

Česká zemědělská univerzita v Praze

Provozně ekonomická fakulta

Katedra informačních technologií



Diplomová práce

Business intelligence

David Mašík

© 2021 ČZU v Praze

ČESKÁ ZEMĚDĚLSKÁ UNIVERZITA V PRAZE

Provozně ekonomická fakulta

ZADÁNÍ DIPLOMOVÉ PRÁCE

Bc. David Mašík

Hospodářská politika a správa
Podnikání a administrativa

Název práce

Business intelligence

Název anglicky

Business intelligence

Cíle práce

Diplomová práce je tematicky zaměřena na problematiku Business Intelligence. Hlavním cílem práce je analýza BI řešení s návrhem vhodného řešení v konkrétní společnosti se zaměřením na finanční řízení a reporting. Dílčími cíli práce je analýza stávajícího stavu ve vybrané společnosti, vypracování přehledu možných řešení Business Intelligence, popis aplikace vhodného řešení a zhodnocení přínosu pro danou společnost se zaměřením na finanční řízení.

Metodika

Metodika řešení problematiky diplomové práce je založena na studiu a analýze odborných informačních zdrojů. Vlastní práce spočívá v analýze BI řešení a návrhu vhodného řešení v konkrétní společnosti z pohledu finančního řízení, analýze stávajícího stavu, vypracování možných řešení Business Intelligence, popis aplikace vhodného řešení a zhodnocení přínosu nového řešení pro danou společnost. Na základě syntézy teoretických poznatků a výsledků praktické části budou formulovány závěry diplomové práce.

Doporučený rozsah práce

60 – 80 stran textu.

Klíčová slova

Business intelligence, Master data management, Data warehouse, ETL

Doporučené zdroje informací

LABERGE, R. Datové sklady Agilní metody a business intelligence. Praha: Computer Press, 2012. ISBN 978-80-251-3729-1.

NOVOTNÝ, O. – POUR, J. – SLÁNSKÝ, D. – ČESKÁ SPOLEČNOST PRO SYSTÉMOVOU INTEGRACI. *Business intelligence : jak využít bohatství ve vašich datech*. Praha: Grada, 2005. ISBN 80-247-1094-3.

SHERMAN, R. Business Intelligence Guidebook: From Data Integration to Analytics. Waltham: Morgan Kaufmann, 2014. ISBN 978-0124114616.

ŠOLIAKOVÁ, L. – FIBÍROVÁ, J. *Reporting*. Praha: Grada, 2010. ISBN 978-80-247-2759-2.

VOCHOZKA, M. – MULAČ, P. *Podniková ekonomika*. Praha: Grada, 2012. ISBN 978-80-247-4372-1.

Předběžný termín obhajoby

2020/21 LS – PEF

Vedoucí práce

Ing. Pavel Šimek, Ph.D.

Garantující pracoviště

Katedra informačních technologií

Elektronicky schváleno dne 29. 7. 2020

Ing. Jiří Vaněk, Ph.D.

Vedoucí katedry

Elektronicky schváleno dne 21. 10. 2020

Ing. Martin Pelikán, Ph.D.

Děkan

V Praze dne 21. 02. 2021

Čestné prohlášení

Prohlašuji, že svou diplomovou práci "Business intelligence" jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou citovány v práci a uvedeny v seznamu použitých zdrojů na konci práce. Jako autor uvedené diplomové práce dále prohlašuji, že jsem v souvislosti s jejím vytvořením neporušil autorská práva třetích osob.

V Praze dne 29.3.2021

Poděkování

Rád bych touto cestou poděkoval vedoucímu této práce Ing. Pavlu Šimkovi, Ph.D. za jeho odborné rady a pomoc v rámci vedení této práce. Dále děkuji všem, kteří cennými podněty a informacemi přispěli k vypracování této práce.

Business intelligence

Abstrakt

Diplomová práce se zabývá optimálním využitím Business Intelligence v rámci finančního řízení daného podniku. Na základě zvolené metodiky je doporučeno nejvhodnější BI řešení pro daný podnik a zároveň je toto nové řešení využito pro tvorbu prototypu manažerského finančního reportu. Na tomto reportu jsou dále demonstrovány a hodnoceny hlavní přínosy nového řešení pro daný podnik. V rámci práce je především zdůrazněn přístup self-service, umožňující časově a nákladově efektivnější přístup k datovým, potažmo finančním analýzám.

Klíčová slova: Business Intelligence, Finanční reporting, Self-service BI, MS Power BI, MS Power Query, ETL, OLAP, Data Warehouse

Business intelligence

Abstract

Diploma thesis is focused on the optimal use of the Business Intelligence with a focus on financial management in a selected company. In accordance with chosen methods there is proposed an optimal BI solution for a company and in the following step is this new solution used for a creation of a new prototype of a managerial financial report. There are demonstrated main positive features of this new BI solution for a company on this report prototype. In the thesis as whole is accented a self-service approach which made possible a time and cost effective access to the data and financial analysis.

Keywords: Business Intelligence, Financial reporting, Self-service BI, MS Power BI, MS Power Query, ETL, OLAP, Data Warehouse

Obsah

1	Úvod	12
2	Cíl práce a metodika	13
2.1	Cíl práce	13
2.2	Metodika	13
3	Teoretická východiska	14
3.1	Business intelligence.....	14
3.1.1	Definice BI	14
3.1.2	Historie BI	14
3.1.3	Užitečnost BI nástrojů	15
3.2	Prvky systému BI.....	16
3.2.1	Technická architektura	17
3.2.2	Datová architektura	19
3.3	Produkční systémy	23
3.4	ETL	24
3.5	Datové sklady	25
3.5.1	Operační datový sklad	27
3.5.2	OLTP a OLAP databáze	27
3.5.2.1	Dimenzionální modelování.....	29
3.5.2.2	Datové modelování	32
3.6	Datové trhy a datový franchising.....	32
3.7	Business Intelligence	34
3.7.1	Reporting	35
3.8	Výběr dostupných BI řešení v ČR	36
3.8.1	IBM.....	36
3.8.2	SAP.....	37
3.8.3	Microsoft	38
3.8.4	Oracle	38
3.8.5	Tableau	39
3.8.6	Qlik.....	39
3.9	Trendy BI.....	40
4	Vlastní práce	42
4.1	Úvod k danému podniku.....	42

4.2	Zhodnocení současného stavu v daném podniku.....	43
4.2.1	Hlavní nedostatky stávajícího řešení - analýzy	44
4.2.2	Hlavní nedostatky stávajícího řešení - infrastruktura	45
4.2.3	Informační a datová architektura	46
4.2.4	Technická a produktová architektura	49
4.3	Definice požadavků na BI řešení	52
4.4	Srovnání BI řešení dostupných na trhu	56
4.5	Výběr BI řešení	65
4.6	Implementace prototypu reportu	67
5	Výsledky a zhodnocení	80
5.1	Zhodnocení implementace	80
5.1.1	Zhodnocení infrastruktury	80
5.1.2	Zhodnocení reportingu	80
6	Závěr	84
7	Seznam použitých zdrojů	86

Seznam obrázků

Obrázek 1 - Vrstvy systému BI	17
Obrázek 2 - Technická architektura BI systému	18
Obrázek 3 –BI datová architektura.....	20
Obrázek 4 - Analytical data architecture	22
Obrázek 5 - Prvky ETL	24
Obrázek 6 - Princip multidimenzionální databáze	29
Obrázek 7 - Schéma hvězdy	30
Obrázek 8 - Schéma sněhové vločky	31
Obrázek 9 - Postup datového franšizingu.....	33
Obrázek 10 - Struktura BI nástrojů	34
Obrázek 11 - Datový tok v dané společnosti.....	43
Obrázek 12 - Datové schéma - ukázka	49
Obrázek 13 - Diagram původní architektury.....	50
Obrázek 14 - Diagram nové architektury	51

Obrázek 15 - Magic Quadrant for Analytics and Business Intelligence Platforms....	57
Obrázek 16 - Power BI desktop	58
Obrázek 17 - Tableau Dashboard	61
Obrázek 18 - Qlik Sense dashboard	63
Obrázek 19 - Struktura tabulek	69
Obrázek 20 - Vstupní data – tabulka faktů.....	70
Obrázek 21 – Vstupní data - dimenzní tabulka	70
Obrázek 22 – Datové schéma #1	72
Obrázek 23 - Datové transformace - příklad #1	73
Obrázek 24 - Datové transformace – příklad #2	74
Obrázek 25 - Vizualizace nového reportu	75
Obrázek 26 - Storyboard reportu.....	76
Obrázek 27 - Výsledný report	77
Obrázek 28- Výsledný report - detail	78

Seznam tabulek

Tabulka 1 - Požadavky a priority	52
Tabulka 2 – Saatyho metoda – srovnávací matice	54
Tabulka 3 - Příklad výpočtu váhy kritéria - Saatyho metoda.....	54
Tabulka 4 – Stanovení vah – Saatyho metoda	55
Tabulka 5 - Power BI specifikace	59
Tabulka 6 - Tableau specifikace.....	61
Tabulka 7 - QlikSense specifikace	63
Tabulka 8 - Zhodnocení řešení.....	65

Seznam použitých zkratk

Termín	Význam
AI	Artificial Intelligence
B2B	Business to business
B2C	Business to customer
BI	Business Intelligence
CRM	Customer Relationship Management
DI	Data Integration
DW	Data Warehouse
EDW	Enterprise Data Warehouse
EIS	Executive Information Systems
ERP	Enterprise Resource Planning
GDPR	General Data Protection Regulation
IoT	Internet of Things
IS	Informační systém
KPI	Key Performance Indicator
MDM	Master Data Management
NLP	Natural language processing
ODS	Operational Data Store
OLAP	Online Analytical Processing
OLTP	Online Transaction Processing
PNL	Profit-and-loss
SaaS	Software as a service
SCM	Supply Chain Management
SOA	System of Analytics
SOI	System of Integration
SOR	System of Record
YTD	Year-to-date

1 Úvod

Současný rychle se rozvíjející svět informačních technologií akcentuje spojení „Business Intelligence“ více, než kdy předtím. Důvodem tohoto faktu není nic jiného než cílení společností na stále vyšší efektivitu svých procesů a tím dosažení také vyššího zisku (Vochozka, a další, 2012). S tímto faktem se neodmyslitelně váže hlavní otázka, a to jak této vyšší efektivity nejlépe dosáhnout. Jednou z odpovědí může být právě implementace Business Intelligence řešení zvyšující efektivitu především analytické a plánovací činnosti těchto společností (Slánský, a další, 2005).

Se zásadními rozhodnutími ve společnosti jsou neodmyslitelně spjaty správné a včasné informace, které se stávají jednou z hlavních konkurenčních výhod dnešního podnikového světa. Přetváření surových dat v uchopitelné informace do reportů pro střední a vyšší management je jedním ze základních principů fungování Business Intelligence systémů. Na základě těchto informací jsou v dynamickém prostředí společnosti tvořeny managementem zásadní rozhodnutí, které mají mnohdy promptní vliv do výsledků dané společnosti. Nejen z pohledu finančního řízení je nezbytnou nutností mít důvěru v tyto informace. Při nesprávné interpretaci či nízké validitě dat může docházet k rozhodnutím, která vedou k nezanedbatelným škodám v rámci dané společnosti.

Základní otázkou, kterou akcentuje (Laberge, 2012) je prvotní analýza účelu a typu Business Intelligence, což znamená přesně definovat podnikové procesy, metody analýz a zaměření souvisejícího reportingu postaveného na multidimenzionálním pohledu na podniková data (Slánský, a další, 2005). V případě této práce je toto zaměření reportingu bráno s patřičným důrazem na finanční řízení dané společnosti.

2 Cíl práce a metodika

2.1 Cíl práce

Diplomová práce je tematicky zaměřena na problematiku Business Intelligence. Hlavním cílem práce je analýza BI řešení s návrhem vhodného řešení v konkrétní společnosti se zaměřením na finanční řízení a reporting. Dílčími cíli práce je analýza stávajícího stavu ve vybrané společnosti, vypracování přehledu možných řešení Business Intelligence, popis aplikace vhodného řešení a zhodnocení přínosu pro danou společnost se zaměřením na finanční řízení.

2.2 Metodika

Práce je rozdělena na dvě hlavní části – teoretickou a praktickou. První teoretická část práce byla vypracována na základě studia odborné literatury a analýzy sekundárních zdrojů. Byl zde obecně proveden úvod do problematiky BI, včetně důvodů proč by společnost měla tomuto segmentu IT věnovat patřičnou pozornost. Dále zde byly popsány jednotlivé prvky BI řešení a související infrastruktury. Nakonec byly uvedeny aktuálně běžná BI řešení na trhu v ČR spolu s aktuální trendy, které se v tomto odvětví vyskytují. V rámci praktické části byl zanalyzován současný stav BI řešení v podniku a byly uvedeny požadavky na zlepšení v rámci implementace BI řešení nového. Z této analýzy potřeb firmy vzešlo několik oblastí zájmu, kterým byly Saatyho metodou párového porovnání přiřazeny jednotlivé váhy. Následně byly dle metodiky společnosti Gartner zvoleny 3 potenciální BI řešení označované jako „lídři“, tak aby byly vzápětí tyto řešení ohodnoceny bodovací metodou a váženým součtem s danými vahami bylo zvoleno nejvhodnější z nich. V následující části byla provedena implementace prototypu manažerského finančního reportu zisku a ztráty v prostředí nového BI řešení a zhodnoceny jeho hlavní přínosy.

3 Teoretická východiska

3.1 Business intelligence

V následujících odstavcích bude věnována pozornost samotnému přiblížení tématu Business Intelligence, tak jak ho vnímají jednotliví autoři a odborníci v daném oboru. Z nalezených podkladů je patrné, že o aktuálnosti tohoto tématu nemůže být sporu. Jedná se o jedno z nejvíce se rozvíjejících odvětví zabývajících se podnikovými systémy poslední dekády. Zásadní proměnou prošla forma BI řešení, kdy se již nejedná o plně unikátní na míru řešení od počátku do konce, ale velká část vychází z již předdefinovaných řešení přizpůsobené zákazníkům v daném segmentu trhu – bankovníctví, pojišťovnictví, telekomunikace a další. (Slánský, a další, 2005) Současně můžeme být svědky posunu ve zpracování dat a výsledném reportingu, kdy z velmi sofistikované branže, kde byla vždy nutná značná a neustálá podpora IT oddělení, jsme ve fázi tzv. self-service, kdy běžný uživatel je schopen běžné analytické úkony provést poměrně intuitivně sám (Ironside Group, 2020). Aktuální trendy v BI jsou blíže specifikovány dále v této práci, následují sekce zabývající se základními parametry běžného BI řešení.

3.1.1 Definice BI

V rámci této práce je pohlíženo na Business Intelligence pohledem mimo jiné autora (Sherman, 2014), tedy jako na komplexní systém aplikací a nástrojů využitých ve svém konečném důsledku především v řízení a rozhodování managementu dané společnosti. V rámci praktické části autor akcentuje využití BI systémů pro především analytickou činnost vedoucí k manažerskému rozhodování a řízení společnosti nikoliv pouze pro základní zpracování dat.

3.1.2 Historie BI

První technické počátky BI systémů se datují do 50. let minulého století, kdy v článku „A Business Intelligence System“ publikovaném IBM analytikem Hans Peter Luhnem, se tento zabývá potenciálem právě BI řešení v různých odvětvích průmyslu pro především

zpracování a zpřehlednění dat podniků či samospráv. Jeho výzkumné metody vedly ve společnosti IBM k vývoji prvních analytických systémů a je mu dnes přezdíváno jako „Otec Business Intelligence“ (Heinze, 2020).

Období jež bývá nazýváno Business Intelligence 1.0 se nachází ke konci 80. let minulého století a prostupuje celými 90. léty až do počátečních let nového tisíciletí. Na počátku tohoto období v roce 1989 se velkou mírou o propagaci BI zasloužila především společnost Gartner. (Slánský, a další, 2005) doplňuje, že již této době byla struktura postavena na ETL nástrojích, jež zásobily zpracovanými daty ze zdrojových systémů datový sklad či datové tržiště následovaná OLAP aplikací či dokonce prezentační vrstvou. Autor též zmiňuje rozvoj především v 90. letech, a to v oblasti webových aplikací či dolování dat, kdy první oblast ovlivnila právě prezentační schopnosti výstupů a dolování dat zase naopak analytickou část BI řešení.

Nastupuje období označované Business Intelligence 2.0, kdy je patrné zaměření tvůrců BI řešení na operační datové sklady (angl. zkr. ODS) a řešení pro online čerpání a dotazování nad daty při vyšším důrazu na kvalitu dat. Dále poté se jedná o směřování vývojářů k jednotné platformě, která obsahuje všechny BI komponenty dané architektury, či rozsáhlá cloudová a mobilní řešení (Heinze, 2020). Předchozí autor se spolu s dalšími (SystemOnline, 2018) a (Sherman, 2014) se shodují mimo jiné na tzv. komoditizaci BI, jež znamená typizace základních funkcionalit systému pro určitý obor a jejich případná následná úprava na míru zákazníkovi. Více o posledních trendech v BI oboru je zmíněno dále v této práci v sekci Trendy.

3.1.3 Užitečnost BI nástrojů

Efektivní systém Business Intelligence definuje (Laberge, 2012) čtyřmi základními pilíři. Takovéto systémy, respektive jejich datové výstupy mají atributy:

- 1) přesné
- 2) aktuální
- 3) vysoká přidaná hodnota
- 4) umožňující správné rozhodování.

Takovýmto datům lze jinak řečeno plně důvěřovat, jsou na pravidelné bázi aktualizovány, mají vysokou hodnotu pro své uživatele a napomáhají svým uživatelům provádět často nelehká rozhodnutí v rámci své kompetence ve společnosti.

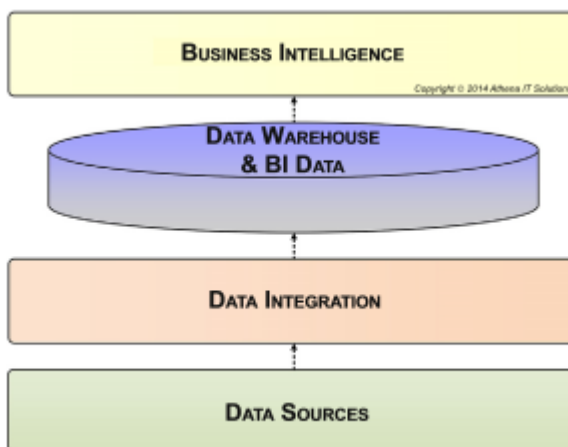
Řešení datového skladu, potažmo BI má dle (Laberge, 2012) nezpochybnitelný vliv na centralizovanost a lepší interpretovatelnost dat. Data z různých systémů mohou používat pro stejný parametr, například výrobku, jinou terminologii a vytvořit z takovýchto dat konsolidovaný report obsahující detailní informace jak o skladu, prodeji, forecastingu výkonů nebo analýz zákaznické spektra je poté téměř nemožné. Pokud k tomuto přidáme nové technologie jako jsou data ze sociálních sítí nebo IoT, stává se integrita dat pro takovýto podnik nutností.

(Laberge, 2012) uvádí, že přes nesporné výhody budování datového skladu a souvisejícího BI řešení jako je především vysoká integrita a spolehlivost dat, je poměrně těžké spočítat návratnost tohoto řešení a je třeba dívat se na takovýto projekt ze širší především strategické perspektivy podniku jako celku. Finální cíl jež zmiňuje současné i (Sherman, 2014) je obecně vyšší konkurenceschopnost podniku, avšak je vždy třeba před implementací tohoto řešení vytvořit validní business case s jasně specifikovanými přínosy BI řešení pro daný podnik.

3.2 Prvky systému BI

Základními prvky související obecně s BI řešením v širším pojetí jsou následující:

Obrázek 1 - Vrstvy systému BI



Zdroj: (Sherman, 2014)

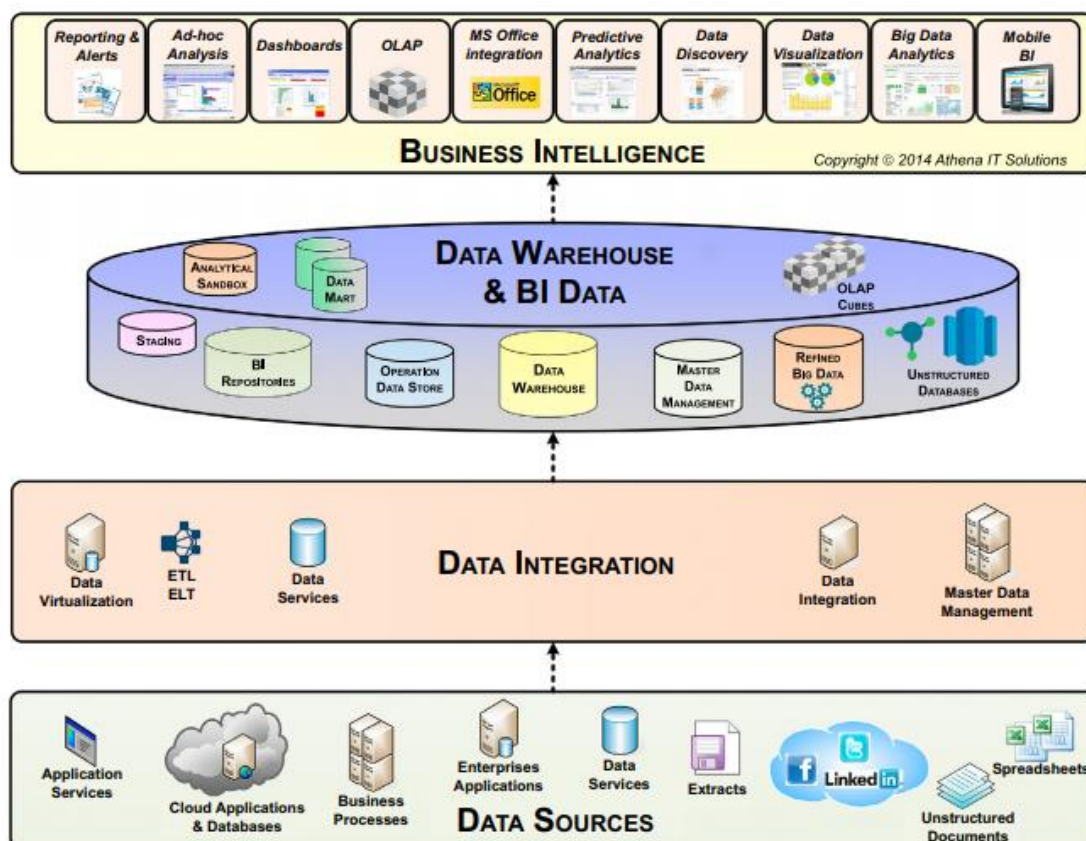
Jednotlivé části jsou do detailu dále rozvedeny v následujících kapitolách. V této chvíli je nutné zmínit rozdílný pohled na pojetí Business Intelligence jako takové. Někteří autoři jako je (Sherman, 2014) považují za Business Intelligence pouze analytickou aplikaci celé architektury podnikového systému. Toto znamená část čerpající data například z datových tržišť, OLAP kostek, ODS a jiných. V této fázi jsou již data konzistentní, agregována do podoby dle business požadavků uživatelů výstupních reportů. Naopak autoři jako je (Slánský, a další, 2005) pohlíží na BI systém jako na celé soukolí podnikového systému fungující od zpracování produkčních dat ETL systémy, dále uložení dat a následně až po výsledný reporting a analýzy. V rámci teoretické části této práce jsou uvedeny základní charakteristiky z komplexního pohledu na BI, avšak v rámci praktické části je již z většiny zaměřena na koncovou analytickou aplikaci BI řešení.

