Vysoká škola ekonomická v Praze

Fakulta informatiky a statistiky



Metodika integrace dat do Google Analytics

DIPLOMOVÁ PRÁCE

Studijní program: Aplikovaná informatika

Studijní obor: Informační systémy a technologie

Autor: Bc. Marek Neufus

Vedoucí diplomové práce: Ing. Libor Gála, Ph.D.

Praha, duben 2019

Prohlášení

Prohlašuji, že jsem diplomovou práci Metodika integrace dat do Google Analytics vypracoval samostatně za použití v práci uvedených pramenů a literatury.

V Praze dne 15. dubna 2019 .....

Bc. Marek Neufus

Poděkování

Rád bych touto cestou poděkoval vedoucímu diplomové práce panu Ing. Liborovi Gálovi, Ph.D. za odborné vedení, věcné rady a věnovaný čas při řešení problematiky této práce.

Abstrakt

Tato diplomová práce vymezuje možnosti zpracování dat aplikací Google Analytics a navrhuje způsob napojení dat z podnikových systémů formou datové integrace. Proces integrace dat do Google Analytics je popsán metodikou, jež je hlavním výstupem diplomové práce. Formalizace navržené metodiky je zajištěna souladem s metodikou MMSP. Motivací pro tuto diplomovou práci je fakt, že v podnikových systémech se obecně nachází celá řada různorodých informací, které lze přiřadit relevantnímu uživateli webové aplikace zaznamenanému analytickým nástrojem Google Analytics. Takto propojená a obohacená data je možno následně zužitkovat pro podporu manažerského rozhodování, optimalizaci webového obsahu, či pro online marketingové aktivity v inzertním systému Google Ads formou přesnějšího cílení a personalizovanější inzerce.

Klíčová slova

Google Analytics, webová analytika, datová integrace, metodika, online marketing

JEL klasifikace

L86, M39

Abstract

This diploma thesis describes the data processing capabilities of Google Analytics and suggests how to connect data from enterprise systems via data integration. The process of data integration into Google Analytics is described by the methodology that is the main output of the thesis. Formalization of the proposed methodology is ensured by compliance with the MMSP methodology. The motivation for this diploma thesis is the fact that business systems generally have a wide range of information that can be assigned to a relevant web application user recorded by the Google Analytics. This linked and enriched data can then be used to support managerial decision-making, content optimization or online marketing activities in the Google Ads advertising system by more accurate targeting and personalized advertising.

Keywords

Google Analytics, web analytics, data integration, methodology, online marketing

JEL Classification

L86, M39

Obsah

Úvod ... 8
1 Rozbor problému ... 9
1.1 Obecná charakteristika problému ... 9
1.2 Technologická charakteristika problému ... 10
1.3 Možné případy užití ... 11
2 Charakteristika komponent GA ... 14
2.1 Datový model Google Analytics ... 14
2.1.1 Rozsah dimenzí a metrik ... 17
2.1.2 Typy požadavků na server ... 19
2.2 Metody integrace ... 21
2.2.1 Datový import ... 22
2.2.2 Rozšířený měřící kód ... 23
2.2.3 Measurement Protocol ... 27
2.2.4 Sumarizace integračních metod ... 29
2.3 Legislativní a etické regulace ... 30
2.3.1 Legislativa ... 30
2.3.2 Samoregulace ... 31
2.4 Shrnutí problému ... 32
3 Charakteristika současného stavu řešení ... 35
3.1 Charakteristika současného poznání dílčích komponent GA ... 35
3.2 Charakteristika současného stavu poznání v globální rovině ... 36
3.3 Shrnutí současného stavu řešení ... 39
4 Metodika diplomové práce ... 40
4.1 Obsahové vymezení metodiky ... 41
4.2 Programové prostředí pro sestavení metodiky ... 42
4.3 Validace platnosti metodiky ... 43
5 Návrh metodiky integrace dat ... 44
5.1 Životní cyklus metodiky ... 44
5.1.1 Fáze Zahájení ... 45
5.1.2 Fáze Rozpracování ... 46
5.1.3 Fáze Konstrukce ... 47
5.1.4 Fáze Zavedení ... 48
5.2 Role metodiky ... 49
5.2.1 Role Systémový integrátor ... 49
5.2.2 Role Datový steward ... 50

5.2.3 Role Webový analytik 51
5.2.4 Role Právník 51
5.2.5 Role Online marketingový specialista 52
5.2.6 Role Databázový specialista 52
5.2.7 Role Vývojář 53
5.2.8 Role Zainteresovaná strana 53
5.3 Disciplíny metodiky a jejich úlohy 54
5.3.1 Disciplína Požadavky a její úlohy 54
5.3.2 Disciplína Řízení projektu a její úlohy 55
5.3.3 Disciplína Dokumentace a její úlohy 56
5.3.4 Disciplína Vývoj a její úlohy 57
5.3.5 Disciplína Testování a její úlohy 57
5.3.6 Disciplína Aplikace a její úlohy 58
5.3.7 Disciplína Řízení změn a její úlohy 58
5.4 Pracovní produkty metodiky 59
5.4.1 Pracovní produkty disciplíny Požadavky 59
5.4.2 Pracovní produkty disciplíny Řízení projektu 60
5.4.3 Pracovní produkty disciplíny Dokumentace 61
5.4.4 Pracovní produkty disciplíny Vývoj 62
5.4.5 Pracovní produkty disciplíny Testování 63
5.4.6 Pracovní produkty disciplíny Aplikace 64
5.4.7 Pracovní produkty disciplíny Řízení změn 64
6 Ukázka užití metodiky 65
6.1 Popis případu užití 65
6.2 Aplikace metodiky 66
6.2.1 Fáze Zahájení 66
6.2.2 Fáze Rozpracování 69
6.2.3 Fáze Konstrukce 71
6.2.4 Fáze Zavedení 73
6.3 Shrnutí aplikace metodiky 74
Závěr 75
Použitá literatura 76
Přílohy I
Příloha A: Vazby mezi prvky navržené metodiky I
Příloha B: Webová verze metodiky V
Příloha C: Zdrojové soubory metodiky VI

Úvod

Podnikání v dnešní digitální době nalézá stále větší oporu v datech, kterých je v podniku a jeho okolí obrovské množství. Moderní podnik tato data zpracovává a na základě výsledků jejich analýz provádí příslušná manažerská rozhodnutí. Hlavním zdrojem těchto dat je typicky zákazník, který se dostává do kontaktu s podnikem celou řadou různých cest. Může se jednat o webové stránky podniku, aplikaci, call centrum, e-mailovou podporu a jiné. Podnik se současně snaží veškerá zákaznická data a požadavky shromažďovat na jednom místě, aby dokázal se zákazníkem v daných požadavcích komunikovat konzistentně a vytvořil kladnou zákaznickou zkušenost. Pro místo, kde se zákaznická data v podniku shromažďují se ustálil pojem CRM (Customer Relationship Management) systém (1).

S rozmachem internetového podnikání se velké množství podnikových aktivit přesouvá do online prostředí. Místem, kde se zákazník (též uživatel) v internetovém prostředí setkává s podnikem, je webová aplikace. Chování uživatelů v podnikové webové aplikaci je žádoucí monitorovat z celé řady důvodů jako optimalizace průchodu aplikací nebo vyhodnocení návratnosti investic do online marketingu. Problematikou pochopení uživatelského chování na internetu se zabývá obor webová analytika. Leaderem v oblasti webové analytiky je nástroj Google Analytics. Přestože Google Analytics již ve svém výchozím nastavení sbírá o uživatelském chování v rámci webových aplikací mnoho dat, existuje způsob, jak tato data obohatit o data z dalších zdrojů. Tímto zdrojem může být i zmíněný CRM systém. Pokud patřičným způsobem data v Google Analytics obohatíme, respektive integrujeme do Google Analytics data z dalšího zdroje, otevřou se dveře k novým případům užití. Tyto případy užití lze chápat ve smyslu přesnějších datových analýz, rozšířených podkladů pro optimalizaci webové aplikace a využití obohacených dat pro přesnější cílení online marketingové inzerce.

Metod integrace dat do Google Analytics je hned několik, přičemž každá z nich má své předpoklady a specifika. Kromě technických náležitostí samotných metod, je potřeba problematiku vnímat v širších souvislostech, a to včetně pohledu na datový model nástroje, či legislativní a etické omezení kladené na integraci těchto dat. A protože dosud neexistuje ucelený dokument, který by procesem integrace dat do Google Analytics provázel, klade si tato diplomová práce této příležitosti využít.

Cílem této práce je návrh metodiky integrace dat do Google Analytics. Pro formalizaci navržené metodiky integrace dat poslouží metodika MMSP. Platnost navržené metodiky bude závěrem práce ověřena na ukázkovém případu užití.

V rámci první a druhé kapitoly je čtenář detailně seznámen se samotným rozborem problému, náčrtem možných budoucích případů užití a komponentami nástroje Google Analytics. Rozbor problému umožní ve třetí kapitole charakterizovat stav současného řešení. Stav současného řešení stanoví pomyslnou úroveň dosavadního poznání o problematice v reálném světě, kam bude navržená metodika zasazena. Před sestavením metodiky integrace dat do Google Analytics je čtenář ve čtvrté kapitole seznámen s metodikou diplomové práce. Kapitola pátá je již věnována samotnému návrhu metodiky, jejíž platnost je validována na ukázkovém případů užití v kapitole šesté.

1 Rozbor problému

Kapitola se věnuje základní charakteristice řešeného problému a možným případům užití. Charakteristika řešeného problému zasazuje problém integrace dat do Google Analytics do širšího kontextu ve vztahu k webové analytice. V další podkapitole je čtenář seznámen s možnými případy užití, jež jsou popsány konkrétními praktickými situacemi integrace dat do Google Analytics a výhodami z takovýchto integrací vyplývajících.

1.1 Obecná charakteristika problému

Základním prvkem této diplomové práce je nástroj Google Analytics. Nástroj Google Analytics (dále také GA) je analytická webová služba se zaměřením na sběr a analýzu dat webového prostředí. Tato data vykreslují charakteristiky internetového uživatele a jeho chování v rámci určité webové stránky, kde je implementován měřící kód nástroje.

Význam analytických služeb tohoto typu narůstá na důležitosti vzhledem k faktu, že v internetovém prostředí se odehrává stále větší objem podnikových aktivit generující pro podniky tržby (2). Zájmem těchto podniků generujících tržby v internetovém prostředí je mít o chování uživatele, který je zároveň zákazníkem, přehled z celé řady důvodů (3):

- Charakteristika uživatele – Snaha o poznání zákazníka (uživatele) a přizpůsobení služby, či produktu jeho potřebám. Díky těmto datům je podnik dále schopen vytvořit segmenty uživatelů se společnými rysy, které mají uplatnění v marketingu.
- Optimalizace webové prezentace – Vygenerování tržeb (splnění určitého cíle) v internetovém prostředí je podmíněno sadou kroků, jež musí uživatel učinit. Sledováním jeho chování lze tyto kroky upravit tak, aby byly uživatelsky přívětivější.
- Vyhodnocení návratnosti investic – V případě, že je skupina uživatelů získána formou placené online inzerce, chceme též vědět, kolik nám tato skupina uživatelů vygenerovala tržeb. Na základě těchto dat lze vyhodnotit návratnost investic.
- Výkon webového obsahu – Součástí mnohých webových prezentací jsou například blogy s články. Na základě sesbíraných dat analytickou webovou službou lze zjistit, kolik uživatelů dokázaly oslovit a kolik z nich podnítily ke splnění cíle.

Všechny z výše uvedených důvodů mají ve výsledku jedno společné a tím je splnění definovaného cíle, který typicky generuje podniku tržby. Přestože mnohé z výše uvedeného dokáže zjistit i celá řada dalších analytických služeb (například Adobe Analytics, Webtrekk, Woopra a jiné), věnuje se tato práce pouze nástroji Google Analytics. Google Analytics byl zvolen, jelikož jeho služeb využívá dle údajů W3techs.com z ledna 2019 55,6% všech webových stránek napříč celým internetem, čímž je na trhu dominantní (4). Jeho vysoká obliba je kromě jiného dána především jeho univerzálním použitím, podporou ze strany společnosti Google a jeho dostupností. Nástroj je totiž ve své základní verzi zcela zdarma.

Typickým uživatelem nástroje Google Analytics je marketingový pracovník podniku. Podnik může v internetovém prostředí generovat tržby například prodejem produktů na svém e-shopu. Primárním cílem e-shopu je dokončená objednávka ze strany internetového uživatele, která podniku generuje tržby (5). Vrátím-li se k důvodům pro využití analytických služeb uvedených výše, pak je lze na příkladu e-shopu snadno demonstrovat. Pokud má podnik na e-shopu nasazenou některou z analytických služeb, může sledovat a vyhodnocovat chování svých uživatelů. Dle uživatelského chování lze dále identifikovat taková místa e-shopu, kde dochází k vysokému procentu odchodů uživatelů a tato místa příslušným způsobem zoptimalizovat. V případě e-shopu se obvykle zaměřujeme na uživatelskou přívětivost průchodu objednávkovým procesem (3). U investic do online marketingové inzerce prostřednictvím inzertních systémů jako jsou Google Ads, Facebook Ads či Sklik, je žádoucí sledovat prostřednictvím analytických služeb návratnost investic, aby podnik provozující e-shop dokázal vyhodnotit, do jaké míry se investice do různých typů online inzerce vyplácí. A závěrem je potřeba sledovat výkon webového obsahu, aby bylo možné odvodit celkové trendy návštěvnosti e-shopu či návštěvnosti jeho produktů (3).

1.2 Technologická charakteristika problému

Veškeré informace, které lze z nástroje Google Analytics vyčíst, vycházejí ze sesbíraných dat v průběhu návštěvy uživatele na měřeném prostředí. V sesbíraných datech je ukryto jádro problematiky řešené touto prací. Nástroj totiž sbírá data primárně javascriptovým měřícím kódem, což je samo o sobě prvním zásadním omezujícím faktorem, ze kterého vyplývá, že nástroj Google Analytics dokáže sesbírat pouze taková data, ke kterým je možno přistoupit skrze technologii JavaScript.

Tato data mohou být dále obohacena dalšími informacemi z různých zdrojů. Prvním z nich jsou data, které má společnost Google z dalších služeb (například: pohlaví, věk nebo zájmy). Tyto informace jsme buď některé ze služeb společnosti Google sdělili a souhlasili s jejím anonymizovaným užitím v rámci ostatních služeb společnosti nebo je vyvodila z našeho chování v rámci využívání Google vyhledávače a návštěv tematických webových prezentací, které jsou součástí partnerské sítě společnosti Google (6).

Jinou možností obohacení sesbíraných dat je propojení nástroje Google Analytics s dalšími službami Google, které toto propojení umožňují. Jedná se o služby jako Google Ads, Adsense, Search Console či Google Play. Ukázkou výhod z tohoto propojení může být demonstrován například na Google Ads. Google Ads je služba, která umožňuje inzerentům vkládat inzeráty do reklamních ploch v rámci inzertní sítě společnosti Google. V případě, že službu Google Ads propojíme s Google Analytics a uživatel klikne na inzerát vytvořený v Google Ads, dokážeme v Google Analytics vyčíst i informace o tom, na jaký typ inzerátu klikl, o jakou kampaň se jedná či přesnou identifikaci prokliknuté reklamy. U jiných systémů, kde lze v rámci online marketingu inzerovat (Sklik, Facebook, Instagram, RTB systémy) toto přímé propojení jako u Google Ads neexistuje. Avšak i u těchto platforem je možné obdobné informace do Google Analytics propsat a to za pomocí UTM parametrů.

Existuje ještě jedna možnost, jak data sesbíraná výchozím javascriptovým kódem dále obohatit. A to tak, že nástroji předáme nějaká vlastní data. V případě e-shopu to mohou být

například informace o zakoupených produktech, o využití slevového kupónu, o přihlášení se k odběru novinek na email, o interakci s applety a další. Můžeme dokonce zajít do takové hloubky, že internetovému uživateli zaznamenanému v Google Analytics dokážeme vhodně přiřadit i takové informace, které s internetovým prostředím přímo nesouvisejí. Podmínkou takového přiřazení však je, že vždy musí existovat určité pojítko mezi zasílanými daty a konkrétním uživatelem v GA. Tímto pojítkem se rozumí přiřazení identifikátoru, který je vygenerován službou Google Analytics, konkrétnímu uživateli v podnikovém CRM systému. Reálné případy užití s tímto způsobem propojení jsou zmíněny v kapitole číslo 1.3. S technickou realizací tohoto propojení je pak čtenář blíže seznámen v kapitole číslo 2.2.

V prostředí Google Analytics existuje hned několik způsobů, jak obohacení sesbíraných dat o vlastní informace zrealizovat. Nicméně každý z nich je vhodný pro jiný typ informací, pro jinou zdrojovou platformu, užívá se různými způsoby a pracuje s různými technologiemi.

Vzhledem k tomu, že se jedná o obohacení dat, které probíhá systémově, automaticky a trvale, pak se zároveň jedná o integraci. Konkrétně se v tomto kontextu jedná o integraci na datové úrovni mezi Google Analytics a dalšími zdrojovými platformami, která probíhá jednosměrně a sice ze zdrojové platformy do Google Analytics (7). Tato diplomová práce si klade za cíl sestavit metodiku integrace dat do Google Analytics, která bude aplikovatelná pro různorodé případy užití tohoto typu integrace. Platnost sestavené metodiky bude závěrem této diplomové práce ověřena na ukázkovém případu užití.

1.3 Možné případy užití

Pro jasnější uchopení problematiky se nabízí zmínit několik možných případů užití integrace dat do Google Analytics, jež budou sestavenou metodikou pokryty. Avšak úvodem bych rád zmínil, co provedením tohoto typu integrace dat vlastně získáme. Existuje několik druhů motivace, které vyplývají z uplatnění nástroje GA:

- Datová analýza – Ze statistik chceme vyčíst údaje o uživatelích a jejich nákupním chování, které následně uplatníme pro manažerské rozhodování. V případě, že výchozí datový model obohatíme o další vlastní data, pak získáme další kritéria pro vyhodnocování a současně další výstupy pro manažerské rozhodování.
- Optimalizace webu – Snahou je přizpůsobení webové prezentace tak, aby byla uživatelsky přívětivá a aby bylo pro uživatele snadné dokončit konverzi. Google Analytics obsahuje ve výchozí podobě řadu dat, které s optimalizací pomohou, avšak můžeme je rozšířit i o další a tím optimalizaci zdokonalit.
- Online marketing – Google Analytics umožňuje propojení s inzertní službou Google Ads. Součástí propojení je i sdílení remarketingových seznamů. Remarketingový seznam je skupina uživatelů (respektive jejich cookies), které náš web už dříve navštívili a na které lze cílit relevantní inzerci. Remarketingové seznamy lze vytvářet na základě celé řady kritérií a pokud Google Analytics poskytneme vlastní data, budeme mít k dispozici další kritéria pro jejich definici (8).

Práce bude ve svém obsahu zmiňovat tři hlavní způsoby, jak lze vlastní data do Google Analytics propsat. Tyto hlavní způsoby (též integrační metody) vycházejí z vývojářské příručky služby Google Analytics (9):

1. Datový import – Jedná se o primitivní způsob integrace dat na základě importu CSV souborů. Google Analytics poskytují předdefinované šablony datového importu pro různé případy užití. Import CSV souborů pak lze zautomatizovat s využitím příslušné API. Více o datovém importu v rámci kapitoly 2.2.1.
2. Rozšířený měřící kód s využitím datové vrstvy – Datová vrstva (dataLayer) umožňuje předávat data mezi webovou aplikací a Google Analytics prostřednictvím javascriptového kódu. Datová vrstva se naplní vlastními daty, jež se odešlou do nástroje Google Tag Manager. Více o integrační metodě v kapitole 2.2.2.
3. Measurement Protocol – Předává data skrze backend, a to formou HTTP požadavku přímo na servery Google Analytics. Výhodou je, že tento HTTP požadavek může být odeslán i mimo web s měřícím kódem, tedy například z podnikového CRM či z jiných podnikových aplikací. V takovém případě je žádoucí data z webu a data z podnikových aplikací propojit za pomoci speciálního identifikátoru. Více o Measurement Protocolu v rámci kapitoly 2.2.3.

Datové importy lze použít pro obohacení stávajících dat v Google Analytics například o vlastní data o uživatelích jako jsou hodnota zákazníka, datum obnovení smlouvy či vyjádření věrnosti zákazníka. Datové importy zároveň umožňují rozšířit informace o konkrétních podstránkách měřené webové aplikace o vlastní parametry jako typ obsahu, autor či kategorie. Prostřednictvím datových importů lze do Google Analytics dále nahrát klíčové metriky z různých inzertních systémů jako náklady za inzerci, počet prokliků a počet zobrazení inzerátů. V neposlední řadě umožňují datové importy modifikovat data z e-commerce modulu pro měření objednávek na e-shopu. V rámci tohoto modulu lze pomocí datového importu vytvářet nové atributy o produktech jako jsou značka produktu, kategorizace produktu či jiné vlastní atributy přímo související s daným produktem. Uplatnění v souvislosti s ecommerce modulem mají datové importy i pro vymazání objednávek ze systému v případě stornované, nevyzvednuté nebo navrácené objednávky.

Rozšířený měřící kód s využitím datové vrstvy má široké uplatnění, pokud chceme do Google Analytics předávat data, která pocházejí přímo z webu. Předpokládejme například e-shop. Na e-shopu probíhají objednávky. Tyto objednávky obsahují informace o zakoupených produktech, použitých slevových kupónech, zvoleném typu dopravy a platby. Takovéto informace lze evidovat do Google Analytics prostřednictvím datové vrstvy. Uplatnění pro tento druh dat posléze nalezneme například v datových analýzách.

