapitole 8. Mezi nejzávažnější neshody patří nejasná datová strategie podniku, na kterou navazuje nedostatečný počet odborných zaměstnanců. Od těchto dvou problémů se odvíjí zbytek neshod, jako například aktuální stav podnikové infrastruktury se vznikem Shadow IT. Autor dále navrhuje diskusi mezi odborníky a vedením podniku na téma digitalizace a datové strategie. Bude také nezbytné obsadit nově vznikající pracovní pozice spojené se změnami, pokud se podnik rozhodne vydat směrem Průmyslu 4.0.

Práce zkoumá provázání procesního přístupu, Průmyslu 4.0 a datových zdrojů. Během průzkumu podniku byl zjištěn funkční přístup k jeho řízení, a to na základě rozhovorů se zaměstnanci. Podnik má jasně danou organizační strukturu a zaměstnanci často

neměli pochopení pro problémy jiných oddělení. Autor práce zmiňuje návaznost procesního přístupu, který prochází podnikem, a jeho datovou stopou. Autor práce přikládá aktuální stav datové infrastruktury vedení podniku a jeho nedostatečnému zájmu ohledně IT. Podniku se aktuálně daří velmi dobře a je jedničkou na českém trhu, čemuž dopomohla i pandemie nemoci COVID-19, protože měla silně negativní dopad na konkurenty. Kvůli pandemii však musel podnik začít řešit digitalizaci, což odhalilo nedostatky v datové infrastruktuře. Autor práce navrhuje řešení, které podnik připraví pro přechod na digitální podnikání a pomůže začít řešit otázky Průmyslu 4.0.

Autor práce předkládá poznatky ze svého průzkumu podniku a poukazuje na problémy, například na nemožnost využití komplexnějších analytických metod ke zpracování dat nebo limity integrace nových technologií. Byl předložen ukázkový příklad přenesení Excel dat do Cloudové databáze. Jedná se o jeden z přístupů, které by podnik mohl využít ke sjednocení svých dat. Nad takto sjednocenými daty lze následně vymáhat datovou politiku a bezpečnost.

Autor práce apeluje na vedení podniku a jejich přístup k digitalizaci a transformaci podniku ve smyslu Průmyslu 4.0. Podniky, které jsou schopny efektivně pracovat se svými daty, mají značnou výhodu nad těmi, které se svým datům náležitě nevěnují. Pokud vedení nezmění svůj přístup k datové základně, mohl by podnik v budoucnu přijít o zdroje i svoji pozici jedničky na českém trhu, protože nebude schopen držet krok se zvyšující se poptávkou a nebude disponovat vhodnými nástroji k efektivnímu řízení a výrobě.

Předmětem dalšího bádání autora bude prohloubení znalostí ohledně návrhu datových architektur. Především se zaměří na průmyslový internet věcí a cloudové služby. Autor věří, že tyto dvě oblasti budou v budoucnu hojně využité i malými a středně velkými podniky. Tuto problematiku by autor chtěl řešit zejména v praxi.

## Seznam použité literatury

- [1] DIEBOLD, Francis X. What's the big idea? "Big Data" and its origins. In: *Significance* [online]. 2021, 18(1), 36-37 [cit. 2021-5-8]. ISSN 1740-9705. Dostupné z: doi:10.1111/1740-9713.01490
- [2] KARAFIÁT, Michal. Big Data Metody zpracování a analýzy velkých dat. [Bakalářská práce.] Zlín: Univerzita Tomáše Bati ve Zlíně, FAI. 2017. [Online.] Dostupné na: <https://theses.cz/id/g8dbsk/fulltext.txt>
- [3] HENDL, Jan. METODY ANALÝZY VELKÝCH DAT [online]. In: MEDSOFT, 31.1.2019, s. 18 [cit. 2021-5-8]. Dostupné z: doi:10.35191/MEDSOFT\_2019\_1\_31\_15\_18
- [4] AL-SAI, Zaher Ali, Rosni ABDULLAH a Mohd heikal HUSIN. Big Data Impacts and Challenges: A Review. In: *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)* [online]. IEEE, 2019, 2019, s. 150-155 [cit. 2021-5-8]. ISBN 978-1-5386-7942-5. Dostupné z: doi:10.1109/JEEIT.2019.8717484
- [5] LANEY, Douglas. 3D Data Management: Controlling Data Volume, Velocity, and Variety [online]. B.m.: META Group. 2001. [cit. 2021-5-8]. Dostupné z: <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>
- [7] LIU, Jia, Tianrui LI, Peng XIE, Shengdong DU, Fei TENG a Xin YANG. Urban big data fusion based on deep learning: An overview. *Information Fusion* [online]. 2020, 53, 123-133 [cit. 2021-5-8]. ISSN 15662535. Dostupné z: doi:10.1016/j.inffus.2019.06.016
- [8] What is big data analytics? IBM [online]. Armonk (New York, United States): IBM, 2021 [cit. 2021-5-8]. Dostupné z: <https://www.ibm.com/analytics/hadoop/big-data-analytics>
- [9] What Does Big Data Mean to You? Microsoft SQL Server Blog [online]. Redmond (Washington, United States): Microsoft SQL Server Blog, c2021, 8. února 2013 [cit. 2021-5-8]. Dostupné z: <https://cloudblogs.microsoft.com/sqlserver/2013/02/08/what-does-big-data-mean-to-you/>