3.2.1 Technická architektura

Technická architektura jako taková je popisována jako souhrn technologií umožňující implementaci BI systému a posléze zajišťující jeho bezproblémový chod v rámci jeho životního cyklu. Obecně je nutné přihlídnout k faktu, že každý BI systém a jeho prvky je možné do velké míry upravit dle představ zákazníka odpovídajícím jeho specifickým

požadavkům. V diagramu níže rozlišuje (Sherman, 2014) základní 4 vrstvy technické infrastruktury zajišťující fungování BI řešení.

Obrázek 2 - Technická architektura BI systému



Zdroj: (Sherman, 2014)

První část nejnižší v diagramu obsahuje tu část infrastruktury kombinující z různých podnikových a externích zdrojů data a provádějící jejich transformaci do jednotné formy, tak aby bylo možné s nimi dále pracovat. Při této transformaci je maximální snaha o čistotu a konzistentnost dat. Tento aspekt zdůrazňuje též (Laberge, 2012), kdy je kvalita dat dle jeho názoru stěžejní.

V druhé fázi následuje uložení takto zregrovaných dat způsobem, kdy je zajištěna maximální optimalizace pro další přístup analytických BI aplikací, které s těmito daty dále pracují. Běžnými znaky těchto uložení, které uvádí (Sherman, 2014), je integrovanost dat, uchování historických dat a skutečnost, že data v uložení jsou read-only, tedy uživatelé je

neupravují. (Laberge, 2012) k tomuto dodává, že právě centralizovanost dat a jejich kvalita dodává dalšímu reportingu a analýzám kýženou spolehlivost.

Třetí částí tohoto soukolí je Business Intelligence, tak jak je povětšinou chápáno v současné době. BI má v této definici především za cíl prezentovat a analyzovat informace, která integrovaná data nabízejí, jejich koncovým uživatelům. Využívá k tomu různé typy analytik, a to i mimo jiné prediktivní formy, kdy za pomoci rozšířené inteligence může uživatel predikovat například budoucí prodeje či stav skladu.

V následujících několika kapitolách se bude práce více dopodrobna zabývat všemi 4 oblastmi, které s BI souvisejí jak bylo uvedeno v předchozím diagramu.

3.2.2 Datová architektura

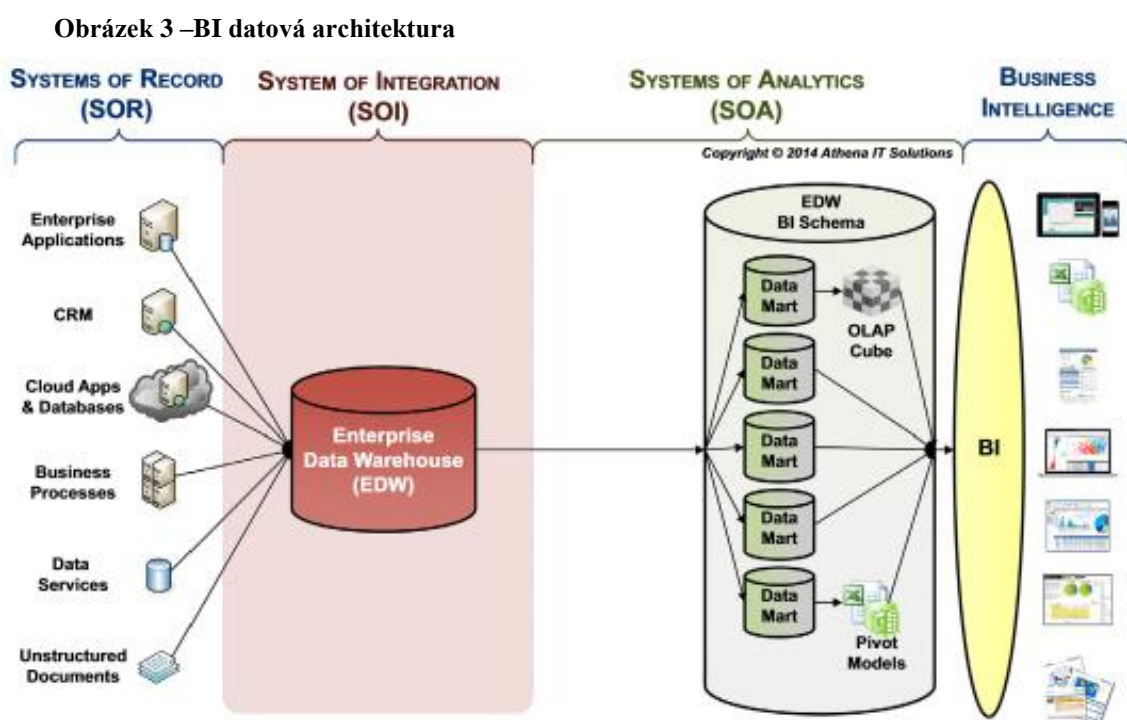
Veskrze všichni analyzovaní autoři se shodují, že pokud je pohlíženo na systém BI v širším pojetí, jsou hlavním základním kamenem samotná podniková data. Tyto data označuje (Laberge, 2012) jako nejdůležitější novodobé podnikové aktivum. V této problematice zmiňuje Sherman (2014) přístup zvaný „5C“

- 1) Čistá (*angl. Clean*) – chybějící či chybné záznamy vedou k poškození integrity dat a zásadně snižují možnost následných analýz.
- 2) Bezesporná (*angl. Consistent*) – data musí být důvěryhodná, všichni uživatelé souvisejících výstupů musí mít důvěru v tyto hodnoty.
- 3) Vyhovující (*angl. Conformed*) – data by měly být analyzována a reportována ve srozumitelné formě pro jejich další uživatele.
- 4) Aktuální (*angl. Current*) – musí být zaručena co nejvyšší aktuálnost dat v systému BI.
- 5) Komplexní (*angl. Comprehensive*) – uživatelé by měli mít k dispozici maximum možných dat a informací aby mohli na základě souvisejících analýz provést kvalifikované rozhodnutí.

S touto definicí se z většiny shodují s Laberge (2012), který taktéž mimo jiné zdůrazňuje vysokou spolehlivost dat, tedy jinak řečeno integritu. Aspekty, které autory k tomuto zjištění autory vedou jsou vesměs nasnadě. Data v systému BI si musíme představit jako například základy u stavby nového domu. Pokud při tvorbě základů dojde k méně či více zásadním

pochybením, mohou být dopady pouze špatné a to ve větší či menší míře, vždy však k neprospěchu věci.

(Sherman, 2014) obecně definuje datovou architekturu jako několik vzájemně propojených oblastí jak znázorňuje diagram níže:



Zdroj: (Sherman, 2014)

První část nazvaná System of Record (angl. zkr. SOR), jak již název napovídá, zahrnuje především zpracování dat z transakčních systémů, tedy kupříkladu objednávkového či účetního systému. V této části je nejdůležitější přesná zmapovanost a důvěryhodnost zdrojů, které chceme aby vstupovaly do podnikového datového skladu (angl. zkr. EDW).

Následuje System of Integration (angl. zkr. SOI), který sestává právě z EDW. V této části probíhá shromáždění zaznamenaných dat a probíhá jejich transformace dle zásad „5C“, tedy s cílem vytvořit a uchovat mimo jiné konzistentnost, čistotu a aktuálnost dat, kdy nakonec jsou tyto data zde také uskladněny.

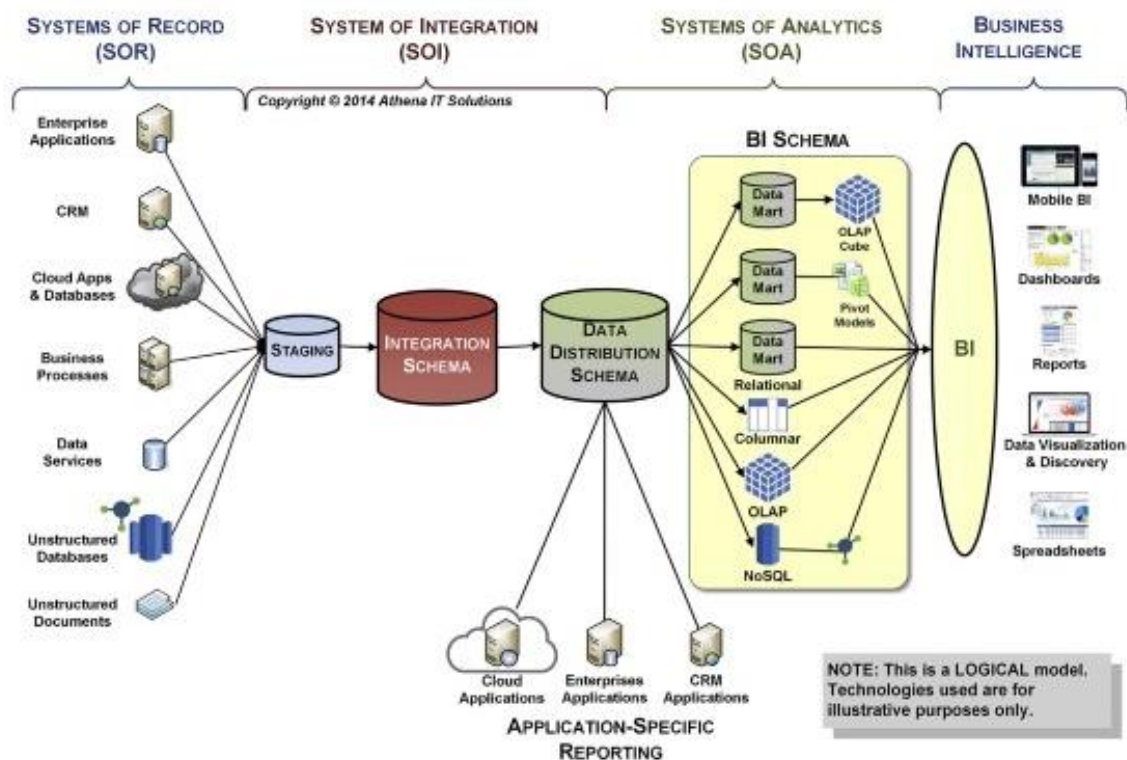
Poslední částí zpracování dat je System of Analytics (angl. zkr. SOA), kde v jednotlivých datových tržištích či OLAP kostkách jsou multidimenzionálně uloženy agregované informace připraveny pro BI analytické nástroje. Dalšími možnými nástroji mohou být ODS či Master Data Management (angl. MDM). Zásadním přínosem do celého systému je zde rychlá responzivita na analytické dotazy do těchto dat bez nutnosti provádět časově náročné přepočty nad zdrojovými (transakčními) daty při každém takovémto dotazu.

V rámci tvorby datové architektury v zásadě (Laberge, 2012) a (Sherman, 2014) vyzdvihují dva základní historické přístupy které vedly ke třetí dnešní nejběžnější formě, jež (Sherman, 2014) nazývá „Analytical data architecture“, či zkratkou „ADA“ a kdy se jedná o kombinace obou přístupů. Jedná se o přístupy, které jsou již desetiletí vzdáleny přístupu, kdy tvořily datovou strukturu pouze jediný EDW či naopak pouze několik samostatně fungujících datových tržišť.

Prvním přístupem je „Enterprise data bus architecture“ od Ralpha Kimballa a tím druhým je „Hub-and-spoke (CIF)“ od Billa Inmona. Základními rozdíly bylo jádro celé architektury, kdy v případě Kimballa se jednalo o datová tržiště a v rámci Inmonova přístupu o EDW. Inmonův přístup je též označován například (Laberge, 2012) jako přístup „shora dolů“, kdežto Kimballův přístup jako „zespodu nahoru“. V prvním případě je brán ohled především na vstupní data ze všech podnikových systémů, z kterých je tvořen datový model. V případě druhém je brán ohled především na požadavky businessu, dle kterých jsou definována jednotlivá datová tržiště. Poměrně zásadním rozdílem byl pohled obou směrů na datovou normalizaci, kdy Inmonův přístup prosazoval nezbytnost implementace tzv. třetí normalizované formy dat (zkr. 3NF), kdežto Kimballův přístup skrze dimenzionální modely pro datová tržiště, ho bere jako postradatelný.

Samotný ADA přístup v sobě kombinuje to nejlepší z dvou výše uvedených přístupů. Diagram níže značí toto schéma:

Obrázek 4 - Analytical data architecture



Zdroj: (Sherman, 2014)

Již výše zmíněná část SOI je v tomto schématu diferencována na několik částí EDW a to část přípravy, část integrační a část distribuční. První část se prvotně snaží o zajištění především referenční integrity, auditu a celkové datové kvality. Integrační část využívá hybridní dimenzionálně normalizovaný datový model v případě, kdy zdrojová data nejsou ve 3NF, kdy může poté využívat data právě v této formě. Funkce této integrační části je především zpracování dat a uchování dat historických. Tento typ modelu umožňuje především efektivní sledování změn v datech a relacemi mezi daty. Distribuční schéma již využívá přístup hybridního dimenzionálního modelu, který je vždy následně transformován pro datové tržiště či kostky OLAP do standardního dimenzionálního modelu. Část SOA, jak již bylo zmíněno, umožňuje především samotným business analytikům využívat BI nástroje pro své analýzy, tak aby na základě nich byly následně přijímána managementem správná rozhodnutí. Jedná se tedy většinou o několik vrstev datových tržišť či kombinace datových

tržišť a OLAP kostek či noSQL databází, které v sobě uchovávají filtrovaná, agregovaná a transformovaná data dle potřeb jednotlivých oddělení a které je možné využívat téměř neomezeně pro jednotlivé analýzy bez nutnosti opakovaně při každé analýze a načtení reportu tyto data znovu agregovat a připravovat do této formy. Bližší detail, mimo jiné ke zmíněné 3NF normalizaci dat a k rozvedení multidimenzionality, je uveden v následujících částech této práce.

3.3 Produkční systémy

Pod produkčními systémy rozumíme soubor transakčních systémů, které společnosti využívají na běžnou agendu. Zdrojovými systémy mohou být mimo jiné podnikové systémy typu:

- 1) ERP
- 2) CRM
- 3) SCM

Bod první se týká v předkladu systému pro podnikového plánování zdrojů, tedy zásoby, nákupy, prodeje, marketing, finance či personalistika. Z druhého bodu se pod zkratkou CRM skrývá řízení vztahů se zákazníky, kdy tento modul může být součástí ERP či se může jednat o samostatnou jednotku. Cílem tohoto modulu je sledovat a udržovat vztah společnosti ke svým zákazníkům formou sledování prodejů jednotlivým podskupinám zákazníků, nákupní trendy jednotlivých skupin či například jaká je schopnost společnosti si daného zákazníka udržet. Poslední bod označuje modul managementu skladu, kdy opět se může jednat jen o součást ERP, ale také o samostatný systém zaznamenávající skladové pohyby. V neposlední řadě se může jednat o data ze sociálních sítí, různých datových služeb, které podnik využívá, například tržních informací, měnových kurzů a mnoha dalších.

Všichni autoři se shodují, že tyto transakční systémy jsou vystavěny formou, kdy podporují především zaznamenání, modifikování a ukládání dat v reálném čase. Tyto systémy nejsou v žádném případě zaměřeny na analytickou funkci. (Laberge, 2012) k tomuto doplňuje bezpodmínečnou nutnost mít vždy kompletně zmapovány všechny zdrojové systémy, které podnik využívá a jasně vydefinované aspekty každého z nich. Zásadním znakem obecně

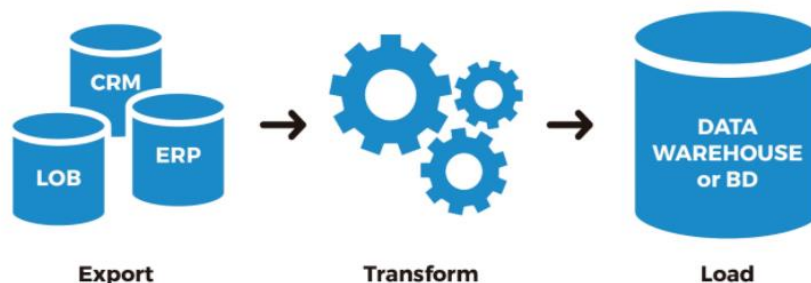
těchto zdrojových systémů je různorodost dat a formy v které jednotlivé systémy informace uchovávají. Úkolem ETL prostředků jež jsou blíže specifikovány dále je transformace a integrace těchto dat ve formě, kdy je zajištěna jejich konzistence a vysoká kvalita (Slánský, a další, 2005)

3.4 ETL

ETL značí anglickou zkratku pro „extract, transform, load“, jež v obecném pojetí se stal synonymem pro datovou integraci a technologii úzce spjatou s datovými sklady. (Laberge, 2012) definuje ETL jako „... extrakci dat ze vstupního zdroje, transformaci těchto dat do příslušného formátu pro načítání a následné načítání dat do cílové databáze.“ . Tuto fázi též jinak nazývá jako fáze pořízení dat (angl. data population).

Jednotlivé stupně této technologie znázorňuje diagram níže:

Obrázek 5 - Prvky ETL



Zdroj: (Bismart, 2020)

V první fázi probíhá extrakce dat pomocí ETL z různorodých systémů podniku. Jak bylo zmíněno v předchozí kapitole, může se jednat o systémy typu ERP, CRM, SCM, IoT, sociální sítě a mnoho dalších zdrojů. V této fázi je těmto různorodým zdrojovým datům dán základní jednotný formát a počíná proces jejich prvotní integrace a čištění.

V navazující fázi číslo dvě se tyto znormalizovaná data transformují takovým způsobem jakým máme vydefinovanou strukturu našeho datového skladu. Zde se jedná již

o práci především vedoucí k dalšímu ošetření dat od duplikátů či jinak přebytečných a nežádoucích dat a jejich následnou agregaci.

V poslední fázi načítání dat přichází na řadu vyčištěná a agregovaná data uložit do datového skladu.

(Laberge, 2012) zmiňuje též další navazující fázi, označovanou jako tzv. distribuce dat. Jedná o proces, kde z centrálního uložení, s daty ve třetí normalizované formě, se data transformují do datových trhů, které mohou být běžně ve formátu tzv. hvězdice či sněhové vločky. Jedná se o poslední krok pro plnou dostupnost agregovaných a čistých dat pro následné analýzy BI aplikacemi. Bližší popis k datovým skladům a datovým trhům je uveden v následujících kapitolách včetně definice výše zmíněných formátů hvězdice a sněhové vločky.

(Sherman, 2014) a (Laberge, 2012) shodně zdůrazňují nutnost odpoutat se od původního vnímání ETL, jako systému jež v rámci dávek (angl. batch) zpracovává většinou přes noc data z podnikových systémů a transformuje je do centrálního datového skladu. Novodobá ETL řešení se přizpůsobila zvyšujícím se datovým objemům, rozmanitosti a rychlosti. (Sherman, 2014) takováto řešení již přímo nenazývá ETL, ale datové integrační služby, které mají komponenty s původem v ETL. Rozdílem je zde mimo jiné vnímání komponent ETL procesu, kdy novodobý přístup mimo jiné zahrnuje do procesu zmiňovanou distribuci dat skrze datové trhy.

3.5 Datové sklady

(Laberge, 2012) definuje datový sklad následujícím způsobem: „Datový sklad (data warehouse) je systém, který umožňuje shromažďovat, organizovat, uchovávat a sdílet historická data.“

Základem datových skladů jsou stále relační databáze, avšak jak uvádí (Sherman, 2014), přicházející nové technologie tuto strukturu nenahrazují, ale jsou spíše jejími komplementy. Zároveň se vyznačují vyšším zaměřením na analytické schopnosti daného řešení. Mezi tyto nové prostředky patří technologie jako OLAP databáze, MPP databáze, datová virtualizace, in-database analytika, in-memory analytika a další. Datové sklady, narozdíl od například datových tržišť, jsou robustním řešením, které má za úkol

shromažďovat veškerá podniková data při zachování jejich konzistence a ukládat jejich dalekosáhlou historii v denormalizované formě, jako je hvězdicové schéma (angl. Star scheme) či schéma sněhové vločky (angl. Snowflake scheme). Dalším souvisejícím pojmem jsou takzvaná datová jezera (angl. Data lakes), které narozdíl od datových skladů slouží povětšinou k uchování nestrukturovaných, či částečně strukturovaných dat z transakčních systémů, jež daná společnost předpokládá, že v budoucnu využije k analytickým účelům (Amazon, 2020).

Jak (Laberge, 2012) zmiňuje nové technologie umožňují rozšířené možnosti indexace, komprese, in-memory a in-database procesování nebo technologii ROLAP.

Základními vlastnostmi datového skladu jsou dle (Sherman, 2014) vrstvy, či schémata – integrační a distribuční. Integrační schéma má za úkol datovou integraci a uchování dat v uložišti při dodržení u dat již dříve zmíněného principu 5C. Distribuční funkce dále odesílá data většinou do datových tržišť a umožňuje odsud čerpat tyto data BI aplikacemi.