V případě webových aplikací, kde je možnost registrace, lze prostřednictvím datové vrstvy rozlišovat registrované a neregistrované uživatele. Toto rozlišení může mít uplatnění například v online marketingu, kde neregistrovaným uživatelům můžeme komunikovat výhody registrace a registrovaným uživatelům naopak prémiové členství či akční nabídky. Další zajímavou informací pro evidování prostřednictvím datové vrstvy může být například dosažené skóre v online hře, počet věrnostních bodů, posledního přihlášení do rozhraní, typ personalizovaného obsahu zobrazeného uživateli a jiné.

Oproti tomu nalezneme celou řadu možných užití, kdy data nejsou k dispozici přímo ve webové aplikaci s měřícím kódem, ale mimo ni. Pak nezbývá nic jiného, než použít další integrační metodu, kterou je Measurement Protocol.

Jako příklad mohou posloužit věrnostní karty obchodních řetězců. V případě, že při vydání věrnostní karty byla vyžadována například emailová adresa a tutéž emailovou adresu zadáme do webových stránek řetězce (například při registraci účtu či přihlášení se k odběru), pak lze vytvořit propojení mezi konkrétní věrnostní kartou a internetovým uživatelem (reprezentovaným formou cookies). Následně lze konkrétnímu uživateli v Google Analytics připisovat informace o aktivitách, které s věrnostní kartou učinil. Například v jaké kamenné pobočce ji použil, jaký typ sortimentu v jeho nákupu převládal, v jaké výši nakoupil a další. Všechny tyto informace lze následně zužitkovat v online inzerci prostřednictvím služby Google Ads či jako cenné podklady pro manažerské rozhodování.

Dalším příkladem mohou být adresované katalogy. Pokud dokážeme zjistit, kterým internetovým uživatelům byl reálně odeslán fyzický katalog, pak můžeme sledovat jejich chování na webu vůči zbytku uživatelů či jim komunikovat odlišný typ inzerce. Obdobně tomu může být v případě mobilních aplikací, kdy nás zajímají informace o tom, zda uživatel má nainstalovanou naší mobilní aplikaci a zda s ní aktivně pracuje.

Cenou informací je také vyhodnocení důležitosti zákazníka skrze Business Intelligence systémy. Zákazníky, které na webu reprezentují jejich cookies, lze třídit do různých skupin. Označení těchto skupin je vhodným typem dat pro provedení integrace do Google Analytics, kde je možno vyhodnocovat skupiny důležitosti dle dalších atributů, jež poskytuje datový model Google Analytics, či každé skupině komunikovat odlišný typ inzerce.

Názorným příkladem užití je i propojení s call centrem, kdy v Google Analytics chceme rozlišovat, zda konkrétní uživatel volal na naše call centrum, jaké problematiky se hovor týkal, či jak často volá. Taktéž se jedná o cenné informace s možností uplatnění jak v datových analýzách, tak přizpůsobení reklamní komunikaci v online marketingu.

Measurement Protocol lze dále využít pro předávání informací, které se staly na webech třetích stran, přičemž tento web třetí strany neumožňuje implementaci měřícího kódu naší služby GA. Tento případ užití nachází důležité uplatnění v případě, kdy se na webu třetí strany odehrává důležitá interakce, která má z hlediska našeho podnikání význam, a proto bychom ji rádi do Google Analytics zaznamenali.

Výše uvedené postupy lze aplikovat na celou řadu typů podniků a řadů různých případů užití. Po technické stránce se jedná o posloupnou sadu kroků, kde platí celá řada zákonitostí. Metodika vypracovaná touto prací si klade za cíl tento proces integrace jasně charakterizovat a vymezit jej tak, aby aplikace metodiky byla snadno uchopitelná pro systémové integrátory s různou úrovní znalostí problematiky a zároveň dobře aplikovatelná v různých typech podnikového prostředí.

2 Charakteristika komponent GA

Tato kapitola je věnována charakteristice vybraných komponent nástroje Google Analytics. Pro charakteristiku byly vybrány takové komponenty, které mají z předmětu této diplomové práce význam, respektive hrají významnou roli v procesu integrace dat do GA a je proto nutné, aby čtenář používající sestavenou metodiku se s těmito komponentami seznámil.

První z těchto komponent je datový model nástroje, na jehož příkladu je rozvedeno členění na dimenze a metriky včetně jejich stěžejních vlastností. V rámci datového modelu je čtenář dále seznámen s druhy požadavků odesílaných do Google Analytics, které lze chápat jako základní jednotku odesílaných dat. Další komponentou jsou integrační metody. Pro potřeby integrace dat rozlišuje Google Analytics 3 základní integrační metody, kterými jsou datový import, rozšířený měřící kód a Measurement Protocol (9). V závěru sekce o integračních metodách je provedena sumarizace rozlišující jednotlivé metody ve vymezených kritériích ve vztahu ke způsobu integrace dat. Třetí a zároveň poslední uvedenou komponentou jsou legislativní regulace, které zásadním způsobem ovlivňují, jaká data a jakým způsobem mohou být integrována.

2.1 Datový model Google Analytics

Sběr dat v nástroji Google Analytics je realizován prostřednictvím asynchronního javascriptového měřícího kódu. Současná verze měřícího kódu je označována jako gtag.js, v terminologii nástroje se o měřícím kódu bavíme zároveň i jako o globální měřící značce.

Výpis 2.1: Globální měřící značka (gtag.js) nástroje Google Analytics

```
1  <!-- Global site tag (gtag.js) - Google Analytics -->
2  <script async src="https://www.googletagmanager.com/gtag/js?id=UA-123456789-1">
3  </script><script>
4  window.dataLayer = window.dataLayer || [];
5  function gtag(){dataLayer.push(arguments);}
6  gtag('js', new Date());
7
8  gtag('config', 'UA-123456789-1');
9  //Prostor pro vlastní příkazy
10
11  </script>
```

Výše umístěný JavaScript by měl být vložen do zdrojového kódu každé ze stránek, které chceme do nástroje Google Analytics měřit. Měřící kód se vkládá na co nejvyšší možnou pozici ve značce <head>. Součástí měřícího kódu je i identifikátor, který stanovuje, do jaké konkrétní služby Google Analytics budou data zaznamenávána. Identifikátor je ve výpisu kódu 2.1 zvýrazněn podtržením.

Měřící kód odesílá na servery nástroje Google Analytics data formou HTTP požadavku. V rámci HTTP požadavku jsou data odesílána metodou GET. V praxi to probíhá tak, že nástroj sesbírá prostřednictvím JavaScript kódu data, která poskládá do parametrů určených pro konkrétní typ dat a parametry následně připojí za URL adresu https://www.google-analytics.com/collect. URL adresa společně se získanými parametry je poté odeslána jako HTTP požadavek metodou GET. Požadavek se typicky odesílá při načtení stránky, alternativně jej lze odeslat i vlastní událostí (viz kapitola 2.1.2).

Výpis 2.2: Ukázka URL adresy odeslaného HTTP požadavku z domovské stránky denik.cz

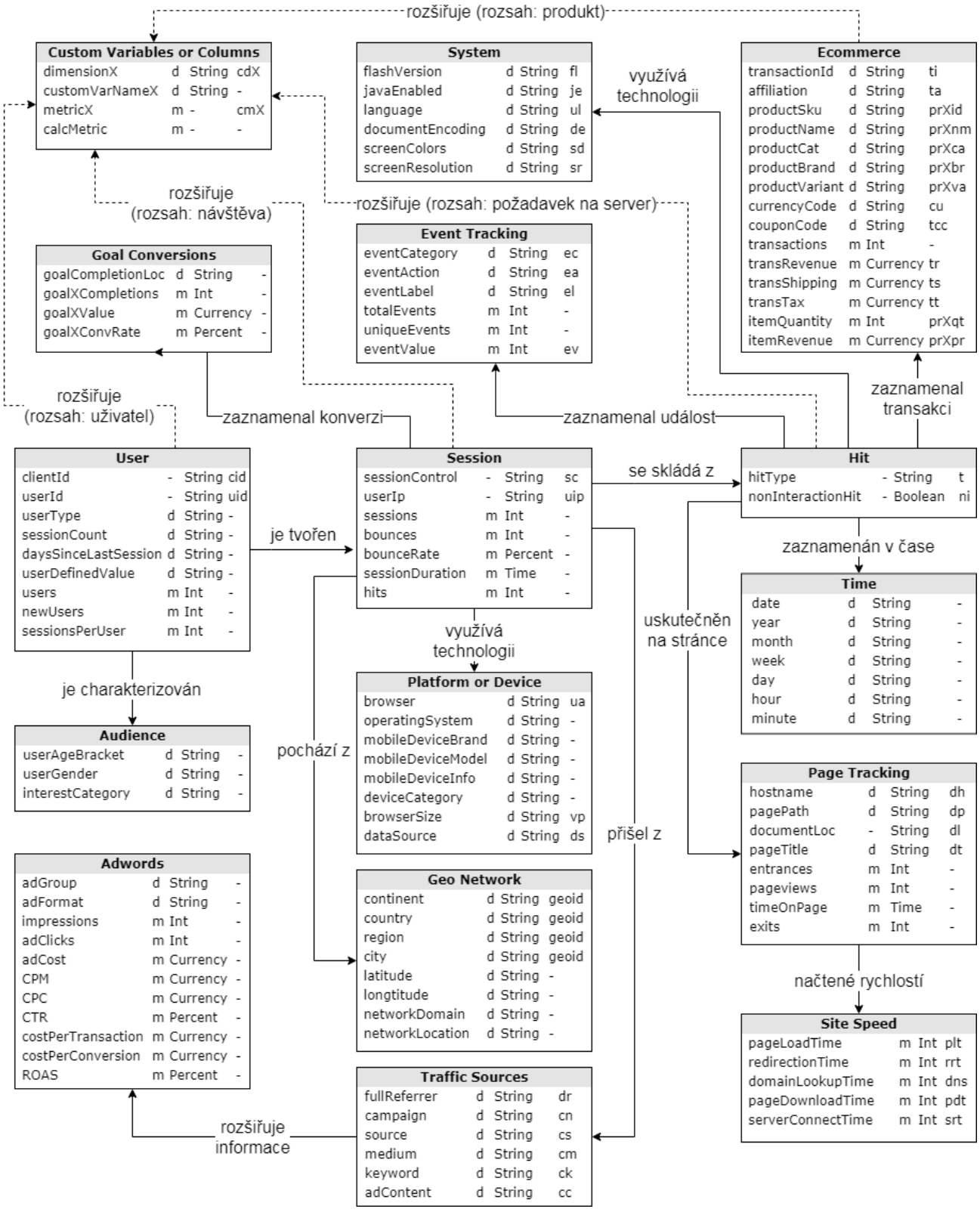
```
1  https://www.google-analytics.com/collect?
2  v=1&_v=j72&a=1460448419&t=pageview&_s=1&dl=https%3A%2F%2Fwww.denik.cz%2F&
3  dp=%2F&ul=cs-cz&de=UTF-8&dt=Den%C3%ADk.cz&sd=24-bit&sr=1024x768&vp=133x680&
4  je=0&_u=QACAAAAB~&jid=&gjid=&cid=249466028.1546956266&tid=UA-2118046-1&
5  _gid=322309157.1547295560&gtm=2wgbc0TGXTRC&z=1376255442
```

Z výše uvedené URL adresy lze samo o sobě vyčíst řadu informací jako URL navštívené stránky (dl=https%3A%2F%2Fwww.denik.cz%2F), titulek stránky (dt=Den%C3%ADk.cz), jazykové nastavení (ul=cs-cz), způsob kódování znaků (de=UTF-8), barevný režim obrazovky (sd=24-bit), rozlišení obrazovky (sr=1024x768), či identifikátor Google Analytics služby (tid=UA-2118046-1). Dekompozice jednotlivých parametrů z HTTP požadavků a jejich uspořádání do náčrtu datového je k dispozici v obrázku 2.1.

Výše uvedené parametry jsou v GA k dispozici jako dimenze a metriky. Dimenzí se v pojetí nástroje rozumí kvalitativní proměnná. Dimenze tak nabývá nečíselných hodnot. Mezi dimenze řadíme například město, URL zobrazené stránky nebo rozlišení obrazovky. Oproti tomu metriky nabývají číselných hodnot a můžeme s nimi provádět běžné matematické výpočty jako průměr nebo součet. Jedná se například o počet návštěv, průměrnou dobu trvání návštěvy, počet zobrazení stránek, počet konverzí či konverzní poměr. Pro sestavení přehledu v nástroji GA, potřebujeme vytvořit platnou kombinaci dimenze s metrikou. Platnou kombinací může být například URL zobrazené stránky a počet zobrazení stránky, město a počet návštěv či rozlišené obrazovky a míra okamžitého opuštění.

Datový model nástroje se skládá z celé řady různorodých dat. Za účelem této práce je jeho poznání klíčové, jelikož si nejprve musíme uvědomit, jaká data se v nástroji nachází, abychom je dokázali vhodně a příslušně obohatit. Dle dostupných informacích o dimenzích, metrikách a parametrech používaných v HTTP požadavcích jsem sestavil datový model reprezentovaný obrázkem číslo 2.1, který sice neobsahuje všechna data dostupná v Google Analytics, ale obsahuje všechna taková data, která jsou z předmětu této práce důležitá.

Datový model z obrázku číslo 2.1 je složen z entit, vztahů, atributů a popisných informací atributů. Každá entita toho modelu je složena ze čtyř sloupců, přičemž první z nich reprezentuje název atributu. Součástí datového modelu je několik atributů, které ve svém názvu obsahují znak „X“ (například atributy *dimensionX*, *customVarNameX* a *metricX* z entity *Custom Variables or Columns*, či atributy *goalXCompletions*, *goalXValue* a *goalXConvRate* z entity *Goal Conversions*). Tento znak uvádí, že těchto atributů může v datovém modelu existovat více v závislosti na definici v rozhraní nástroje Google Analytics. Znak „X“ je následně nahrazen číselným indexem dle definice z rozhraní nástroje.



Obrázek 2.1: Datový model Google Analytics
Zdroj: vlastní zpracování v nástroji Draw.io na základě (13) a (35)

Druhý sloupec udává informaci o tom, zda atribut v prostředí Google Analytics reprezentuje dimenzi (d) či metriku (m). V datovém modelu existují případy, kdy atribut není z rozhraní nástroje dostupný v podobě dimenze ani v podobě metriky, pak je zde uvedena pomlčka. Tato informace je důležitá, protože dimenze a metriky se chovají odlišně a zároveň nám pomůže rozklíčovat platnou kombinaci dimenzí a metrik, která souvisí s tzv. rozsahem (scope). Rozsahu je věnována samostatná kapitola 2.1.1.

Třetí sloupec udává informaci o datovém typu atributu jako String (řetězec znaků), Int (číselná hodnota), Currency (hodnota uvedená v měně), Percent (procentuálně vyjádřená hodnota), Time (časové vyjádření) a Boolean (proměnná nabývající hodnot 0 nebo 1). V případě entity *Custom Variables or Columns* se u atributů *metricX* a *calcMetric* nachází místo datového typu pomlčka. Jedná se o vlastní metriky, které je nejprve prostřednictvím rozhraní nutno vytvořit a definovat. Součástí definování je i volba datového typu. Tento atribut tak může nabývat datového typu Int, Currency nebo Time. A to v závislosti na tom, jaký datový typ při definici vlastního atributu bude zvolen.

Čtvrtý sloupec udává informaci o názvu parametru, který je uveden v rámci URL adresy odesílaného HTTP požadavku (viz výpis 2.2). Atributy, které tento parametr v modelu uvedený mají, lze odesílat skrze HTTP požadavek. Oproti tomu atributy, které mají v datovém modelu uvedenou pouze pomlčku, jsou buď nějakým způsobem odvozené od ostatních, nebo Google Analytics neumožňují jejich předání prostřednictvím požadavku. Ukázkou odvozeného atributu je například atribut *sessions* z entity *Session*. Atribut udává výpočet o tom, kolik návštěv bylo zaznamenáno. Z podstaty věci se tato informace do HTTP požadavku neuvádí, protože ji GA dopočítají. Atributem, který se přes HTTP požadavek nepředává je pak například i atribut *userGender* z entity *Audience*. Jedná se o pohlaví uživatele, které Google nezjišťuje z odeslaného požadavku, ale ze svých vlastních zdrojů (6).

2.1.1 Rozsah dimenzí a metrik

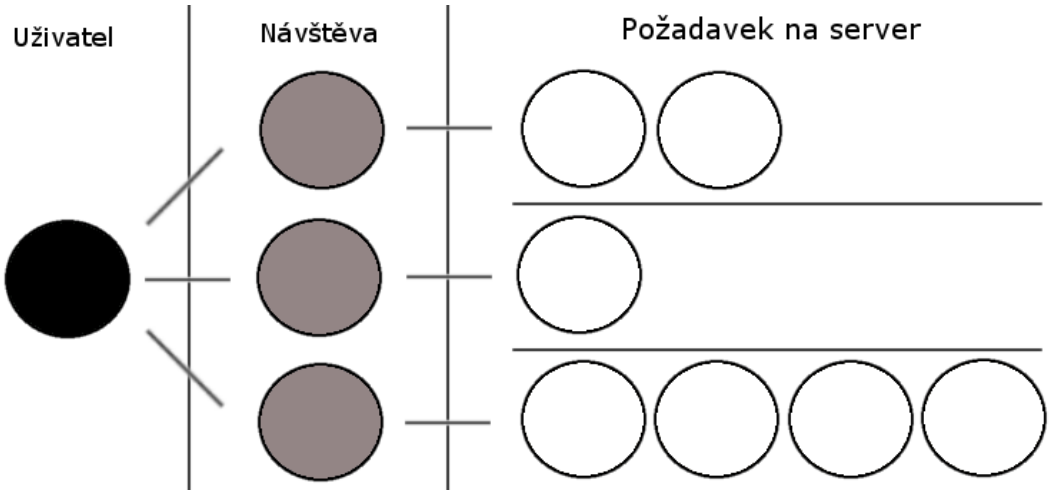
Důležitou součástí datového modelu nástroje Google Analytics jsou tzv. rozsahy (scope). Platnou kombinaci dimenze a metriky lze obvykle vytvořit, když dimenze i metrika náleží stejnému rozsahu. GA rozlišuje 4 rozsahy: požadavek na server, návštěva, uživatel a produkt (10). Základním je požadavek na server (hit). Jedná se o již zmíněný HTTP požadavek. Je to tedy sada dat odesílána měřícím kódem na servery Google.

Dalším rozsahem je návštěva (session). Návštěvou se rozumí skupina požadavků na server, jež proběhne během stanovené časové lhůty a za stanovených podmínek. Návštěva vždy začíná odesláním požadavku na server a končí jedním z následujících způsobů (11):

- Od posledního odeslaného požadavku na server uběhlo alespoň 30 minut
- V průběhu trvání návštěvy nastala půlnoc
- Uživatel na stránku vstoupil znovu z jiného kanálu

Kanálem se v tomto smyslu rozumí skupina dimenzí, která rozlišuje odkud uživatel vstoupil na webovou stránku. Tato skupina dimenzí dokáže rozlišit, že uživatel přišel například z organického vyhledávání Google, z placené bannerové inzerce, z konkrétní sociální sítě a jiné. Tyto informace jsou součástí odesílaného požadavku na server a jsou získávány

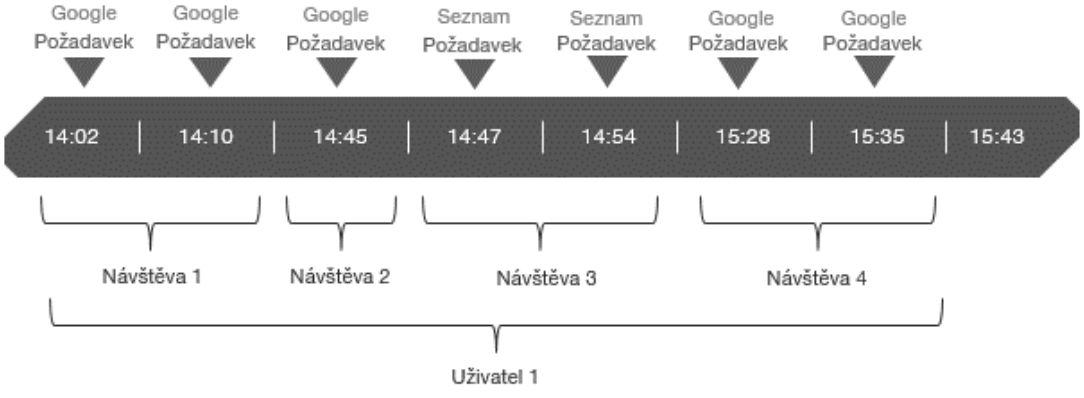
měřícím kódem z HTTP referreru, či z UTM parametrů (12). Vztah mezi uživatelem, návštěvou a požadavkem na server je demonstrován obrázkem 2.2.



Obrázek 2.2: Schéma vztahu mezi uživatelem, návštěvou a požadavkem na server
Zdroj: vlastní zpracování

Jiným rozsahem je uživatel (user). Uživatel je v Google Analytics skupina návštěv, která připadá jednomu internetovému uživateli. Přiřazení návštěv uživatelům probíhá primárně na základě cookie souboru uloženém v prohlížeči.

Pro úplné dokreslení vztahu mezi uživatelem, návštěvou a požadavkem na server poslouží časová osa na obrázku 2.3. Tato časová osa reprezentuje jednoho uživatele, který v daném období provedl na webové stránce měřené nástrojem Google Analytics 7 požadavků na server a 4 návštěvy. První a čtvrtá návštěva byly ukončeny časovou odmlkou delší než 30 minut, druhá a třetí návštěva byly ukončeny vstupem z jiného kanálu. Kanál na obrázku reprezentuje informace Google/Seznam.



