

**M A S A R Y K O V A
U N I V E R Z I T A**

FILOZOFICKÁ FAKULTA

KATEDRA INFORMAČNÍCH STUDIÍ A KNIHOVNICTVÍ

Projekt diplomové práce

Zavádění datového skladu v rámci Masarykovy Univerzity

Magisterská práce

Pavel Dytrych

527565

Vedoucí práce:

Ing. Filip Janovič, PhD., MBA

Anglický název tématu: Implementation of the data warehouse within Masaryk University

1. (Výzkumný) problém teoreticko-aplikační DP

V rámci diplomové práce se budu věnovat analýze a evaluaci procesů zavádění služby datového skladu na Ústavu výpočetní techniky Masarykovy univerzity v Brně a jeho praktickému využití ke tvorbě reportů určených k optimalizaci chodu univerzity.

Práce bude rozdělena do dvou částí, a to části teoretické a aplikační. Teoretická část bude obsahovat ujasnění a ukotvení potřebné terminologie týkající se datových skladů a managementu služeb a dále se bude věnovat samotným procesům zavádění a údržby datového skladu jako služby dle ITSM, respektive dle frameworku ITIL (Carlidge, Hanna, Rudd, Macfarlane & Rance, 2007).

Aplikační část se pak bude věnovat analýze, evaluaci a případné optimalizaci procesu návrhu a implementace nového případu využití datového skladu, a to z pohledu uživatele, na jehož podnět je tento uživatelský případ vytvářen.

2. Přehled literatury, výzkumný kontext, popř. projekty

a) Teoretické ukotvení (výchozí koncepty, teorie)

Carlidge, A., Hanna, A., Rudd, C., Macfarlane, I., & Rance, S. (2007). *An Introductory Overview of ITIL® V3* [Online]. London, United Kingdom: The UK Chapter of the itSMF. Retrieved from https://itil.it.utah.edu/itilv3/docs/itSMF_ITILV3_Intro_Overview.pdf

Tato kniha obsahuje stručný úvod do frameworku ITIL v jeho třetí verzi. Jejím obsahem je ukotvení pojmů, jako je „Service management“ a popis samotného frameworku ITIL, z jeho pohledu se pak příručka věnuje problematice designu služeb, jejich provozování a kontinuálnímu zlepšování. Součástí příručky je také seznam souvisejících standardů a dalších zdrojů, které se týkají například „best practices“ v rámci frameworku ITIL.

Matula, J. (2017). *Informační management: normy, frameworky a nejlepší praxe v řízení služeb IT (ITSM)*. V Opavě: Slezská univerzita, Filozoficko-přírodovědecká fakulta v Opavě, Ústav bohemistiky a knihovnictví.

Autor v této knize poskytuje ucelený náhled na problematiku managementu v IT sektoru v rámci ITSM. Kniha nabízí ucelený souhrn vývoje IT managementu včetně jeho historického vývoje a příbuzných témat. Kniha se zaměřuje na zejména normu ISO 2000, třetí verzi frameworku ITIL a odlehčený standard FitSM, u kterých rozebírá jednotlivé části a uvádí praktické informace o jejich využití, a to včetně příkladů „best practices“.

Inmon, W. H. (2005). *Building the Data Warehouse* (4th.). Hoboken (New Jersey): Wiley.

Kniha, ve svém čtvrtém vydání, se věnuje designu datových skladů, a to od úplného začátku, kdy popisuje postupnou evoluci samotného datového skladu včetně problematiky přirozeného, a tedy nedesignovaného vývoje, přes jeho samotný návrh a jeho důležité aspekty, jako je struktura a granularita dat nebo metodika přístupu a pokračuje dále až po finální technické řešení, včetně distribuovaných datových skladů, jejich nasazení a například propojení s klasickým webem.

Kimball, R., & Ross, M. (2013). *The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling* (3rd.). Hoboken (New Jersey): Wiley.

Kniha se primárně zabývá designem samotné struktury a modelováním dat se zaměřením na schéma datové hvězdy. Součástí knihy je přehled designových technik, určených k návrhu datových schémat., které jsou následně ukázány a vysvětleny na velkém množství příkladů, typických pro jejich použití. Součástí třetího vydání jsou také kapitoly, které se týkají ETL procesu, tedy samotného zpracování dat a také kapitola týkající se analýzy tzv. „velkých dat“.

Kratochvíl, T. (2014). Principy tvorby a užití dashboardů. *Systémová Integrace*, (21), 104-122. Retrieved from <https://docplayer.cz/6122466-Principy-tvorby-a-uzitidashboardu.html>

Autor se v tomto článku zabývá návrhem informačního dashboardu v systému SAP Dashboards, který není pro diplomovou práci relevantní, teoretická část se však zabývá návrhem dashboardu z pohledu business intelligence a je přenositelná i na jiné platformy. Článek kategorizuje dashboardy dle jejich užití, definuje důležité ukazatele a věnuje se i motivacím a výhodám, které kvalitně vytvořený dashboard přináší.