(Laberge, 2012) rozlišuje tyto dvě zmíněné vrstvy ještě podrobněji a přidává též navazující analytické vrstvy a to:

- 1) Vrstva získávání dat
- 2) Vrstva centralizovaných dat
- 3) Vrstva distribuce dat
- 4) Výkonnostní vrstva
- 5) Vrstva uživatelské prezentace

V tomto schématu je ve vrstvě získávání dat přítomen prvek ETL viz předchozí kapitola. Navazující vrstva centralizovaných dat obsahuje samotné centrální úložiště datového skladu, obsahující veškerá podniková data. Vrstva distribuce dat má obdobnou strukturu jako vrstva získávání dat. Následuje výkonnostní vrstva, zahrnující oblast datových trhů či obdobných struktur typu OLAP kostky apod. V poslední řadě je zmíněna vrstva uživatelské prezentace, jež obsahuje povětšinou BI vizualizační nástroj a slouží samotným podnikovým uživatelům. Poměrně novým řešením je posun datového skladu do cloudového prostředí, což poskytuje mimo jiné společnosti úspory v té formě, že nemusí sama řešit výstavbu své vlastní hardwarové a softwarové infrastruktury a také nabízí vyšší flexibilitu, co se týče výpočetního výkonu a mnoha dalších aspektů. (Oracle, 2020).

3.5.1 Operační datový sklad

Běžné označení je zde pro tento prvek BI struktury anglická zkratka ODS (angl. Operation Data Store). Jedním z nejčastějších důvodů implementace ODS do datové infrastruktury podniku, bývá nedostatečná kapacita zpracování datového objemu datovým skladem případně nedostatečná frekvence aktualizace dat, kterou vyžaduje operativní reporting.

(Sherman, 2014) uvádí nepřenosť využívání označení ODS, které se často zjednodušeně užívá pro vše co není datový sklad či datové tržiště. Pro upřesnění uvádí (Sherman, 2014) následující znaky ODS:

- 1) integruje data z různých zdrojů
- 2) umožňuje tvořit konsolidované reporty
- 3) obsahuje přednastavený reporting či integrační řešení pro podnikový systém
- 4) používá stejné datové schéma jako zdrojové transakční systémy

Primární funkcí ODS byla vždy kompilace dat z různorodých zdrojů a to v reálném čase. Jedná se o nástroj operativního reportingu, který se nezabývá transformováním dat a jejich čištěním. ODS pracuje s operativními daty a nezabývá se uchováváním dat historických. (Sherman, 2014) zmiňuje spíše historickou praxi, kdy ODS bylo často používáno pouze v kombinaci s EDW. V rámci BI bylo možno získat data jak z EDW tak z ODS, či jen z jednoho systému, což bohužel hlavně v rámci ODS vedlo k získání často nekonzistentních dat z důvodu jen základní či žádné úpravy a transformovanosti dat. Novodobá praxe je kombinovat ODS a EDW s datovými tržišti, což vyhovuje i novodobým náročným BI aplikacím. ODS má často úzké businessové zaměření jako je finanční plánování, kdy ODS může BI aplikacím zabývajících se právě tímto aspektem businessu, poskytnou data, které získá z datového skladu, ale zároveň po jejich zpracování je může do ODS vrátit a ODS je rozdistribuuje zpět do datových skladišť, kde zůstanou uloženy.

3.5.2 OLTP a OLAP databáze

Jednou z možností jak zachovat požadovanou vysokou granularitu dat, a tedy dodržet zachování vysoké podrobnosti v nich, může být úložiště provozních dat. (angl. „Operational

data store“), tak jak bylo definováno v minulé kapitole. Možnou nevýhodou může být povětšinou pouze krátkodobé uložení dat v takovémto uložišti, a je tedy často preferován přístup, kdy jsou i tyto data ukládána v rámci samotného BI řešení. (Laberge, 2012)

Na tomto místě je nasnadě připomenout základní rozdělení typů informací s kterými dnešní společnosti v rámci svého informačního systému pracují. Jedná se o rozdělení na:

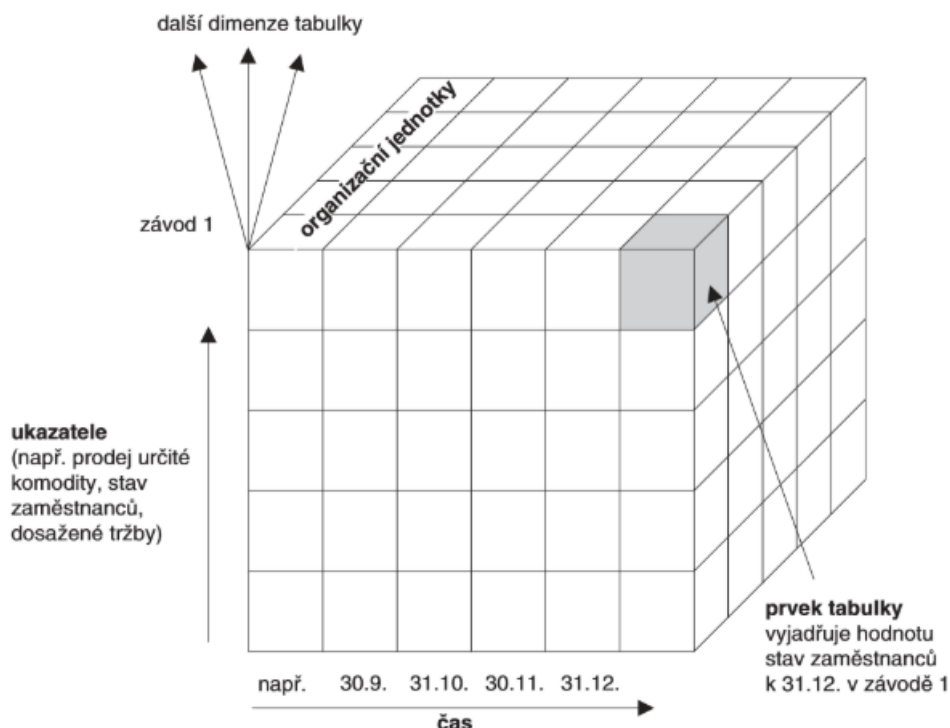
- 1) operativní informace
- 2) analytické informace

Operativní informace můžeme definovat jako běžné informace o každodenních transakcích, které se ve společnosti uskutečňují. Příkladem takovýchto transakcí mohou být transakce účetní. Zpracování těchto informací v reálném čase je jednou ze základních úloh OLTP systémů. Jak doplňuje (Slánský, a další, 2005) tyto data jsou obecně označována jako produkční, primární či zdrojová.

Informace analytické, dle (Slánský, a další, 2005) pracují právě s informací z bodu 1). S tímto typem informací jsou neodlučitelně spjaty OLAP systémy. (Slánský, a další, 2005) definuje OLAP jako „informační technologii založenou na koncepci multidimenzionálních databází, kdy hlavním principem je několikadimenzionální tabulka umožňující pružně měnit jednotlivé dimenze a měnit tak pohledy uživatele na modelovanou ekonomickou realitu.“ Princip multidimenzionality je u OLAP stěžejní, kdy data jsou uložena v multidimenzionálních databázích. Tyto databáze jsou dále specifické dle (Slánský, a další, 2005) víceúrovňovou agregací dat, možností časového srovnávání dat či tvorby predikcí. Z výše řečeného je patrné, že výstupní funkcí OLAP systémů je dále zmiňovaný reporting a s ním pevně spojené řízení společnosti jako takové.

OLTP databáze (transakční) neumožňuje multidimenzionální ani agregovaný pohled na produkční data, což je pro jejich analýzu základní předpoklad. Na druhou stranu OLAP databáze právě tuto základní multidimenzionální funkcionalitu zvládají a je to dokonce jejich již výše zmíněná primární funkce, kdy uživatel má možnost měnit jednotlivé dimenze a tím i pohled na tyto agregovaná analytická data (Slánský, a další, 2005).

Obrázek 6 - Princip multidimenzionální databáze



Zdroj: (Slánský, a další, 2005)

Na obrázku výše jsou patrné základní dimenze, kterými jsou čas a ekonomické ukazatele. Jednotlivé dimenze obsahují takzvané prvky dimenzí – například pro proměnnou čas to mohou být jednotlivé dny. Protnutí jednotlivých dimenzí dostáváme prvek multidimenzionální databáze. Velmi častým jevem je rozřazení prvků dimenze do skupin a podskupin v rámci vydefinované hierarchické struktury.

Hierarchická struktura ve spojení s agregační funkcí OLAP databáze umožňuje v předstihu napočítávat hodnoty v těchto agregacích a při analytickém využití dat již není nutné znovu hodnoty pro jednotlivé agregace napočítávat. (Slánský, a další, 2005)

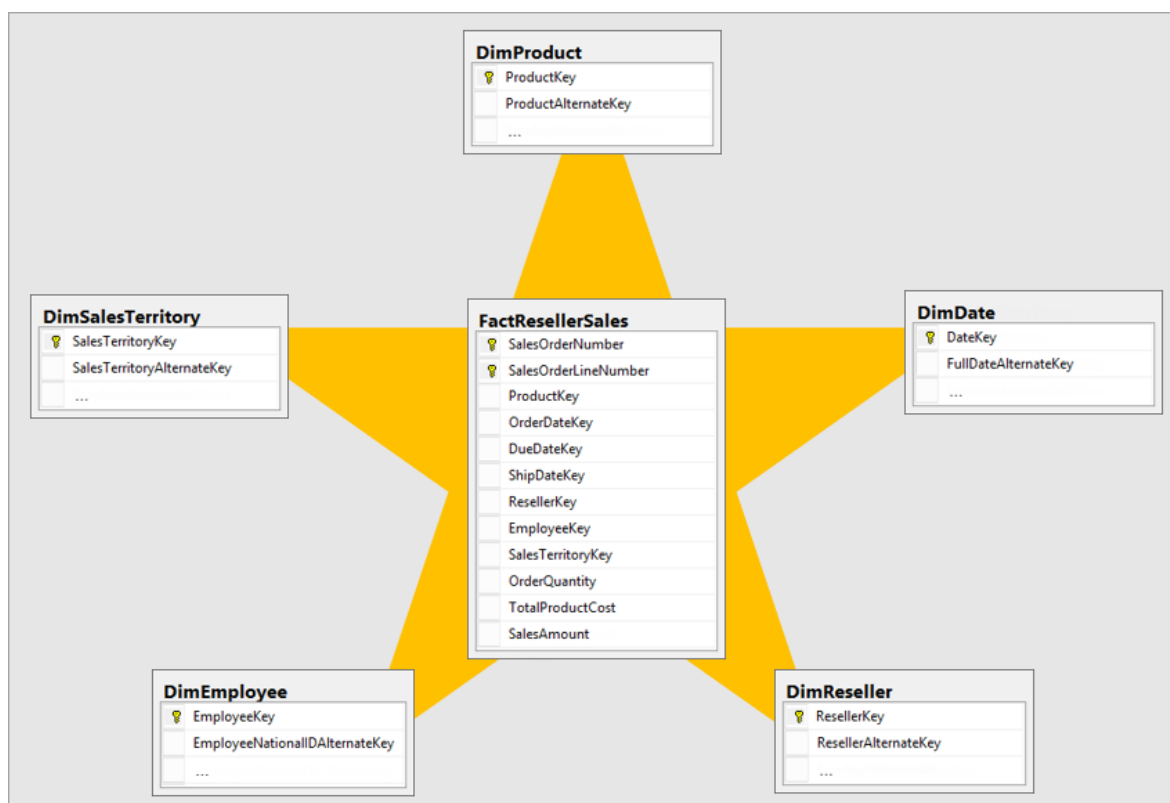
3.5.2.1 Dimenzionální modelování

V rámci relační databáze rozeznáváme 2 přístupy k budování datového modelu a práci s multidimenzionalitou.

- 1) schéma hvězdy (angl. „star scheme“)
- 2) schéma sněhové vločky (angl. „snowflake scheme“)

Schéma hvězdy je postaveno na 2 typech tabulek. Prvním typem jsou tabulky dimenzí a druhým jsou tabulky faktů. Dimenzní tabulky popisují jednotlivé sekce daného segmentu podnikání a její součástí bývá tabulka obsahující data o produktech, tabulka obsahující data o zaměstnancích či tabulka obsahující definice časového období. Na druhou stranu tabulky faktů obsahují události či pozorování, které nastaly a souvisí věcně právě s definovaným segmentem podnikání viz dimenzní tabulky. Může se jednat o uskutečněné objednávky, teploty, vyrobené výrobky a nespočet dalších událostí, které ovlivňují daný business a ve výsledku je budou business uživatelé vyžadovat v agregované a sumarizované formě dle definic v dimenzních tabulkách. Rozlišení tabulky dimenzní a tabulky faktů je stanoveno na základě nastavení vztahové kardinality, kdy u vztahu „one-to-many“ vždy dimenzní tabulka má na své straně „one“. Toto schéma je zobrazeno v diagramu níže, kdy uprostřed se jedná o tabulku faktů:

Obrázek 7 - Schéma hvězdy

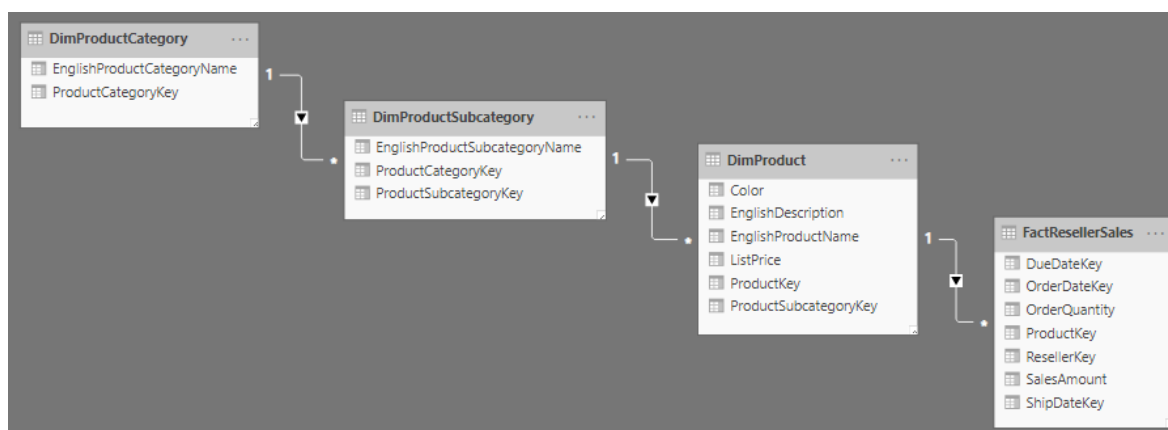


Zdroj: (Microsoft Corporation, 2020)

Tabulky faktů jsou spojeny s tabulkami dimenzními skrze klíčové dimenzní sloupce, kdy tyto sloupce určují též dimenzionalitu tabulek faktů – mohou jimi být například sloupec s číselným označením produktu a druhý sloupec označující časový úsek, čímž se z tabulky stává stabilka se dvěma dimenzemi – dimenze produkt a dimenze čas, tedy granularita daných dat. (Microsoft Corporation, 2020)

Schéma sněžové vločky je v mnohém podobné schématu hvězdy, avšak zásadní rozdíl je v pohledu na dimenzní tabulky. Dimenzní tabulky jsou v tomto schématu dále rozlišovány sub-dimenzními tabulkami a je tak tvořena podrobnější a více rozvětvená struktura zprostředkující agregovaný náhled na data z tabulky faktů. Využití této rozvětvené struktury vyzvihuje (Oracle, 2020) například u společností, které mají mnoho velmi specifických produktů, které nemají mnoho společných atributů a proto pro přehlednost a lepší práci s agregací těchto dat je výhodné produkty rozlišit do takovýchto sub-tabulek. Diagram níže zobrazuje jak může vypadat právě rozvětvení jedné z dimenzních tabulek obsahující kategorii produktu.

Obrázek 8 - Schéma sněžové vločky



Zdroj: (Microsoft Corporation, 2020)

3.5.2.2 Datové modelování

Obecně bývá datové modelování nazýváno EDM (angl. Enterprise data model). Základními cíli EDM dle (Laberge, 2012) je analýza a modelování co největšího množství dat na kterém lze vytvořit a následně pracovat se standardizovanými datovými strukturami ustálit mimo jiné podnikovou terminologii. EDM má výhody především v omezení redundance a komplikovanosti dat, kdy při kategorizaci jednotlivých oblastí činností podniku je v rámci této kategorizace vnesen mezi data určitý řád. Datové modelování a datové profilování jde spolu ruku v ruce při práci na přípravě dat pro BI využití. Datové profilování se zaměřuje na zdrojová data a jejich porozumění a datové modelování se naopak zaměřuje na definici cílové podoby do které data v rámci celého procesu chceme transformovat. (Sherman, 2014)

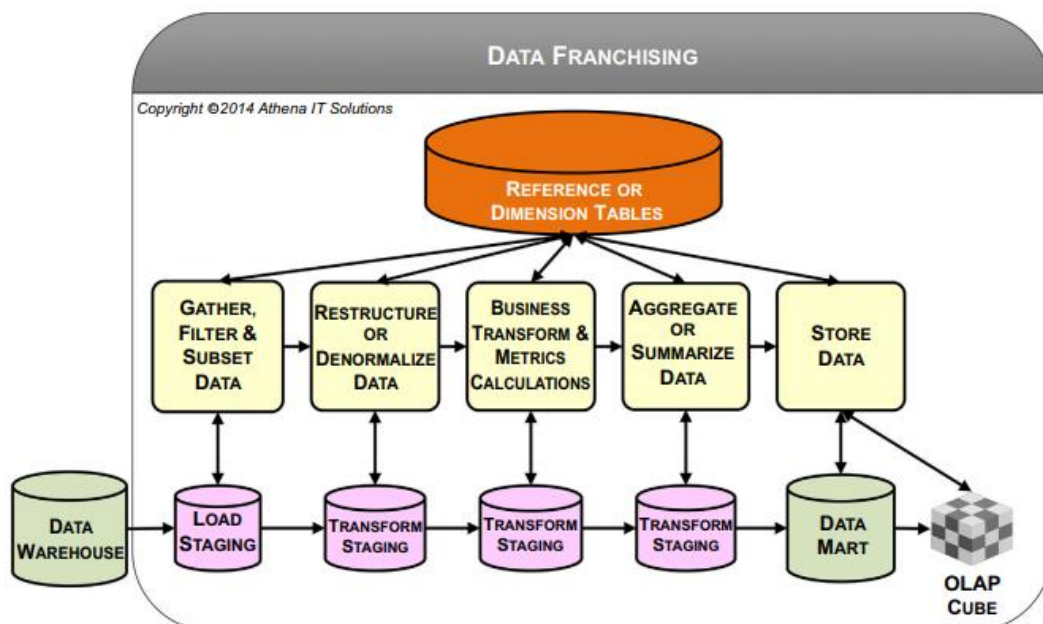
S datovým modelování neodmyslitelně souvisejí tzv. metadata. Metadata obsahují několik subkategorií jako jsou – datové definice, ETL datové mapování či metadata BI aplikací. S těmito kategoriemi souvisí nutnost patřičné dokumentace metadat aby business uživatelé měli možnost získat informaci odkud se daná finální data berou, jak byly transformovány a jak je tedy možno využít. Z druhého pohledu je takováto dokumentace nepostradatelná pro IT profesionály spravující celý systém a provádějící jeho údržbu a úpravy v něm. Pokud se zmiňujeme o metadatach, je nutné též zmínit tzv. Master data management (angl. zkr. MDM). Jedná se základní seznamy s primárními referenčními daty podniku – obsahují kupříkladu seznamy zákazníků, dodavatelů, zaměstnanců a mnoho dalších seznamů z celé oblasti daného podniku. Jedná se o základní pilíř podnikového managementu informací jako takových. (Sherman, 2014)

3.6 Datové trhy a datový franchising

Podnikový datový sklad poskytuje v rámci své distribuční funkce data do datových tržišť, jež ve své struktuře datových tržišť a sub-datových tržišť mají za cíl podporovat analytické business prvky, tedy následné BI aplikace. Důležitým krokem pro zvýšení relevantnosti výstupních dat pro samotnou BI analytickou část je krok datového franšizingu. V tomto kroku jsou data z datového skladu transformována do formy dle analytických požadavků divizí a to v několika krocích. Připomeňme pouze, že tyto data by již měla

splňovat přístup „5C“ a tedy měla by se vyznačovat konzistentností, čistotou a vysokou kvalitou. Tyto kroky obsahují získání dat z datového skladu, následné provedení restrukturalizace a transformace, aby následně byla tato data agregována a sumarizována do podoby co nejvíce analyticky využitelných dat businesssem a také rychlejší responzivitě na dotazy BI analytickými nástroji. Tyto výstupní data jsou uložena právě v datovém tržišti či OLAP kostce, jež jsou většinou specificky vytvořeny pro daný segment businessu, například finanční řízení. Tato transformace je obecně níže znázorněna v diagramu:

Obrázek 9 - Postup datového franšizingu



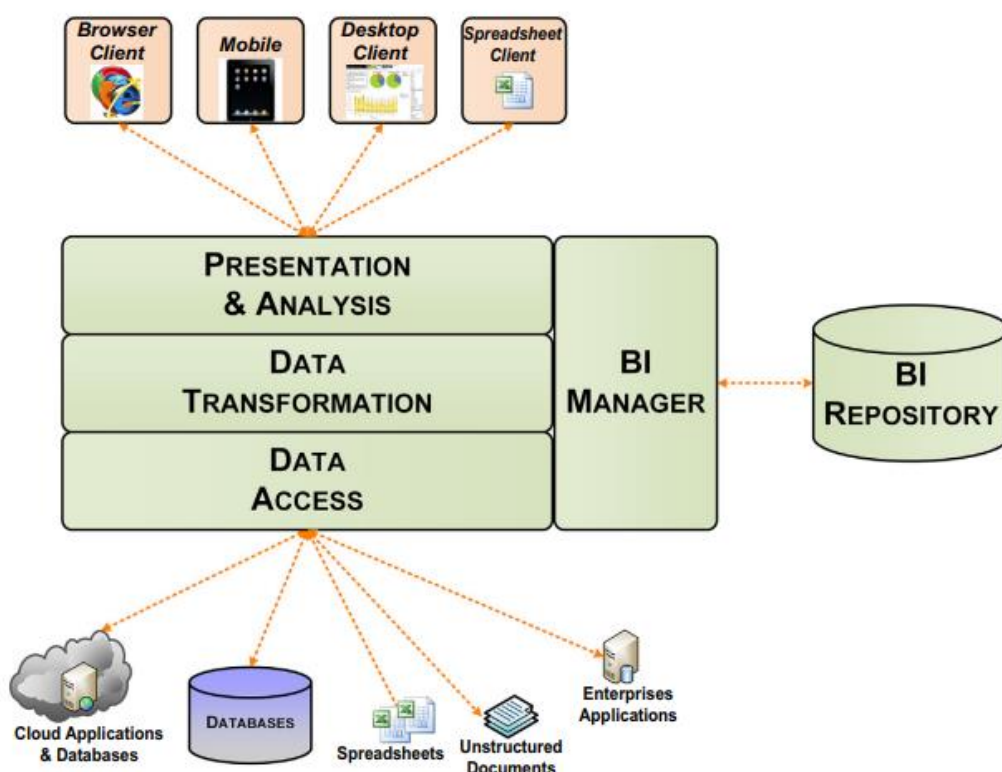
Zdroj: (Sherman, 2014)

Hlavním cílem této datové transformace je zvýšení využitelnosti těchto dat pro další analýzy. V rámci tohoto procesu mohou být přidána též externí data z externích zdrojů. Příkladem který uvádí (Sherman, 2014), může být externí rating společností, kdy můžeme k našim kupříkladu datům z faktur přiřadit potenciál včasné splatnosti dle souvisejícího ratingu dané společnosti.