Obrázek 2.3: Časová osa odeslaných požadavků na server
Zdroj: vlastní zpracování

Obrázek 2.3 dále poukazuje na zajímavost, která vyplývá z principu měření skrze zmíněné požadavky na server. Každý odeslaný požadavek na server s sebou nese časovou značku,

tedy informaci o tom, kdy byl odeslán. S časovými značkami jednotlivých požadavků mimo jiné pracuje například metrika Doba trvání návštěvy.

Doba trvání návštěvy je definována jako rozdíl časových značek posledního požadavku na server a prvního požadavku na server dané návštěvy (11). Při pohledu na čtvrtou návštěvu v obrázku 2.3 je vidět, že požadavek byl odeslán požadavek v 15:28 a v 15:35, uživatel stránku ale reálně opustil až v 15:43. V 15:43 už ale žádný požadavek odeslán nebyl, jelikož požadavek na server se ve výchozím nastavení odesílá pouze při načtení stránky, které proběhlo naposledy v 15:35. Znamená to tedy, že uživatel posledních 8 minut prohlížel stránku, aniž by ji znovu načetl nebo se přesunul na další podstránku, čímž by došlo k novému načtení. Doba trvání čtvrté návštěvy pak byla zaznamenána jako 7 minut, přestože reálně trvala 15 minut.

Posledním z výše čtyř uvedených rozsahů je rozsah produkt. Jedná se o speciální rozsah, který byl vytvořen pro modul elektronického obchodu (ecommerce). Modul elektronického obchodu je určený primárně e-shopům. Prostřednictvím tohoto modulu mohou e-shopy evidovat informace o uskutečněných objednávkách (v terminologii nástroje o transakcích). Transakce se skládají z produktů. Produkt může nést řadu datových atributů (značka, barva, velikost či cena), pro které je vymezen rozsahu produkt.

Rozhraní nástroje Google Analytics dále umožňuje definovat vlastní dimenze a metriky. Vlastní dimenze a metriky alokují v datovém modelu nový prostor pro plnění vlastními daty. Při definici každé dimenze či metriky je nutno zvolit jeden ze čtyř zmíněných rozsahů v závislosti na tom, jaká data budou do Google Analytics v rámci vlastních dimenzí a metrik předávána. Způsobů předávání dat do vlastních dimenzí a metrik je hned několik. Více o nich v rámci kapitoly 2.2.

2.1.2 Typy požadavků na server

Měřící kód Google Analytics odesílá informace na své servery jako HTTP požadavky (v terminologii nástroje se jedná o hit) za pomoci metody GET. Požadavků se rozlišuje několik typů, jedná se o požadavky typu pageview, screenview, event, transaction, item, social, exception a timing (13). Z hlediska předmětu této práce jsou klíčové především požadavky typu pageview a event.

Pro odesílání požadavku typu pageview postačuje základní implementace měřícího kódu, kterou se rozumí nasazení základního měřícího kódu Google Analytics do značky <head> každé z měřených podstránek. Měřící kód následně začne při každém načtení stránky odesílat požadavek typu pageview, který s sebou nese celou řadu dat. Požadavek tohoto typu může přenášet data, jež jsou v datovém modelu na obrázku číslo 2.1 reprezentována jako tabulky *User*, *Session*, *Hit*, *Platform or Device*, *Geo Network*, *Traffic Sources*, *Page Tracking*, *Site Speed*, *System* nebo *Custom variables or Columns* a zároveň mají v datovém modelu na obrázku 2.1 uvedený parametr pro HTTP požadavek, jež se nachází ve čtvrtém sloupci každé z uvedených tabulek. Ukázkou odeslaného požadavku typu pageview je výpis číslo 2.2 v rámci kapitoly 2.1.

Dalším zmíněným typem požadavku je typ event. Tento typ požadavku využívá skutečnosti, že měřící kód Google Analytics je asynchronní. Požadavek event se totiž standardně

neodesílá při načtení stránky, ale odesílá se, pokud nastane určitá událost, kterou chceme v Google Analytics zaznamenat. Touto událostí lze chápat například stisknutí klíčového prvku stránky, odeslání formuláře, přehrání videa, využití vyhledávání, interakce s vybranými prvky stránky, stažení souboru, přihlášení se do administrace a jiné. Pokud nastane událost, jež chceme do Google Analytics zaznamenat, spustí se ve zdrojovém kódu příkaz znázorněný ve výpisu 2.3.

Výpis 2.3: Syntaxe příkazu pro odeslání požadavku typu event do Google Analytics (14)

```
gtag('event', <action>, {
  'event_category': <category>,
  'event_label': <label>,
  'value': <value>
});
```

Ve výpisu 2.3 se nacházejí čtyři proměnné, které jsou znázorněny jako <action>, <category>, <label> a <value>. Jedná se o proměnné, které slouží k identifikaci události. Proměnné tak můžeme nahradit libovolnými vlastními hodnotami. Tyto čtyři proměnné jsou též znázorněny v datovém na obrázku číslo 2.1 v rámci tabulky „Event Tracking“ jako atributy *eventCategory*, *eventAction*, *eventLabel* a *eventValue*. Požadavek typu event předává stejná data jako požadavek typu pageview s tím rozdílem, že navíc předává ještě čtyři výše uvedené proměnné.

Posledním typem požadavku, který je vhodné z hlediska této diplomové práce zmínit jsou požadavky týkající se ecommerce, tedy požadavky typu transaction a item. Tento typ požadavku je důležitý pro internetové obchody, jelikož umožňuje předávat data o dokončených transakcích včetně dat o produktech, jež byly v rámci transakce zakoupeny. Z hlediska datového modelu na obrázku číslo 2.1 tento typ požadavku odesílá stejná data jako požadavek typu pageview a navíc odesílá i data znázorněná tabulkou *Ecommerce*. Nutno podotknout, že význam požadavku typu ecommerce klesá, protože nová verze modulu pro měření internetového obchodu v rámci Google Analytics označovaná jako Enhanced Ecommerce umožňuje předávání dat o transakcích a zakoupených produktech prostřednictvím požadavku typu pageview.

Součástí parametrů, jež jsou obsaženy v HTTP požadavku, je i povinný parametr *t* (viz tabulka *Hit* na obrázku 2.1 nebo ve výpisu 2.2), tento parametr reprezentuje typ odesílaného požadavku. Může tedy nabývat výše uvedených hodnot (například pageview či event) a slouží jednak k identifikaci, jaký typ dat je v rámci požadavku předáván, tak jako příkaz pro Google Analytics jak daná data zpracovat.

Volitelným atributem, který definuje zpracování požadavku nástrojem Google Analytics je dále parametr *ni* (viz atribut *nonInteractionHit* v tabulce *Hit* na obrázku 2.1), který může nabývat pouze hodnot 1 (true) nebo 0 (false). S tímto parametrem je úzce svázána metrika Míra okamžitého opuštění (Bounce Rate), tato metrika udává procentuální podíl návštěv, jež provedly s webovou stránkou pouze jednu interakci. Typicky se jedná o první požadavek typu pageview a následný odchod z webu. Parametr *ni* pak udává informaci, zda daný požadavek má, či nemá vstoupit do výpočtu metriky Míra okamžitého opuštění.

2.2 Metody integrace

Předchozí kapitoly vymezily, jaká data existují ve výchozím datovém modelu nástroje Google Analytics a jakým způsobem jsou výchozí data do nástroje předávána. Nicméně primárním zaměřením této diplomové práce je integrace dat do Google Analytics z jiných zdrojů. Jedná se tedy o taková data, která výchozím způsobem nejsou do nástroje předávána a je proto potřeba nalézt takový způsob integrace, který toto předávání dat umožní. Dle Gály se „*integrací v podnikové informatice rozumí akt nebo proces, kterým kombinujeme, propojujeme a spojujeme různorodé zdroje podnikové informatiky do vyššího celku, přičemž komponenty dohromady spolupracují a sdílejí data bez znatelného zdržení a koordinují svojí funkcionalitu tak, že se taková kombinace komponent jeví uživateli jako jednotný systém*“ (1). Různorodé zdroje, jež jsou uvedeny v předchozí definici, odpovídají v této diplomové práci zdrojovým datových systémů (typicky CRM systémy), odkud jsou data v příslušné podobě a příslušnou metodou předávána do Google Analytics. Google Analytics lze posléze chápat jako vyšší celek, jelikož s novými integrovanými daty se značně rozšiřují jeho možnosti užití (viz kapitola 1.3).

Integrace, jež je předmětem této diplomové práce, je tedy na datové úrovni, bavíme se proto o tzv. integraci dat. Komunikace je asynchronní, jelikož zdrojový systém nečeká na odpověď a v komunikačním procesu pokračuje neblokovaně. Dále lze tuto integraci chápat jako jednosměrnou, tedy že data plynou ze zdrojového systému do Google Analytics a nikoliv opačně. Integrace probíhá v modelu push. Zdrojové systémy totiž data v požadovaném formátu odesílají samy a Google Analytics je zpracuje ve chvíli, kdy je obdrží. Google Analytics se nedoptává, zda jsou ve zdrojových systémech nová data určená k integraci. Transformací dat do požadovaného formátu na straně zdrojových systému se tato diplomová práce již nezabývá, čtenáře s požadovaným formátem pouze seznamuje. Důvodem je skutečnost, že možných zdrojových systémů existuje nepřeberné množství a není proto z hlediska rozsahu práce možné, zabývat se transformací dat do požadovaného formátu pro každé z existujících řešení, jež jsou na trhu k dispozici.

Datovým zdrojem pro integraci dat mohou být:

- Data, která má k dispozici měřená webová aplikace
- Data z vlastních podnikových systémů
- Data ze dalších externích systémů

Dále rozlišujeme integraci z hlediska rozsahu dimenzí a metrik (viz kapitola 2.1.1):

- Data na úrovni požadavku na server
- Data na úrovni návštěvy
- Data na úrovni uživatele
- Data na úrovni produktu

Integraci lze dále rozlišit dle toho, odkud se data na servery odesílají:

- Skrze požadavek měřícího kódu webové aplikace
- Skrze požadavek odeslaný z podnikového systému
- Skrze import údajů přímo do služby Google Analytics

K integraci lze využít jednu ze třech dostupných integračních metod (9):

- Datový import
- Rozšířený měřící kód
- Measurement Protocol

Každá z integračních metod má určitá specifika, předpoklady, je vhodná pro jiný typ prostředí a pro jiný typ dat. Jednotlivé metody jsou detailně rozebrány a charakterizovány v následujících podkapitolách.

2.2.1 Datový import

Datový importem se v prostředí Google Analytics rozumí textový soubor (typicky ve formátu CSV), který je skrze administrační rozhraní odeslán na servery Google Analytics, kde je následně zpracován a spojen s již sesbíranými daty. Textový soubor lze vložit do administračního rozhraní manuálně nebo automaticky prostřednictvím, pro tento účel vyvinutého, API označovaného jako Google Analytics Management API.

Google Analytics data odeslaná formou datového importu následně páruje na základě dimenze, jež je pro daný import označena jako klíč. Jakmile je nalezena shoda mezi hodnotou klíčové dimenze datového importu a hodnotou téže dimenze v již nashromážděných datech, dojde k propojení.

Datových importů Google Analytics rozlišuje několik typů, přičemž každý z nich je určen pro jiný účel, pracuje s jinými klíčovými dimenzemi a umožňuje odesílat jiná data. Jedná se celkem o 8 různých typů datových importů rozdělených do 3 skupin (15):

- Hit Data Import
    - Refund Data – Data o navrácených platbách v rámci transakcí modulu pro elektronický obchod. Klíčem je dimenze ID transakce (viz atribut *transactionId* z tabulky *Ecommerce* na obrázku 2.1) a umožňuje nahrát data jako skladové číslo produktu, cenu produktu, tržby a vrácené množství.
- Extended Data Import
    - User Data – Rozšiřuje sesbíraná data o uživatelích o vlastní údaje, například o hodnotu zákazníka, datum obnovení smlouvy či vyjádření věrnosti zákazníka. Klíčem je dimenze ID uživatele (viz atribut *userId* z tabulky *User* v datovém modelu GA dostupném na obrázku 2.1).
    - Campaign Data – Pokud označujeme kampaň vlastním kódem, lze prostřednictvím tohoto typu importu předat údaje o tom, co kód znamená. Klíčem je tedy kód kampaně a umožňuje nahrát data jako zdroj, médium, název kampaně, klíčové slovo či obsah reklamy.
    - Geographical Data – Umožňuje vytvořit vlastní geografické členění dat na základě jednoho ze zvolených klíčů, kterými jsou identifikační číslo města, kód ISO země, kód subkontinentu nebo číslo oblasti.

    - Content Data – Jedná se o data rozšiřující URL adresy ze sesbíraných dat o vlastní informace jako typ obsahu, autor či kategorie. Klíčem datového importu je tzv. Page Path, tedy část URL adresy umístěné za doménou (viz atribut *pagePath* z tabulky *Page Tracking* na obrázku 2.1).
    - Product Data – Rozšiřuje informace o produktech shromážděných modulem pro elektronický obchod. Umožňuje uvést data jako značka produktu, kategorie nebo vlastní údaje související s produktem. Klíčovou dimenzí je skladové číslo produktu (SKU), respektive jeho identifikátor (viz atribut *productSku* v tabulce *Ecommerce* na obrázku 2.1).
    - Custom Data – Jde o vlastní import s libovolným účelem. Klíčovou dimenzí může být například název hostitele, číslo aplikace, vlastní dimenze, jazyk uživatele, operační systém, prohlížeč, rozlišení obrazovky a řada dalších.
- Summary Data Import
    - Cost Data – Data o nákladech souvisejících s marketingovými kampaněmi z inzertních nástrojů mimo Google. Klíčové dimenze jsou zde hned tři: datum, médium a zdroj (viz atributy *source* a *medium* z tabulky *Traffic Sources* a atribut *date* z tabulky *Time* na obrázku 2.1). Je možno nahrát data o ceně, kliknutí a zobrazení (viz atributy *adCost*, *adClicks* a *impressions* z tabulky *Adwords* na obrázku 2.1).

Výše uvedené typy datových importů znázorňují širokou pestrost možných dat, které lze tímto způsobem nahrát. Výhodou datových importů oproti metodám integrace z následujících kapitol je relativní jednoduchost, kdy pro jednorázový datový import není nutná znalost programování, protože požadovaný soubor lze snadno zkonstruovat textovým editorem nebo tabulkovým procesorem. V případě, že k nahrávání importů využijeme API, pak lze celý proces dokonce zautomatizovat. Další výhodou je možnosti importu velkého množství dat najednou, což další metody integrace neumožňují.

2.2.2 Rozšířený měřící kód

Přestože javascriptový měřící kód sám o sobě sbírá velké množství různorodých dat, můžeme jej modifikovat, aby odesílal navíc taková data, která ve výchozím nastavení zaznamenat nedokáže. Podmínkou je, že tato data musí měřená webová stránka měřícímu kódu určitým způsobem předat. Zároveň nutno zvážit, v jakém rozsahu metrik a dimenzí (viz kapitola 2.1.1) chceme, aby byla data v Google Analytics zpracována.

Implementace skrze zdrojový kód

Základním způsobem, jak do GA předat data nad rámec výchozích jsou události, tedy typ požadavku event (viz kapitola 2.1.2). Události jsou primárně určeny pro interakce uživatele s obsahem a jsou měřeny asynchronně. Nejde tedy o integraci v pravém slova smyslu. Nicméně jedná se o způsob, jak snadno do GA předat taková data, která přímo vyplývají z chování uživatele v rámci měřené webové stránky. Vzhledem k tomu, že událost je specifickým typem požadavku, pak z toho vyplývá, že dimenze a metriky, které událost charakterizují (tedy dimenze kategorie, akce a událost a metrika hodnota události) jsou rozsahem na úrovni požadavku na server. Události jako takové proto nejsou vhodné pro data, kde požadujeme rozsah dimenzí a metrik jiný než rozsah požadavek na server.

Dalším způsobem, jak do datového modelu Google Analytics nahrát vlastní data, je využít některou ze stávajících dimenzí. Měřící kód totiž umožňuje přepsat naměřenou hodnotu vybraných existujících dimenzí vlastní hodnotou, respektive vlastní proměnnou. Toto přepsání umožňují například dimenze, které jsou v datovém modelu na obrázku 2.1 znázorněny jako atributy tabulky *Traffic Sources*, atributy *hostname*, *pagePath* a *pageTitle* tabulky *Page Tracking*, či atributy *clientId* a *userId* tabulky *User*. Ukázku měřícího kódu (viz výpis 2.1) rozšířenou o přepsání existujících dimenzí Název stránky a Stránka (atribut *pageTitle* a *pagePath* z tabulky *Page Tracking* na obrázku 2.1) je výpis 2.4.

Výpis 2.4: Globální měřící značka (gtag.js) rozšířená o přepis hodnot dimenzí

```
<!-- Global site tag (gtag.js) - Google Analytics -->
<script async src="https://www.googletagmanager.com/gtag/js?id=UA-123456789-1">
</script><script>
window.dataLayer = window.dataLayer || [];
function gtag(){dataLayer.push(arguments);}
gtag('js', new Date());

gtag('config', 'UA-123456789-1', {
  'page_title' : 'Domovská stránka',
  'page_path': '/homepage'
});
</script>
```

Poslední možností, jak předat vlastní data prostřednictvím měřícího kódu je využití vlastních dimenzí a metrik. V datovém modelu na obrázku 2.1 se jedná o atributy *dimensionX* a *metricX* z tabulky *Custom Variables or Columns*. Novou vlastní dimenzi či metriku je nejprve nutno založit a definovat prostřednictvím administračního rozhraní Google Analytics, kde se dimenze či metrika pojmenuje a určí její rozsah (viz kapitola 2.1.1). V případě metrik se dále určí i typ formátování (číslo, měna nebo čas) a volitelně i extrémní hodnoty (minimum a maximum), které budou nástrojem Google Analytics akceptovány. Po založení bude dimenzi, či metrice přidělen číselný index.

Výpis 2.5: Příkaz pro mapování dimenze a metriky s následným odesláním hodnot

```
 gtag('config', 'UA-123456789-1', {
  'custom_map': {
   'dimension<index>': 'dimension_name',
   'metric<index>': 'metric_name'
   }
 });

 gtag('event', '<action>', {
  'dimension_name': dimension_value,
  'metric_name': metric_value
 });
```

V měřícím kódu je pak potřeba aktualizovat config rozšířením o mapu vlastních dimenzí a metrik (tzv. custom_map), která mapuje index dimenze či metriky s názvem proměnné. Následně se v rámci příkazů pracuje již jen se zde definovaným názvem proměnné, ke kterému se přiřazuje hodnota proměnné, jež je odeslána na servery Google Analytics. Nastavení probíhá prostřednictvím následujícího javascriptového příkazu ve výpisu číslo 2.5. Z výše uvedeného výpisu si lze povšimnout, že po namapování dimenzí a metrik se hodnoty odesílají prostřednictvím typu požadavku event (viz kapitola 2.1.2).

Google Tag Manager a datová vrstva

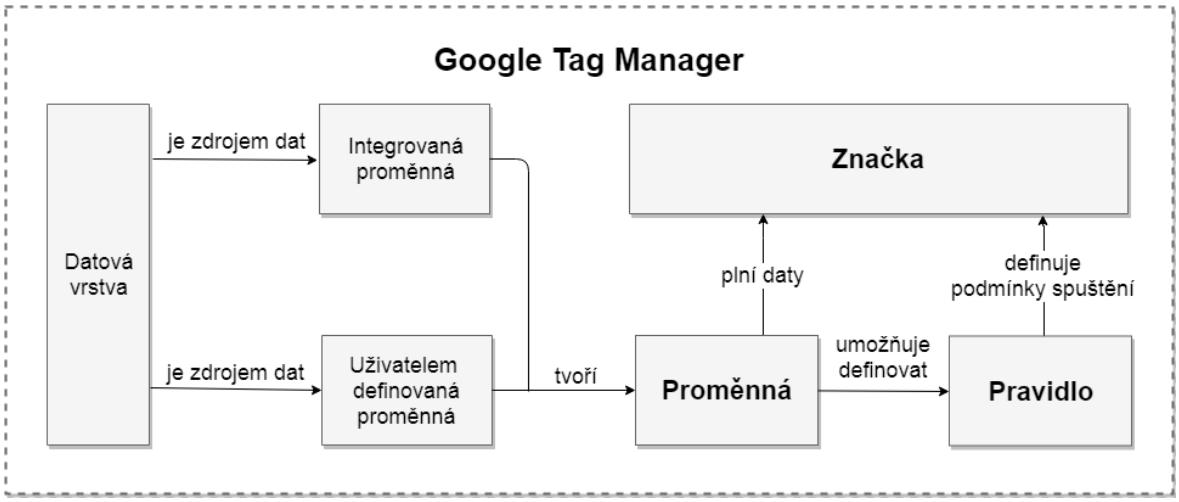
Alternativním přístupem k modifikaci měřícího kódu jsou systémy pro správu měřících kódů. Leaderem systémů pro správu měřících kódů je produkt od společnosti Google označovaný jako Google Tag Manager. Dle údajů k 23.2.2019 od W3Techs.com 98,5 % všech webových stránek, které nějaký systém pro správu měřících kódů používají, používají právě Google Tag Manager. Google Tag Manager (zkráceně GTM) zároveň používá 20,4 % ze všech webových stránek, které W3Techs.com do svých statistik zahrnuje. Z historických dat je patrné, že popularita nástroje v čase citelně roste. Například k datu 1.2.2018 nástroj Google Tag Manager používalo dle statistik W3Techs.com pouhých 10,5 % webových stránek (16).