Augutter, C. (2013). *ITIL Lifecycle Essentials: Your Essential Guide for the ITIL Foundation Exam and Beyond*. Cambridgeshire, UK: IT Governance Publishing.

Kniha popisuje kompletní životní cyklus IT služby z pohledu frameworku ITIL. Autor v knize postupně prochází všechny fáze, kterými musí nově nasazovaná služba projít a do detailu je rozebírá. Nad standardních pět etap vývoje nové služby tak, jak jsou popsány v ITIL přidává kniha ještě několik dalších, které se věnují například automatizovanému managementu služeb anebo školení samotného managementu služeb.

Polaine, A., Løvlie, L., & Reason, B. (2013). *Service design: from insight to implementation / by Andy Polaine, Lavrans Løvlie and Ben Reason*.

Kniha se nejprve věnuje rozdílům mezi pojmy služba a produkt, poté se věnuje praktickým aspektům návrhu služeb, autoři detailně popisují několik konkrétních metod, využívaných k jejich návrhu a implementaci. Součástí knihy je i několik případových studií, na kterých je samotný proces ukázán prakticky, jako je například redesign metra v norském Oslu.

b) Výzkumné články, studie, projekty

F. Lapura, E. V., Fernandez, J. K. J., Pagatpat, M. J. K., & Dinawanao, D. D. (2018). Development of a University Financial Data Warehouse and its Visualization Tool. *Procedia Computer Science*, 135, 587-595. <https://doi.org/10.1016/j.procs.2018.08.229>

Článek se zabývá projektem zavádění datového skladu na univerzitě v Iliganském Technologickém Institutu. Článek je z části věnován důvodům a motivaci zavádění datového skladu v rámci této instituce, dále pak technologickému řešení a výsledné efektivitě v doručování informací. Obecně je v článku řešen velmi podobný případ, jako je hlavní náplní této diplomové práce.

Kovacevic, S., & Orucevic, F. (2010). Using ITILv3 methodology for implementing new email services in operator for producing and distributing electrical energy. *Proceedings Of The Iti 2010, 32Nd International Conference On Information Technology Interfaces, Information Technology Interfaces (Iti)*, 2010, 147-154. Retrieved from <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5546379>.

Konferenční článek, který se zabývá projektem zavádění nové emailové služby dle pravidel frameworku ITIL v3 ve společnosti distribuující elektřinu. Článek postupně prochází všech pět základní etap zavádění nové služby dle ITIL jako je například design služby, tranzice anebo kontinuální a popisuje, jak byla služba zaváděna. V článku jsou uvedeny převážně finální výstupy spolu s původními problémy, která má nová služba řešit.

c) Diplomové práce

Guoth, M. (2019). *Návrh a implementace datového skladu pro analýzu objemných heterogenních dat* (Diplomová práce). Brno. Retrieved from <https://is.muni.cz/th/t026i/>

Diplomová práce se věnuje praktickému návrhu a implementaci prototypu datového skladu pro reálný projekt ANALÝZA, který byl řešen ve spolupráci s Policií ČR. Autor v

práci navrhuje samotnou strukturu a koncepci datového skladu, architektonické a technické řešení a dotýká se i samotné vizualizace a analýzy dat. Výsledkem práce je tedy funkční prototyp datového skladu, umožňující nahrávání a získávání souborů.

Špringerová, P. (2020). *Analýza IT služeb Masarykovy univerzity z pohledu jejich vlastníků* (Diplomová práce). Brno. Retrieved from <https://is.muni.cz/th/n22js/>.

Autorka se v diplomové práci zabývá analýzou kvality IT služeb na Ústavu výpočetní techniky MUNI z pohledu jejich vlastníků a ITSM, konkrétně pomocí frameworků ITIL a COBIT. Kromě teoretické části se věnuje kvalitativnímu výzkumu mezi vlastníky služeb, s cílem zmapovat problémy a příležitosti související s poskytováním služeb.

Dedrlé, T. (2015). *Analýza a dokumentace procesů provozu služeb v IT organizaci podle ITIL* (Diplomová práce). Brno. Retrieved from <https://is.muni.cz/th/w2996/>

Diplomová práce se zabývá implementací frameworku ITIL. Z teoretického hlediska práce věnuje popisu samotného frameworku a mapuje procesy, které framework zavádí, z praktického hlediska pak práce ukazuje praktické nasazení ITIL pro službu IP telefonie, hodnocení již existujících služeb a následné vyhodnocení zlepšení kvality služeb a spokojenosti zákazníků.

3. Opodstatnění tématu, význam a dopad diplomové práce

Všechny velké organizace, společnosti anebo instituce, jako je Masarykova univerzita vyprodukuje v rámci svého chodu velké množství surových dat, která mohou být různá, co se týče jak jejich charakteru, tak i výstupních formátů.