3.7 Business Intelligence

Standardní rolí BI aplikací je zvolit, načerpat a transformovat data z infrastruktury uložení a poskytnout tyto data v žádané formě BI front-end aplikacím pro další analýzy a vizualizace. Níže je uveden diagram obecné struktury BI řešení jako takového.

Obrázek 10 - Struktura BI nástrojů



Zdroj: (Sherman, 2014)

Jak je ze schématu patrné, v první fázi BI aplikace vstupuje do infrastruktury uložení – například datového tržiště a tyto data odsud převezme do své paměti. Následuje datová transformace, kdy je zároveň z načerpaných dat povětšinou pomocí SQL dotazů zvolen užší výběr dat pro další analýzy. Fáze prezentace a analýz je finální vrstva BI systému, s kterou interaguje koncový uživatel, jež využívá a tvoří vizualizace a reporty.

Za zmínku zde stojí dosud nezmíněné tzv. BI Repository – jedná se o operativní databázi, ukládající data související s analýzami jako jsou reporty, vizuály, workflow, kontrolní listy přístupů a mnoho dalších provozních dat BI systému.

Jak zmiňuje (Sherman, 2014), BI aplikace jsou jedinou součástí celé informační a technické infrastruktury s kterou přichází koncový uživatel do styku. (Sherman, 2014) zde naráží na fakt, že právě tato skutečnost vede v obecném povědomí k situacím, kdy celá informační struktura v podniku související s BI aplikacemi, včetně ETL a souvisejících součástí infrastruktury je označována právě názvem koncové BI analytické aplikace.

3.7.1 Reporting

Z pohledu reportingu je nutné se především zaměřit na jeho smysl a význam, jaký bude mít pro své koncové uživatele. Ve svém důsledku nahlížíme na tyto koncové uživatele jako dvě samostatné skupiny. První skupinou jsou takzvaní interní uživatelé, kdy se jedná především o management podniku na všech stupních řízení. Tento reporting bývá poměrně robustní a měl by mít svá specifika, přesně definovaná výkonem funkce svých uživatelů v rámci managementu dané společnosti. Hlavním cílem interního reportingu je zajištění informací pro své uživatele z řad managementu, které vedou k rozhodnutím týkající se především hodnotového řízení hlavní výdělečné činnosti dané společnosti. (Fibírová, a další, 2010, s. 12)

Druhá, neméně početná část uživatelů jsou externí uživatelé. Mezi tyto uživatele patří zaměstnanci, státní orgány, spolupracující společnosti. V závislosti na právní formě společnosti má v neposlední řadě také zájem o reportované informace investor v dané společnosti, který nefiguruje v řídicí funkci, a tedy externí reporting je jedna ze zásadních možností jak může sledovat výkonnost své investice. Zásadními kritérii pro externí reporting je úplnost, srovnatelnost a spolehlivost. (Fibírová, a další, 2010, s. 21)

Z pohledu Business Intelligence je z mého pohledu nezbytná provázanost a ucelenost informací v obsažených systémech tak, aby byly splněny kritéria použitelnosti daných reportů jejich uživateli. Zmíněná ucelenost má především význam v rámci interpretovatelnosti reportovaných informací. V rámci Business Intelligence jako souboru nástrojů a aplikací vyznačující se svou provázaností s ostatními aplikacemi IS je

bezpodmínečně nutné dbát na zachování integrity a celkové kvality obsažených dat a výsledných informací pro možnost vytvoření kvalitního reportingu.

3.8 Výběr dostupných BI řešení v ČR

Český trh nabízí velkou řadu možností BI řešení pro společnosti, a to od spíše lokálních dodavatelů až po zvučná jména jako je SAP, Microsoft či Oracle. Ze seznamu sestaveném (CIO Business World, 2020) je patrné, že mnoho společností se zabývá pouze dílčími částmi systému BI, jako je například poskytování datového skladu pro své zákazníky. Lokální dodavatelé z drtivé většiny spoléhají na hardware a částečně i software větších společností zmíněných výše.

Na základě průzkumu (SystemOnline, 2018), jsou řešení na českém trhu rozlišena na pomyslné 3 kategorie.

- 1) Kompletní BI řešení
- 2) BI nástroje integrované do ERP
- 3) Samostatné dílčí nástroje (databázové platformy, integrační nástroje, analytické nástroje, reporting)

V prvním případě se jedná o dodávku kompletního podnikového manažerského systému. V případě druhém se jedná povětšinou o subdodávku uceleného analytického balíku BI nástrojů pro aktuální ERP, které podnik využívá. Nakonec v poslední skupině se jedná o dodávku již jen jednotlivých subkomponent pro BI řešení v podniku. Může se jednat například jen samostatně o databázové platformy, ETL či komponenty reportingu (SystemOnline, 2018).

3.8.1 IBM

Americká společnost IBM je technologický gigant jehož počátky sahají až do roku 1911. BI platformou nabízenou IBM je IBM Cognos BI. Tento produkt nabízí uživatelům běžné možnosti kompletního BI řešení, tedy automatickou přípravu dat v reálném čase, analýzy dat a reporting včetně vizualizací. V rámci analýz IBM prezentuje rozšířené možnosti umělé inteligence (angl. AI) jež je v produktu zpracována. Tato funkcionality umožňuje uživatelům efektivněji objevovat nové spojitosti, vazby či trendy v datech a

napomáhá tak například přesnějším forecastingu nad daty. Cloudové řešení pro uložení a zpracování dat v rámci tohoto BI řešení uvádí společnost pod názvem IBM Cloud Pak™. Zajímavostí tohoto řešení je uživatelské prostředí, kdy aplikace je ovládána skrze webový prohlížeč, narozdíl od řešení SAP či Microsoft, kdy mimo tuto volbu je možné ovládat prostředí v desktopové aplikaci. (IBM, 2020)

IBM Cognos BI je platformou cílenou především na obchodní a výrobní společnosti a to v oblasti reportingu a souvisejících analýz. Dle (SystemOnline, 2018) se mezi přední zákazníky využívající aktuálně tuto platformu řadí například Komerční Banka.

3.8.2 SAP

SAP je tradiční německou IT společností zabývající se především ERP řešením pro své zákazníky. V rámci BI se rozhodla představit vlastní BI řešení, nadstavbu právě na obecně užívané ERP řešení jako je SAP S/4HANA. Související BI nadstavba je označena jako SAP Business Objects BI v aktuální nabízené verzi 4.3. Jedná se o řešení, které se vyznačuje svou komplexitou, kdy se jedná, jak bylo zmíněno již výše, o integrovanou část celopodnikového řešení od této společnosti. Samotný analytický software s kterým uživatelé pracují na svém zařízení a tvoří zde samotné reporty je označen v tomto řešení jako SAP Crystal Reports.

Jednou z posledních novinek v rámci tohoto řešení je produkt SAP BusinessObjects Cloud. SAP v rámci tohoto produktu zdůrazňuje možnosti především kombinace cloudového zdroje dat a on-premise (lokálních) zdrojů v rámci SAP architektury a efektivnějších možnostech analýz dat bez nutnosti duplikace či přesouvání dat mezi cloudem a lokálními zdroji. (SAP, 2020) Poměrně důležitou poznámkou je informace od SAP, který avizuje u posledního release verze 4.3. garantovanou podporu produktu SAP BusinessObjects BI až do roku 2027. (SAP, 2020)

Dle dat (CIO Business World, 2020) je povětšinou toto řešení aplikováno ve střední a malých podnicích (SMB). Největšími referenčními projekty tohoto BI řešení je implementace ve společnosti ČSOB či ČD Cargo.

3.8.3 Microsoft

V posledních letech asi nejvíce propagované řešení pro širokou paletu uživatelů nabízí společnost Microsoft se svým produktem Power BI. Toto řešení v aktuální podobě představuje především desktopovou aplikaci Power BI, skrze kterou se uživatel připojuje na jednotlivé datové zdroje a tvoří z těchto dat vizualizace a reporty pro další analýzy. Propagovanou variantou společností Microsoft je kombinace Power BI a cloudové služby Microsoft Azure. Pod tímto označením se skrývá cloudové řešení od společnosti Microsoft, nabízející mimo jiné externí výpočetní výkon, datový management skrze různé typy běžných databází např. MySQL či PostgreSQL, nebo také zálohování dat společnosti. Paleta aplikací a řešení v rámci Azure je dále velmi obsáhlá s mnoha aplikacemi, kdy za zmínku stojí oblast AI, počítačového učení, blockchain nebo i obecně rozšířených predikcí a analýz. Mimo tuto cloudovou službu lze obecně vyzdvihnout také fakt, že Microsoft zde těží ze spojení a vysoké kompatibility Power BI se svými produkty z rodiny Office 365 (Microsoft Corporation, 2020).

Dle (CIO Business World, 2020), jsou hlavním cílem Power BI střední společnosti a velké korporáty (až 80% uživatelů). Zároveň uvádí hlavní uživatele v ČR, kdy se mimo jiné jedná o společnost Škoda Auto.

3.8.4 Oracle

Dalším zvučným jménem ve světě BI je americká společnost Oracle. Hlavním mnohaletým obchodním záměrem společnosti je vývoj a implementace relačních databází. Tato společnost se v BI prostředí prezentuje produktem jednoduše nazvaným Oracle Business Intelligence obsahující, podobně jako konkurence, jednotnou platformu pro čerpání a následnou analýzu dat z různých podnikových a externích zdrojů.

V rámci BI řešení Oracle vyzdvihuje tzv. Oracle Analytics a to buď ve formě cloudového řešení, nebo naopak on-premise (lokální) řešení přímo u zákazníka. Oracle v rámci této služby propaguje podobně jako IBM strojové učení a umělou inteligenci, která má poskytnout uživatelům rozšířené možnosti při analýze dat a forecastingu. (Oracle Corporation, 2020).

Toto BI řešení je často zajímavé již pro zákazníky Oracle využívající infrastrukturu od této společnosti, jako je databázové a serverové řešení Oracle Exadata (Oracle Corporation, 2020). Průzkum (CIO Business World, 2020) uvádí, zastoupení SMB společností v rámci zákaznického spektra pro toto řešení přibližně 80%.

3.8.5 Tableau

Společnost Tableau je jednou z nejmladších ve výběru, se svým založením v roce 2003 ve Spojených státech. Společnost se zabývá především BI produktem, nazvaným stejně jako společnost samotná. Z uživatelského pohledu se jedná o především desktopovou aplikaci Tableau Desktop, ve které uživatel připojuje zdroje, definuje reporty a především tvoří vizualizace dat nad kterými je poté možné aplikovat související analýzy. Pro sdílení reportů a dat je možno využít služby Tableau Online, což je ve výsledku cloudová služba podobná řešením ostatních výrobců, avšak nabízející oproti například Power BI razantně menší paletu nástrojů, jak bude dále rozvedeno v praktické části této práce. (Tableau, 2020)

Mezi společnosti využívající Tableau v ČR patří mimo jiné Komerční banka či AirBank (Inekon Systems, 2020).

3.8.6 Qlik

Posledním vybraným BI řešením je od původem švédské společnosti QlikTech, jež datuje svůj vznik v roce 1993. Nejnovější BI řešení této společnosti s názvem QlikSense je rozšířenou analytickou verzí základního řešení QlikView, nabízející možnosti hloubkových analýz za použití rozšířené inteligence (angl. augmented intelligence) a nativním cloudovým rozhraním. Uživatelské prostředí opět sestává především z desktopové aplikace, kde uživatel definuje vše od zdrojů až po výsledné vizualizace a reporty. (Qlik, 2020)

V ČR se mezi zákazníky této platformy řadí společnosti jako je Zentiva, Equa Bank, Vodafone a další. (KOMIX s.r.o.)

3.9 Trendy BI

Vesměs všichni autoři se shodují, že za poslední dekádu proběhla v rámci BI revoluce. Podniková data znásobují svou velikost a to jak v celkovém objemu tak také v detailu, který jsme schopni zachytit. Efektivní vizualizace dat se stávají jedním ze základních kamenů analýz podnikovými uživateli a stejně tak již tvorba analýz a reportů není výhradou pouze IT oddělení ale i business analytiků v rámci jejich každodenní práce. Tento aspekt označován jako self-service je jedním z nejvyzdvihovanějších trendů poslední doby a je jedním ze stavebních kamenů v praktické části této práce. Jedná se o fakt kdy business analytik je schopen sám vytvářet pracovní modely nad vstupními daty a nálezat v nich tak nové souvislosti a tvořit předpoklady pro rozvoj daného reportingu společnosti.

Jedním z dalších hlavní trendů je využití AI a strojového učení při analýzách a reportingu. Odehrává se zde přechod z pouze reportovacích aplikací, předávající statická historická data na aplikace obsahující širší analytické a vizualizační možnosti právě za použití AI a strojového učení (DataPine, 2020). Z tradičních reportovacích aplikací se tak stávají aplikace, které jsou schopny v reálném čase analyzovat data a zprostředkovat svým uživatelům užitečné informace nad rámec běžného vnímání daných dat. Hledání trendů, anomálií či forecasting je tak tímto posunut na pomyslnou další úroveň. V zájmu poslední doby se objevuje značně prediktivní analytika navazující a rozšiřující pojem dolování dat (angl. data mining), které se zabývá spíše historickými daty. Jak již název napovídá prediktivní analytika indikuje budoucí vývoj daných proměnných s určitou úrovní pravděpodobností, kdy s velkými objemy dat a kontinuálním učením chybovost modelů má klesající trend. (SystemOnline, 2018)

Dalším zásadním trendem v rámci BI řešení je dle autorů stále vyšší využití cloudových služeb a také konceptu „Software jako služba“, anglicky označovaného zkratkou „SaaS“, v tomto kontextu tedy také označovaného jako SaaS BI. Tato produktová sestava umožňuje podnikovým uživatelům využívat skrze cloud značně rošířené BI možnosti o další související aplikace a také poskytuje dynamické možnosti výkonu při souvisejících výpočtech. Zároveň daná společnost šetří náklady za budování vlastní robustní hardwarové a softwarové infrastruktury. Často se též objevuje kombinace základního tzv. on-premise řešení a cloudu, kdy cloud poskytuje dodatečnou flexibilitu pro například analytický

výpočetní výkon. (Microsoft Corporation, 2020) Kladný aspekt SaaS řešení je velice patrný v současné době pandemie COVID-19, kdy toto řešení poskytuje vysokou flexibilitu ve výpočetním výkonu a v přístupu k datům či následným analýzám jako takovým. Samostatnou podsložkou tohoto přístupu lze uvést též mobilní BI, kdy se již dnes jedná o poměrně běžný aspekt BI řešení, jež obsahuje také možnosti analýz a připojení skrze různá mobilní zařízení. (DataPine, 2020)

Posledním vybraným trendem je automatizace dat, jež je též nazývána ve své rozvinuté formě v posledních letech jako tzv. hyperautomatizace. BI přineslo mnoho možností automatizovat zdroje informací a reporting business dat. Pokud k tomuto přidáme též AI prvky inteligentního reportingu, dostáváme mocný nástroj, kdy se hlavní zaměření analytiků posouvá na samotnou analýzu dat a ne na jejich sestavování a reporting. Nehledě v tomto ohledu také na fakt, že znásobující se množství dat nebude schopen člověk jednoduše sám o sobě ani zpracovat a zajisté též ne v téměř bezchybné kvalitě. Velkým přínosem je v tomto případě fakt, že nová BI řešení jsou schopny poskytnou mocné nástroje pro analýzy a reporting řadovému business uživateli, který nemusí mít rozsáhlé znalosti IT. (DataFloq, 2020)

4 Vlastní práce

4.1 Úvod k danému podniku

Z důvodů souvisejících mimo jiné s obchodním tajemstvím autor anonymizoval některá fakta, především název samotné společnosti a také názvy software, které nejsou přímo podstatné pro zkoumanou problematiku v této práci.

Analyzovaná společnost, která poskytla autorovi potřebnou součinnost při zpracování analýz se již od svého založení pohybuje v rámci odvětví energetiky, kde hlavní činností je obchod s energetickými komoditami jako je plyn, elektřina a uhlí. Obchodování probíhá jak na bázi velkoobchodní tak také maloobchodní, kdy dodává energie koncovým zákazníkům. Společnost samotná se v posledních letech blíží mimo jiné svým objemem aktiv a obratem spíše kategorii korporátu. Daná společnost za poslední dekádu zaznamenala razantní růst ve svých aktivitách. Toto nastalo především za přispění vstupu silného investora, jenž kapitálově umožnil silný rozvoj nových aktivit týkající se jak zmiňovaného retailového odvětví tedy B2C a také B2B.

Rychlý růst společnosti znamenal razantní nárůst rozsahu dat a nutných implementací systému pro kvalitní správu účetnictví, finanční reporting, skladové hospodářství a další oblasti. Postupně byl pro tyto účely implementován transakční software v němž probíhá veškeré zpracování především účetních dat. Dále byl v počátcích implementován BI software jenž měl primárně sloužit k analýzám agregovaných účetních dat, reportingu, forecastingu a také budgetingu.

V mnoha ohledech byl výše zmíněný SW dostačující po několik let, avšak s rostoucími obchodními aktivitami rostl násobně také objem transakcí i dat. Výsledkem byla postupná expozice slabých míst v nastavených postupech a infrastruktuře, a to především v oblasti právě BI. V této oblasti se jednalo o poddimenzovanost řešení, kdy nebylo počítáno s tak značným nárůstem vstupních dat, uživatelů, reportů a souvisejících požadavků. Tyto problémy byly dále umocněny rozdílností požadavků jednotlivých divizí podniku, kdy každá z divizí má poměrně specifické zaměření a tedy i požadavky.

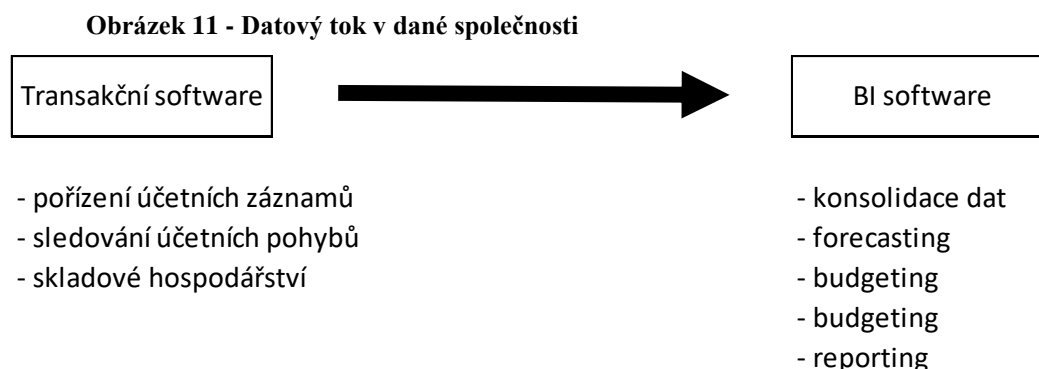
Následující praktické části se zabývají požadavky společnými skrze divize podniku, ale i specifickými požadavky se zaměřením na divizi obchodu s komoditami. Tato divize se

vyznačuje rychlým rozvojem v oblasti nákupu a prodeje různých energetických komodit (plyn, elektrina, uhlí a dalších komodit. Pro tyto transakce využívá specifický obchodní software, který není součástí analýzy v rámci této práce. Tato práce, jak již bylo zmíněno v úvodu, se zaměřuje naopak na finanční řízení a to především z pohledu analýz a nákladového controllingu právě za využití BI aplikace a související infrastruktury.

4.2 Zhodnocení současného stavu v daném podniku

Tato část práce rozvádí a blíže se zabývá zhodnocením stavu a souvisejících parametrů dosavadního BI softwaru z pohledu finančního řízení obchodní divize s komoditami. Je rozdělena na dvě související subsektory, kde první z nich je zaměřena na infrastrukturu blízkou ETL a DW, kdežto druhá část je zaměřena na bližší analýzu a pohled na BI analytickou a reportingovou část. Analýza proběhla z pohledu finančního řízení a analýz využívajících tento SW a souvisejících komponent. Jedná se tedy o uživatele, který potřebuje naplno využít spolehlivosti a vysokého výkonu daného řešení. Cílem pro uživatele je stav, kdy může bez velkých překážek získat jednoduše potřebná data ve vhodném formátu a dále se zabývat navazujícími analýzami, a to v ideálním případě skrze rozšířené analytické možnosti a aplikace daného BI řešení.

Zjednodušená struktura úkonů, které v rámci dosavadního řešení jsou prováděny, je uvedena v diagramu níže.



Zdroj: Vlastní zpracování

Jak je z tohoto diagramu patrné, původně nastavená infrastruktura je poměrně přímočará. Jak transakční tak původní BI software pracuje na jednotlivě implementovaných MS SQL databázích, kdy v obou případech se jednalo o on-premise řešení. Takto stanovená infrastruktura je v dynamickém, rychle se rozvíjejícím prostředí nových aktivit a požadavků náchylná k problémům, které jsou blíže rozvedeny dále.