Google Tag Manager lze chápat jako modul (v terminologii Google označovaný jako kontejner), prostřednictvím kterého lze spravovat měřící kódy (včetně měřícího kódu GA) na daném webu, kde je Google Tag Manager naimplementovaný. Výhod z používání Google Tag Manageru vyplývá hned celá řada:

- Měřící kódy mohou spravovat i jiné zainteresované strany než pouze programátoři s přístupem ke zdrojovému kódu, například pracovníci marketingu.
- Všechny měřící kódy jsou přehledně na jednom místě. Je proto snadné se rychle zorientovat ve stávajícím nastavení a vyvarovat se případných chyb.
- Pokud nástroj obsahuje pro daný typ měřícího kódu vlastní šablonu, je zde záruka toho, že měřící kód bude nástrojem automaticky aktualizován do aktuální verze.
- Nástroj obsahuje propracovaný systém pro testování měřících kódů, který zásadním způsobem pomáhá eliminovat chyby při implementaci kódů.
- Každá změna, která se propíše do ostrého prostředí je zaevidována v historii změn, kde lze snadno dohledat kdy, kým a jaké změny byly provedeny.

Google Tag Manager se skládá ze třech základních komponent. Jedná se o měřící kódy (v terminologii GTM označované jako značky), pravidla a proměnné. Všechny tyto tři komponenty jsou spolu úzce provázány. Značka totiž musí mít v GTM prostřednictvím pravidla definováno, na jaké části webu a za jakých podmínek se má spustit. Značka s sebou zároveň přenáší řadu informací, které čerpá z proměnných.

Vazba existuje i mezi pravidly a proměnnými. Pravidla jsou definována určitými hodnotami, kterých mohou proměnné nabývat. Proměnné rozlišujeme na 2 základní druhy – integrované (například Page URL, Referrer, Click Classes, Event) a proměnné definované uživatelem (například proměnná datové vrstvy, vlastní kód JavaScript, soubor cookie první strany). Vazby mezi jednotlivými komponenty nástroje Google Tag Manager demonstruje následující obrázek.



Obrázek 2.4: Schéma komponent nástroje Google Tag Manager
Zdroj: Vlastní zpracování

V obrázku číslo 2.4 je dále znázorněn prvek pojmenovaný jako datová vrstva. Z hlediska předmětu této diplomové práce se jedná o velmi důležitý prvek nástroje Google Tag Manager. Weber definuje datovou vrstvu jako komunikační objekt mezi webovou stránkou a GTM obsahující další data nad rámec těch, které je GTM schopno zachytit prostřednictvím integrovaných proměnných. Datovou vrstvu lze chápat jako centralizovaný seznam dat, ke kterému může GTM snadno přistupovat a současně je tento seznam dat uživatelům prohlížejících si web skrytý (17). Na základě dat z datové vrstvy je možno definovat proměnné, které jsou následně užity pro plnění dat v rámci značek, nebo pro definici podmínek spuštění v rámci pravidel.

Z technického hlediska se jedná o objekt vytvořený javascriptovým kódem, který může být plněn libovolnými páry klíčů a hodnot. Existují dva základní přístupy k definici datové vrstvy ve zdrojovém kódu. Tím prvním je definice datové vrstvy nad měřícím kódem nástroje Google Tag Manager. Podmínkou tohoto přístupu je, že zasílaná data do datové vrstvy musí být známa už při načítání konkrétní webové stránky. Ukázkou je script z výpisu 2.6, který prostřednictvím datové vrstvy popisuje přihlášeného uživatele, který odebírá novinky na email a v téže chvíli se nachází stránce s nastavením svého profilu.

Výpis 2.6: Script definující datovou vrstvu při načtení webové stránky

```
<script>
    dataLayer = [{
      'userStatus': 'logged-in',
      'newsletterSubscription': 'active',
      'contentType': 'profile-settings'
    }];
</script>
<!-- Begin Google Tag Manager Tracking Code -->
...
<!-- End Google Tag Manager Tracking Code-->
```

Druhým přístupem k definici datové vrstvy ve zdrojovém kódu webové stránky je využití metody dataLayer.push. Tato metoda umožňuje naplnit datovou vrstvu proměnnými asynchronně, tedy kdykoliv, kdy uživatel webovou stránku prohlíží a interaguje s ní, aniž by bylo nutné webovou stránku znovu načíst. Metoda dataLayer.push je proto vhodná pro typ informací, které při načítaní stránky ještě nejsou známy. Jde tedy o informace, které vyplynou až z uživatelova chování. Typicky se jedná o nějakou informaci o tom, že nastala určitá událost (například odeslání formuláře, stažení souboru či scrollování stránkou). Tato událost může být dokreslena dodatečnými vlastnostmi. Názorným příkladem je výpis kódu 2.7, kde uživatel v rámci webové aplikace sestavil vlastní vůz na míru. Jakmile konfiguraci vozu dokončil, předaly se do datové vrstvy informace o tom, že konfigurace byla dokončena, typ nakonfigurovaného vozu, model vozu a jeho barva.

Výpis 2.7: Script definující datovou vrstvu při interakci s obsahem metodou dataLayer.push

```
<script>
   dataLayer.push({
     'event': 'carConfigurationDone',
     'carType': 'sedan',
     'carModel': 'Superb',
     'carColor': 'silver'
    });
</script>
```

Na základě předchozího výpisu kódu jsme schopni skrze Google Tag Manager nastavit například značku pro odeslání událostí do Google Analytics (viz kapitola 2.1.2, požadavek typu event). Značka bude definována pravidlem s podmínkou, že proměnná *event* nabyla hodnoty *carConfigurationDone*. Událost se následně odešle do Google Analytics vždy, když se v datové vrstvě objeví proměnná *event* s hodnotou *carConfigurationDone*. K předání informací o typu, modelu nebo barvě lze využít proměnných Google Analytics událostí, kterými jsou *category, action, label* a *value* (viz kapitola 2.1.2).

Obdobně lze informaci v datové vrstvě využít například za účelem plnění vlastních dimenzí či metrik. Google Tag Manager totiž velmi snadno umožňuje zvolit konkrétní informaci z datové vrstvy a vložit ji do značky na místo a v podobě, jaké konkrétní vlastní dimenze nebo metrika vyžaduje. S vlastními dimenzemi a metrikami je nutno zvážit jejich rozsah (viz kapitola 2.1.1). Například barva nakonfigurované vozu z výpisu kódu 2.7 dává smysl pouze v rozsahu požadavek na server. Barva vozu totiž není informace, která by charakterizovala návštěvu nebo uživatele. A zároveň se nebavíme o modulu pro elektronický obchod, proto není možný ani rozsah na úrovni produktu.

2.2.3 Measurement Protocol

Přestože doposud měl každý pojem v terminologii nástroje Google Analytics český ekvivalent, v případě Measurement Protocolu tomu tak není. Measurement Protocol je dle Waisberga *„flexibilní metoda odesílání dat do Google Analytics napřímo, bez potřeby využít JavaScript, SDK nebo jiného mechanismy sběru dat*“ (18). V kapitole 2.1 této

diplomové práce byl čtenář seznámen s podobou HTTP požadavků obsahující data, která odesílá javascriptový měřící kód nástroje Google Analytics na své servery. Smyslem metody Measurement Protocol je sestavit a odeslat obdobný HTTP požadavek na straně backendu měřené webové aplikace, či aplikace třetí strany. Google Analytics takto odeslaný požadavek přijme a zpracuje. Z podstaty věci Measurement Protocol umožňuje sbírat data do Google Analytics nejen z různých typů aplikací, ale i z různých typů zařízení. Tedy takových zařízení, která jsou nějakým způsobem připojena k internetu a jsou schopna sestavit a odeslat HTTP požadavek ve stanoveném formátu.

Skutečnost, že díky metodě Measurement Protocol může Google Analytics zpracovávat v podstatě jakákoliv data a odkudkoliv je z pohledu předmětu této diplomové práce velmi přínosná. Avšak je zde druhý úhel pohledu a sice, že klasický javascriptový kód sbírá velmi mnoho různorodých dat už ze své podstaty, a ne všechna tato data lze získat v aplikacích a zařízeních třetích stran za účelem sestavení HTTP požadavku pro Measurement Protocol. Uvažujme například čipový turniket, který je schopen potřebný HTTP požadavek sestavit a zároveň se podívejme na datový model Google Analytics na obrázku 2.1. Většinu dat obsažených v datovém modelu Google Analytics z čipového turniketu z podstaty věci nezískáme (například internetový prohlížeč, rozlišení obrazovky nebo název zobrazené webové stránky). Možnosti využití takto sesbíraných omezených dat z čipového turniketu nebudou proto příliš široké, jako tomu je u dat sesbíraných měřícím kódem. Z turniketových dat bychom při správné integraci mohli vyčíst například počty průchodů, či při využití vlastních dimenzí i některé rozšiřující informace související s průchodem turniketu (například typ turniketu, či typ čipové karty). Nevýhodou však je, že zde zcela chybí přesah do online marketingu. V případě čipového turniketu, který odesílá HTTP požadavky na servery Google Analytics nejsme schopni dále tyto informace zužitkovat pro účely online marketingu, jelikož zde chybí pomyslné pojítko s reálným internetovým uživatelem.

Případem, kdy je data ze systémů třetích stran, zaslaných metodou Measurement Protocol, možné využít pro účely online marketingu je, když se podaří zmíněné pojítko mezi systémem a internetovým uživatelem v Google Analytics nalézt. V Google Analytics existuje pro každého uživatele anonymní unikátní identifikátor. Tento identifikátor je exponován v cookie pojmenované jako _ga. Příkladem struktury identifikátoru uloženého v cookie _ga je identifikátor „*GA1.2.1022484576.1547118686*“.

Nyní jako systém schopný odesílat HTTP požadavky přes Measurement Protocol uvažujme například CRM systém v obchodním řetězci. Tento obchodní řetězec vydává věrnostní karty. O věrnostní kartu je možné zažádat prostřednictvím jednoduchého formuláře na webových stránkách, kde je potřeba zadat emailovou adresu a způsob vyzvednutí věrnostní karty (vybraná pobočka, nebo poštovní adresa). Do CRM systému se automaticky zaznamená emailová adresa, anonymní identifikátor z _ga cookie a přidělený identifikátor věrnostní karty. Pokud zákazník provede nákup v libovolné pobočce obchodního řetězce s použitím věrnostní karty, může CRM systém odeslat informace o provedené transakci do nástroje Google Analytics s uvedením anonymního identifikátoru z _ga cookie uloženého v CRM databázi. Odeslaná transakce se v Google Analytics připáruje ke konkrétnímu uživateli, v důsledku čehož, je možné na něj cílit personalizovanou inzerci prostřednictvím systém Google Ads. Příkladem personalizovaného inzerátu v online marketingu může být třeba sdělení „Děkujeme za Váš nákup v pobočce Roztoky u Prahy. Podělte se s námi o Vaši

zkušenost!“, který se bude zobrazovat uživateli, který provedl nákup s věrnostní kartou v pobočce Roztoky u Prahy.

Jiným případem užití, kde zmíněné pojítko můžeme nalézt, je například internetový obchod, který disponuje vlastním call centrem. Internetový obchod má k dispozici CRM systém, kde kromě identifikátoru objednávky eviduje i anonymní identifikátor z _ga cookie. Pokud zákazník zavolá na call centrum a stěžuje si na konkrétní objednávku, operátor call centra k objednávce do CRM systému poznamená, že zde proběhla stížnost. CRM systém odešle HTTP požadavek obsahující informaci o stížnosti a zároveň obsahující identifikátor z _ga cookie uložený v CRM systému. V Google Analytics se informace o stížnosti připáruje ke konkrétnímu uživateli. Tomuto uživateli lze následně komunikovat personalizované sdělení prostřednictvím inzertního systému Google Ads, které může mít například následující podobu „Je nám líto, že Váš poslední nákup neproběhl bez problémů. Jako omluvu jsme pro Vás připravili 50% slevu na další nákup.“

Problémem vyplývajícím z přechozích dvou případů užití je, že anonymní identifikátor Google Analytics je vázán na cookie. Na cookie je zároveň vázána následná inzerce prostřednictvím inzertního systému Google Ads. Po smazání cookes tak není možné personalizovaného cílení z předchozích případů užití využít. Pokud uživatel cookie smaže, pak mu při nové návštěvě webové stránky bude vygenerován nový anonymní identifikátor do _ga cookie. Je tedy vhodné nastavit pro CRM, kde se identifikátory z _ga cookie ukládají, takové mechanismy, aby identifikátor uživatele byl vždy pokud možno co nejaktuálnější.

2.2.4 Sumarizace integračních metod

Předmětem kapitoly 2.2 byla charakteristika základních integračních metod dat do nástroje Google Analytics. Každá z uvedených metod je svými vlastnostmi odlišná od zbývajících. Odlišnosti se nacházejí v celé řadě kritérií od míry automatizace přes použité technologie až po vhodnost užití metody pro konkrétní typ dat. Zmíněné integrační metody jsou základními stavebními prvky pro integraci dat do Google Analytics a tvoří jádro navržené metodiky, která je výstupem této diplomové práce.

Z hlediska předmětu této diplomové práce je proto klíčové mít zřetelný přehled o vzájemných odlišnostech charakterizovaných integračních metod tak, aby pro konkrétní typ integrace byla zvolena vhodná integrační metoda. Pro přehledné odlišení jednotlivých integračních metod byla do práce začleněna tato kapitola, která jednotlivé integrační metody sumarizuje v tabulce číslo 2.1, a to za pomoci kritérií formát odesílaných dat, míra automatizace, časové hledisko odeslání a zpracování dat, potřeba programovacích znalostí, typ dat vhodných pro konkrétní integrační metodu, možnost následného smazání dat, vazba integrační metody na cookies a místa, odkud byla data odeslána.

Tabulka 2.1: Sumarizace integračních metod (zdroj: vlastní zpracování)

	Datový import	**Rozšířený měřící kód**	**Measurement Protocol**
Datový formát	CSV soubor	JS script	HTTP požadavek
Míra automatizace	Manuální import (automatizovaně API)	Automatizované	Automatizované
Časové hledisko	Zpětně k existujícím uživatelským relacím	V reálném čase dle interakcí uživatele	V čase odeslání HTTP požadavku
Programovací znalosti	Ne (jen v případě API)	Ano (GTM usnadní)	Ano
Typ dat	Taková, pro které je dostupný některý z osmi datových importů	Taková, která jsou známa při interakci uživatele s webem	Taková, která nevyplývají z interakcí uživatele s webem
Smazání dat	Datový import lze snadno smazat	Odeslaná data nelze smazat	Odeslaná data nelze smazat
Využívání cookies	Ne	Ano	Ne (volitelně ID)
Místo odeslání dat	Administrace (v případě API libovolný systém)	Web s měřícím kódem	Libovolný systém

2.3 Legislativní a etické regulace

Na data se obecně vztahuje celá řada různorodých legislativních i etických regulací. Pro potřeby této práce jsou důležité takové legislativní regulace, které mají platnost na území České republiky. Společnost Google se působením na českém trhu zavazuje k jejich dodržování, a to včetně služby Google Analytics nevyjímaje. Kromě legislativních regulací se společnost Google zavazuje k dodržování vlastních zásad ochrany soukromí, které v mnohém rozšiřují a upřesňují stávající platnou legislativu. Existující regulace se však netýkají pouze společnosti Google, nýbrž i jejich uživatelů. V souvislosti s chováním uživatelů nástrojů Google Analytics a Google Ads je vhodné zmínit i pojem etika, který lze v tomto kontextu vnímat jako morální rámec zásad pro nakládání s informacemi.

2.3.1 Legislativa

O legislativních regulacích se bavíme zejména v souvislosti s osobními daty (též osobní údaje). Zásadním právním předpisem vymezující oblast osobních údajů a jejich regulaci je obecné nařízení Evropské Unie GDPR (plným názvem General Data Protection). Jedná se o právní předpis, který v české legislativě k datu 25.5.2018 nahradil dosud platný zákon o ochraně osobních údajů (č. 101/2000 Sb.) (19). Definicí pojmu osobní údaj dle GDPR *„se rozumí jakákoliv informace, vztahující se k subjektu údajů.“* (20). Subjektem údajů je pak

dle definice GDPR *„člověk (fyzická osoba), ke které se vztahují libovolné informace“* (20) Obecně lze osobní údaj vnímat jako jakýkoliv údaj, který nějakým způsobem identifikuje konkrétní fyzickou osobu (například číslo občanského průkazu, jméno, rodné číslo, emailová adresa či bydliště). (21) GPDR stanovuje, že k nakládání s osobními údaji musíme mít právní důvod, veřejný zájem nebo jednoznačný souhlas dotčené osoby. Dozor nad dodržováním legislativních regulací nakládání s osobními údaji (tedy i GDPR) provádí na území České republiky Úřad pro ochranu osobních údajů (zkráceně ÚOOÚ) (22).

Jinou legislativní regulací s přímým dopadem na proveditelnost integrace stanovenou touto prací je zákon o elektronických komunikacích (č. 127/2005 Sb.). Zákon mimo jiné zahrnuje evropskou směrnici č.2009/136/EC, která je známá jako „cookie law“. Jedná se o nařízení, které vymezuje povinnosti pro uživatele a zpracovatele souborů cookies, respektive údajů v nich obsažených. V souvislosti s tímto nařízením se uplatňuje tzv. opt-in přístup souhlasu nakládání s cookies, tedy že provozovatel webové stránky musí uživatele s použitím souboru cookies seznámit a uživatel má možnost s jejich použitím nesouhlasit (23).

Legislativu řešené problematiky lze nalézt i v Listině základních práv a svobod, která v článku 10 uvádí, že *„každý má právo na ochranu před neoprávněným shromažďováním, zveřejňováním nebo jiným zneužíváním údajů o své osobě“* (24). Problematikou neoprávněného nakládání s osobními údaji se dále zabývá trestní zákoník ve svém paragrafu číslo 180, který mimo jiné definuje tresty pro osoby, které neoprávněným nakládáním s osobními údaji způsobí dalším stranám právní újmu (25).

V souvislosti s faktem, že Google je původem americká společnost a současně nakládá s osobními údaji osob Evropské unie, musela se zavázat k dodržování směrnice označované jako „EU-US Privacy Shield“, která je regulačním rámce pro toky osobních údajů mezi USA a EU za účelem ochrany těchto údajů a transparentnosti nakládání s nimi (26).

2.3.2 Samoregulace

Samoregulace se obvykle uplatňuje v odvětví, kde příliš nezasahuje stát. I přesto je snahou zainteresovaných stran daného odvětví vymezit určitá „pravidla hry“. Zainteresované strany tak odvětví regulují samy. Oblast nakládání daty (především těch osobních) je bezesporu státem velmi regulovanou a hlídanou oblastí, avšak na samoregulace se lze dívat i tak, že stávající legislativu mohou určitým způsobem rozšiřovat nad rámec legislativních vymezení. Večerková uvádí, že samoregulace rozhodně nemůže právní regulace žádným způsobem nahradit, avšak může je vhodným způsobem doplnit a rozšířit (27).

Se samoregulací úzce souvisí etika a etické kodexy. Samoregulační etické kodexy ve vztahu k nakládání s daty nalezneme například v oblasti marketingového výzkumu. Etickým kodexem marketingového výzkumu je Mezinárodní kodex ICC/ESOMAR, jehož uplatňování na českém trhu vyžaduje po svých členech Sdružení agentur pro výzkum trhu a veřejného mínění (SIMAR). Kodex mimo jiné dbá na transparentní shromažďování osobních údajů, ochranu osobních údajů před neoprávněným přístupem a etické chování datových výzkumníků (28). Obdobným etickým kodexem je kodex americké Association of National Advertisers, který klade velký důraz na etické nakládání s daty v oblasti online marketingu (29). Jinou asociací disponující samoregulačním etickým kodexem pro oblast

nakládání s daty v oblasti webové analytiky je Digital Analytics Association, která stanovuje 5 základních pilířů: ochrana soukromí, transparentnost, uživatelská kontrola, edukace a přijetí zodpovědnosti (30). Společnost je Google se svou službou Google Analytics certifikovaným členem Digital Analytics Association a zavazuje se tak k dodržování všech pilířů etického kodexu této asociace.

Omezení nakládání s údaji ve službě Google Analytics

Přestože společnost Google disponuje vlastními Zásadami ochrany soukromí, které informují veřejnost o nakládání s údaji v rámci služeb Google, bavíme-li se konkrétně o službě Google Analytics, je potřeba vymezit určitá další omezení nad rámec zmíněných zásad. Hlavním omezením ve směru nakládání s údaji ve službě Google Analytics je zákaz importu jakýchkoliv údajů, které mohou umožnit identifikaci konkrétní osoby (například emailová adresa či jméno) nebo trvalých identifikátorů zařízení, tedy takových unikátních identifikátorů, které nelze resetovat. Dotčení uživatelé zároveň musí být plně informování o tom, jaké údaje a jakým způsobem jsou do služby Google Analytics importovány a současně musí mít k dispozici možnost odhlášení se od tohoto nakládání s údaji. Neposlední podmínkou je, že správce služby Google Analytics (tedy osoba zodpovědná za import údajů) musí disponovat veškerými právy pro tento typ nakládání s údaji. V případě užití importovaných dat pro potřeby marketingové inzerce, je dále nutno splňovat požadavky vyplývající ze Zásad pro inzertní funkce Google Analytics (31).

Od roku 2010 zároveň služba Google Analytics nabízí funkci tzv. anonymizace IP adres. IP adresy mohou být v některých zemích označovány jako osobní údaj, a proto mohou podléhat řadě regulací pro nakládání s osobními údaji. Funkce anonymizace IP adres nahrazuje poslední oktet nulou a až poté ji služba Google Analytics zpracuje. Plná IP adresa se pak nikdy na servery společnosti Google nedostane. V důsledku anonymizace IP adres však může být zhoršena přesnost geografických přehledů služby Google Analytics (32).