V případě Masarykovy univerzity se typicky jedná o data provozního charakteru, jako jsou data pocházející z monitoringu uživatelů, vytíženosti jednotlivých budov a kanceláří, dále pak data finančního rázu a v neposlední řadě také data týkající se výkonu samotné organizace, jako je například úspěšnost studentů, citovanost publikací a jiné.

Zpracování těchto dat může poskytnout velmi cenný náhled na funkčnost celé organizace, jejich jednotlivých částí a následně pak pomoci s optimalizací interních procesů.

Vzhledem k tomu, že množství takto generovaných dat je v rámci instituce velikosti Masarykovy univerzity enormní a jednotlivé zdroje dat nejsou většinou nijak standardizovány, je jejich zpracování značně komplikované.

Za účelem zjednodušení a zlepšení možností analýzy těchto dat, je momentálně v rámci Masarykovy univerzity zaváděna služba datového skladu, který má za hlavní cíl data sjednotit na jedno místo, strukturalizovat, vyčistit a následně zpracovat do vhodných reportů, které poskytnou uživatelům souhrnný přehled o dané problematice.

Samotná diplomová práce se bude věnovat analýze služby takového rozsahu, a to jak z hlediska managementu služeb dle frameworku ITIL, tak z hlediska technického a tvorby samotného analytického nástroje.

Výstupem diplomové práce tak bude souhrnný popis nové služby v rámci instituce velikosti Masarykovy univerzity, a to jak z hlediska jejího provozu dle ITSM, tak i z hlediska technického.

4. (Výzkumné) otázky, záměr a cíle

Hlavním cílem týkajícím se teoretické části diplomové práce je, zdali je služba datového skladu IT služeb Masarykovy univerzity provozována optimálně a efektivně a zda je možné ji dále optimalizovat z pohledu ITSM.

Cílem aplikační části je pak analýza technického řešení datového skladu a jeho následné využití v podobě návrhu a vytvoření uživatelského případu, využívajícího službu datového skladu, která se bude věnovat zpracování provozně – ekonomických dat.

5. Aplikační část/projekt

Aplikační část se bude věnovat zpracování konkrétního případu využívajícího datový sklad. Zpracovávaný požadavek se bude zabývat reportem vizualizujícím provozní data.

Report bude postaven na datech, která jsou uložena v datovém skladu, jako primární zdroj těchto slouží informační systém ISEP, ze kterého se data nahrají do datového skladu a provedou se potřebné úpravy a transformace do schématu datové hvězdy (ETL proces).

Z takto zpracovaných dat se následně vytvoří report, který bude mít podobu dynamického reportu v PowerBI a bude vycházet z požadavků uživatel.

Report bude vytvářen ve spolupráci s týmem správců v čele s Bc. Filipem Švaříčkem a ve spolupráci s konzultantem Ing. Jindřichem Zechmeisterem.

6. Časový harmonogram

Vzhledem k tomu, že se diplomová práce bude dělit na teoretickou a aplikační část, bude prvotním úkolem sepsání potřebné teorie a ukotvení pojmů, které budou potřeba jak v pokračování teoretické, tak v aplikační části.

Tato teoretická příprava by měla být hotová během února roku 2023 s tím, že by ve stejném období bylo vhodné začít s prací i na samotném mapování procesu zavádění datového skladu a sběru požadavků na jeho využití.

Samotná aplikační část, která se týká implementace jednoho konkrétního případu využití datového skladu. V tomto případě se bude jednat o provozní data, jejich zpracování a vytvoření výsledného reportu. Tato část by v ideální případě měla začít nejpozději na přelomu druhého a třetího čtvrtletí 2023 a celkové dokončení diplomové práce na podzim téhož roku.

7. Zdroje a vybavení

Hlavním konzultantem diplomové práce bude Ing. Jindřich Zechmeister z Ústavu výpočetní techniky Masarykovy univerzity a dále pak Bc. Filip Švaříček, který má správu služby datového skladu na starosti.

Pro tvorbu samotného reportu pak budou využita anonymizovaná data ze systému ISEP, přístup k datovému skladu, nástroj pro práci s daty v jazyce SQL a analytický nástroj PowerBI.

8. Seznam použité literatury

Datawarehouse [Online]. Retrieved January 15, 2023, from <https://it.muni.cz/sluzby/datawarehouse>

Carlidge, A., Hanna, A., Rudd, C., Macfarlane, I., & Rance, S. (2007). *An Introductory Overview of ITIL® V3* [Online]. London, United Kingdom: The UK Chapter of the itSMF. Retrieved from https://itil.it.utah.edu/itilv3/docs/itSMF_ITILV3_Intro_Overview.pdf

Inmon, W. H., & . (2005). *Building the Data Warehouse* (4th.). Hoboken (New Jersey): Wiley.

Augutter, C. (2013). *ITIL Lifecycle Essentials: Your Essential Guide for the ITIL Foundation Exam and Beyond*. Cambridgeshire, UK: IT Governance Publishing.

~~9. Zařazení do profilace~~