4.2.1 Hlavní nedostatky stávajícího řešení - analýzy

Prvním zásadním nedostatkem, který limituje businessové uživatele v jejich práci je nízký výpočetní výkon BI řešení. Nízký výkon se týká odezvy jednotlivých dotazů, na kterých jsou postaveny uživatelské reporty s konsolidovanými daty. Vykreslení dat pro obsáhlejší report za období delší než několik měsíců je otázkou několika desítek vteřin až minut. Po bližším zkoumání a analýze byl nalezen fakt, že BI aplikace při každém načtení, updatu či změně parametrů v rámci reportu, vždy kontaktuje přímo zdrojovou transakční databázi a hledá zde jednotlivé transakce, které následně agreguje a poté vykresluje do daného reportu. V případě zvyšujícího se počtu reportů a počtu uživatelů, kteří často v tu samou chvíli provádějí načítání svých reportů se dostáváme do situace, kdy výkon dotazů rapidně klesá a řešení se stává uživatelsky problémové.

Další problém je poměrně zásadní, a týká se stanovených agregací v rámci dosavadního BI řešení. Zpracovaná data v BI aplikaci nenabízejí potřebný detail transakcí, který by mohl být využit pro další analýzy. Vykreslená data v reportech jsou tedy již značně agregována a nelze jednoduše získat plný detail jednotlivých transakcí včetně všech parametrů z transakčního systému. V rámci návazného reportingu to znamená častou potřebu dodatečných analýz přímo z transakční databáze a jejich zpracování mimo BI řešení.

Další problém se také týká reportingu, a to možností vizuálních prvků. Stávající BI aplikace má pouze velmi omezené možnosti v rámci této sekce. V kombinaci s prvním a druhým zmíněným problémem navíc použitelnost těchto vizuálních prvků je značně omezena. Příkladem může být prvek velmi podobný kontingenční tabulce, kde vizuální zpracování samotného prvku je poměrně základní bez možnosti větších úprav podobně jako jsou uživatelé zvyklí v aplikaci MS Excel. Dále je i u tohoto prvku patrná vysoká odezva zobrazení jednotlivých položek a návazných sub-subpoložek, kdy i přes pomalé vykreslení dat, není možné zobrazit data v plném detailu až na jednotlivé účetní transakce. Nabídka

další vizuálních prvků jako jsou grafy či dashboardy je poměrně strohá a neodpovídá požadavkům na reporting obchodní divize.

Další problém, souvisí se současnou BI aplikací jako takovou, je v oblasti přístupnosti a možností kolaborace mezi uživateli. Současné řešení poskytuje pouze možnost desktopové aplikace bez širší možnosti kolaborace mezi uživateli, sdílení reportů či nastavení notifikací. Chybí zde tedy mimo jiné přístup z webového rozhraní či mobilní aplikace.

Poměrně zásadním problémem s kterým se analytici potýkají je chybějící přístup k analýzám formou self-service. Business uživatelé nemají možnost tzv. datového sandboxu jenž jiné aplikace poskytují a umožňují tak uživatelům vytvářet své analýzy a zkušební reporty ve kterých jsou schopni zkoumat veškeré aspekty zdrojových dat a hledat kupříkladu nové souvislosti mezi nimi. Procházení dat v rámci dosavadního řešení, v samotné databázi, stanovování agregací, úprava hierarchií je uživatelsky značně náročné a z povahy řešení i značně náchylné k zásahu do globálních struktur dosavadního BI řešení.

Z výše zmíněného vyplývá, že oblast analýz a reportingu je u stávajícího řešení poměrně nedostačující a nedostahuje potenciálu na který by chtěla společnost cílit. Bylo zjištěno, že problémy současného řešení v této oblasti souvisí s řadou aspektů od nízkého výpočetního výkonu až po uživatelskou nepřívětivost v rámci možností vizualizací či kolaborace. V neposlední řadě velmi omezené analytické možnosti BI řešení vedou logicky k nutnosti analýz mimo toto řešení v substitučních řešeních jako je tradiční MS Excel a další.

4.2.2 Hlavní nedostatky stávajícího řešení - infrastruktura

V této části bude uveden výčet toho nejzásadnějšího co souvisí přímo s technickou a datovou infrastrukturou, kdy tato část je již doménou spíše IT oddělení než samotných business uživatelů a autor v rámci této práce dále uvede ty nejzásadnější zjištění, které svou analýzou ve spolupráci právě s IT oddělením dosáhl a které mají neodmyslitelný vliv také dále do analýz, reportingu a celkové spolehlivosti a výkonnosti BI řešení.

Jedním z hlavních problému co se týče samotné infrastruktury je chybějící plnohodnotné DW řešení. Databáze z transakčního SW a BI aplikace jsou propojeny napřímo a jak bylo zmíněno v předchozí kapitole, při každé aktualizaci reportů se BI aplikace dotazuje přímo do databáze transakčního SW. Efektivita a dopady tohoto řešení

byly též nastíněny v předchozí kapitole, kdy se jednalo o velmi slabou výkonnost celého řešení a negativní uživatelskou zkušenost.

Souvisejícím problémem byl určen chybějící logický datový model a související datový a business slovník. Jak je uvedeno v teoretické části této práce, jedná se o základní stavební kámen celé infrastruktury, její efektivní výstavby a dalšího rozvoje. Bez datového modelu uvedeném v patřičné dokumentaci není na první pohled patrné jak jsou na sebe procesy a datové transformace navázány a je nutná zevrubná investigace těchto procesů pokud se jedná o další rozvoj či přebírání zodpovědnosti za danou infrastrukturu novým pracovníkem. Znalost struktury tak mají pouze pracovníci, kteří byli u zrodu této struktury a v dalším koloběhu a fluktuaci pracovníku může být tato znalost podstatně redukována. Stejně tak je problémem chybějící standardizované názvosloví, tedy business slovník. Bez tohoto podkladu je opět poměrně složité, především v rozsáhlejších projektech mající prostup skrze různé divize, dojít rychle a efektivně k jasné specifikaci kroků a být si plně jist, že všichni uživatelé definující další kroky, myslí stejným názvoslovím stejnou věc. Obecně lze tedy konstatovat, že byl nalezen problém s dokumentací, celkovým zmapováním a patřičně kvalitním zaznamenáním jednotlivých procesů v dané infrastruktuře. Společnost se v rámci tohoto jevu vystavuje úměrně vyššímu operačnímu riziku, kdy se může v průběhu let znalost vytrácet a je nutné při zaškolování nových zaměstnanců vynakládat vysoké úsilí a především čas na plné pochopení všech principů a použitých technologií.

Jedním z dalších problémů související také s DW byla označena chybějící historizace dat. Jedná se o problém při kterém jsou záznamy v databázi přímo přepisovány, nikoliv zneplatněny, kupříkladu zvoleným příznakem. Tento fakt především limituje možnosti auditování změn, které v databázi proběhly a mohly mít určitý dopad, který je třeba analyzovat.

4.2.3 Informační a datová architektura

Tato část požadavků, jak již název napovídá, se soustředí na identifikaci parametrů týkajících se toku dat a informací v rámci nově zaváděného BI řešení nahrazující řešení původní. Identifikace v této části počíná v oblasti dat zdrojových, tedy především transakčních systémů a dalších doplňkových databází a ostatních různorodých zdrojů. Dále se zabývá datovými transformacemi a práci s daty v rámci DW řešení, aby nakonec řešil

otázku dalších transformací a agregací pro využití takto zpracovaných dat k následným analýzám především business uživateli.

Mezi výstupy v této části analýz pro sestavení požadavků pro nové BI řešení, na nichž se autoři (Laberge, 2012), (Sherman, 2014) a další shodují patří v tomto případě:

- 1) Zmapování zdrojových databází a zdrojů
- 2) Specifikace Master data managementu (MDM)
- 3) Specifikace čištění a transformace dat (ETL)
- 4) Specifikace uchování dat (DW)
- 5) Specifikace agregací a rozdělení dat pro následné analýzy

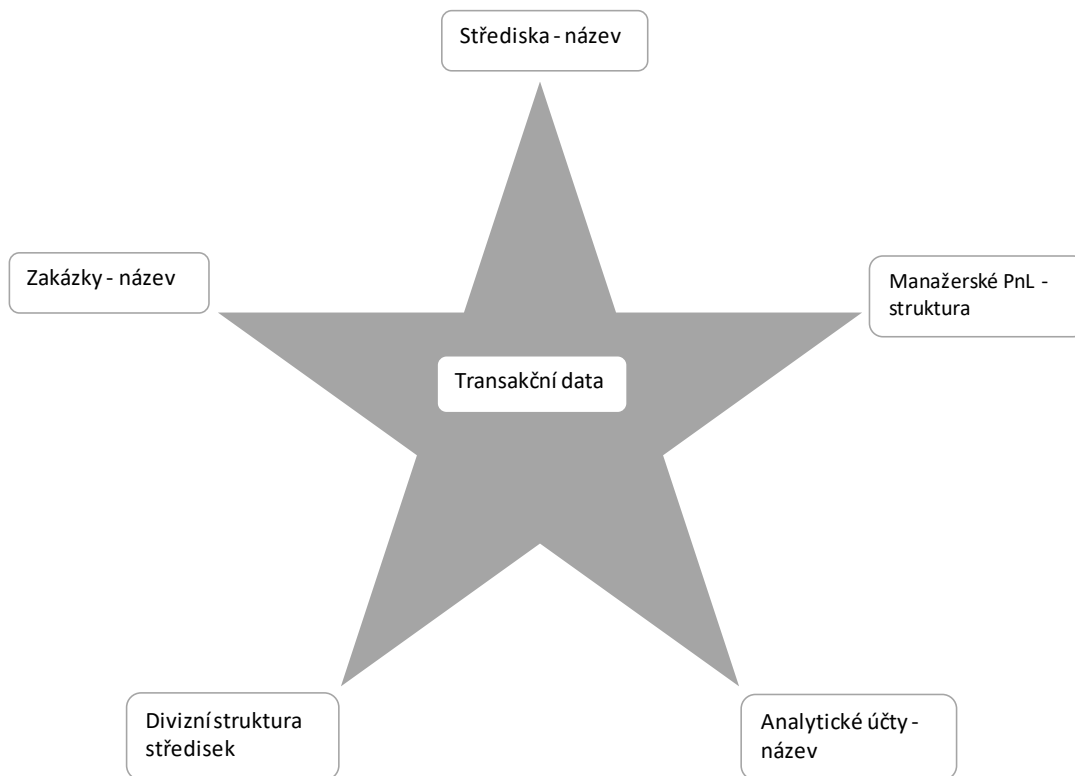
V rámci výše zmíněného bodu 1) a 2) je obsažen krok profilování dat v aktuálně dostupných zdrojových databázích, dále zjišťování jejich struktury následované analýzou samotných prvků a dat v daných zdrojových databázích. V detailním pohledu se jedná o analýzu samotných tabulek a jejich obsahu, tedy jednotlivých záznamů. Zásadním požadavkem v této části je zaměření se na dříve zmíněné tzv. 5C dat, tedy důraz na jejich kvalitu, čistotu, konzistenci, aktuálnost a celkovou relevanci pro další analýzy. V rámci tohoto kroku tedy bylo sledováno na kolik relevantní a kvalitní data se ve zdrojových databázích nacházejí a je možno jejich plné využití pro nové BI řešení. V rámci dosavadního řešení byla kvalita dat zdrojového transakčního systému na vysoké úrovni kvality a konzistentnosti. Jediný zásadní problém v rámci potřeb dalších finančních analýz byly některé chybějící parametry u starších záznamů, a to především tzv. identifikační číslo zakázky. Tento parametr je zásadní pro další reconciliaci dat a především finanční nákladový controlling. Tento parametr sdružuje jednotlivé nákladové skupiny, které jsou definovány v rámci manažerského výkazu zisku a ztráty. V rámci analýz k novému BI řešení byly definovány některé nové dimenzní tabulky, jako je nová struktura manažerského výkazu zisku a ztráty obchodní divize, nebo tabulka korekcí, obsahující opravy některých záznamů z transakčního SW. Jedná se o řešení, kdy není narušena integrita původní databáze, a tedy oprava daných transakcí je prováděna mimo tuto zdrojovou databázi v rámci připojené samostatné dimenzní tabulky.

V navazujících krocích se dále zkoumají jednotlivé tabulky ve vztahu mezi sebou a jsou mimo jiné hledány primární klíče v rámci definice vztahů mezi jednotlivými tabulkami.

Další krok 3) transformující data do DW bude obsahovat v novém BI řešení již souhrn filtrací, kdy v rámci aktuálního řešení probíhá například běžná filtrace středisek a účtů až na základě vytvořených dotazů v rámci reportu koncové BI aplikace. Toto v návaznosti na následné agregace dat má za následek již dříve zmíněné problémy s odezvou jednotlivých dotazů. V rámci nového řešení je tato filtrace dat prováděna v rámci ETL rozhraní, kdy do DW se ukládají pouze relevantní data použitelná k následným analýzám business uživatelů, čímž je kladen důraz na efektivitu architektury. Takto jsou zpracovávána pouze relevantní data a jsou tak šetřeny související zdroje, ať již výpočetní výkon, nebo místo v DW. V rámci ETL zpracování dat byly též zakomponovány IT oddělením auditní prvky, které při významných událostech při zpracování dat, kdy by byla ohrožena integrita výsledných dat notifikuje IT správce.

V rámci kroku 4) bylo dosavadní řešení definováno jako dvě samostatné databáze, jedna pro transakční software a druhá pro BI software. V rámci nového BI řešení je toto nahrazeno a doplněno o plnohodnotnou DW strukturu, včetně souvisejícího ETL a analytického agregačního řešení, a to formou datových tržišť. Co se týče architektury samotného DW jedná se mimo jiné o tradiční soubor faktových a dimenzních tabulek obsahující širší kontext k transakčním datům. Obecně v rámci analýz datové architektury bylo shledáno dané datové schéma nejbližší odpovídající dimenznímu schématu hvězdice. Níže je uveden diagram, který zobrazuje výčet některých základních prvků dimenzní datové struktury.

Obrázek 12 - Datové schéma - ukázka



Zdroj: Vlastní zpracování

Navazuje poslední krok 5) zabývající se především agregacemi a rozdělením dat do jednotlivých datových tržišť. Daná tržiště byla specifikována na základě působnosti jednotlivých divizí a jejich specifických požadavků na data, které má datové tržiště obsahovat. Dále byly brány v potaz analytické požadavky jednotlivých divizí a tedy související agregace zpracovávaných a uchovávaných dat, opět s cílem co nejvíce zefektivnit danou strukturu zpracování dat.

4.2.4 Technická a produktová architektura

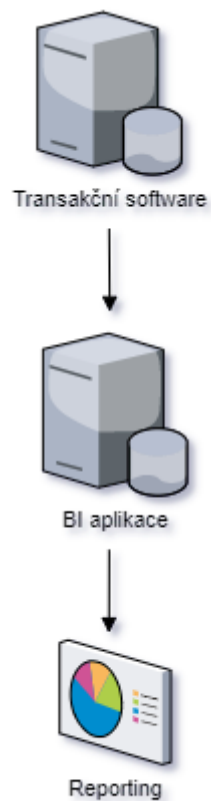
Technická architektura a její specifikace přímo navazuje na specifika definována v oblasti informací a dat v předchozí kapitole. Níže jsou uvedeny některé základní otázky, které byly v rámci analýz řešeny:

- 1) On-premise / cloud přístup k technickému řešení

- 2) Bezpečnost řešení
- 3) Prodejce HW a SW řešení
- 4) Technické možnosti HW a SW řešení
- 5) Výpočetní výkon a možnosti optimalizace
- 6) Podpora

Základní přístup k technické architektuře v případě zkoumaného podniku je založen na přesunu od původního BI řešení založeném v podstatě na transakčním software a na druhé straně BI aplikací bez významných prvků mezi těmito dvěma systémy, které by tok dat zefektivnily. V diagramu níže je uveden původní přístup.

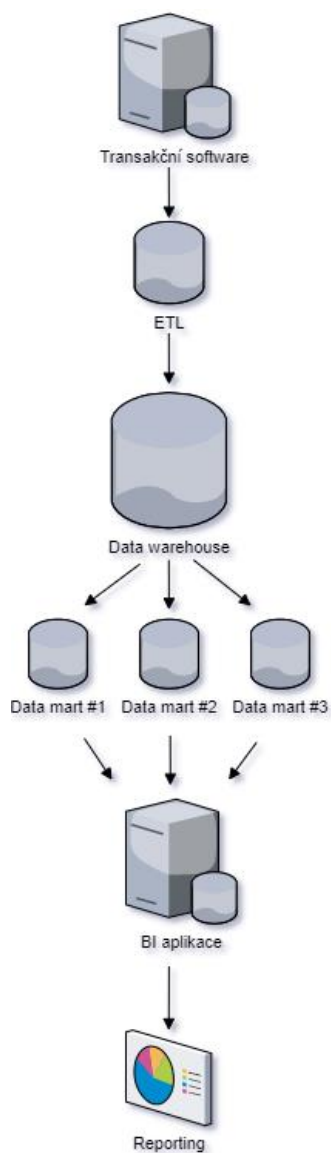
Obrázek 13 - Diagram původní architektury



Zdroj: Vlastní zpracování, (Diagrams.net, 2021)

Je patrné, že toto uspořádání architektury není ideální a vede mimo jiné k zásadním problémům v oblasti výkonnosti především analytické části BI řešení a BI aplikace samotné. Diagram následující níže, obsahuje nový návrh struktury, která odpovídá aktuálním standardům a požadavkům jednotlivých business divizí.

Obrázek 14 - Diagram nové architektury



Zdroj: Vlastní zpracování, (Diagrams.net, 2021)

Z diagramu je zřejmé, že do architektury přibylo několik dalších prvků zabývajících se transformacemi, uložením a distribucí kvalitních agregovaných dat. Tyto data jsou následně určeny k využití v rámci analýz a reportingu v rámci koncové BI aplikace.

Z pohledu analýzy potřeb nového BI řešení byly výše zmíněné body 1) až 2) poměrně silně propojeny. Z pohledu cloudového řešení byly brány v potaz otázky regulace ze strany veřejných autorit a také pohled samotné společnosti na bezpečnost a uložení svých dat. V rámci pravidel autorit se jedná především o GDPR, tedy tzv. Obecné nařízení o ochraně osobních údajů, jímž se společnost zpracovávající data o svých zákaznících musí závazně řídit. Dále byly brány v potaz operační procedury a vize samotného podniku s ohledem na bezpečnost, důvěrnost a dostupnost uchovávaných dat. V neposlední řadě byly v rámci analýz týkající se obecně SW a HW brány v potaz aspekty jako je výkon, cenová politika, reputace či technická podpora.

4.3 Definice požadavků na BI řešení

V této části práce jsou definovány stěžejní požadavky především na analytickou a reportingovou část BI řešení, tedy ze většinové části na koncovou BI aplikaci jako takovou. Jsou zde také zohledněny další parametry jež by mělo řešení neodmyslitelně splňovat v kontextu BI řešení jako celku včetně ETL a DW komponent. Priority jsou uvedeny jak z obecné roviny tak z roviny finančního řízení v rámci daného podniku. Část priorit těchto požadavků vychází ze zkušeností s původním BI řešením, kdežto zbylá část požadavků vychází z aktuálního tzv. best-practice, které se na aktuálním BI trhu projevuje.

Tabulka 1 - Požadavky a priority

Oblast	Požadavek
Datové analýzy	- vysoká kompatibilita s datovými zdroji - self service přístup
Vizuály	- kvalitní, vysoce upravitelné vizuály - vysoký výkon a rychlá odezva při aktualizaci dat a vizuálů - jednoduchá možnost kolaborace na projektech a sdílení dalším uživatelům

Rozšířená analytika	- možnosti AI a strojového učení, prediktivních analýz
Platformy	- přístup z různých prostředí k projektům a reportům (webové, mobilní)
Trénink / Podpora	- efektivní trénink, dostatek podkladů k tréninku
Cenová politika	- přijatelná cena samotných licencí / uživatel
Rozšířenost	- znalostní základna
Reputace	

Zdroj: Vlastní zpracování

Z metodického hlediska jsou níže stanoveny jednotlivým kritériím výběru jejich váhy, tak aby dále v kombinaci s bodovou metodou bylo možné provést vážený součet těchto hodnot pro celkové srovnání vybraných řešení a vybrání toho nejideálnějšího.

Stanovení vah je provedeno tzv. Saatyho metodou, kdy se jedná o metodu kvantitativního párového srovnání. Každé kritérium v matici porovnáváme proti ostatním kritériím tak, že stanovujeme poměr preferencí na škále 1-9 jednotlivým kritériím. Hodnotu 1 mají rovnocenně preferovaná kritéria, naopak hodnotu 9 pár kdy kritérium je absolutně preferováno před druhým porovnávaným. Vždy jsou porovnávány parametry na vertikální ose proti parametrům na ose horizontální.

Tabulka 2 – Saatyho metoda – srovnávací matice

	K1	K2	K3	K4	K5	K6	K7	K8
K1	1	3	4	6	6	7	9	9
K2	1/3	1	3	5	6	6	7	8
K3	1/4	1/3	1	2	1/3	2	6	4
K4	1/6	1/5	1/2	1	1/4	1/2	2	3
K5	1/6	1/6	3	4	1	5	4	8
K6	1/7	1/6	1/2	2	1/5	1	1/2	4
K7	1/9	1/7	1/6	1/2	1/4	2	1	3
K8	1/9	1/8	1/4	1/3	1/8	1/4	1/3	1

Zdroj: Vlastní zpracování

Ze srovnávací matice je pro každé kritérium níže spočten geometrický průměr a následně je dle poměru jednotlivých kritérií na součtu těchto průměrů stanovena váha jednotlivých kritérií. Příklad výpočtu je uveden v tabulce níže.

Tabulka 3 - Příklad výpočtu váhy kritéria - Saatyho metoda

Kritérium	Geometrický průměr	Váha
K1	$\sqrt[8]{1 \times 3 \times 4 \times 6 \times 6 \times 7 \times 9 \times 9} = 4.72$	$\frac{4.72}{12.50} = 0.377$
...
Celkem	12.50	1

Zdroj: Vlastní zpracování

Na základě výše uvedeného výpočtu dostáváme výslednou váhu každého kritéria, kdy jejich součet činí 1.