2.4 Shrnutí problému

Kapitola číslo 2 rozpadla problematiku na dílčí klíčové komponenty, jež jsou pro realizaci návrhu metodiky integrace dat do Google Analytics zásadní. Z těchto komponent lze vyvodit několik otázek souvisejících s dosavadním přístupem k současnému poznání těchto komponent ve vztahu k řešenému tématu práce. Zmíněné otázky jsou v rámci této kapitoly vyvozeny a souhrnně uvedeny v její spodní části. Hledání odpovědí na tyto otázky lze zároveň chápat jako mapování stávající úrovně poznání problematiky. Odpověďmi na tyto otázky se bude zabývat kapitola 3 Charakteristika současného stavu řešení.

Kromě úvodního vymezení nástroje Google Analytics a náčrtu možných případů užití byl čtenář seznámen s první klíčovou komponentou, kterou je stávající datový modelem Google Analytics. Vzhledem k absenci oficiální dokumentace datového modelu nástroje jsem na základě ověřených zdrojů sestavil datový model vlastní. Otázkou však je, jaké přístupy k odkrytí datového modelu nástroje doposud existují a jak datový model GA charakterizují?

Dílčí komponentou datového modelu, jsou dimenze a metriky. Dimenzí a metrik v datovém modelu Google Analytics existuje celá řada, případně lze definovat dimenze a metriky vlastní. Při zvažování užití dimenzí a metrik (ať už stávajících, či nově definovaných) je potřeba vždy zvážit jejich rozsah (viz kapitola 2.1.1). Z rozboru dimenzí a metrik vyplynulo, že je potřeba stanovit takový proces, který dokáže bezpečně určit, jakou dimenzi či metriku pro řešení daného případu užití zvolit.

Další komponentou nástroje jsou integrační metody. V této práce byly popsány 3 integrační metody, kterými jsou datový import, rozšířený měřící kód a Measurement Protocol. Každá z integračních metod má svůj specifický způsob realizace integrace dat. Problémem však je, že se musíme jednoznačně rozhodnou pro takovou integrační metodu, která je pro daný případ užití nejvhodnější. Jak se na rozhodnutí pro určení konkrétní integrační metody dívá současná úroveň poznání?

Jako poslední komponentu lze uvažovat legislativní aspekty problematiky. Před samotnou realizací integrace dat proto vždy musíme zmapovat, zda je daný případ užití z hlediska právních, případně i etických stránek věci proveditelný.

Tabulka 2.2: Dotazy mapující úroveň poznání komponent GA (zdroj: vlastní zpracování)

Dílčí komponenta Google Analytics	**Dotaz mapující úroveň poznání**
Datový model	Jaké existují přístupy k odkrytí datového modelu GA? Jak existující přístupy datový model GA charakterizují?
Dimenze a metriky	Jak je popsán proces volby vhodné dimenze a metriky pro integraci? Jak široce jsou tyto popsané procesy volby platné?
Integrační metody	Podle jakých kritérií je volena vhodná integrační metoda? Jak je ověřována proveditelnost procesu integrace pro daný případ užití?
Legislativa	Jak je pohlíženo na právní a etické náležitosti integrace dat do GA?

Kromě pohledu na úroveň poznání ve vztahu k dílčím komponentám řešené problematiky, existuje ještě jedna rovina pohledu, a sice globální. K formalizaci vymezení pohledu na globální úroveň problematiky dobře poslouží soupis dimenzí metodiky MMDIS (plným názvem Multidimensional Management and Development of Information Systems). Metodika MMDIS celkem vymezuje 12 dimenzí (33). Tyto dimenze lze v tomto kontextu chápat jako ucelenou šablonu pro systematický přístup k identifikaci klíčových dotazů.

Pro každou z dimenzí je vymezen dotaz mapující úroveň současného poznání ve vztahu k předmětu práce. Přestože některé z těchto dimenzí mohou mít v určitých bodech překryv s dotazy v tabulce 2.2., bude na následující dotazy nahlíženo výhradně na globální úrovni řešené problematiky, tedy s určitou dávkou nadhledu. Odpovědi budou posléze zmapovány v rámci kapitoly 3 Charakteristika současného stavu řešení.

Tabulka 2.3: Dotazy mapující úroveň poznání v globální rovině (zdroj: vlastní zpracování dle (33))

Dimenze dle metodiky MMDIS	**Dotaz mapující úroveň poznání**
Funkce a procesy	Do jaké míry je popsán sled činností integrace dat do GA?
Data a informace	Jaká přístupy pro správu dat s ohledem na GA existují?
Organizační a legislativní aspekty	Jak je popsán proces zacházení s daty během integrace? Jaké náležitosti musejí data pro integraci splňovat?
Lidské zdroje	Jak jsou definovány role, jež do procesu integrace zasahují?
Software	Jaké programové vybavení je předpokladem integrace?
Hardware	Jaké existují požadavky na technické vybavení pro realizaci integrace?
Uživatelské rozhraní	Jak je nahlíženo na složitost realizace integrace dat do GA? Do jaké míry jsou postupy integrace dat do GA standardizované?
Bezpečnost	Jak stávající úroveň poznání hledí na bezpečnostní aspekty?
Ekonomické a finanční aspekty	Jaká je časová a finanční náročnost realizace integrace dat do GA na základě již existujících a zrealizovaných případů užití?
Metody	Jaké existují dostupné techniky a nástroje pro integraci dat do GA?
Dokumentace	Do jakého detailu je zpracována dokumentace pro realizaci integrace?
Management	Existují doporučení z praxe pro řízení zdrojů spotřebovávaných během integrace dat do GA? Pokud ano, jaká?

3 Charakteristika současného stavu řešení

Záměrem této kapitoly je zmapování, do jaké míry již byly poznány jednotlivé aspekty potřebné k sestavení metodiky integrace dat do Google Analytics, jež je cílem této práce. Charakteristiku současného stavu řešení lze zároveň chápat jako popis existující hladiny současného poznání, na jejíž základech tato diplomová práce dále staví.

V rámci předchozí kapitoly byl čtenář podrobně seznámen s řešeným problémem. Řešený problém byl rozložen na dílčí komponenty, přičemž každá z těchto komponent je pro sestavení cíle této práce zásadní, je proto na místě provést analýzu jejich současného poznání. V předchozí kapitole byly současně pro každou z komponent vymezeny dotazy mapující současnou úroveň poznání, které jsou ve vztahu k předmětu práce relevantní.

Cílem práce je sestavení metodiky. Metodiku definuje Buchalcevová jako *„souhrn metod a postupů pro realizaci určitého úkolu“* (34). Již samotná definice napovídá, že sestavení metodiky vyžaduje určitou míru nadhledu nad danou problematikou. Proto byl závěrem předchozí kapitoly vymezen i pohled globální, formalizovaný dimenzemi metodiky MMDIS. K jednotlivým globálním dimenzím byly, stejně jako v případě dílčích komponent, vymezeny relevantní dotazy mapující současnou úroveň poznání ve vztahu k tématu práce.

Vymezené dotazy mapující současnou úroveň poznání v kapitole 2.4 poslouží pro potřeby této kapitoly jako osnova, která je zároveň klíčem k systematickému poznání současného stavu řešení v náležitém rozsahu potřebném pro konstrukci hlavního výstupu této práce.

3.1 Charakteristika současného poznání dílčích komponent GA

První komponentou je datový model. Smyslem datového modelu Google Analytics pro potřeby této práce je rozklíčování struktury dat, se kterými nástroj pracuje. Oficiální dokumentace k datovému modelu Google Analytics v současnosti neexistuje. Pro pochopení datového modelu mohou však posloužit oficiální zdroje Dimension & Metrics Explorer (35) a Measurement Protocol Parameter Reference (13). Tyto zdroje jsou v podstatě souhrnným seznamem skupin dat, se kterými Google Analytics pracují. O sestavení datového modelu se následně pokusil Drkušić, který rozčlenil datové dimenze do 4 oblastí: Location dimensions, System dimensions, Personal & acquisition dimensions a Page & sessions (36). Drkušić však ve svém návrhu záměrně opomíjí metriky. Společný pohled na dimenze a metriky mohou nabídnout nástroje sloužící pro export dat z GA do nástrojů třetích stran. Takovým příkladem je dokumentace společnosti GoodData Corporation, která pro svůj nástroj CloudConnectDesigner sestavila balíček s návrhem datového modelu Google Analytics (37). Přestože tento balíček pracuje s dimenzemi i metrikami nástroje GA, neobsahuje bohužel zcela všechna data, kterými GA disponuji. Balíček obsahuje pouze taková data, která Google Analytics umožňuje exportovat a zároveň taková, která SW CloudConnect společnosti GoodData Corporation dokáže zpracovat. Na základě těchto poznatků jsem se rozhodl sestavit datový model vlastní (viz obrázek 2.1), který je dále popsán v kapitole číslo 2.1.

Další komponentou jsou dimenze a metriky. K nalezení příslušných dimenzí a metrik dobře poslouží sestavený datový model na obrázku číslo 2.1. K ověření, zda je daná dimenze s konkrétní metrikou platnou kombinací, lze ověřit za pomoci nástroje Dimension & Metrics Explorer (35). Selig za pomoci Google Analytics Core Reporting API sestavil úplný seznam všech dimenzí a metrik, které lze za pomoci této API nalézt (38). Rozbor platnosti určitých kombinací dimenzí a metrik včetně příslušných příčin rozebírá ve svém článku Hall (39). Platnost kombinace dimenzí a metrik je mimo jiné úzce provázána s rozsahem dimenzí a metrik. Podrobným rozborem rozsahu dimenzí a metrik, a to včetně přesahu do vlastních dimenzí a metrik se zabývá Rosche (40).

Třetí z uvedených komponent jsou integrační metody. Integrační metody byly identifikovány 3: datový import, rozšířený měřící kód a Measurement Protocol. Každá z uvedených integračních metod má oporu v oficiální dokumentaci společnosti Google. Kromě oficiální dokumentace existuje dále celá řada případů užití pro každou z integračních metod. Zde se určitě přísluší uvést knihu Google Analytics Integrations od Daniela Wasiberga, která zmíněné integrační metody rozebírá a u každé z nich uvádí možné případy užití (18). Přestože jsou metody v knize dobře popsané, chybí zde rozhodovací cyklus volby relevantní integrační metody pro daný případ užití. Zdrojem, který zmíněné integrační metody do jisté míry porovnává je oficiální příručka pro developery od společnosti Google, která se na integrační metody dívá ve vztahu integrace dat ze CRM systémům za účelem tvorby remarketingových publik pro službu Google Ads (9).

Poslední zmíněnou dílčí komponentou, která do integrace dat do nástroje Google Analytics velmi zásadně promlouvá, je platná legislativa a podmínky užití nástroje Google Analytics stanovené společností Google. Legislativní omezení byly vymezeny v rámci kapitoly 2.3. Avšak existují další zdroje, které se problematikou integrace dat ve vztahu ke konkrétní legislativě zaobírají detailněji. Velmi aktuálním tématem je soulad dat v Google Analytics vůči GDPR. Tento soulad detailně rozebírá Sharma, který sestavil seznam nutných podmínek pro splnění souladu dat v Google Analytics vůči nařízení GDPR (41). Postupy, jak se při integraci dat do Google Analytics vyvarovat případnému konfliktu s platnými podmínkami Google a současnou legislativou pro zpracování osobních údajů (anglicky personally identifiable information, zkráceně PII), uvádí ve svém článku Paul Koks (42).

3.2 Charakteristika současného stavu poznání v globální rovině

První z globálních dimenzí dle metodiky MMDIS jsou funkce a procesy. Z hlediska funkcí a procesů nás ve vztahu k integraci dat do Google Analytics zajímá popis posloupnosti jednotlivých kroků integrace. V obecné rovině jsou procesy integrace dat, především ve vztahu k technologiím Business Intelligence, popsány velmi detailně. Procesy datové integrace společně s běžně užívanými vzory datové integrace v oblasti Business Intelligence rozebírá ve své publikaci Sherman (43). Dayal a kolektiv se dále zabývají návaznostmi datových toků v průběhu integrace ve vztahu k systému ETL (extract, transform, load), přičemž jednotlivé složky ETL systému jsou dále detailně charakterizovány svou vnitřní procesní strukturou (44). Posloupnost kroků integrace dat do Google Analytics na globální úrovni dosud zpracována nebyla. Avšak vnitřní posloupnost procesů jednotlivých integračních metod popisuje ve své knize Waisberg (18).

V souvislosti s datovou dimenzí, tedy se sběrem a správou dat z různých zdrojů s následným využitím pro online marketingové aktivity, se ustálil pojem DMP (Data Management Platform). S určitou mírou abstrakce lze mezi DMP systémy a nástrojem Google Analytics nalézt jisté podobnosti. Tato podobnost se nachází ve schopnosti integrovat okolní datové zdroje či v možnostech využití sesbíraných dat pro online marketingové aktivity. Problematikou nakládání dat v DMP systémech, a to včetně datové integrace z různých online i offline datových zdrojů, se zabývá Elmeleegy a kolektiv (45). Otázkou, zda Google Analytics může systémy DMP suplovat se zabývá Van Der Vurst, který došel k závěru, že Google Analytics sice může suplovat určité funkcionality DMP systému, avšak plnohodnotně jej prozatím nahradit nedokáže (46).

Globální dimenze zabývající se organizačními a legislativními aspekty má přímou souvislost s legislativními a etickými regulacemi, tedy kapitolou číslo 2.3. Charakteristiku ve smyslu současného stavu poznání legislativy ve vztahu k tématu práce pak rozebírá kapitola 3.1 v rámci svého posledního odstavce. Z hlediska obecné organizace nakládání s daty v databázích existuje přímá souvislosti se SŘBD (Systém řízení báze dat), jehož anglickým ekvivalentem je DBMS (Database Management Systém). SŘBD vytváří pomyslnou mezivrstvu mezi daty v databázích a uživatelem. Uživatelův přístup k těmto datům je touto mezivrstvou řízen a zároveň umožňuje uživateli s daty pouze takové operace, které jsou platné. Problematikou architektur SŘBD systémů se zabývá Härder (47). SŘBD může díky svým omezujícím podmínkám zamezit výskytu nevalidních dat v datovém zdroji, jež je s Google Analytics integrován. V případě, že by k výskytu nevalidních dat v datovém zdroji došlo, mohla by integrace zdroje s Google Analytics selhat.

Další globální dimenzí jsou lidské zdroje, respektive pracovní role vstupující do procesu integrace dat do Google Analytics. Na straně zdrojových podnikových systémů bude pro realizaci integrace zásadní role datového stewarda. Rouse popisuje datového stewarda jako osobu, která má dokonalý přehled o podnikových datech a zároveň pomáhá nalézt podnikovým uživatelům taková data, která podpoří jejich business aktivity (48). Zastřešující rolí se širokou znalostí problematiky integrace dat je pak typicky systémový integrátor, jehož roli definuje Gartner jako *„osobu specializující se na implementaci, plánování, koordinaci, řízení, testování a zlepšování výpočetních operací“* (49). Systémový integrátor má k dispozici vývojáře, jež integraci technicky realizují. Na straně systému Google Analytics pak s daty nakládá role webového analytika. V případě přesahu Google Analytics do online marketingového prostředí, pak zde dále vstupuje dále role online marketingového specialisty, jež data prakticky zužitkovává pro online marketingové aktivity.

Z hlediska softwarové dimenze nás zajímá dostupné programové vybavení pro usnadnění integrace dat. Software přímo určený pro integraci dat do Google Analytics na současném trhu neexistuje. Existuje však řada způsobů, jak konektor pro integraci dat do Google Analytics relativně snadno sestavit, a to formou middleware. Microsoft definuje middleware jako skrytou vrstvu umožňující komunikaci a správu dat mezi distribuovanými aplikacemi prostřednictvím sítě (50). Názornou ukázkou middleware v praxi jsou tzv. webhooks, jež umožňují využití integrační metody Measurement Protocol. S těmito webhooks pracují například softwary jako JIRA, GitHub, Zendesk, Bamboo či Pingdom. Reálnou aplikací webhooks za účelem integrace dat do GA ze softwaru JIRA se zabývá v článku Barnes (51). Na trhu lze zároveň nalézt řadu aplikací, které importují data z různých marketingových

systémů do GA. Jedná se například o Analytics Importer, OWOX BI Pipeline či Supermetrics Uploader. Z hlediska jednotlivých integračních metod dat do GA pak existuje pro každou z těchto metod řada podpůrných nástrojů. V případě datového importu existují CSV validátory ověřující, zda je připravený soubor správně strojově čitelný. U integrační metody přes rozšířený měřící kód pak mohou být nápomocné doplňky do webových prohlížečů, které sledují odchozí požadavky z webu. Příkladem je doplněk do prohlížeče Google Chrome WASP.inspector. Poslední integrační metodou je Measurement Protocol. Pro sestavení a validaci požadovaného požadavku pro potřeby Measurement Protocolu sestavil Google pomocný nástroj Hit Builder.

Další globální dimenzí je hardware. Google Analytics žádné nároky na hardwarové vybavení pro realizaci integrace dat nemají. Důvodem je, že Google Analytics je poskytován jako služba v modelu SaaS (Software-as-a-Service). Potřebnou technickou infrastrukturu tak zajišťuje poskytovatel služby, tedy společnost Google. Nicméně určité nároky na technickou infrastrukturu pro realizaci integrace dat do Google Analytics mohou vzniknout na straně zdroje dat. Do hardwarové globální dimenze lze zároveň zařadit takový hardware, který je možné přímo integrovat s GA prostřednictvím některé z dostupných integračních metod. Praktickým případem takovéhoto hardware jsou IoT zařízení společnosti Lantronix, konkrétně produktová řada síťových prvků PremierWave a produktová řadu senzorů xSenso. Tato zařízení dokáží v reálném čase předávat data do GA (52).

V globální dimenzi věnující se uživatelské rozhraní ve vztahu k integraci dat do Google Analytics je tématem uživatelská složitost procesu a jeho znovupoužitelnost. Každá z jednotlivých integračních metod dat do GA má svůj proces velmi dobře popsán, přičemž některé prvky integračních metod se dokonce opírají o již existující standardy. Příkladem je Measurement Protocol pracující s metodami protokolu HTTP/1.1. Díky popisu jednotlivých metod a využití běžných standardů je uživatelská obsluha těchto metod relativně snadná. Neexistuje však žádný standard, který by určil, jaká data jsou pro jakou integrační metodu vhodná. V případě specifických skupin dat, kde dosud neexistuje předloha, je proto nutno pečlivě zvážit alternativy a zvolit takovou možnost, která je pro daný případ nejvhodnější.

Bezpečnost a ochranu dat klade společnost Google mezi své priority. Google má v tomto směru uvedené transparentní informace v rámci své oficiální nápovědy nástroje Google Analytics, kde se současně odkazuje na platnou legislativu týkající se ochrany osobních údajů. Mangold dokonce uvádí, že společnost Google přísně řídí přístup k zákaznickým datům i v rámci svých zaměstnanců. Mangold dále popisuje šifrování přenosu dat na servery nástroje Google Analytics za pomoci bezpečnostního mechanismu HTTP Strict Transport Security (HSTS) (53). Problematikou bezpečnosti integrace dat v širších souvislostech, především na straně potenciální zdrojů dat pro integraci dat do GA, se zabývá Allen, který ve svém textu popisuje bezpečnostní rizika a aktuální obavy z integrace dat (54).

Ekonomické a finanční aspekty jsou globální dimenzí, která hledí na problematiku skrze pracnost vyjádřenou časem a penězi. Pokud se integrace dat pojme komplexně, můžeme se opřít o fáze životního cyklu datové integrace, které popisuje Holman. Holman rozlišuje 7 fází životního cyklu, kterými jsou přístup, objevování, čištění, integrace, doručení, zpracování a monitoring (55). Jestliže se na integraci dat do GA budeme dívat pohledem projektu, pak jeho časovou i finanční náročnost stanovíme za pomoci technik odhadu

pracnosti projektu. Vahalík rozlišuje tyto kategorie technik odhadu pracnosti softwarového projektu: expertní odhady, analogie a odhady založené na modelu (56).

Poslední tři globální dimenze podle metodiky MMDIS jsou označovány jako metodicko-organizační. První z těchto dimenzí jsou metody. Ve smyslu předmětu práce se v této dimenzi zaměřuji především na techniky integrace. Guaraw rozvádí 4 základní techniky datové integrace, kterými jsou manuální integrace, integrace skrze middleware, virtualizace dat a integrace přes datové sklady (57). V obecnější rovině se na problém dívá Levy, který uvádí, že klíčovou technikou pro integraci dat je srozumitelný jazyk, jež popisuje obsah veškerých datových zdrojů. Díky popisu veškerých datových zdrojů lze posléze řešit možné datové překryvy mezi zdroji či nesourodost datových model a schémat (58). Ve vztahu k integraci dat do Google Analytics jsou v kontextu metod a technik myšleny především integrační metody, se kterými byl čtenář seznámen v rámci kapitoly 2.2.

Dimenze dokumentace charakterizuje stav stávající dokumentace k integraci dat do GA. Ve vztahu k dílčím integračním metodám a jednotlivým prvkům integrace je stavu dokumentace zcela dostačující. Dokumentace nástroje se dělí do dvou hlavních částí. První částí je uživatelská nápověda dostupná z URL adresy support.google.com/analytics/. Uživatelská nápověda Google Analytics je sadou článků popisující řadu základních funkcí nástroje včetně náčrtu jednotlivých integračních metod. Druhou částí dokumentace je příručka pro vývojáře dostupná z URL adresy developers.google.com/analytics/, která se na problematiku dívá ve větším technickém detailu, poskytuje ukázky programovacích kódů, popis dostupných API a nabízí pokročilé nástroje pro testování funkcí Google Analytics.