- [10] Drive innovation with Google Cloud smart analytics solutions. Google Cloud [online]. Mountain View (California, United States): Google Cloud, c2021 [cit. 2021-5-8]. Dostupné z: <https://cloud.google.com/solutions/smart-analytics>
- [11] ZIKOPOULOS, Paul, Chris EATON, Dirk DERROOS, Tom DEUTSCH a George LAPIS, SIT, Steven, ed. *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*. 1. USA: The McGraw-Hill Companies, 2012. ISBN 978-0-07-179053-6.
- [12] HARIRI, Reihaneh H., Erik M. FREDERICKS a Kate M. BOWERS. Uncertainty in big data analytics: survey, opportunities, and challenges. In: *Journal of Big Data* [online]. 2019, 6(1) [cit. 2021-5-8]. ISSN 2196-1115. Dostupné z: [doi:10.1186/s40537-019-0206-3](https://doi.org/10.1186/s40537-019-0206-3)
- [13] COX, Michael a David ELLSWORTH. *Managing Big Data for Scientific Visualization* [online]. 1. května 1997, 17 [cit. 2021-5-8]. Dostupné z: [https://www.researchgate.net/profile/David-Ellsworth-2/publication/238704525\\_Managing\\_big\\_data\\_for\\_scientific\\_visualization/links/54ad79d20cf2213c5fe4081a/Managing-big-data-for-scientific-visualization.pdf](https://www.researchgate.net/profile/David-Ellsworth-2/publication/238704525_Managing_big_data_for_scientific_visualization/links/54ad79d20cf2213c5fe4081a/Managing-big-data-for-scientific-visualization.pdf)
- [14] KRAMER, Oliver. *Genetic Algorithm Essentials*. Cham: Springer International Publishing, 2017. *Studies in Computational Intelligence*. ISBN 978-3-319-52156-5.
- [15] VINCENT, L. Big data – buzzword s obřím a nenaplněným potenciálem. *Mawrick: Magazín KPMG v České republice* [online]. Praha 8: KPMG Česká republika, 2020, 28.2.2020 [cit. 2020-11-20]. Dostupné z: <https://www.marwick.cz/data/big-data-buzzword-s-obrim-a-nenaplnenym-potencialem>
- [16] CODD, E.F. A relational model of data for large shared data banks. In: *Communications of the ACM*. Červen 1970, **13**(6), 377-387. [cit. 2021-5-8]. Dostupné z: [doi:https://doi.org/10.1145/362384.362685](https://doi.org/10.1145/362384.362685)
- [17] CHEN, P. The entity-relationship model—toward a unified view of data. In: *ACM Transactions on Database Systems*. Březen 1976, **1**(1), 9-36. [cit. 2021-5-8]. Dostupné z: [doi:https://doi.org/10.1145/320434.320440](https://doi.org/10.1145/320434.320440)

- [18] CODD, E. F. Relational database. *Communications of the ACM* [online]. 1982, 25(2), 109-117 [cit. 2021-5-8]. ISSN 0001-0782. Dostupné z: doi:10.1145/358396.358400
- [19] BERG, Kristi L., Tom SEYMOUR a Richa GOEL. History Of Databases. In: *International Journal of Management & Information Systems (IJMIS)* [online]. 2012, 17(1), 29-36 [cit. 2021-5-8]. ISSN 2157-9628. Dostupné z: doi:10.19030/ijmis.v17i1.7587
- [20] GARTNER INC. Big Data. *Gartner* [online]. Stamford: Gartner, c2020 [cit. 2020-12-19]. Dostupné z: <https://www.gartner.com/en/information-technology/glossary/big-data>
- [21] DRUŽSTEVNÍ ZÁVODY DRAŽICE - STROJÍRNA S.R.O. Výroční zpráva za rok 2019. 1. Benátky nad Jizerou, Dražice 69: Družstevní závody Dražice - strojírna, 2020. [online] Dostupné z: <https://or.justice.cz/ias/ui/vypis-sl-detail?dokument=61818792&subjektId=443416&spis=99233>
- [22] CHOLLET, François. *Deep learning v jazyku Python: Knihovny Keras, TensorFlow*. Praha 7: Grada publishing, 2018. ISBN 978-80-247-3100-1.
- [23] SIMMONS, A.B. a S.G. CHAPPELL. Artificial intelligence-definition and practice. In: *IEEE Journal of Oceanic Engineering* [online]. 1988, 13(2), 14-42 [cit. 2021-5-8]. ISSN 03649059. Dostupné z: doi:10.1109/48.551
- [24] VESELOVSKÁ, Kateřina. Analýza nestrukturovaných dat je klíčovou kompetencí. In: *Systemonline: S přehledem ve světě informačních technologií* [online]. Brno, 2017, 2017 [cit. 2020-12-27]. Dostupné z: <https://www.systemonline.cz/business-intelligence/analyza-nestrukturovanych-dat-je-klicovou-kompetenci.htm?mobilelayout=false>
- [25] SHENG, Jie, Joseph AMANKWAH-AMOA, Zaheer KHAN a Xiaojun WANG. COVID-19 Pandemic in the New Era of Big Data Analytics: Methodological Innovations and Future Research Directions. In: *British Journal of Management*. 1467-8551.12441. ISSN 1045-3172. [cit. 2021-5-8]. Dostupné z: doi:10.1111/1467-8551.12441