Tabulka 4 – Stanovení vah – Saatyho metoda

Kritérium	Popis kritéria	Geometrický průměr	Váha
K1	Datové analýzy	4.72	0.377
K2	Vizualizace	3.17	0.253
K3	Rozšířená analytika	1.13	0.090
K4	Platformy	0.58	0.046
K5	Trénink / Podpora	1.64	0.132
K6	Cenová politika	0.56	0.045
K7	Rozšířenost	0.46	0.037
K8	Reputace	0.24	0.019
celkem		12.50	1.00

Zdroj: Vlastní zpracování

Jak je patrné z tabulky výše, priority jsou především kladeny na oblast datové analýzy. Zde se jedná především o rychlou odezvu a vysokou flexibilitu v rámci zpracování dat z různých datových zdrojů. Druhou v pořadí je oblast vizualizací, kdy odezva vizuálů při změně vstupního parametru formou filtru či změně formy vizuálu by měla být promptní. Obecně v rámci oblastí datové analýzy a vizualizací by se mělo jednat o řešení, které umožňuje v maximální možné míře přístup self-service, tedy že mnoho úkonů je schopen provádět samotný businessový uživatel a nepotřebuje k nim bezpodmínečně kooperovat s IT oddělením. Jedná se o důležitý faktor, který odlehčuje IT oddělení a na druhé straně poskytuje určitou volnost v analýzách a zpracování reportů jednotlivým business uživatelům, povětšinou analytikům. Samozřejmě tento přístup musí jít ruku v ruce s dostatečným nastavením bezpečnostních vrstev, tak aby nebyla ohrožena integrita řešení a business uživatel nemohl zasahovat do páteřních struktur daného řešení. Posledním z vysoce hodnocených požadavků patří dostupnost tréninku a celková složitost tréninku pro uživatele daného řešení. Toto je silně korelováno s rozšířeností daného řešení, kdy je k dispozici velká znalostní základna a často tak uživatelé mohou najít odpověď na řešení svého problému v online webovém prostředí.

4.4 Srovnání BI řešení dostupných na trhu

Tato část práce navazuje na předešlou kapitolu zabývající se technickou a produktovou architekturou a dále rozvádí především volbu nejvhodnější BI aplikace pro daný podnik. Pro analýzu a srovnání aktuálních řešení byla zvolena z velké části metodika společnosti Gartner, jako uznávaného subjektu na poli analýz mimo jiné právě BI prostředí a hodnocení řešení, které se zde uplatňují. Tato společnost zpracovává na pravidelné bázi report zabývající se platformami business intelligence a vytváří na základě jejich srovnání tzv. Magický čtyřúhelník (angl. „Magic quadrant“). V rámci této analýzy dává společnost Gartner důraz na aktuálnost charakteristik a také uživatelského vnímání daných řešení jako je jednoduchost užívání a způsob zpracování analytického workflow.

Na základě nejaktuálnějšího reportu této agentury, byly analyzovány 3 nejvýznamější BI řešení. Tyto řešení jsou označeny v rámci magického čtyřúhelníku jako lídři odvětví BI aplikací, jak je patrné z obrázku níže.

Obrázek 15 - Magic Quadrant for Analytics and Business Intelligence Platforms



Zdroj: (Gartner, 2020)

Jednotlivé hodnocené oblasti byly na základě dostupných zdrojů a požadavků podniku definovány následovně:

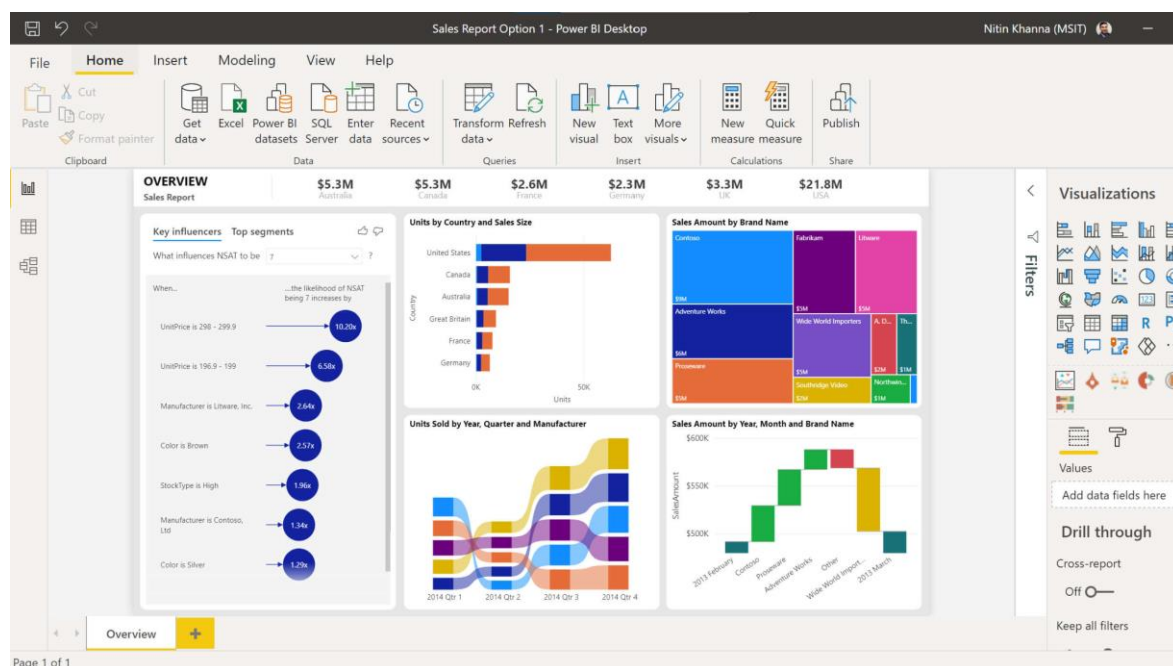
- 1) Datové analýzy
- 2) Vizualizace
- 3) Rozšířená analytika
- 4) Platformy

- 5) Trénink / Podpora
- 6) Cenová politika
- 7) Rozšířenost
- 8) Reputace

Ke každé výše uvedené části jsou uvedeny kladné a záporné vlastnosti daného řešení a je jim na základě dostupných hodnocení analytických společností a hodnocení z pohledu společnosti samotné přiřazen počet bodů od 1 do 10.

S poměrně značným náskokem je patrné, že lídrem v oblasti BI řešení je aktuálně společnost Microsoft, se svým produktem Power BI. Dle (Gartner, 2020) se jedná o vizionářský produkt s obrovským tržním dosahem skrze skupinu produktů Microsoft Office.

Obrázek 16 - Power BI desktop



Zdroj: (Microsoft, 2021)

Tabulka 5 - Power BI specifikace

Microsoft Power BI	klady	zápory	hodnocení (1-10)
Datové analýzy	- vysoká kompatibilita s datovými zdroji - datový sandbox (Power Query)	- datový sandbox nepodporuje přímé SQL dotazy	10
Vizualizace	- množství vizualizací - kvalitní design	- komplexnost vizualizací	8
Rozšířená analytika	- Quick Insights - Azure Machine Learning - NLP	- Power BI Pro cloud service	8
Platformy	- desktop aplikace - webové rozhraní - mobilní aplikace		10
Trénink / Podpora	- on site - online		10
Cenová politika	- free verze - 9.99USD / měsíc / uživatel - pro verze		8
Rozšířenost	- velmi rozšířené		10
Reputace	- značně kladná		9

Zdroj: Vlastní zpracování, (Capterra, 2021), (G2, 2021), (Gartner, 2021), (SelectHub, 2021)

Toto řešení je jedno z nejrozšířenějších a v rámci datové analýzy se vyznačuje vysokou konektivitou. Dále také obsahuje vlastní datový transformační nástroj Power Query kdy v rámci self-service jsou business uživatelé schopni sami intuitivním způsobem propojit různorodé datové zdroje a dále na nich provádět transformace. Tímto může business analytik velmi efektivně nacházet nová řešení a kontext pro další analýzy. Jedná se o kvalitně zpracovaný prvek, který byl několik let před uvedením Power BI představen uživatelům jako doplněk MS Excel a již v této své podobě se jednalo o silný nástroj extenzivně rozšiřující možnosti zpracování a transformace dat v rámci MS Excel. V posledních verzích MS Excel byly mnohé funkcionality implementovány přímo do samotné standardní aplikace a uživatelé tak již nemusí instalovat tento doplněk dodatečně.

Co se týče vizualizací, aplikace nabízí nepřehledné množství vizualizací, které lze libovolně navzájem kombinovat a propojovat. Tyto vizualizace jsou na vysoké úrovni zpracování a poskytují opravdu rozsáhlé možnosti úprav mnoha parametrů jednotlivých vizuálů. Řešení nabízí také rozsáhlé možnosti nastavení práv jednotlivým uživatelům a umožňuje tak značně efektivní kolaboraci v rámci reportingu jako takového. Nevýhodou v tomto případě může být v některých částech až velká komplexita při sestavení a nastavení správného fungování daných vizuálů, k čemuž je třeba delší čas patřičného tréninku.

Zásadním kladem tohoto řešení je vysoká kvalita a rozšířené možnosti tréninků. Jelikož se jedná o velmi rozšířené řešení, je poměrně snadné najít množství kvalitních školitelů tohoto řešení a zároveň existuje velká členská základna řešící různé problémy v online prostředí, ze kterých mohou uživatelé též čerpat dané znalosti. Vysokou přidanou hodnotu této části dokazuje dále velmi kvalitně zpracovaná online platforma na stránkách vydavatele obsahující interaktivní návody včetně video školení, které uživatele provádí různými funkcionalitami v poměrně vysokém detailu.

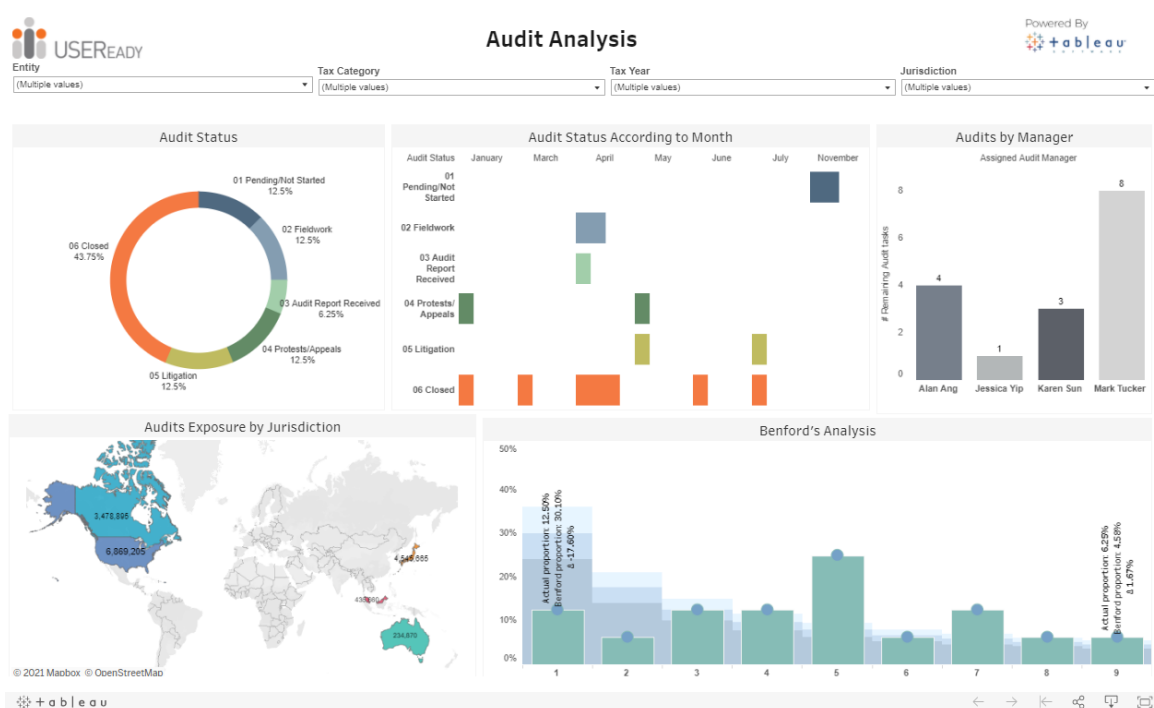
V neposlední řadě je u tohoto řešení vysoká kompatibilita jak na vstupu u zdrojů tak na výstupu, kdy řešení podporuje většinu běžně dostupných platform, včetně mobilní aplikace či webového rozhraní.

Co se týče rozšířených analytických možností řešení, tak se jedná o soubor AI služeb jako je analýza textu, obrázků a dalších prvků rozšířené analytiky. Bohužel jak zmiňuje (Gartner, 2020), jsou vesměs všechny tyto služby spojeny s Power BI Pro cloud service, což může být bráno jako nevýhoda z pohledu dodatečných nákladů. Za zmínku stojí také fakt,

že Microsoft toto řešení nabízí v rámci multicloudové architektury v rámci přístupu SaaS, avšak pouze v rámci vlastní cloudové služby MS Azure, kdy s touto službou je též spojena většina prvků výše zmíněné rozšířené analytiky a služby Pro cloud.

Druhým v pomyslném pořadí lídrů dle (Gartner, 2020) je společnost Tableau se svým stejnojmenným produktem.

Obrázek 17 - Tableau Dashboard



Zdroj: (Tableau)

Tabulka 6 - Tableau specifikace

Tableau	klady	zápory	hodnocení (1-10)
Datové analýzy	<ul style="list-style-type: none"> - vysoká kompatibilita s datovými zdroji - rozšířená datová analytika přípravy dat (Prep Builder) 		9

Vizualizace	- množství vizualizací - návrh vizualizací dle typu dat	- pomalý update a vykreslování dat	6
Rozšířená analytika	- Prep Builder - NLP		9
Platformy	- Desktop aplikace - webové rozhraní - mobilní aplikace		10
Trénink / Podpora	- on site - online	- složité pokročilejší analytické funkce	6
Cenová politika	- 14 dní zkušební verze - 70USD / uživatel / měsíc		4
Rozšířenost	- nadprůměrná		8
Reputace	- značně kladná		9

Zdroj: Vlastní zpracování, (Capterra, 2021), (G2, 2021), (Gartner, 2021), (SelectHub, 2021)

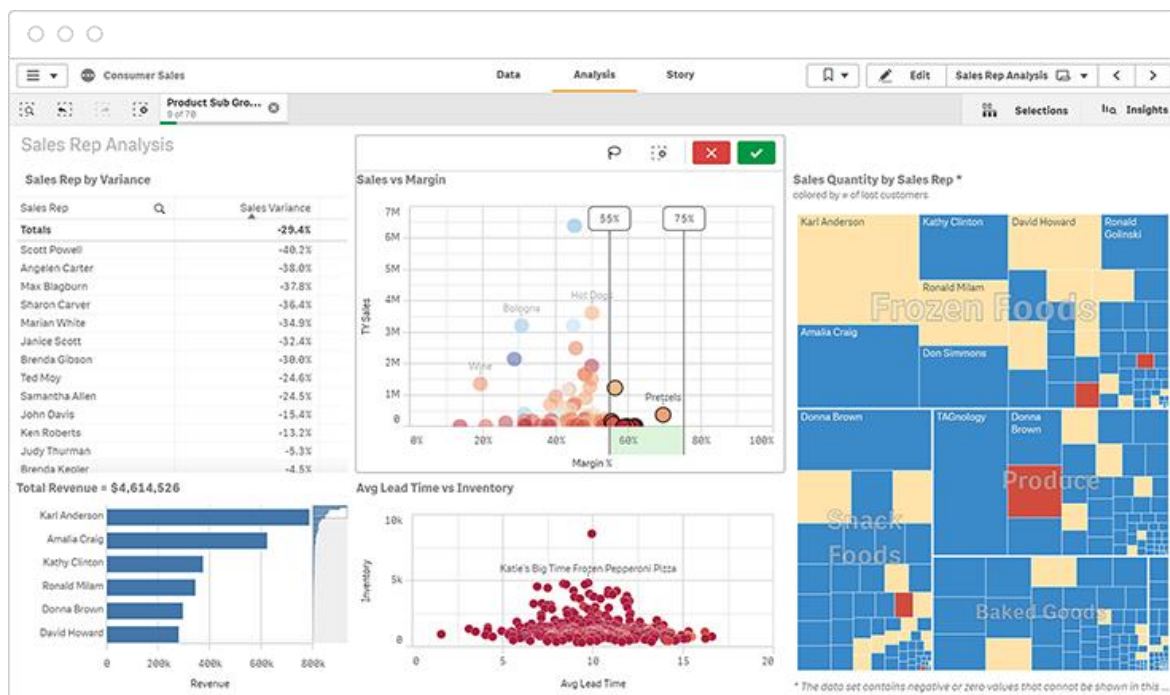
Jedná se o další z nejrozšířenějších BI řešení, které má široký záběr na trhu. Vlastnosti produktu jsou ve značné míře podobné těm uvedeným u produktu společnosti Microsoft. Přesto jsou zde patrné rozdíly a to například v rámci vizualizací. Tableau v rámci vizualizací oproti Power BI nabízí možnosti automatických návrhů template dashboardu na základě rozpoznané struktury dat. Na druhou stranu platforma v rámci vizualizací trpí problémy s výkonem při aktualizacích a vykreslení dat. (G2, 2021)

V rámci dalších srovnávaných aspektů je patrný rozdíl v náročnosti tréninku uživatelů, kdy dle referencí mimo jiné (SelectHub, 2021) uvádí jednoduchost v užívání základních vizualizací a zpracování dat avšak po poměrně krátké chvíli při hlubší analýze dat a rozšířené využití vizuálů uživatel naráží na velký skok v náročnosti obsluhy a nutného dalšího tréninku.

Posledním silně patrným aspektem je rozdíl v ceně a nastavení zkušební verze, kdy uživatel má pouze omezenou platnost zkušební verze 14 dní a poté již software nemůže využívat bez zakoupení platebního plánu. Stejně tak je oproti řešení společnosti Microsoft značně vyšší cena za uživatele a licenci v rámci základní verze řešení.

Posledním z analyzovaných řešení v rámci této práce, u něhož je zvažována vhodnost pro daný podnik je produkt společnosti Qlik s názvem Qlik Sense.

Obrázek 18 - Qlik Sense dashboard



Zdroj: (Qlik)

Tabulka 7 - QlikSense specifikace

QlikSense	klady	zápory	hodnocení (1-10)

Datové analýzy	- vysoká kompatibilita s datovými zdroji	- omezená customizace	6
Vizualizace	- nadprůměrná kvalita vizualizací - vysoká rychlost zpracování dat	- problémy při kolaboraci	7
Rozšířená analytika	ML a AI analytické add-iny		6
Platformy	- desktop aplikace - mobilní aplikace		7
Trénink / Podpora	- on site - on-line	- nedostatečná podpora a vysoká náročnost	5
Cenová politika	30USD / uživatel / měsíc		6
Rozšířenost	- nadprůměrná		8
Reputace	- značně kladná		8

Zdroj: Vlastní zpracování, (Capterra, 2021), (G2, 2021), (Gartner, 2021), (SelectHub, 2021)

Toto řešení má znatelně nižší rozšířenost než předchozí dvě zmíněné. Zaměření má Qlik podobný jako Microsoft, kdy na rozdíl od Tableau je obecně rozsah jejich softwarových řešení pro společnosti poměrně široký a dosahuje daleko za hranice pouze BI.

Oproti předchozím řešením Qlik dle dostupných referencí postrádá rozmanitost customizací v rámci zpracování dat a tvorby reportů. V dalším případě jak uvádí (SelectHub, 2021) řešení vykazuje značné problémy v rámci kolaborativní práce s výstupy, mimo jiné jejich sdílení a exportování. Oproti především Microsoft řešení je zde patrný také značný

nedostatek v podpoře tréninku uživatelů, kdy mimo jiné (G2, 2021) uvádí náročnost a nízkou úroveň především on-line tréninků pro tento software.

Na závěr, jak je patrné ze srovnávací tabulky, jsou i zde vyšší náklady na licenci pro jednotlivé uživatele oproti řešení společnosti Microsoft, avšak ne až tak rapidně vyšší než Tableau.

Naopak značnou výhodou tohoto řešení, které zmiňuje mimo jiné (Gartner, 2020), je přístup multicloudového řešení, kdy narozdíl od Power BI není toto řešení omezeno na pouze službu MS Azure a nabízí tak svým uživatelům větší svobodu volby, jakou službu si sami zvolí.

4.5 Výběr BI řešení

Autor v této části shrnuje analýzu jednotlivých řešení dle stanovených parametrů z předchozích kapitol a zhodnocuje výslednou volbu jež daná společnost v rámci BI systému provedla. Cílem této části je především nalézt odpověď na otázku zda odpovídá výběr řešení provedený společností kontextu předchozích analýz provedených autorem.

V tabulce níže jsou uvedeny jednotlivé aspekty a bodové ohodnocení v rámci jednotlivých řešení. K jednotlivým bodovým ohodnocením je zde přiřazena patřičná váha, dle definiční tabulky na počátku této kapitoly.

Tabulka 8 - Zhodnocení řešení

Kritérium	Popis kritéria	Váha	Power BI	Tableau	Qlik
K1	Datové analýzy	0.377	10	9	6
K2	Vizualizace	0.253	8	6	7
K3	Rozšířená analytika	0.090	8	8	6
K4	Platformy	0.046	10	10	7
K5	Trénink / Podpora	0.132	10	6	5
K6	Cenová politika	0.045	8	4	6
K7	Rozšířenost	0.037	10	8	8
K8	Reputace	0.019	9	9	8
Skalární součin			9.2	7.5	6.3

Zdroj: Vlastní zpracování

Z tabulky jako celku je patrná převaha Power BI téměř ve všech zkoumaných oblastech a také v konečném skalárním součinu jednotlivých vah a bodového ohodnocení.

Z pohledu datových analýz získává nejvíce bodů řešení společnosti Microsoft a to především díky důrazu na self-service přístup, kdy řešení nabízí vysokou kompatibilitu s většinou běžně užívaných zdrojů. Dále nabízí značně intuitivní transformační nástroj Power Query, užitečný především v rámci extenzivní datové analýzy zdroje a získání nového pohledu na zpracovávaná data. Tento zmíněný transformační nástroj skrývá též výhodu pro pokročilé uživatele MS Excel, kteří s tímto nástrojem jsou často již seznámeni a tedy práce s ním v rámci Power BI je pro ně značně jednoduchá a intuitivní.