Poslední z dimenzí je dimenze managementu. V rámci této dimenze se na problematiku díváme ve vztahu k řízení zdrojů spotřebovávaných během integrace dat do GA. Pokud se na integraci dat do GA díváme jako na jakýkoliv jiný IS/ICT projekt, lze aplikovat doporučené techniky řízení zdrojů z oblasti řízení IS/ICT projektů. Schwalbe v oblasti řízení IS/ICT projektů popisuje celkem devět oblastí poznatků vedoucích k naplnění cíle projektu. Každá oblast poznatků je charakteristická svými metodami. Jedná se o řízení integrace, rozsahu, času, nákladů, kvality, lidských zdrojů, rizik, komunikací a obstarávání (59).

3.3 Shrnutí současného stavu řešení

Jednotlivé aspekty dílčích komponent i v globální roviny již byly nějakým způsobem řešené. Nicméně v rámci charakteristiky současného stavu řešení na úrovni několika zkoumaných dimenzí poznání, stanovených dle metodiky MMDIS, bylo zjištěno, že sice existují relativně dobře popsané řešení platné pro řešenou oblast jako celek, avšak dosud nebyly aplikovány konkrétně na problematiku integrace dat do GA. Zároveň bylo zjištěno, že dosud neexistuje ucelená metodika integrace dat do GA, která by celou problematiku zaobalila a čtenáře celým procesem realizace integrace dat do GA provedla. Na základě takto identifikované příležitosti si diplomová práce klade za cíl ucelenou metodiku sestavit.

4 Metodika diplomové práce

Hlavním výstupem této diplomové práce je metodika pojmenovaná jako Metodika integrace dat do Google Analytics, která je adaptací metodiky MMSP (Metodika pro Malé Softwarové projekty). Metodiku MMSP sestavila Rejnková v rámci své diplomové práce na VŠE v Praze za účelem výuky kurzu softwarového inženýrství. Díky svým vlastnostem avšak lze metodiku MMSP využít i pro řešení reálných projektů z praxe. Rejnková dále uvádí, že metodika MMSP je lokalizací a přizpůsobením open source metodiky Open Unified Proces (zkráceně OpenUP) verze 1.5.1.1, která je součástí projektu Eclipse Process Framework (60). Metodiku společně s praktickými ukázkami aplikace na konkrétních řešených případech sepsala v knižní podobě Buchalcevová, která uvádí, že metodika MMSP dále rozšiřuje metodiku OpenUP o některé prvky z dalších metodik, především z metodik agilních (61).

Rejnková zároveň zmiňuje, že metodika MMSP je určena výhradně pro projekty, jež jsou charakteristické následujícími vlastnostmi (60):

- *„délka projektu do 6 měsíců,*
- *do 10 členů týmu,*
- *projekt resp. vyvíjený systém není kritický“*

Metodika MMSP je vymezena prvky metodického obsahu, kterými jsou životní cyklus, fáze, iterace, aktivity, úlohy, disciplíny, role a pracovní produkty. Jednotlivé prvky jsou vzájemně úzce provázány. Provázanost prvků vychází již z jejich definic. Tabulka 4.1 uvádí popis každého ze zmíněných prvků metodického obsahu metodiky MMSP, jak jej uvádí Rejnková (60), případně následně upravuje a rozšiřuje Buchalcevová (61).

Tabulka 4.1: Popis prvků metodického obsahu MMSP (zdroj: vlastní zpracování dle (60) a (61))

Prvek metodického obsahu MMSP	**Popis prvku**
Životní cyklus	*„Posloupnost fází, iterací, aktivit a úloh.“* (60)
Fáze	*„Část životního cyklu tvořená iteracemi.“* (61)
Iterace	*„Opakovatelná sekvence aktivit.“* (61)
Aktivita	*„Množina souvisejících úloh.“* (61)
Úloha	*„Popisuje, co dělat, aby bylo dosaženo stanovených cílů.“* (61)
Disciplína	*„Skupina úloh, které se vztahují k určité oblasti zájmu.“* (60)
Role	Osoba, která provádí úlohy a vytváří pracovní produkty (60).
Pracovní produkt	V průběhu projektu vytvářený nebo modifikovaný výstup (60).

4.1 Obsahové vymezení metodiky

K obsahovému vymezení Metodiky integrace dat do Google Analytics poslouží kritéria kategorizace metodik tak, jak je vymezuje Buchalcevová. Jedná se tedy o kritéria zaměření metodiky, rozsah metodiky, váha metodiky, typ řešení, doména a přístup k řešení (62). Z hlediska zaměření metodiky se jedná o metodiku projektovou, respektive zaměřenou pouze na určitou oblast, kterou je integrace dat do GA.

Rozsah metodiky vymezuje Buchalcevová jako fáze, role a dimenze, které jsou metodikou pokryty (62). Metodika integrace dat do GA v tomto ohledu pokrývá fáze vymezené metodikou MMSP, tedy zahájení, rozpracování, konstrukce a zavedení (60). Počátek zahajovací fáze je vymezen záměrem realizace integrace dat a koncem zaváděcí fáze se rozumí zužitkování integrovaných dat. Role metodiky vycházejí z kapitoly 3.2 této práce, respektive z její části popisující dimenzi lidských zdrojů, kde byly identifikovány role datového stewarda, systémového integrátora, vývojáře, právníka, webového analytika, databázového specialisty a online marketingového specialisty. Vymezením dimenzí dle kritéria rozsahu metodiky se pak metodika omezí na dimenze technologie, data/informace a funkce/procesy vymezené metodikou MMDIS (33).

Z pohledu kritéria váha metodiky dle Buchalcevové je navržená metodika spíše rigorózní, tedy s přesně popsanými prvky. Dle kritéria typ řešení lze tuto metodiku charakterizovat jako integraci existujících řešení. Z hlediska kritéria doména má metodika přesah do oblastí Business Intelligence, Customer Relationship Management, Enterprise Application Integration a e-commerce. Posledním kritériem je kritérium přístup k řešení. Buchalcevová uvádí, že přístup k řešení se aplikuje pouze v případě vývoje nového SW, což není obsahem této metodiky. Kritérium přístup k řešení proto v kategorizaci zohledněno není (62).

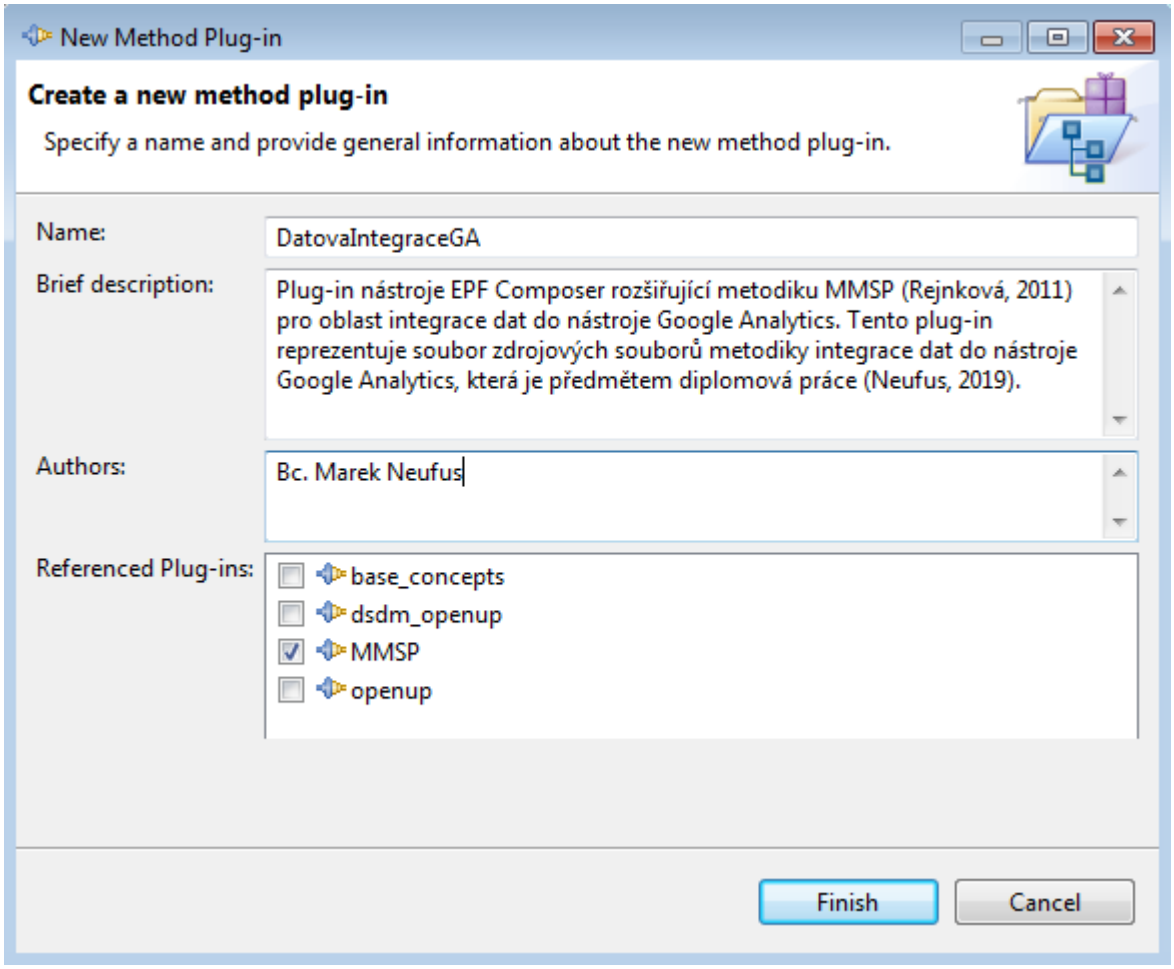
Tabulka 4.2: Kategorizace metodiky integrace do GA (zdroj: vlastní zpracování na základě (62))

Kritérium	Hodnocení
Zaměření metodiky	Projektová metodika v oblasti datové integrace.
Rozsah: Fáze	Dle MMSP: zahájení, rozpracování, konstrukce, zavedení (60).
Rozsah: Role	Datový steward, systémový integrátor, vývojář, webový analytik, právník, databázový a online marketingový specialista.
Rozsah: Dimenze	Funkce/procesy, data/informace a software dle MMDIS.
Váha metodiky	Převážně rigorózní s agilními prvky.
Typ řešení	Integrace existujících řešení.
Doména	BI, CRM, EAI a e-commerce.
Přístup k řešení	Nezávislé na přístupu k řešení.

4.2 Programové prostředí pro sestavení metodiky

Rejnková vytvořila samotnou metodiku MMSP v nástroji EPF Composer ve formě plug-inu k metodice OpenUP, přičemž diplomová práce Rejnkové slouží jako dokumentace metodiky MMSP obsahující popis předdefinovaných prvků (60). Výstupní podobou metodiky vytvořené v nástroji EPF Composer je webová stránka společně s přiloženými zdrojovými soubory. Pro zobrazení takto vytvořené metodiky doporučuje komunita vyvíjecí nástroj využívat webový prohlížeč Internet Explorer nebo Mozilla Firefox (63). Pro sestavení Metodiky integrace dat do Google Analytics je využit nástroje EPF Composer ve své poslední verzi, tedy verzi 1.5.2 vydané dne 5.6.2018. Samotná metodika MMSP je ve své aktuální verzi dostupná na webových stránkách Zlepšování softwarových procesů (Software Process Improvement) při VŠE v Praze dostupných z URL adresy spicenter.vse.cz/metodiky/.

Samotná příprava prostředí pro sestavení metodiky spočívala ve stažení aktuální verze nástroje EPF Composer (verze 1.5.2) a aktuální verze zdrojových souborů knihovny MMSP (verze 1.0). Po otevření knihovny MMSP jsem v projektu založil nový plug-in reprezentující Metodiku integrace dat do Google Analytics. Tento plug-in, v projektu pojmenovaný jako „DatovaIntegraceGA“, odkazuje na již existující metodiku MMSP.



Obrázek 4.1: Vytvoření plug-inu pro metodiku integrace dat do GA
Zdroj: vlastní zpracování

Tímto krokem bylo vytvořeno programové prostředí připravené k naplnění obsahem. V rámci plnění metodiky obsahem lze vytvořit a odstraňovat jednotlivé prvky či převzít prvky původní metodiky a v nově vytvořeném prostředí je modifikovat. S prvky a vazbami mezi prvky Metodiky integrace dat do Google Analytics bude čtenář seznámen v kapitole 5 Návrh metodiky integrace dat. K seznámení se s výslednou metodikou může čtenář také využít přílohu této diplomové práce, kde se nachází metodika v podobě webové prezentace sestavená přímo nástrojem EPF Composer, kde jsou jednotlivé vazby mezi použitými prvky metodiky provázány hypertextovými odkazy (viz příloha B).

K diplomové práci jsou dále přiloženy zdrojové soubory výsledné metodiky (viz příloha C), které v nástroji EPF Composer umožňují další úpravy. Výslednou metodiku lze díky přiloženým zdrojovým souborům v budoucnu snadno přizpůsobit změnám v používaných technologiích, metodách či jiných aspektech mající dopad na datovou integraci dat do GA. Vzhledem k tomu, že navržená metodika se zabývá pouze jednosměrnou integrací, tedy integrací z podnikových systémů do GA, určitě lze metodiku obsahově rozšířit i o datové integrace mající opačný nebo obousměrný datový tok. Alternativně lze navrženou metodiku adaptovat a rozšířit na oblast datové integraci mezi libovolnými jinými aplikacemi.

Vztahy mezi prvky navržené metodiky jsou zároveň znázorněny v tabulkách dostupných v příloze A této diplomové práce. Příloha A obsahuje celkem 4 tabulky, kde každá z tabulek odpovídá jedné z fází životního cyklu metodiky. Příloha A tak může posloužit pro rychlou orientaci se v rozsahu metodiky, či pro nalezení konkrétních prvků, jejichž hlubší význam je dohledatelný v textu této diplomové práce a v přiložené webové verzi metodiky.

4.3 Validace platnosti metodiky

Platnost navržené metodiky bude demonstrována na ukázkovém případu užití z praxe. Čtenář bude na tomto případu užití proveden celým procesem integrace dat do Google Analytics na základě navržené metodiky v jejím vymezeném obsahu. Ověřením platnosti navržené metodiky se zabývá kapitola číslo 6.

Ukázkový případ užití pro ověření platnosti metodiky bude demonstrován na fiktivním maloobchodním řetězci potravin, jehož marketingoví pracovníci se rozhodli spustit online marketingovou kampaň s cílem informovat zákazníky o věrnostní kartě. Za tímto účelem budou na základě integrovaných dat vytvořeny uživatelské segmenty v nástroji Google Analytics, na které lze cílit online marketingovou inzerci skrze inzertní systém Google Ads. Procesem integrace dat do Google Analytics za účelem vytvoření uživatelských segmentů bude využito navržené metodiky z kapitoly číslo 5.

5 Návrh metodiky integrace dat

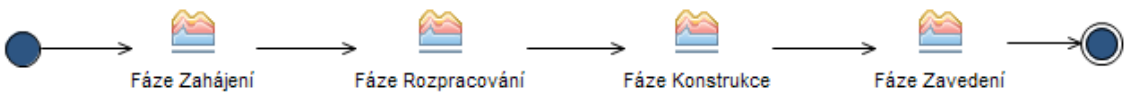
Páteřním prvkem metodiky je životní cyklus, který se skládá ze čtyř fází. Fáze v Metodice integrace dat do Google Analytics zůstávají totožné jako v případě metodiky MMSP s tím rozdílem, že je přizpůsoben jejich význam a obsah řešené problematice. Metodika tak po vzoru MMSP rozlišuje fáze Zahájení, Rozpracování, Konstrukce a Zavedení (60). Každá z fází je současně definována jako iterace, která je zakončená milníkem. Metodika tak zároveň pracuje se čtyřmi iteracemi a čtyřmi milníky. Každá z iterací je dále rozdělena na sadu souvisejících úloh, které jsou v terminologii metodiky nazývány aktivity. Rozpadem aktivit se dostaneme k nejmenším prvkům, jež jsou metodikou vymezeny, a sice k úlohám. Metodika integrace dat do Google Analytics rozlišuje celkem 53 úloh rozdělených do 14 aktivit. Avšak nutno podotknout, že některé z úloh i aktivit se napříč iteracemi mohou opakovat. Opakují se především o úlohy, které danou iteraci plánují, řídí a následně vyhodnocují. Reálně tak metodika popisuje 42 unikátních úloh a 10 unikátních aktivit.

Dalším prvkem jsou dle MMSP role. Metodika integrace dat do Google Analytics vymezuje projektový tým čítající celkem 8 rolí, jedná se o role: Systémový integrátor, Datový steward, Webový analytik, Právník, Online marketingový specialista, Databázový specialista, Vývojář a Zainteresovaná strana. Role jsou charakteristické svými schopnostmi, dovednostmi, pravomocemi a zodpovědnostmi. Kromě toho role vykonávají zmíněné úlohy a současně jsou zodpovědné za pracovní produkty (60). Pracovní produkty Metodiky integrace dat do Google Analytics taktéž vycházejí z MMSP. Některé pracovní produkty Metodika integrace dat do Google Analytics z MMSP převzala a přizpůsobila řešené problematice, další nově zavedla (například pracovní produkt Datový model) a některé nezohledňuje. Posledním prvkem metodiky jsou disciplíny, které nejsou nic jiného, než množina úloh vztažená ke konkrétní oblasti zájmu (60). Smyslem disciplín je úlohy přehledně a logicky kategorizovat.

Přiložená metodika ve formě webového prezentace dále obsahuje podstránky s pokyny, které čtenáři vysvětlují význam metodiky, popisují řešenou problematiku, vysvětlují jak s metodikou zacházet a předkládají souvislý seznam použité literatury. Význam těchto podstránek s pokyny spočívá v tom, aby byl čtenář schopen s metodikou nakládat alespoň v základním rozsahu, aniž by musel nejprve studovat tuto diplomovou práci.

5.1 Životní cyklus metodiky

Životní cyklus Metodiky integrace dat do Google Analytics je složen ze čtyř fází, jež jsou totožné s MMSP. Jedná se o fáze Zahájení, Rozpracování, Konstrukce a Zavedení (60).



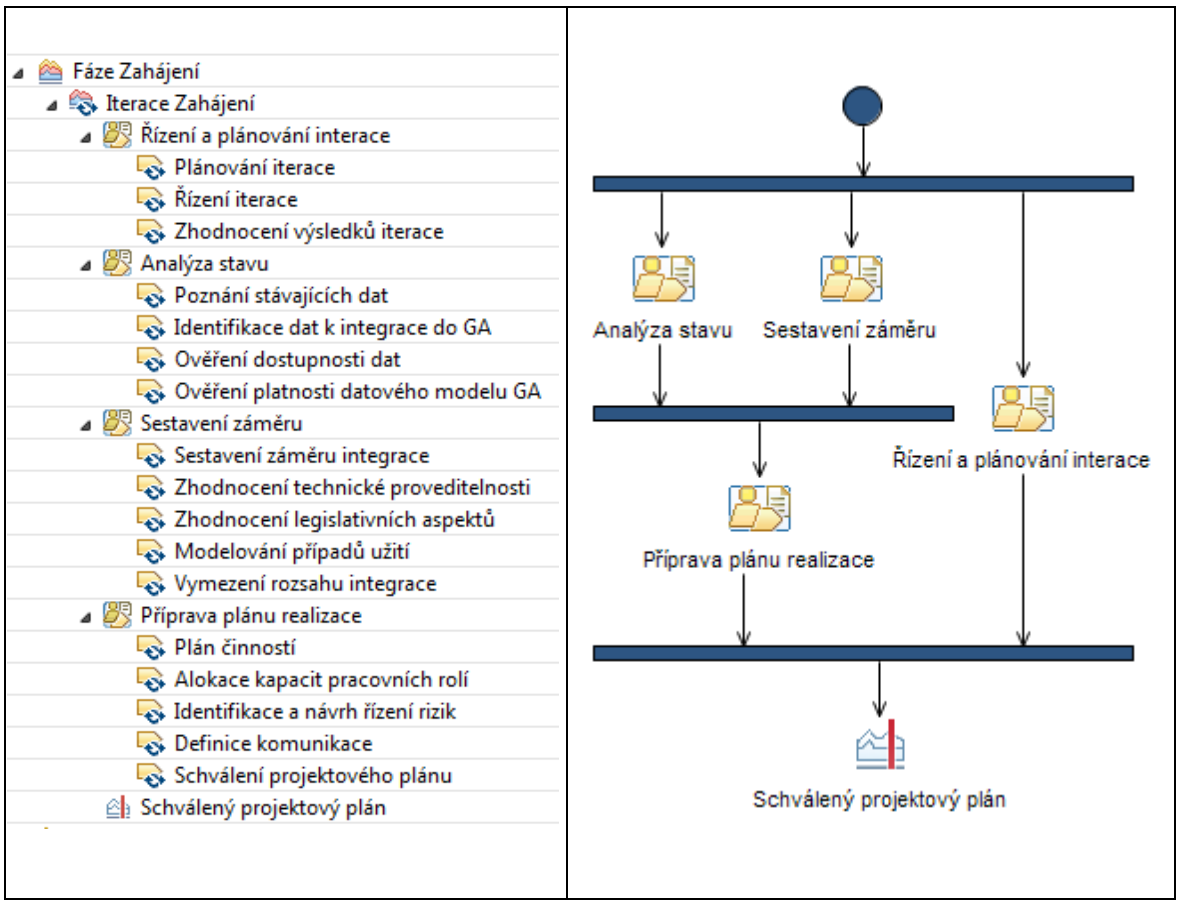
Obrázek 5.1: Životní cyklus metodiky integrace dat do Google Analytics
Zdroj: vlastní zpracování na základě (60) v nástroji EPF Composer

Přestože počet, pojmenování a posloupnost fází životního cyklu Metodiky integrace dat do Google Analytics je zcela totožný s metodikou MMSP, svým obsahem se jednotlivé fáze vzájemně zásadně odlišují. S obsahem jednotlivých fází je čtenář seznámen v průběhu nadcházejících kapitol. Důvodem obsahových odlišností jednotlivých fází je, že metodika MMSP je obecnou metodikou pro vývoj menších softwarových projektů (60), avšak Metodika integrace dat do GA má jednoznačně a úzce vymezenou oblast platnosti, jež je popsána kapitolou 4.1 Obsahové vymezení metodiky.