- [26] AL-ABASSI, Abdulrahman, Hadis KARIMIPOUR, Hamed HADDADPAJOUH, Ali DEHGHANTANHA a Reza M. PARIZI. Industrial Big Data Analytics: Challenges and Opportunities. In: *Handbook of Big Data Privacy*. Cham: Springer International Publishing, 2020, 2020-03-19, s. 37-61. ISBN 978-3-030-38556-9. Dostupné z: doi:10.1007/978-3-030-38557-6\_3
- [27] AL-SAI, Zaher Ali, Rosni ABDULLAH a Mohd heikal HUSIN. Big Data Impacts and Challenges: A Review. In: *2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)* [online]. IEEE, 2019, 2019, s. 150-155. ISBN 978-1-5386-7942-5. [cit. 2021-5-8]. Dostupné z: doi:10.1109/JEEIT.2019.8717484
- [28] OUSSOUS, Ahmed, Fatima-Zahra BENJELLOUN, Ayoub AIT LAHCEN a Samir BELFKIH. Big Data technologies: A survey. In: *Journal of King Saud University - Computer and Information Sciences* [online]. 2018, s. 431-448. ISSN 13191578. [cit. 2021-5-8]. Dostupné z: doi:10.1016/j.jksuci.2017.06.001
- [29] TALHA, M., A. ABOU EL KALAM a N. ELMARZOUQI. Big Data: Trade-off between Data Quality and Data Security. In: *Procedia Computer Science* [online]. 2019, s. 916-922. ISSN 18770509. [cit. 2021-5-8]. Dostupné z: doi:10.1016/j.procs.2019.04.127
- [30] BORODO, Salisu Musa, Siti Mariyam SHAMSUDDIN a Shafaatunnur HASAN. Big Data Platforms and Techniques. In: *Indonesian Journal of Electrical Engineering and Computer Science* [online]. 2016, s. 191-200. ISSN 2502-4760. [cit. 2021-5-8]. Dostupné z: doi:10.11591/ijeecs.v1.i1.pp191-200
- [31] DAI, Hong-Ning, Hao WANG, Guangquan XU, Jiafu WAN a Muhammad IMRAN. Big data analytics for manufacturing internet of things: opportunities, challenges and enabling technologies. In: *Enterprise Information Systems* [online]. 2020, s. 1279-1303. ISSN 1751-7575. [cit. 2021-5-8]. Dostupné z: doi:10.1080/17517575.2019.1633689
- [32] Iniciativa Průmysl 4.0. *Ministerstvo průmyslu a obchodu* [online]. Praha 1: odbor 01200, c2005-2020, 14.7.2017 [cit. 2021-03-07]. Dostupné z: <https://www.mpo.cz/cz/rozcestnik/ministerstvo/aplikace-zakona-c-106-1999->

sb/informace-zverejnovane-podle-paragrafu-5-odstavec-3-zakona/-iniciativa-prumysl-4-0--230485/

[33] KAGEMANN, Henning, Wolfgang WAHLSTER, Johannes HELBIG a Joaquín BLASCO, HELLINGER, Ariane, Veronika STUMPF a Christian KOBSDA, ed. *Recommendations for implementing the strategic initiative INDUSTRIE 4.0: Final report of the Industrie 4.0 Working Group* [online]. In: *Acatech*. 8 duben 2013, s. 82 [cit. 2021-03-07]. Dostupné z: <https://en.acatech.de/publication/recommendations-for-implementing-the-strategic-initiative-industrie-4-0-final-report-of-the-industrie-4-0-working-group/download-pdf?lang=en>