V rámci sekce vizualizací by bylo řešení Power BI a Tableau značně srovnatelné ve svých možnostech, avšak jako velkým problémem hodnotí autor problém s rychlostí vykreslování dat u Tableau, kdy tento problém je pro zkoumanou společnost zásadní a to především vzhledem k problémům s předchozím řešením.

Co se týče možnosti rozšířené analytiky byly ohodnoceny nejvíce body produkty Power BI a Tableau jenž shodně umožňují dodatečné získání aplikací pro rozšířenou analytiku skrze dodatečně hrazenou službu – v rámci Power BI se jedná o aplikace v rámci MS Azure a v rámci Tableau o aplikace v rámci Tableau Server či Tableau CRM.

V rámci dostupnosti z různých platform jsou nejlépe hodnoceny možnosti produktu Power BI a Tableau, kdy je možné k datům přistupovat jak z desktopové aplikace, webového rohraní tak i z aplikace mobilní.

Z pohledu náročnosti tréninku, tedy průměrně nutnému času pro získání běžné znalosti pro business analytiku a dostupnosti studijních zdrojů, je nejlépe hodnoceno řešení společnosti Microsoft. Je tomu tak z důvodu značné intuitivnosti daného řešení, rozsáhlé znalostní databáze samotné společnosti na webu včetně video návodů a v neposlední řadě značným rozsahem uživatelské základny umožňující jednak značnou odladěnost řešení samotným vydavatelem skrze zpětnou vazbu, tak také znalostní základnu v rámci různých fór, kde uživatelé danou problematiku řeší. V souhrnu tedy má nový uživatel k dispozici poměrně intuitivní řešení a k tomu k dispozici obrovské množství oficiálních a neoficiálních materiálů, které mu umožňují s řešením kvalitně a rychle pracovat.

V rámci cenové politiky vychází nejvýhodněji, co se týče měsíčního poplatku za základní licenci, pro jednotlivého uživatele řešení Power BI s částkou necelých 10USD. Tato částka je samozřejmě diskutovatelná a nemusí být ani konečná, v kontextu dalších služeb této společnosti jako je MS Azure a dalších, které dané BI řešení mohou značně obohatit o další možnosti z pohledu analýz a automatizace. Kladnému hodnocení této oblasti též pomáhá zdarma zkušební verze, kdy oproti konkurenci není omezena časově, ale pouze rozsahem možností především kolaborace, kdy v tomto případě nelze například sdílet projekty a reporty mezi uživateli. Pro prvotní osvojení funkcí a zkoušce vlastností řešení co se týče především datových analýz a vizuálů je to dle autorova názoru řešení Power BI víc než dostatečné.

Z posledního zkoumaného hlediska rozšířenosti a reputace má opět mírný náskok společnost Microsoft. Toto je možné mimo jiné díky dosahu balíku aplikací Microsoft Office, kdy jak bylo zmíněno s některými prvky Power BI řešení již mohli uživatelé Microsoft Excel přijít do styku v rámci doplňků této aplikace a jedná se tedy o poměrně značnou výhodu.

Na základě výše provedených analýz a zhodnocení byl shledán jako nejvhodnější produkt Power BI, společnosti Microsoft. Autor musí konstatovat, že tato volba odpovídá řešení, které společnost na základě svých vlastních analýz vybrala a implementace daného produktu je aktuálně ve svém průběhu. Řešení Power BI tedy bude hlavní součástí finančního reportingu sledované obchodní divize daného podniku.

4.6 Implementace prototypu reportu

V rámci této části práce bude uveden autorův návrh finančního reportu, který by měl obsáhnout základní bázi finančního reportingu skrze manažersky definovaný výkaz zisku a ztráty.

Cílem tohoto reportingu v rámci finančního plánování je především tzv. komparativní analýza dat. V rámci finančního řízení srovnáváme předpoklad nákladů a výnosů oproti aktuálním datům z účetnictví za již uzavřená období. V rámci především nákladového controllingu analyzujeme odchylky od plánovaných a předpokládaných hodnot a hledáme příčiny daného jevu. Stejně tak se posléze stávají tyto data podkladem pro vytváření rozpočtu na budoucí období. Kvalitně sestavený finanční reporting s dostatečným detailem a

srozumitelností poskytuje řídicím pracovníkům potřebné informace pro učinění správných rozhodnutí a vedoucí ve svém důsledku ke kladnému dopadu na výsledek hospodaření dané společnosti.

Pro autorovo vytvoření reportu manažerského výkazu zisku a ztráty (angl. zkr. PnL) bylo využito desktopové Power BI aplikace, poskytující i v rámci bezplatné verze plnohodnotné možnosti transformace dat a tvorby vizuálů, což autor hodnotí značně kladně. Jednotlivé datové transformace a tvorba vztahů mezi jednotlivými tabulkami jsou vytvářeny skrze tzv. datové transformace (v této práci též označováno jako „datový sandbox“), využívající aspekty jazyku Power Query M.

Každý prototyp reportingu má také svůj životní cyklus, kdy nejdůležitějšími milníky jsou:

- 1) stanovení cílů prototypu
- 2) stanovení termínů a odpovědností
- 3) vyhodnocení výsledného prototypu

V návaznosti na cíl 2) byly stanoveny související termíny a odpovědnosti tak aby prototyp byl k dispozici v dostatečném předstihu před implementací finálního řešení a mohly tak být zapracovány získané poznatky a návrhy. Nakonec je provedeno na konci této kapitoly zhodnocení prototypu a jeho přínosu. Co se týče prvního bodu, tedy cílu prototypu, ty jsou obecně definovány (Sherman, 2014) následovně:

- 1) prozkoumání a zvalidování vstupních dat a také samotných výstupů (prozkoumat při stavbě prototypu datové možnosti řešení)
- 2) získat zpětnou vazbu od budoucích uživatelů a dle jejich podnětů upravit zadání a požadavky na budoucí reporting

Počátkem pro stavbu modelu bylo prozkoumání a validace vstupních dat z jednotlivých systémů. Primárním vstupem pro finanční reporting jsou účetní data, které jsou tvořeny v rámci transakčního software. Z těchto dat je možné v plném detailu získat pohled na účetní transakce za uplynulá období. V této databázi se nacházejí data historická, pro komparativní analýzu dat však je třeba mít k dispozici též data týkající se přepokládaného vývoje hodnot tedy angl. forecastu a také dlouhodobě plánovaných hodnot

tzv. budgetu. Tyto data v rámci daného prototypu jsou použita z původního BI software, jenž umožňoval tvorbu scénářů a konsolidaci dat v rámci vlastní integrované databáze. Avšak jak bylo zmíněno v dřívějších pasážích vykazoval problém přílišné agregace bez možnosti například sledovat náklady dle nákladových skupin – tzv. zakázek a mnoho dalších dříve zmíněných problémů především v oblasti výkonnosti řešení. V rámci rozvoje a potřebného vyššího detailu forecastingu potažmo plánování je v rámci implementace nového řešení plánováno z většiny upuštění od tohoto původního zdroje a nahrazení standardizovanými tabulkovými vstupy a cloudovou službou MS OneDrive. Výsledkem má být vyšší detail plánovaných dat a možnost poměrně snadné úpravy struktury vstupů při zachování integrity dat a zachování aspektu nízké ceny.

V rámci prvotního zkoumání datových zdrojů byly po načtení dat do datového modelu nalezeny následující datové struktury.

Obrázek 19 - Struktura tabulek



Zdroj: Vlastní zpracování, Power BI

V prvotní fázi zkoumání a implementace prototypu se v modelu nacházely dvě faktové tabulky „ACT“ a „PLAN_FCST“ a dále několik dimenzních tabulek. Níže je uveden úryvek z výše uvedených dat. Následuje vzorek dat z faktové tabulky.

Obrázek 20 - Vstupní data – tabulka faktů

POSTING_DATE	VAT_DATE	DOCUMENT_TYPE	DOCUMENT_NO	G_L_ACCOUNT_NO	DESCRIPTION	DEBIT_AMOUNT	CREDIT_AMOUNT	AMOUNT	DOCUMENT_DATE
pondělí 31. prosince 2018	pondělí 31. prosince 2018		CZ/2181009	518910	parkovné	20	0	20	pondělí 31. prosince 2018
pondělí 18. prosince 2017	pondělí 18. prosince 2017		CZ/2171018	518910	parkovné	20	0	20	pondělí 18. prosince 2017

Zdroj: Vlastní zpracování, Power BI

Dále je uveden vzorek dat z tabulky dimenzní, v tomto případě uvádějící rozlišení struktury výkazu a ztráty.

Obrázek 21 – Vstupní data - dimenzní tabulka

LEVEL_1	NAME_1	LEVEL_3	NAME_3	LEVEL_6	NAME_6	Pnl_ocekavka	Pnl_ocekavka_suma
5	Náklady	568	Ostatní finanční náklady	568710	odpis půjčky - při postoupení	Other costs	Other costs
5	Náklady	568	Ostatní finanční náklady	568300	poplatky	Other costs	Other costs
5	Náklady	568	Ostatní finanční náklady	568200	poplatky - BCP, SCP	Other costs	Other costs
5	Náklady	568	Ostatní finanční náklady	568700	odpis finančního majetku	Other costs	Other costs
5	Náklady	568	Ostatní finanční náklady	568800	poplatky k dluhopisu	Other costs	Other costs
5	Náklady	518	Ostatní služby	518512	výkony od	Other costs	Other costs
5	Náklady	518	Ostatní služby	518900	zřizovací výdaje do 60 tis.Kč	Other costs	Other costs
5	Náklady	518	Ostatní služby	518302	nájemné leasing	Other costs	Other costs
5	Náklady	518	Ostatní služby	518507	geofyzikální práce	Other costs	Other costs
5	Náklady	518	Ostatní služby	518740	ostatní služby - nedaňové	Other costs	Other costs

Zdroj: Vlastní zpracování, Power BI

Z prvotního zkoumání datových zdrojů vyvstalo několik problémů, které byly třeba řešit. Prvním problémem byly chybějící faktové tabulky, které by obsahovaly formu nově nastaveného PnL reportu. V rámci tohoto prototypu byly tyto tabulky definovány a načteny skrze aplikaci MS Excel, potažmo MS OneDrive. Jedná se o několik tabulek faktů obsahující důležité informace přiřazující jednotlivé transakce do položek PnL a to na základě jak analytického účtu tak definované zakázky. V potaz je také vždy bráno středisko na které je transakce účtována.

Další problémy vyvstaly v rámci dat jako takových, kdy bylo v rámci datových transformací nutné provést jejich patřičnou úpravu a očištění.

Základními úpravami bylo odstranění sloupců, které nemají žádnou potenciální souvislost s dalšími transformacemi a následnými analýzami. Dále byly upraveny

v jednotlivých tabulkách formáty některých sloupců, které měly stanoven neurčitý formát, který by působil problém v rámci dalších transformací – příkladem je sloupec datumu v transakčním software, který neměl daný formát stanoven. V rámci dat z původního BI software byl u položky datum značný problém, kdy nebyl datum definován v běžném formátu. Definice zde byla formou českého názvu měsíce v jednom sloupci a číselném označení roku ve sloupci druhém. Toto bylo vyřešeno přes pomocnou dimenzní tabulku přiřazující jednotlivým česky označeným měsícům, jejich pořadí v roce a po jejím připojení následné vytvoření nového sloupce slučující tyto sloupce do běžné formy datumu. Transformace dat probíhala také v rámci dimenzních tabulek, kdy některé sloupce měly u některých řádků více prvků, oddělených středníkem – toto bylo vyřešeno skrze funkci rozložení textu do sloupců a využití funkcionalit tradičně spojených s kontingenční tabulkou – funkce převést sloupce na řádky a naopak.

Pro co nejjednodušší a nejrychlejší zpracování dat v rámci tohoto prototypu byla zvolena dále cesta doplnění FCST_PLAN tabulky do tabulky ACT, tak aby všechny scénáře které chceme sledovat byly uvedeny v jednom data setu. Toto nemusí být nejideálnější řešení v rámci produkčního řešení, avšak v rámci daného prototypu byl tento způsob vyhodnocen jako dostačující. Před samotným sloučením těchto dvou tabulek byly provedeny úpravy jako změna názvů totožných sloupců na stejný a dále například přidání dimenze „scénář“ do tabulky z transakčního software, kde tato dimenze nebyla uvedena a nebylo by tedy možné rozlišit o jaký scénář se jedná. Stejně tak byl mimo jiné vytvořen v datech v ACT a FCST_PLAN tabulkách sloupec „konec měsíce“ pro následnou lepší orientaci v datech při jejich filtrování v rámci reportů. Mírná nevýhoda této transformace je fakt, že nelze definovat, které sloupce z jednotlivých tabulek chceme sledovat – je tedy nutné po tomto sloučení tabulek smazat ručně větší množství nepotřebných sloupců.

Níže je uvedeno schéma po úpravách zmíněných příklady výše.

Obrázek 24 - Datové transformace – příklad #2

Advanced Editor

act_plan_fcst_asset_subtotals

```
let
Source = Table.NestedJoin(map_pnl_asset, {"id"}, act_plan_fcst_asset, {"pnl_def_asset"}, "act_plan_fcst_asset", JoinKind.FullOuter),
#"Expanded act_plan_fcst_asset" = Table.ExpandTableColumn(Source, "act_plan_fcst_asset", {"POSTING_DATE_EOM", "SCENAR", "AMOUNT_adj_allocated"}, {"act_plan_fcst_asset.POSTING_DATE_EOM",
"act_plan_fcst_asset.SCENAR", "act_plan_fcst_asset.AMOUNT_adj_allocated"}),
#"Removed Columns2" = Table.RemoveColumns(#"Expanded act_plan_fcst_asset",{"pnl_category", "subtotal", "subtotal_ids", "order"}),
#"Filtered Rows - only with pnl id" = Table.SelectRows(#"Removed Columns2", each ([id] <> null)),
#"Pivoted Column" = Table.Pivot(Table.TransformColumnTypes(#"Filtered Rows - only with pnl id", {"id", type text}), "cs-CZ", List.Distinct(Table.TransformColumnTypes(#"Filtered Rows - only with pnl id",
"act_plan_fcst_asset.AMOUNT_adj_allocated", List.Sum),
#"Replaced Value" = Table.ReplaceValue(#"Pivoted Column",null,0,Replacer.ReplaceValue,{"20", "2", "18", "18", "26", "22", "25", "11", "7", "4", "21", "9", "17", "19",
"14", "13", "3", "8", "1", "5", "6", "12", "15", "16", "23", "24", "27", "28"}),
#"Changed Type1" = Table.TransformColumnTypes(#"Replaced Value",{"20", type number}, {"2", type number}, {"18", type number}, {"18", type number},
{"16", type number}, {"22", type number}, {"25", type number}, {"11", type number}, {"7", type number},
{"4", type number}, {"21", type number}, {"9", type number}, {"17", type number}, {"19", type number}, {"14", type number}, {"13", type number}, {"3", type number},
{"8", type number}, {"1", type number}, {"5", type number}, {"6", type number},
{"12", type number}, {"15", type number}, {"16", type number}, {"23", type number}, {"24", type number}, {"27", type number}, {"28", type number})),
#"Added Custom" = Table.AddColumn(#"Changed Type1", "6_new", each [2]+[3]+[4]+[5]),
#"Changed Type3" = Table.TransformColumnTypes(#"Added Custom",{"6_new", type number}),
#"Added Custom2" = Table.AddColumn(#"Changed Type3", "12_new", each [6_new]+[7]+[8]+[9]+[10]+[11]),
#"Changed Type4" = Table.TransformColumnTypes(#"Added Custom2",{"12_new", type number}),
#"Added Custom3" = Table.AddColumn(#"Changed Type4", "15_new", each [13]+[14]),
#"Changed Type5" = Table.TransformColumnTypes(#"Added Custom3",{"15_new", type number}),
#"Added Custom4" = Table.AddColumn(#"Changed Type5", "16_new", each [12_new]+[15_new]),
#"Changed Type6" = Table.TransformColumnTypes(#"Added Custom4",{"16_new", type number}),
#"Added Custom5" = Table.AddColumn(#"Changed Type6", "23_new", each [17]+[18]+[19]+[20]+[21]+[22]),
#"Changed Type7" = Table.TransformColumnTypes(#"Added Custom5",{"23_new", type number}),
#"Added Custom1" = Table.AddColumn(#"Changed Type7", "24_new", each [16_new]+[23_new]),
#"Changed Type8" = Table.TransformColumnTypes(#"Added Custom1",{"24_new", type number}),
#"Added Custom6" = Table.AddColumn(#"Changed Type8", "27_new", each [24_new]+[23]),
#"Changed Type9" = Table.TransformColumnTypes(#"Added Custom6",{"27_new", type number}),
#"Added Custom7" = Table.AddColumn(#"Changed Type9", "28_new", each [27_new]-[21]),
#"Changed Type10" = Table.TransformColumnTypes(#"Added Custom7",{"28_new", type number}),
#"Removed Columns" = Table.RemoveColumns(#"Changed Type10",{"6", "12", "15", "16", "23", "24", "27", "28"}),
#"Renamed Columns" = Table.RenameColumns(#"Removed Columns",{"6_new", "6"}, {"12_new", "12"}, {"15_new", "15"}, {"16_new", "16"}, {"23_new", "23"}, {"24_new", "24"}, {"27_new", "27"}, {"28_new", "28"}),
#"Unpivoted Columns" = Table.UnpivotOtherColumns(#"Renamed Columns", {"act_plan_fcst_asset.POSTING_DATE_EOM", "act_plan_fcst_asset.SCENAR"}, "Attribute", "Value"),
#"Renamed Columns1" = Table.RenameColumns(#"Unpivoted Columns",{"Attribute", "Pnl_ID"}, {"Value", "AMOUNT_adj_allocated"}),
#"Changed Type" = Table.TransformColumnTypes(#"Renamed Columns1",{"Pnl_ID", type number}),
#"Merged Queries" = Table.NestedJoin(#"Changed Type", {"Pnl_ID"}, map_pnl_asset, {"id"}, "map_pnl_asset", JoinKind.LeftOuter),
#"Expanded map_pnl_asset" = Table.ExpandTableColumn(#"Merged Queries", "map_pnl_asset", {"subtotal", "order"}, {"map_pnl_asset.subtotal", "map_pnl_asset.order"}),
#"Filtered Rows1" = Table.SelectRows(#"Expanded map_pnl_asset", each ([map_pnl_asset.subtotal] = true)),
#"Changed Type2" = Table.TransformColumnTypes(#"Filtered Rows1",{"AMOUNT_adj_allocated", type number}),
#"Renamed Columns2" = Table.RenameColumns(#"Changed Type2",{"act_plan_fcst_asset.SCENAR", "SCENAR"}, {"act_plan_fcst_asset.POSTING_DATE_EOM", "POSTING_DATE_EOM"},
{"map_pnl_asset.subtotal", "subtotal"}, {"Pnl_ID", "pnl_def_asset"}),
#"Removed Columns1" = Table.RemoveColumns(#"Renamed Columns2",{"map_pnl_asset.order"})
```

Zdroj: Vlastní zpracování, Power BI

Po jednotlivých datových transformacích nastala fáze tvorby definovaných

vizualizací. Na základě analýz a diskuze v rámci obchodního oddělení byla stanovena struktura reportu tak, aby bylo možno splnit daný cíl prototypu reportu. Tímto cílem je již dříve zmíněné prozkoumání a zvalidování transformovaných dat z jednotlivých zdrojů a získání následné zpětné vazby.

Vizualizace níže již obsahuje plně detail skrze středisko, zakázku až na samotnou účetní transakci. Dalším prvkem je položka scénáře, kdy se tento parametr definuje v rámci drop-down boxu a všechny vykreslená data v následných sloupcích rozdělených dle jednotlivých měsíců patří tomuto scénáři. Vykreslení dat ve sloupcích „cost_center“ až po „description“ obsahující detail transakce je definováno formou blížící se podobě kontingenční tabulky, kdy uživatel rozkliknutím jednotlivých položek má možnost se podívat na detail dané PnL položky až po úroveň jednotlivé transakce.

Obrázek 25 - Vizualizace nového reportu

PnL Overview Detailed									
Year									
Scenario									
Division									
P&L	kCZK	cost_center	order_id	order_name	account	description	Scenario MM/YYYY	Scenario ...	Scenario MM/YYYY
	PnL item 1								
	PnL item 2								
	PnL item 3								
	PnL item 4								
	PnL item 5								
	Total 1								
	PnL item 7								
	PnL item 8								
	PnL item 9								
	PnL item 10								
	PnL item 11								
	Total 2								
	PnL item 13								
	PnL item 14								
	Total 3								
	Total 4								
	PnL item 17								
	PnL item 18								
	PnL item 19								
	PnL item 20								
	PnL item 21								
	PnL item 22								
	Total 5								
	Total 6								
	PnL item 25								
	PnL item 26								
	Total 7								
	Grand Total								

Zdroj: Vlastní zpracování, MS Excel

Dále je níže uvedena tato vizualizace s naznačeným ovládáním daného reportu, tedy tzv. storyboardem.

Obrázek 26 - Storyboard reportu

PnL Overview Detailed									
Year									
Scenario									
Division									
P & L	kCZK	cost_center	order_id	order_name	account	description	Scenario MM/YYYY	Scenario ...	Scenario MM/YYYY
	PnL item 1								
	PnL item 2								
	PnL item 3								
	PnL item 4								
	PnL item 5								
	Total 1								
	PnL item 7								
	PnL item 8								
	PnL item 9								
	PnL item 10								
	PnL item 11								
	Total 2								
	PnL item 13								
	PnL item 14								
	Total 3								
	Total 4								
	PnL item 17								
	PnL item 18								
	PnL item 19								
	PnL item 20								
	PnL item 21								
	PnL item 22								
	Total 5								
	Total 6								
	PnL item 25								
	PnL item 26								
	Total 7								
	Grand Total								

Zdroj: Vlastní zpracování, MS Excel

Z výše uvedených návrhů je patrná přehlednost a jednoduchost v obsluze při zachování plného detailu zkoumaných dat. Důraz na minimalismus je kladen i v samotném barevném rozlišení kdy je report uzpůsoben tak aby byly data dobře čitelná a nebyly v něm žádné rušivé prvky. Obecně využití dimenze a metriky rozlišují transakce v rámci PnL položek, mimo jiné, dle jednotlivých divizí, měsíců a let. Zde je patrné, že v produkčním prostředí bude zcela jistě dále požadována značně rozsáhlá sada reportů, ve kterých bude mimo jiné možnost provést srovnání jednotlivých scénářů a to v plném detailu. V této fázi prototypu reportu však toto není vyžadováno a tento report jsou dostačující pro zkoumání nad daty.