5.1.1 Fáze Zahájení

Jedná se o první fázi životního cyklu. Fáze se skládá ze čtyř aktivit, kterými jsou Řízení a plánování iterace, Analýza stavu, Sestavení záměru a Příprava plánu realizace. Celou fázi uzavírá milník Schválený projektový plán. Předmětem této fáze je poznání stávajícího stavu dostupných dat použitelných k integraci do Google Analytics a sestavení projektového záměru vymezující rozsah integrace.

Aktivity Analýza stavu a Sestavení záměru běží souběžně, jelikož záměr je podnětem k nalezení patřičných dat v rámci analýzy stavu, ale analýza stavu je zároveň potřebnou pro zhodnocení a kompletaci záměru. Projektový záměr je následně podkladem pro přípravu projektového plánu. Z projektového plánu pak vychází veškerý další postup v rámci dalších fází. Schválením projektového plánu fáze Zahájení končí.



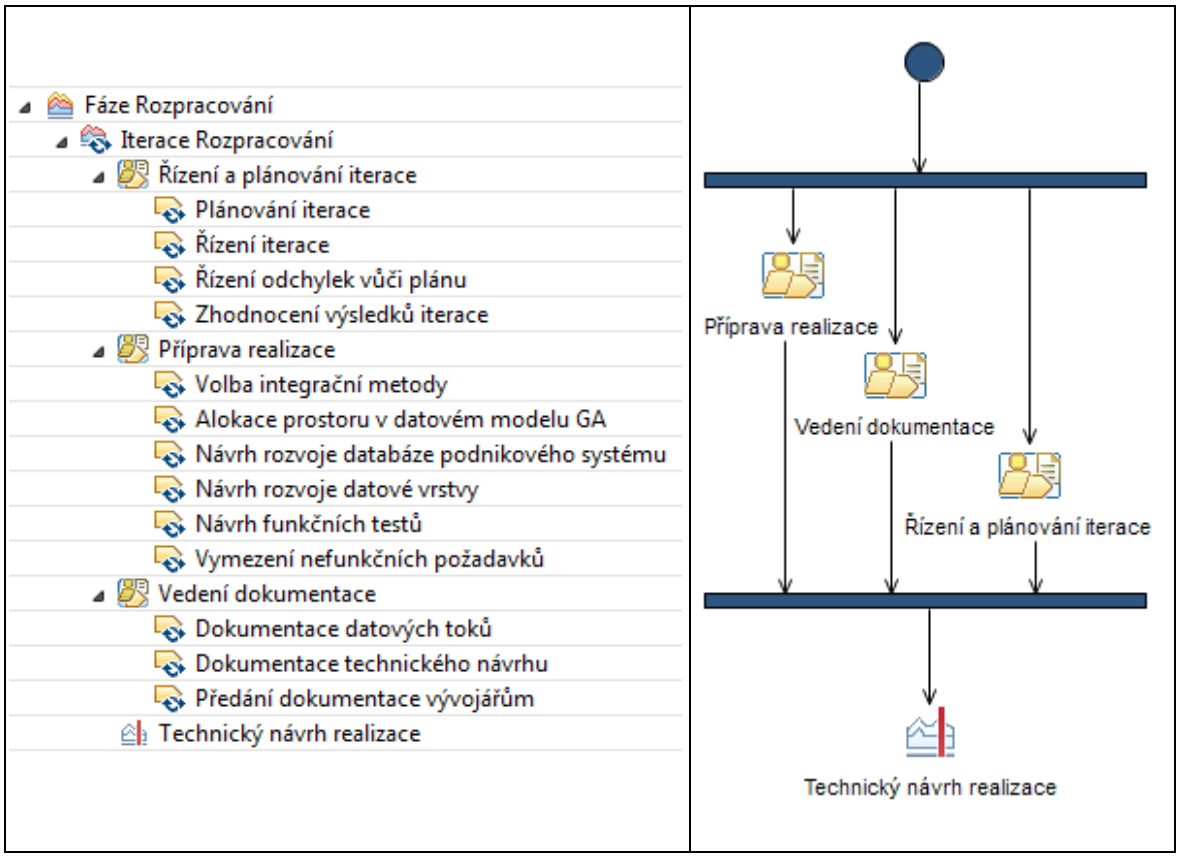
Obrázek 5.2: Fáze zahájení
Zdroj: vlastní zpracování v nástroji EPF Composer

5.1.2 Fáze Rozpracování

Fáze Rozpracování má z předchozí fáze k dispozici jasně vymezený záměr a rozsah integrace dat do Google Analytics, což umožňuje postoupit o krok dále, a sice směrem k technickým náležitostem integrace. Z hlediska této fáze je zásadní aktivitou Příprava realizace, která volí vhodnou integrační metodu, alokuje prostor v datovém modelu GA, připravuje technické prostředí pro realizaci integrace a vymezuje požadavky pro testování.

Souběžně jsou realizovány i zbylé dvě aktivity, kterými jsou Vedení dokumentace a Řízení a plánování iterace. Aktivita Vedení dokumentace je aktivitou doprovodnou, která veškeré náležitosti eviduje do příslušného typu dokumentace. Účelem této aktivity je, aby proces integrace byl snadno dešifrovatelný i pracovníky, kteří se na projektu v době jeho realizace nepodíleli, avšak mohli by v budoucnu určité informace o projektu potřebovat. Součástí dokumentace se předpokládá úplný popis datových toků z podnikových systému do nástroje Google Analytics a technický návrh realizované integrace. Tento obsah dokumentace je následně podkladem pro realizaci vývoje v rámci další fáze.

Poslední aktivitou je Řízení a plánování iterace. Jedná se o aktivitu, která se nachází napříč všemi fázemi. Jejím smyslem je danou iteraci řídit tak, aby byla zrealizována ve stanoveném čase, se stanovenými náklady a v požadované kvalitě. Oproti fázi Zahájení je aktivita Řízení a plánování iterace v této fázi rozšířena o jednu novou úlohu, kterou je Řízení odchylek vůči plánu. Smyslem této úlohy je zachování kompatibility s projektovým plánem formou řízení odchylek, tedy porovnáváním skutečných hodnot vůči naplánovaným.



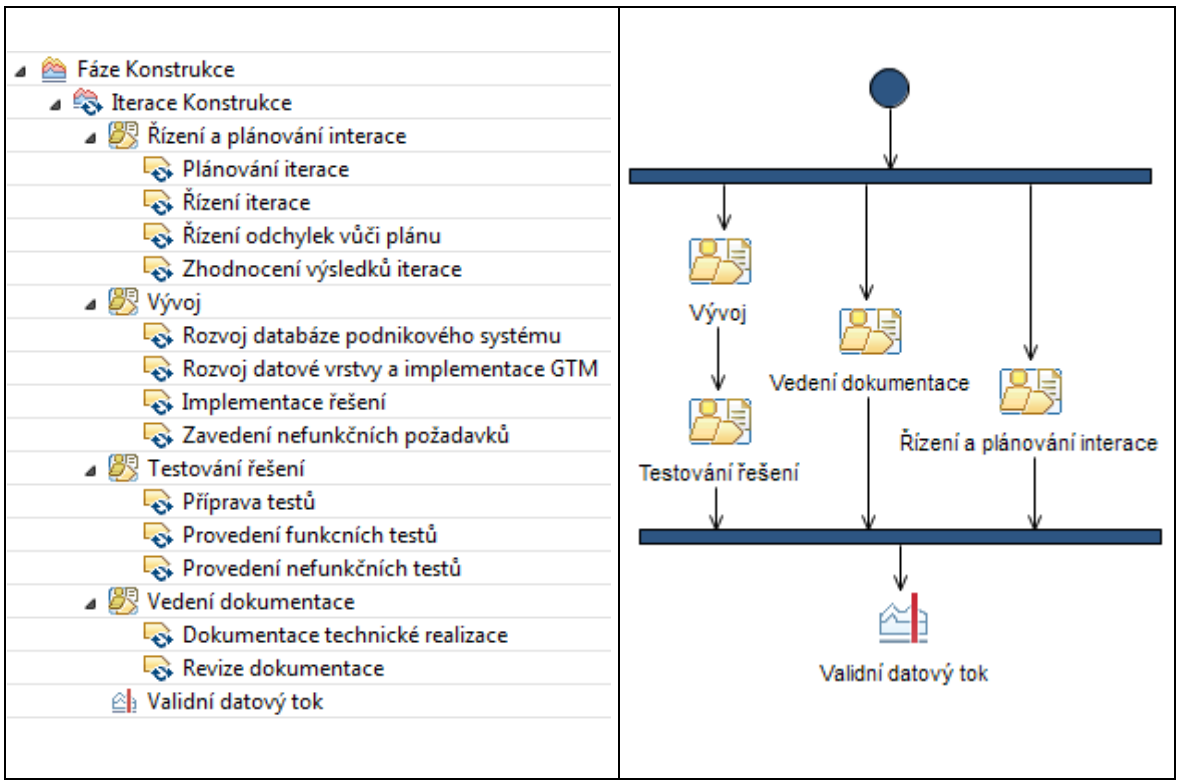
Obrázek 5.3: Fáze Rozpracování
Zdroj: vlastní zpracování v nástroji EPF Composer

5.1.3 Fáze Konstrukce

Předmětem fáze Konstrukce je primárně samotná technická realizace integrace. Konstrukce má k dispozici z předchozích fází podrobný technický návrh a dokumentaci navržených datových toků. Na základě těchto podkladů je v aktivitě Vývoj nejprve rozvinuto databázové prostředí o potřebné prvky (netýká se manuálního datového importu). A v případě integrační metody prostřednictvím rozšířeného měřícího kódu je v této fázi dále připravena datová vrstva a implementován měřící kód nástroje Google Tag Manager. Posléze probíhá již samotná implementace řešení (včetně funkčních požadavků), které je následně rozvinuto o navržené nefunkční požadavky. Těmito nefunkčními požadavky se dle Šilhavého rozumí například zabezpečení, výkonost, bezpečnost nebo spolehlivost (64).

Další neméně důležitou aktivitou konstrukční fáze je aktivita Testování řešení, která svým úvodem testy nejprve navrhuje a připravuje vhodné testovací prostředí. Následují úlohy zaměřené na testování funkčních požadavků, tedy takových funkcí, které jsou podmínkou správného fungování (64). Podmínkami správného fungování lze v tomto případě chápat například bezchybný datový tok v souladu s dostupnými specifikacemi či požadované uložení dat v příslušných místech datového modelu GA. Poslední úloha aktivity Testování řešení se následně zabývá nefunkčními požadavky, kde se jedná především o bezpečnost datového toku, řízení přístupu k datům či pravidelnost datového toku.

Součástí fáze Konstrukce je dále aktivita Řízení a plánování iterace, která obsahuje stejné úlohy a plní stejnou funkci jako tomu bylo v případě fáze Rozpracování, a aktivita Vedení dokumentace, kde je celý proces realizace zaznamenán a následně zrevidován. Milníkem fáze je Validní datový tok, tedy zrealizovaná a funkční integrace dat.



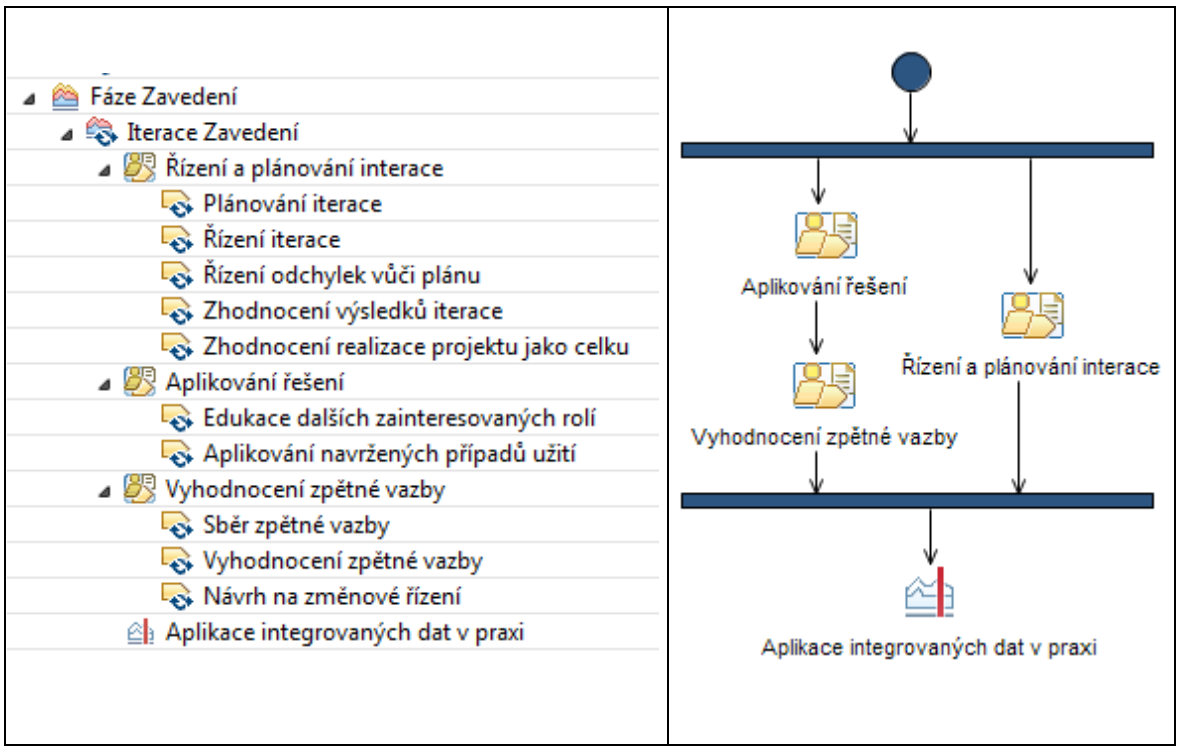
Obrázek 5.4: Fáze Konstrukce
Zdroj: vlastní zpracování v nástroji EPF Composer

5.1.4 Fáze Zavedení

Finální fází životního cyklu této metodiky je fáze Zavedení. Fáze je složena z aktivit Řízení a plánování iterace, Aplikování řešení a Vyhodnocení zpětné vazby. Aktivita Řízení a plánování iterace je vůči předchozím dvěma fázím rozšířena o jednu novou úlohu, kterou je Zhodnocení realizace projektu jako celku. Tato úloha si klade za cíl vyhodnotit realizaci projektu vůči záměru projektu. V terminologii projektového řízení bychom se bavili o závěrečné zprávě projektu (59).

Další aktivitou je Aplikování řešení. Smyslem aktivity je nejprve edukovat zainteresované role o nových možných případech užití na základě provedené integrace a následně tyto případy užití zrealizovat v praxi. Zdrojem případů užití je pracovní produkt Případy užití, který byl navržen již ve fázi Zahájení v rámci úlohy Modelování případů užití. Aplikace řešení v praxi v kontextu integrace dat do Google Analytics se rozumí využití pro pokročilé datové analýzy, optimalizaci webového obsahu či využití v online marketingu formou remarketingových publik v inzertním systému Google Ads.

Poslední aktivitu této fáze je Vyhodnocení zpětné vazby. Zpětná vazba je v tomto kontextu vnímána jako sběr uživatelských připomínek ke zrealizované integraci, lze na ni nahlížet jako na jistou formu uživatelského testování. Zpětná vazba je po sesbírání konsolidována a vyhodnocena. Zpětná vazba, jež je vyhodnocena jako předmětná, je následně podkladem pro návrh na změnové řízení, které v případě schválení spouští celý projekt znovu od začátku s modifikovanými vstupními parametry. Fázi Zavedení ukončuje milník Aplikace integrovaných dat v praxi, který lze chápat jako posouzení systémového integrátora, že projekt jako celek splnil úvodem vytyčený cíl. Tímto milníkem je celý projekt ukončen.



Obrázek 5.5: Fáze Zavedení
Zdroj: vlastní zpracování v nástroji EPF Composer

5.2 Role metodiky

Role vykonávají stanovené úlohy a jsou zodpovědné za přiřazené pracovní produkty (60). Jako role nejsou v kontextu metodik MMSP a OpenUP chápány konkrétní osoby, ale pouze obecný pracovní popis s charakteristikou potřebných dovedností a schopností. V praxi tak může jedna osoba zastupovat více rolí, nebo jedna role může být zastupována více osobami. Z agilního pojetí metodiky OpenUP vyplývá, že je kladen vysoký důraz na interakce mezi rolemi a jejich vzájemnou spolupráci. Přímo metodika OpenUP tvrdí, že *„Nikdo nevytvoří skvělý software sám. Ale tým, který spolupracuje, může vytvořit mimořádné věci“* (65).

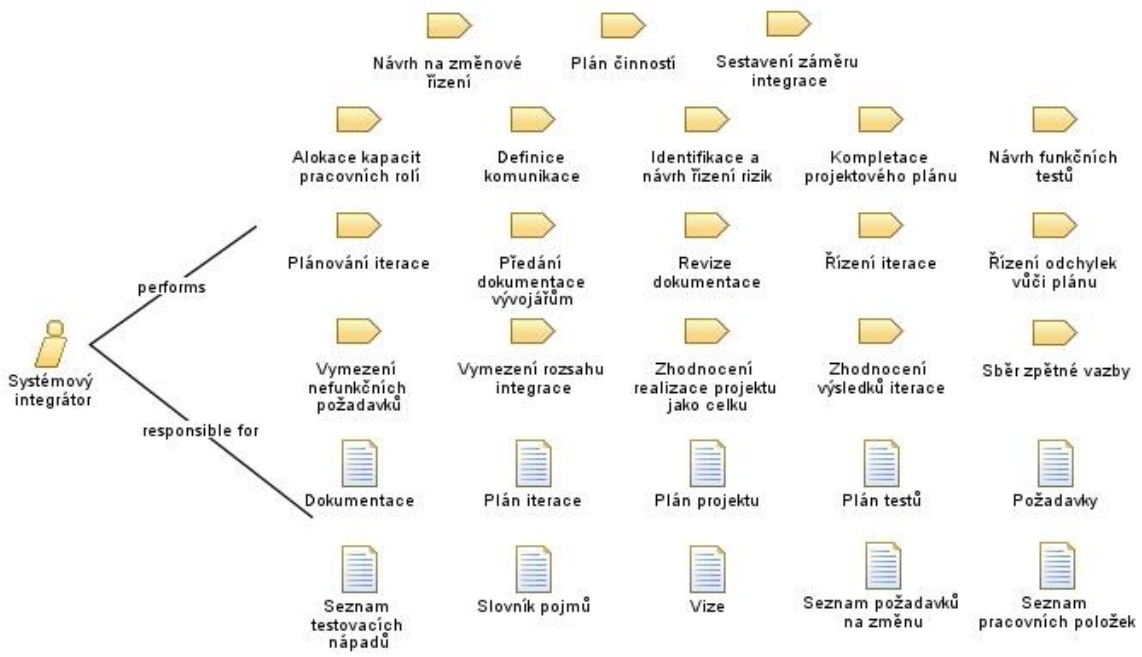
Metodika integrace dat do Google Analytics vymezuje celkem 8 rolí, které jsou postupně představeny v rámci nadcházejících podkapitol. Jedná se o role Systémový integrátor, Datový steward, Webový analytik, Právník, Online marketingový specialista, Databázový specialista, Vývojář a Zainteresovaná strana. Role mohou pocházet z jednoho interního projektového týmu, ale mohou být ve vztahu vzájemně spolupracujících subdodavatelů služeb. Takovýmto vztahem je například konzultační marketingová agentura (s rolemi jako Webový analytik, Online marketingový specialista či Právník) ve vztahu ke klientské straně (reprezentující role Systémového integrátora či Zainteresované strany) a dále ve vztahu k dodavateli webové aplikace (s rolemi Datový steward, Databázový specialista či Vývojář). Některé z uvedených rolí se prolínají napříč všemi fázemi, jiné vstupují pouze do některých úloh. Vzhledem k tomu, že některé role vstupují do projektu jen velmi okrajově, lze je obsadit i zcela nezávislou, externě postavenou osobu s cílem vykonat pouze danou úlohu.

5.2.1 Role Systémový integrátor

Role Systémového integrátora v rámci této metodiky vykonává nejvíce úloh a je zodpovědná za nejvíce pracovních produktů ze všech vymezených rolí. V kontextu klasického řízení projektů ji lze připodobnit projektovému manažerovi s tím rozdílem, že Systémový integrátor je osobou, která do problematiky integrace dat vidí ve větším technickém detailu a má bohaté zkušenosti se systémovou integrací. Jedná se o roli, která celý projekt integrace koordinuje a má dokonalý přehled o průběhu během jeho celého trvání. Předpokladem role je schopnost pracovat s manažerskými technikami tak, aby celý projekt byl zdárně doveden do svého cíle a hlubokou úroveň orientace se v řešené problematice. Systémový integrátor je zároveň hlavní kontaktní osobou pro zadavatele projektu, ale i pro veškeré další zúčastněné projektové role.