[34] LIAO, Yongxin, Fernando DESCHAMPS, Eduardo de Freitas Rocha LOURES a Luiz Felipe Pierin RAMOS. Past, present and future of Industry 4.0 - a systematic literature review and research agenda proposal. In: *International Journal of Production Research* [online]. 2016, 55(12), 3609-3629 [cit. 2021-03-07]. ISSN 0020-7543. Dostupné z: doi:10.1080/00207543.2017.1308576

[35] KUSIAK, Andrew. Smart manufacturing. In: *International Journal of Production Research* [online]. 2017, 56(1-2), 508-517 [cit. 2021-5-8]. ISSN 0020-7543. Dostupné z: doi:10.1080/00207543.2017.1351644

[36] HELU, Moneer, Thomas HEDBERG a Allison BARNARD FEENEY. Reference architecture to integrate heterogeneous manufacturing systems for the digital thread. In: *CIRP Journal of Manufacturing Science and Technology* [online]. 2017, 19, 191-195 [cit. 2021-03-07]. ISSN 17555817. Dostupné z: doi:10.1016/j.cirpj.2017.04.002

[37] MONOSTORI, L., B. KÁDÁR, T. BAUERNHANSL, et al. Cyber-physical systems in manufacturing. In: *CIRP Annals* [online]. 2016, 65(2), 621-641 [cit. 2021-03-07]. ISSN 00078506. Dostupné z: doi:10.1016/j.cirp.2016.06.005

[38] CUI, Yesheng, Sami KARA a Ka C. CHAN. Manufacturing big data ecosystem: A systematic literature review. In: *Robotics and Computer-Integrated Manufacturing* [online]. 2020, 62 [cit. 2021-03-07]. ISSN 07365845. Dostupné z: doi:10.1016/j.rcim.2019.101861

[39] GOASDUFF, Laurence. Gartner Top 10 Trends in Data and Analytics for 2020. Gartner [online]. Stamford, Connecticut, USA: *Gartner*, c2021, 19 října 2020 [cit.

2021-03-07]. Dostupné z: <https://www.gartner.com/smarterwithgartner/gartner-top-10-trends-in-data-and-analytics-for-2020/>

[40] MARR, Bernard. The 4 Biggest Trends In Big Data And Analytics Right For 2021. *Forbes* [online]. Jersey City: Forbes Media, c2021, 22 února 2021 [cit. 2021-03-07]. Dostupné z: <https://www.forbes.com/sites/bernardmarr/2021/02/22/the-4-biggest-trends-in-big-data-and-analytics-right-for-2021/>

[41] FISCHER, Marcus, Florian IMGRUND, Christian JANIESCH a Axel WINKELMANN. *Strategy archetypes for digital transformation: Defining meta objectives using business process management* [online]. 2020, 57(5) [cit. 2021-03-12]. ISSN 03787206. Dostupné z: doi:10.1016/j.im.2019.103262

[42] ŠMÍDA, Filip. *Zavádění a rozvoj procesního řízení*. Liberec, 2005. Disertace. Technická univerzita v Liberci, Hospodářská fakulta. Vedoucí práce Josef Sixta.

[43] Microsoft Azure: Azure SQL Database. Microsoft Azure [online]. Redmond (Washington, United States): Microsoft, c2021 [cit. 2021-5-8]. Dostupné z: <https://azure.microsoft.com/cs-cz/products/azure-sql/database/>

[44] Microsoft Azure. Microsoft Azure [online]. Redmond (Washington, United States): Microsoft, c2021 [cit. 2021-5-8]. Dostupné z: <https://azure.microsoft.com/cs-cz/free/sql-on-azure/>

[45] Microsoft Docs: SQL Docs. Microsoft Docs: SQL Docs [online]. Redmond (Washington, United States): Microsoft, c2021, 29.1.2021 [cit. 2021-5-8]. Dostupné z: <https://docs.microsoft.com/en-us/sql/connect/odbc/download-odbc-driver-for-sql-server?redirectedfrom=MSDN&view=sql-server-ver15>

[46] Norma ČSN EN IEC 60812: 2019 (010675) Analýza způsobů a důsledků poruch (FMEA a FMECA). Praha: UNMZ, 2019

[47] PELANTOVÁ, Věra a Jiří HAVLÍČEK. *Integrace a systémy managementu: [monografie]*. Liberec: Technická univerzita v Liberci, 2014. ISBN 978-80-7494-164-1.



[48] RAHUL, Kumar a Rohitash Kumar BANYAL. Data Life Cycle Management in Big Data Analytics. In: *Procedia Computer Science* [online]. 2020, 173, 364-371 [cit. 2021-5-8]. ISSN 18770509. Dostupné z: doi:10.1016/j.procs.2020.06.042

[49] WEBBER, Larry a Michael WALLACE. *Quality control for dummies*. 1. Indianapolis, Indiana: Wiley Publishing, c2007. ISBN 9780470069097.