Dále je vyobrazena zmíněná vizualizace v samotném prostředí Power BI.

Obrázek 27 - Výsledný report

Scenario 1.1.2021 1.2.2021

Skutečnost ▾

Year	2021	
PnL	January	February
<input type="checkbox"/> Gas revenue		
<input type="checkbox"/> End customers gas revenue		
<input type="checkbox"/> Cost of sales		
<input type="checkbox"/> IFRS revaluation		
<input type="checkbox"/> Gas trading margin		
<input type="checkbox"/> Gas storing costs		
<input type="checkbox"/> Transportation costs		
<input type="checkbox"/> Broker fees		
<input type="checkbox"/> Trading fees		
<input type="checkbox"/> FX gains and losses		
<input type="checkbox"/> Gas trading margin incl. cos...		
<input type="checkbox"/> Gas flexibility revenue		
<input type="checkbox"/> SoS revenue		
<input type="checkbox"/> Gas flexibility margin		
<input type="checkbox"/> TOTAL TRADING MARGIN		
<input type="checkbox"/> Staff costs		
<input type="checkbox"/> IT expenses		
<input type="checkbox"/> Overhead costs		
<input type="checkbox"/> Trading services		
<input type="checkbox"/> Depreciations		
<input type="checkbox"/> Other costs		
<input type="checkbox"/> Administrative expenses		
<input type="checkbox"/> OPERATING PROFIT		

Zdroj: Vlastní zpracování, Power BI

Níže je uveden detail vizualizace, kde je patrné jak je možné v novém reportu získat plný detail jednotlivých transakcí.

Obrázek 28- Výsledný report - detail

Scenario 1.2.2021 1.2.2021
 Skutečnost

PnL	order_id	account	Year	description	2021
					February
Other costs	105342	501220		...	-166,00
		513720	
		518101	
		518405	
		548770	
		602900	
	105348	518303	
		518409	
		518910	
		829900	
	105360	518910	
	105396	518820	
		521100	
		521200	
		521250		...	-2 140,00
		521410	
		524100		...	-1 140,00
		524200	
		527100		...	-14 507,00

Zdroj : Vlastní zpracování, Power BI

Tento výsledný report byl následně podroben zkoumání daných dat, v především v rámci jejich relevance v rámci definovaného PnL rozlišení. Jedna z částí zkoumání bylo zjištění, zda všechny transakce jsou přiřazeny do PnL položky, která logicky dává smysl v rámci finančního reportingu. Zde bylo nalezeno několik nesouladů plynoucích například z účtování některých účetních transakcí na jiné analytické účty než bylo v rámci definice předpokládáno. Naopak k některým transakcím byla duplicitně přiřazena definice skrze jak zakázkové číslo, tak skrze analytický účet, kdy se vždy jednalo o jinou PnL položku – příkladem byly transakce se směnným kurzem CZK/EUR či CZK/USD. Zde bylo nutné určit, která z definice této zakázky je více relevantní a má smysl ji zachovat. Dalším

problémem byly například chybějící zakázky u transakcí, které měly mít zakázku definovanou.

Obecně prototyp reportu splnil očekávání, které na něj byla kladena. Byl vytvořen dostatečně včas, tak aby poznatky mohly být zakomponovány do produkční verze reportingu nového BI řešení. Dále prototyp splnil očekávání co se týče možnosti a výkonu vizualizací a souvisejících analýz. V neposlední řadě byl splněn předpoklad zachování detailu až na jednotlivé transakce, kdy související aktualizace a vykreslení nových dat se vyznačuje rychlou odezvou a přepočtem.

5 Výsledky a zhodnocení

5.1 Zhodnocení implementace

5.1.1 Zhodnocení infrastruktury

Zhodnocení implementace nového BI řešení lze zhodnotit pouze částečně a to z důvodu ne zcela dokončené realizace řešení v rámci daného podniku ve chvíli finalizace této práce. Přesto stěžejní části nového řešení, jako byla změna datové architektury a souvisejícího zpracování dat, nebo také implementace samotného reportingu, byla ve své finální fázi. Finální výsledky a dopad do finančního řízení můžeme tedy pouze předpokládat v závislosti na aktuálním stavu projektu. Z dosavadního průběhu a aktuálního stavu řešení je patrný razantní posun k možnostem do budoucna flexibilnějšího a efektivnějšího finančního řízení. Z pohledu infrastruktury je patrný posun k vyššímu výkonu a rozšířeným technickým možnostem. Vysoká kompatibilita nového řešení dává možnost i do budoucna pracovat s daty z téměř jakéhokoliv zdroje. Dále přechod z velké části ke cloudovému řešení znamená oproti původnímu, z velké části on-premise řešení, vyšší flexibilitu ve výpočetním výkonu, tak také ve správě infrastruktury. Především využití možnosti variabilního výkonu se ukázalo jako značně efektivním řešením v kontextu potřeb sledovaného podniku. Namísto nutnosti zakoupení, konfigurace a pravidelné obměny on-premise řešení, které je nutné dimenzovat na maximální potenciální zátěž, bylo využití cloudového řešení vyhodnoceno jako rámcově efektivnější. V neposlední řadě také dává nové řešení rozšířené možnosti správy podnikovému IT oddělení, bez nutnosti komunikace veškerých kroků s externím správcem on-premise řešení, jak tomu bylo u řešení původního. Obecně hlavní přínos nově nastavené infrastruktury je znatelně vyšší výkon výpočtů a vykreslení dat při zachování razantně vyššího detailu v podkladových datech.

5.1.2 Zhodnocení reportingu

Z pohledu nákladového řízení a obecně finančního řízení je zde nejpatrnější posun v rámci detailu zpracovávaných dat. Obecně je zjevné, že nové řešení Power BI splňuje požadavky, jež jsou z pohledu nového finančního a nákladového řízení na něj kladeny. Dále

co se týče prototypu reportu, sestavení bylo možné v rámci self-service přístupu samotným business analytikem. Tento přístup byl doceněn mimo jiné z důvodu realizace značné časové úspory v rámci projektu. Tato úspora plyne z faktu, že transformace nad daty byl schopen provádět přímo business analytik, který má značné porozumění daným datům a chápe je v kontextu finančního řízení. Je zde zjevný rozdíl oproti běžnému přístupu, kdy business analytik není plně seznámen s podkladovými daty a možnostmi transformací a je odkázán na zprostředkované informace od pracovníků IT. V tomto případě mohl business analytik samostatně analyzovat podkladová data a formovat vlastní transformace bez obsáhlých znalostí práce s databázemi a mohl tedy poskytnout k realizaci IT oddělení kvalitní a odladěné zadání. Bezpochyby produkční řešení by mělo splňovat vyšší požadavky co se týče sestavení transformací a technického zázemí. Právě v tomto kontextu má IT odborník větší prostor právě na to, se soustředit na tyto technické aspekty a nikoliv na doúpravy původního zadání transformování dat a jejich výstupu. Zadání transformací a definice požadovaného výstupu je totiž již odladěna business analytikem a minimalizuje se tím nutnost extenzivního odladování zadání a aplikace změn. Toto se také dle názoru autora v rámci tohoto prototypu podařilo a zadání pro produkční řešení BI reportingu bylo ve vyšší kvalitě a byla tímto maximalizována časová a peněžní efektivnost v rámci dané implementace. Dále řešení vykresluje data v požadované kvalitě a detailu a má velmi rychlou odezvu při změnách v rámci transformací a vizuálů..

Řešení dále je hodnoceno jako uživatelsky velmi přívětivé, kdy mnoho funkcionalit vizuálů a datových transformací je skrze „drag-and-drop“ přístup a je značně intuitivní. K uživatelské přívětivosti též přispívá značná rozšířenost řešení a tedy dostupnost velkého množství informací v rámci on-line prostředí. Nad rámec tohoto samotný výrobce poskytuje velmi kvalitní dokumentaci na svém webu včetně videonávodů, které novým uživatelům značně pomáhají při orientaci v aplikaci. Kladně hodnocena je dále vyšší přístupnost, kdy mimo desktopovou aplikaci je k dispozici též webová rozhraní nebo mobilní aplikace. Přes toto všechno je však bráno jako hlavní pracovní prostředí, prostředí desktopové případně webové.

Dále je mimo jiné kladně hodnocena vyšší variabilita vizualizačních prvků, kdy jejich výběr je oproti původnímu řešení značně vyšší a je tedy možné tvořit větší počet kvalitních vizuálů odpovídajícím lépe rozmanitým účelům daných reportů.

V rámci dalšího rozvoje je nasnadě několik prvků z aktuálních trendů v rámci BI řešení, které se na trhu vyskytují. Jedním z prvků je automatizace či tzv. hyperautomatizace, kdy může být řeč v rámci rodiny Microsoft o řešení MS Power Automate. V aktuální infrastruktuře je související automatizace načítání dat či zasílání notifikací spravována skrze vícero externích aplikací na základě historických implementací. Z pohledu finančního reportingu je několik aspektů na něž v budoucnu bude kladen důraz a mají potenciál dalšího rozvoje. Prvním aspektem je rozvoj forecastingu a budgetingu tak, aby již nebyly ve značné míře třeba pro přípravu těchto dat pomocné tabulky především v rámci MS Excel. V současném řešení již je na značně kvalitní úrovni zpracováno a automatizováno zpracování dat z účetního (transakčního) software. V oblasti forecastingu a budgetingu však jsou stále další možnosti pro další zlepšení. Jedním z důvodů, proč tomu tak je, je mimo jiné prozatím nepropojení obchodního software, sledující pohyby skladů jednotlivých komodit a dalších ukazatelů, díky kterým je možné předpokládat další finanční vývoj celé divize a tvořit právě její forecast či budget. Možností v této fázi je mnoho, ať již cesta rozvoje současného transakčního software, který je též z rodiny Microsoft, nebo možnosti samostatného „customizovaného“ řešení, které by bylo plně uzpůsobeno požadavkům obchodního oddělení.

Dalším prvkem budoucího rozvoje by mohla být vyšší frekvence aktualizace dat z jednotlivých systémů. Tento bod je však poněkud diskutabilní. Přesto, že je tento aspekt poměrně často zmiňován a real-time aktualizace jsou označovány za formu trendu, z pohledu finančního reportingu v rámci daného podniku to takto vnímáno není. Důvodem je workflow, které souvisí se zadáváním účetních záznamů, kdy do chvíle úplné závěrky posledního měsíce nejsou data za tento měsíc plně prokazatelná a validní v kontextu PnL výkazu. V kontextu tohoto je bráno aktuálně za plně dostačující dávková aktualizace dat v jednodenní frekvenci.

Z pohledu nových zdrojů bylo již výše zmíněno propojení nových databází jako je obchodní databáze obsahující data o jednotlivých obchodech, stavech zásob a mnoha dalších prvků, jež by usnadnila ve značné míře finanční plánování a případně i nákladové řízení. Z pohledu nového řešení Power BI je zde patrná výhoda při propojení nových zdrojů a to z pohledu vysoké kompatibility řešení tak také možností self-service. Toto by znamenalo

vyšší efektivitu v rámci prozkoumání nového zdroje business analytiky a využití maximálního potenciálu, zároveň při minimalizaci nákladů a času při dané implementaci.

V neposlední řadě stojí za zmínku možnosti rozšířené analytiky, které je v rámci řešení Power BI do značné míry spjato s cloudovou službou MS Azure, která samozřejmě přichází také s dodatečnými náklady na dané řešení. V rámci této služby jsou nabízeny možnosti dodatečného výpočetního výkonu současně s prvky strojového učení, AI analýzou, IoT a mnoha dalších prvků a aplikací.

6 Závěr

Důležitost podkladových dat a znalost možností jejich transformací a rozšířených analýz je jedna ze stěžejních znalostí, která se stává předpokladem pro efektivní a kvalitní finanční analýzy dnešní doby. V rámci této práce je patrná v tomto směru akcentace trendu sbližování finančního a datového řízení.

Hlavním výsledkem této práce je zhodnocení vybraných řešení dostupných na trhu dle zvolené metodiky. Z analýzy potřeb firmy vzešlo několik oblastí zájmu, kterým byly Saatyho metodou párového porovnání přiřazeny jednotlivé váhy. Následně byly dle metodiky společnosti Gartner zvoleny 3 potenciální BI řešení označované jako „lídři“, tak aby byly vzápětí tyto řešení ohodnoceny bodovací metodou a váženým součtem s danými vahami bylo zvoleno nejvhodnější z nich. Z dostupných řešení bylo vybráno řešení Power BI společnosti Microsoft. Přidanou hodnotou pro danou společnost je poskytnutá zpětná vazba k jejich finálnímu výběru nového BI řešení. Bylo zjištěno, že výběr, který se ukázal jako ideální z pohledu výše zmíněné metodiky odpovídá i konečnému výběru dané společnosti.

Dalším výsledkem této práce je vytvoření prototypu nového finančního reportu v kontextu možností nově aplikovaného BI řešení. Díky nabytým teoretickým znalostem z oblasti dimenzního modelování a databázového vnímání dat byly optimálně vybrány vstupy a provedeny následné transformace pro co nejoptimálnější výstup odpovídající vysokým analytickým nárokům v rámci finančního řízení. Zpracování prototypu reportu dále vedlo k efektivnějšímu průběhu implementace produkčních finančních reportů, kdy zadání poskytnuté zpracovateli nového řešení bylo zpracováno již na základě obsáhlých datových analýz a zkoumání. Zadání tedy bylo značně odladěno z pohledu finančního řízení a výstupy byly již při prvotních implementacích ve vysoké kvalitě, což vedlo k úspoře času a prostředků při jeho implementaci. Toto mohlo být dosaženo důrazem nového řešení na tzv. self-service, kdy business analytik je schopen provést rozsáhlou analýzu dat a související transformace v rámci intuitivního prostředí bez rozsáhlé znalosti SQL či jiného dotazovacího jazyka.

Obecným přínosem nového řešení pro daný podnik je mimo jiné nový finanční reporting ve vyšším detailu poskytující možnosti agilnějšího finančního řízení, snížení

zbytných nákladů a v konečném důsledku vedoucí k realizaci vyššího zisku společnosti. Tomuto nové řešení dopomáhá výrazným zlepšením v definovaných oblastech jako je výkon při aktualizacích dat, výrazně vyšší detail zpracovaných transakcí i ve finálních reportech, self-service přístupem, rozsáhlými možnostmi vizualizací, rozhraní na platformách jako webový prohlížeč či mobilní zařízení a související nové možnosti kolaborace mezi uživateli. Rychlá odezva jednotlivých vizualizací a také celého řešení v kombinaci se self-service možnostmi, kdy business analytik je schopen sám provádět rozsáhlé datové průzkumy a následné transformace dává společnosti v oblasti finančního řízení znatelně agilnější možnosti. Tato výhoda přetrvává i do budoucna při připojování nových zdrojů a jejich souvisejících analýz, kdy tyto nové možnosti pomáhají šetřit čas a náklady při implementaci nového zdroje.

Potenciálními budoucími kroky v rámci daného podniku a nového BI řešení je mimo jiné připojení dalších obchodních systémů. Tento krok by dále zefektivnil forecasting a budgeting v rámci divizního plánování, kdy by se zvýšila spolehlivost dat a jejich transformací ruku v ruce s nižší časovou náročností celého procesu. Nové BI řešení též otevírá možnosti v oblasti rozšířené analytiky a strojového učení a to v rámci kombinace s cloudovou službou MS Azure či další rozšířenou automatizací v rámci finančního reportingu skrze kupříkladu MS Automate.

7 Seznam použitých zdrojů

Amazon. 2020. Data Warehouse Concepts. [Online] 05. 12 2020. <https://aws.amazon.com/data-warehouse/>.

Bismart. 2020. What Do We Do? - ETL. [Online] 2020. <https://blog.bismart.com/en/what-do-we-do-etl>.

Capterra. 2021. Business Intelligence Software. [Online] 15. 01 2021. <https://www.capterra.com/business-intelligence-software/#top-20>.

CIO Business World. 2020. Business Intelligence a datové sklady - dodavatelé řešení. [Online] 2020. <https://katalog-ict-reseni.businessworld.cz/prehledy/business-intelligence-a-datove-sklady-vypis-vsech-spolecnosti/>.

DataFloq. 2020. Intelligent Reporting - How Big Data and AI Are Changing the Way We Create and Analyze Reports. [Online] 05. 09 2020. <https://datafloq.com/read/intelligent-reporting-big-data-ai-change-reports/6647>.

DataPine. 2020. Top 10 Analytics And Business Intelligence Trends For 2021. [Online] 25. 11 2020. <https://www.datapine.com/blog/business-intelligence-trends/>.

Diagrams.net. 2021. Draw.io. [Online] 2021. <https://app.diagrams.net/>.

Fibírová, Jana a Šoljaková, Libuše. 2010. *Reporting*. Praha : Grada, 2010. 978-80-247-6708-6.

G2. 2021. [Online] 20. 01 2021. <https://www.g2.com/categories/analytics-platforms>.

Gartner. 2021. <https://www.gartner.com/reviews/market/analytics-business-intelligence-platforms>. [Online] 25. 01 2021. <https://www.gartner.com/reviews/market/analytics-business-intelligence-platforms>.

—. 2020. Magic Quadrant for Analytics and Business Intelligence Platforms. [Online] 1. 2 2020. <https://www.gartner.com/doc/reprints?id=1-1XYUYQ3I&ct=191219&st=s>.

Heinze, Justin. 2020. History of Business Intelligence. *BetterBuys*. [Online] BetterBuys, 2020. <https://www.betterbuys.com/bi/history-of-business-intelligence/>.

IBM. 2020. IBM Cognos Analytics. [Online] 2020. <https://www.ibm.com/products/cognos-analytics>.

ICT Revue. 2017. BI v odpovědích dodavatelů podnikových řešení v ČR. [Online] 2017. https://ictrevue.ihned.cz/c3-65759900-0ICT00_d-65759900-bi-v-odpovedich-dodavatelu-podnikovych-reseni-v-cr.

Inekon Systems. 2020. Tableau. [Online] 2020. <https://www.inekon-systems.cz/tableau/>.

Ironside Group. 2020. Is “Self-Service” Analytics the New “Big” Data? [Online] 2020. <https://www.ironsidegroup.com/2019/02/19/self-service-analytics-new-big-data/>.

KOMIX s.r.o. Qlik - Naši zákazníci. [Online] <https://www.analyzyareporting.cz/nasi-zakaznici/>.

Laberge, Robert. 2012. *Datové sklady Agilní metody a business intelligence*. Praha : Computer Press, 2012. 978-80-251-3729-1.

Microsoft Corporation. 2020. Azure products. [Online] 2020. <https://azure.microsoft.com/en-us/services/>.

—, **2020.** Power BI and Office. [Online] 2020. <https://powerbi.microsoft.com/cs-cz/power-bi-and-office/>.

—, **2020.** Software jako služba. [Online] 05. 10 2020. <https://azure.microsoft.com/cs-cz/overview/what-is-saas/>.

—, **2020.** Understand star schema and the importance for Power BI. [Online] 06. 12 2020. <https://docs.microsoft.com/cs-cz/power-bi/guidance/star-schema>.

—, **2020.** What is Power BI? [Online] 2020. <https://powerbi.microsoft.com/cs-cz/what-is-power-bi/>.

Microsoft. 2021. Power BI. [Online] 13. 02 2021. <https://powerbi.microsoft.com/en-us/desktop/>.

Oracle Corporation. 2020. Business Intelligence (BI) - Discover New Insights. [Online] 2020. <https://www.oracle.com/cz/business-analytics/business-intelligence/>.

—, **2020.** Oracle Exadata. [Online] 2020. <https://www.oracle.com/cz/engineered-systems/exadata/>.

Oracle. 2020. Star and Snowflake Schemas. [Online] 01. 12 2020. https://www.oracle.com/webfolder/technetwork/tutorials/obe/db/10g/r2/owb/owb10gr2_gs/owb/lesson3/starandsnowflake.htm.

—. 2020. What is a data warehouse. [Online] 06. 12 2020.
<https://www.oracle.com/cz/database/what-is-a-data-warehouse/>.

Qlik. 2020. Data Analytics Products. [Online] 2020.
<https://www.qlik.com/us/products/analytics-products>.

—. Qlik Sense® | Data Analytics Platform. [Online] [Citace: 30. 01 2021.]
<https://www.qlik.com/us/products/qlik-sense>.

SAP. 2020. SAP BusinessObjects BI 4.3 Now Available. [Online] 2020.
<https://blogs.sap.com/2020/06/16/sap-businessobjects-bi-4.3-now-available/>.

—. 2020. SAP BusinessObjects Business Intelligence suite. [Online] 2020.
<https://www.sap.com/products/bi-platform/technical-information.html#integration>.

SelectHub. 2021. Best Business Analytics Software. [Online] 05. 01 2021.
<https://www.selecthub.com/business-analytics-tools/>.

Sherman, Rick. 2014. *Business Intelligence Guidebook - From Data Integration to Analytics*. Waltham : Morgan Kaufmann, 2014. 978-0124114616.

Slánský, David, Pour, Jan a Novotný, Ota. 2005. *Business Intelligence - Jak využít bohatství ve vašich datech*. Praha : Grada, 2005. 80-247-1094-3.

SystemOnline. 2018. Využití nástrojů a aplikací Business Intelligence v podnicích ČR. [Online] 2018. <https://m.systemonline.cz/business-intelligence/vyuziti-nastroju-a-aplikaci-bi-v-podnicich-cr.htm>.

Tableau. 2020. Tableau Desktop. [Online] 2020.
<https://www.tableau.com/products/desktop>.

—. Tableau for Finance Analytics. [Online] [Citace: 19. 01 2020.]
<https://public.tableau.com/en-us/s/tableau-finance>.

Vochozka, Marek a Mulač, Petr. 2012. *Podniková ekonomika*. Praha : Grada, 2012. 978-80-247-4372-1.