Systémový integrátor vykonává celkem 15 úloh, mezi kterými jsou primárně úlohy řídící. Jedná se tak o úlohy, které řídí a plánují jednotlivé iterace, sestavují záměr a plán projektu, či definují způsob komunikace a spolupráce v projektovém týmu. Kromě řídících úloh vykonává Systémový integrátor úlohy související s dokumentací, respektive jejím vedením, kompletací a revizemi. Ostatně Systémový integrátor je za dokumentaci zodpovědný.

Další úlohy Systémového integrátora spočívají v řízení požadavků na integraci dat a to jak v počáteční fázi, kdy je vymezen rozsah a účel integrace, tak v poslední fázi, kdy se sestavují návrhy na změny na základě zpětné vazby od uživatelů. S řízením požadavků úzce souvisí i testování. Systémový integrátor na základě požadavků sestavuje kritéria testování tak, aby výsledná integrace byla v plně v souladu s těmito požadavky a plně funkční.



Obrázek 5.6: Role Systémový integrátor
Zdroj: vlastní zpracování v nástroji EPF Composer

5.2.2 Role Datový steward

Role Datového stewarda se zabývá správou a kategorizací dat za pomoc meta údajů na straně podnikového systému, tedy zdrojového systému. Díky tomu má o těchto datech dokonalý přehled. Datový steward je zároveň schopen posoudit dostupnost dat na straně zdrojového systému a posoudit tak vhodnost využití dat pro účely integrace. Tyto znalosti jsou využity především při identifikaci dat vhodných pro integraci dat do Google Analytics za účelem podpory business aktivity podniku.

Datový steward se díky hlubokým znalostech dat na straně zdrojového systému také podílí na dokumentaci, respektive jejích částech týkajících se datových toků a technické realizace návrhu integrace. Kromě toho je Datový steward dále zodpovědný za část pracovního produktu Datový model, která popisuje datový model zdrojového systému. Součástí pracovního produktu Datový model je dále datový model nástroje Google Analytics, za jehož správnost je zodpovědná role Webového analytika.

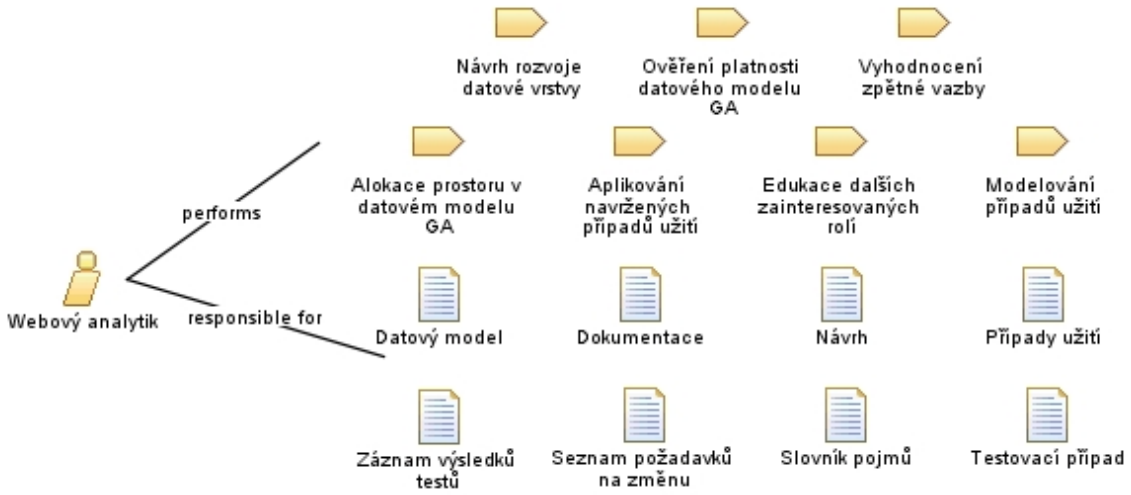


Obrázek 5.7: Role Datový steward
Zdroj: vlastní zpracování v nástroji EPF Composer

5.2.3 Role Webový analytik

Webový analytik je rolí, která má podrobné znalosti o fungování a možnostech nástroje Google Analytics. Ověřuje tak aktuálnost datového modelu GA a alokuje v něm příslušný prostor pro datovou integraci. Součástí popisu práce role Webového analytika je dále návrh a aplikace navržených případů užití. Jelikož velmi dobře zná možnosti nástroje Google Analytics, tak je zároveň schopen namodelovat případy užití s integrovanými daty podporující různorodé podnikové aktivity. Některé z namodelovaných případů užití dokonce později sám může realizovat (například datové analýzy).

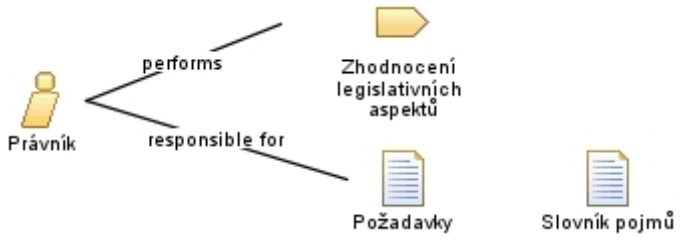
Webový analytik se dále podílí na technickém návrhu realizace integrace formou konzultací, respektive konzultuje vhodnou integrační metodu pro daný typ dat a technický návrh ze svého úhlu pohledu zaznamenává do příslušného oddílu dokumentace. Kromě toho se dále podílí na části testovacího procesu, kde ověřuje, že příslušná data jsou v Google Analytics po realizaci integrace dostupná v potřebné podobě tam, kde dostupná být mají.



Obrázek 5.8: Role Webový analytik
Zdroj: vlastní zpracování v nástroji EPF Composer

5.2.4 Role Právník

Právník je pouze okrajovou rolí, která záměr integrace zhodnocuje z hlediska legislativních aspektů, na jejichž základě navrhuje úpravy požadavků integrace tak, aby byly se současnou platnou legislativou v souladu. Právník dále rozšiřuje slovník použitých pojmů o relevantní právnickou terminologii, která byla v průběhu projektu použita.

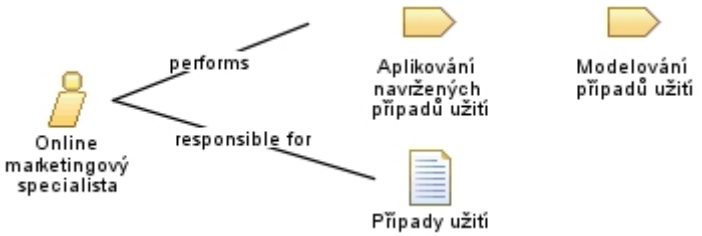


Obrázek 5.9: Role Právník
Zdroj: vlastní zpracování v nástroji EPF Composer

5.2.5 Role Online marketingový specialista

Online marketingový specialista je odborníkem na oblast marketingové komunikace, kde zná možnosti běžně používaných inzertních systémů a v rámci těchto systémů spravuje reklamní kampaně. V nástroji Google Analytics se Online marketingový specialista orientuje, protože skrze něj vyhodnocuje úspěšnost a návratnost spravovaných kampaní. Online marketingový specialista je tak schopen posoudit, jaká data z nástroje Google Analytics mu pomohou lépe vyhodnocovat úspěšnost kampaní a jaká data mohou jím spravované kampaně nějakým způsobem vylepšit.

Typickým příkladem vylepšení inzerce je sestavení pokročilých remarketingových publik na základě integrovaných dat v nástroji Google Analytics, které je možné naimportovat do inzertního systému Google Ads a lépe tak zacílit komunikaci na relevantnější uživatele. Právě tento typ případů užití, kdy dochází k propojení integrovaných dat v Google Analytics s marketingovou komunikací, tato role navrhuje a později aplikuje v praxi.



Obrázek 5.10: Role Online marketingový specialista
Zdroj: vlastní zpracování v nástroji EPF Composer

5.2.6 Role Databázový specialista

V případě volby integrační metody Measurement Protocol se nabízí ukládat unikátní ID uživatele vygenerované nástrojem Google Analytics (získané z cookie _ga) do databáze zdrojového systému (též podnikového systému). K tomu, aby byl identifikátor patřičně do databáze zaznamenán a zároveň byl využitelný pro potřeby implementace integrační metody Measurement Protocol, je nejprve potřeba databázi rozvinout. Právě návrhem rozvoje a samotným rozvojem databáze podnikového systému se zabývá role Databázový specialista. Veškeré změny, jež Databázový specialista v rámci rozvoje databáze navrhne, jsou nejprve zaznamenány do pracovního produktu Návrh. Následná realizace je pak předmětem pracovního produktu Implementace.

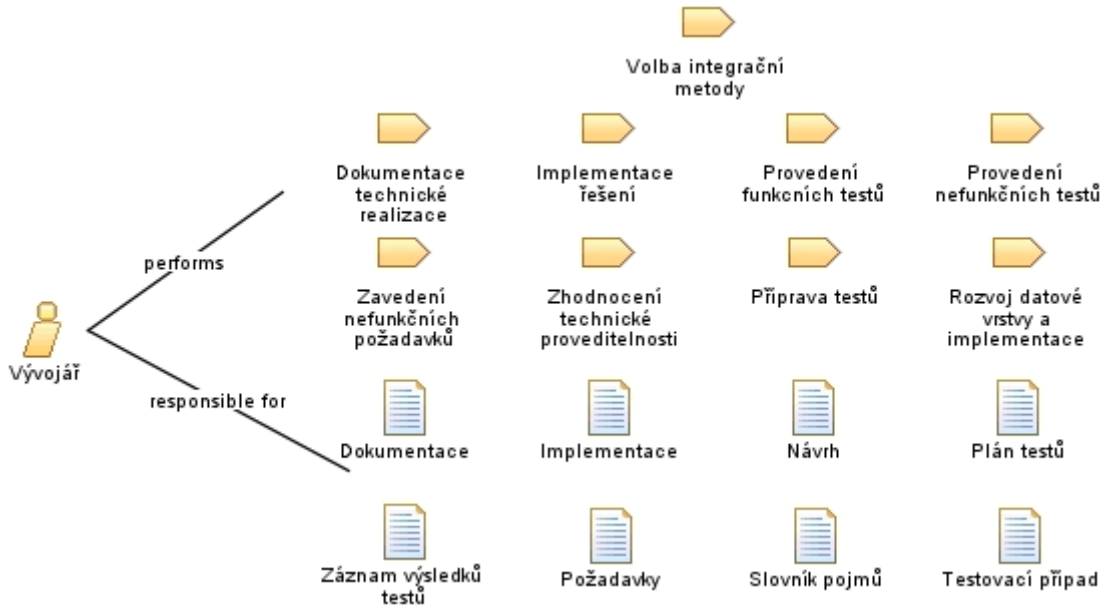


Obrázek 5.11: Role Databázový specialista
Zdroj: vlastní zpracování v nástroji EPF Composer

5.2.7 Role Vývojář

Vývojář je v kontextu této metodiky rolí, která je schopna provádět úpravy ve zdrojovém kódu zdrojového systému (tedy v podnikovém systému, který zahrnuje i webovou aplikaci měřenou kódem Google Analytics) a připravené návrhy integrace dat do Google Analytics zrealizovat, respektive implementovat řešení. Vzhledem ke znalostem o zdrojového systému poskytuje Vývojář v rané fázi projektu cenné konzultace týkající se technické proveditelnosti integrace. Vývojář je dále rolí, která implementované řešení, společně s rolí Webového analytika, otestuje na základě předem stanovených požadavků Systémovým integrátorem.

Vývojář se dále podílí na řadě pracovních produktů. Jedním z nich je Dokumentace, kterou rozšiřuje o informace týkající se samotné technické realizace. Vývojář dále rozvíjí Slovník pojmů, kam zaznamenává vysvětlení použité terminologie pro pochopení dokumentace a zajištění konzistentnosti komunikace mezi členy projektového týmu.



Obrázek 5.12: Role Vývojář
Zdroj: vlastní zpracování v nástroji EPF Composer

5.2.8 Role Zainteresovaná strana

Tato role přímo vychází z metodiky MMSP. Jedná se o veškeré zainteresované subjekty, které mají na úspěšné realizaci projektu svůj zájem. Právě realizace integrace dat do GA nějakým způsobem tyto zainteresované subjekty ovliňuje a proto je důležité takovou roli v projektu zohlednit. Správná komunikace směrem k Zainteresované straně je základním předpokladem úspěchu projektu. Jako Zainteresovanou stranu je možné chápat například potenciální uživatele či investory (60). Pro Zainteresovanou stranu je kontaktní osobou Systémový integrátor, se kterým je Zainteresovaná strana v aktivním kontaktu dle agilních praktik metodiky OpenUP, které mimo jiné udávají aktivní spolupráci Zainteresované strany a kontaktním bodem řešeného projektu jako předpoklad úspěchu (65).

5.3 Disciplíny metodiky a jejich úlohy

Nejpočetněji zastoupeným prvkem metodiky jsou úlohy. Metodika integrace dat do Google Analytics obsahuje celkem 42 unikátních úloh. Tyto úlohy jsou pro přehlednost uskupeny do množin týkajících se konkrétní oblasti zájmu, které jsou v terminologii metodiky nazývány disciplíny (60). Disciplíny Metodiky integrace dat do Google Analytics vycházejí z metodiky MMSP. Oproti MMSP pracuje tato metodika navíc s disciplínami Dokumentace a Aplikace, a naopak nezohledňuje disciplínu Architektura. Metodika integrace dat do Google Analytics tak pracuje celkem se sedmi disciplínami, kterými jsou Aplikace, Dokumentace, Požadavky, Řízení projektu, Řízení změn, Testování a Vývoj. Význam jednotlivých disciplín je představen v rámci nadcházejících podkapitol. U každé z disciplín je zároveň stručně charakterizován význam úloh, které jsou do dané disciplíny zařazeny.

5.3.1 Disciplína Požadavky a její úlohy

Smyslem úloh disciplíny Požadavky je jednoznačné vymezení rozsahu realizované integrace a požadavků, které jsou na integraci kladeny. Disciplína si zároveň klade za cíl posoudit, zda a do jaké míry je navržená integrace proveditelná z hlediska legislativních a technických aspektů. Úlohy této disciplíny primárně modifikují pracovní produkt Požadavky, který jasně vymezuje rozsah a záměr realizované integrace. Sestavený záměr je později podkladem pro pracovní produkt Plán projektu, kterým se zabývají úlohy disciplíny Řízení projektu. Z úloh disciplíny Požadavky však vyplývá vymezení, rozsah a záměr integrace pouze v koncepční rovině, nikoliv na úrovni technického návrhu. Technickým návrhem se totiž zabývají především úlohy z disciplíny Vývoj a zaznamenávají jej úlohy disciplíny Dokumentace.

Je důležité, aby definované požadavky byly jasné, srozumitelné a především přesné. Požadavky jsou v rámci úloh disciplíny rozděleny do dvou podmnožin. První podmnožinou jsou požadavky funkční, které vymezuje úloha Vymezení rozsahu integrace a druhou množinou jsou požadavky nefunkční stanovené úlohou Vymezení nefunkčních požadavků.

Tabulka 5.1: Úlohy disciplíny Požadavky

Úloha	Charakteristika
Sestavení záměru integrace	Formování vize a prvotních požadavků zamýšlené integrace na základě analýzy současného stavu dostupných dat.
Zhodnocení legislativních aspektů	Posouzení navrženého záměru z hlediska souladu s platnou legislativou, etiky a podmínek společnosti Google.
Zhodnocení technické proveditelnosti	Posouzení navrženého záměru z hlediska technických možností na straně zdrojového systému a na straně GA.
Vymezení rozsahu integrace	Upřesnění požadavků na základě legislativního a technického posouzení. Jednoznačné vymezení funkčních požadavků.
Vymezení nefunkčních požadavků	Vymezení požadavků, které nemají přímý vliv na funkčnost integrace, ale je nutné je zohlednit (například bezpečnost).

5.3.2 Disciplína Řízení projektu a její úlohy

Úlohy disciplíny Řízení projektu mají hned dvě roviny. Tou první rovinou je sestavení pracovního produktu Plán projektu, který je klíčovým podkladem pro úspěšný průběh celého projektu integrace dat do Google Analytics. Úlohy této disciplíny v rámci Plánu projektu alokují kapacity jednotlivých pracovních rolí nebo stanovují způsob komunikace včetně vymezení používaných komunikačních prostředků a kanálů v průběhu realizace celého projektu. Součástí Plánu projektu je dále identifikace potenciálních rizik s návrhem na jejich ošetření a popisem způsobu vypořádání se s nimi v případě, že by reálně nastaly. Podstatnou součástí Plánu projektu je dále plánování činnosti, které lze chápat jako jakýsi časový harmonogram s přiřazenými rolemi a očekávanými výstupy.

Druhou rovinou jsou úlohy, které jsou spojené s řízením a plánováním jednotlivých iterací. Mezi tyto úlohy patří kromě úloh Plánování iterace a Řízení iterace patří i Zhodnocení výsledků iterace. Tyto tři úlohy jsou součástí každé z iterací v průběhu celého životního cyklu metodiky. Od druhé fáze životního cyklu metodiky je do této roviny nově zařazena úloha Řízení odchylek vůči plánu. V poslední fázi životního cyklu je vyhodnocena realizace celého projektu úlohou Zhodnocení realizace projektu jako celku.

Tabulka 5.2: Úlohy disciplíny Řízení projektu

Úloha	Charakteristika
Plánování iterace	Identifikace úloh určených k realizaci v dané iteraci a jejich uspořádání do harmonogramu v souladu s Plánem projektu.
Řízení iterace	Řízení a monitoring naplánovaných úloh v rámci dané iterace za účelem jejich úspěšného dokončení v souladu s Požadavky.
Řízení odchylek vůči plánu	Plán projektu byl stanoven v první fázi životního cyklu, realita se během projektu může od plánu odchylovat a je nutné ji řídit.
Zhodnocení výsledků iterace	Úloha si klade za cíl posoudit, zda bylo dosaženo vymezených cílů a vyvodit doporučení pro zlepšení procesu v další iteraci.
Plánování činností	Rozpadne projekt na činnosti, které naplánuje v čase a popíše jejich očekávané výstupy. Vše zaznamená do Plánu projektu.
Alokace kapacit pracovních rolí	Na základě naplánovaných činností je odhadnuta pracnost pro jednotlivé role, které jsou následně k činnostem alokovány.
Identifikace a návrh řízení rizik	Identifikace možných rizik, jež mohou projekt ohrozit. Součástí je i návrh na jejich řízení a návrh řešení krizového scénáře.
Definice komunikace	Vymezení způsobu komunikace a komunikačních kanálů, které budou respektovány všemi rolemi během realizace projektu.
Kompletace projektového plánu	Revize navrženého Plánu projektu a jeho doplnění o náklady a rezervy. Následně probíhá schválení navrženého plánu.
Zhodnocení realizace projektu jako celku	V závěrečné fázi je celý projekt vyhodnocen. Jde o zhodnocení reality vůči pracovním produktům Požadavky a Plán projektu.

5.3.3 Disciplína Dokumentace a její úlohy

Disciplína Dokumentace pracuje se dvěma pracovními produkty. Jedná se o pracovní produkty Datový model a Dokumentace. Úlohy vztahující se k pracovnímu produktu Datový model se nacházejí v první fázi životního cyklu metodiky. Jejich smyslem je na jedné straně poznat aktuální stav datového modelu a stav dostupných dat na straně zdrojového systému. Na druhé straně se jedná o posouzení možností integrování dostupných dat z hlediska možností nástroje Google Analytics a ověření aktuálnosti návrhu datového modelu nástroje Google Analytics, který je k metodice i této diplomové práci přiložen.

Další úlohy této disciplíny se týkají pracovního produktu Dokumentace, respektive tento pracovní produkt vytvářejí a dále rozšiřují o nové sekce. Ve fázi Rozpracování je do pracovního produktu Dokumentace zaznamenán návrh datových toků ze zdrojového systému do alokovaných prostor datového modelu GA. V této fázi je dále navržen a zaznamenán technický návrh celé integrace dat. Dokumentace obsahující návrh datových toků a technický návrh je předána vývojářům jakožto podklad pro implementaci řešení. Vývojář pracovní produkt Dokumentace dále rozšíří o sekci popisující technickou realizaci, respektive popis samotné implementace řešení. Poslední úlohou této disciplíny je Revize dokumentace rolí Systémového integrátora.

Tabulka 5.3: Úlohy disciplíny Dokumentace

Úloha	Charakteristika
Poznání stávajících dat	Prozkoumání současných databázových struktur podnikového systému a popsání nalezených dat metainformacemi.
Identifikace dat k integraci do GA	Posouzení, zda zvolený typ dat je technicky možné do GA integrovat některou z dostupných integračních metod.
Ověření dostupnosti dat	U zvolených dat je v rámci této úlohy posouzeno, zda a do jaké míry jsou data pro integraci v podnikovém systému přístupné.
Ověření platnosti datového modelu GA	Před nalezením prostoru pro data v datovém modelu GA je potřeba ověřit aktuálnost přiloženého datového modelu.
Dokumentace datových toků	Integrovaná data tečou ze zdrojového systému do příslušného prostoru datového modelu GA. Tento tok je úlohou popsán.
Dokumentace technického návrhu	Pro implementaci řešení je žádoucí sestavit technický návrh integrace dat do GA, který je hlavním podkladem pro Vývojáře.
Předání dokumentace vývojářům	Předmětem úlohy je schválení navrženého technického návrhu a jeho následné předání vývojářům za účelem implementace.
Dokumentace technické realizace	Vývojář v průběhu implementace řešení zaznamenává proces a použité nastavení do pracovního produktu Dokumentace.
Revize dokumentace	Dokumentace je zrevidována Systémovým integrátorem, který může ostatní role požádat o upřesnění jednotlivých sekcí.

5.3.4 Disciplína Vývoj a její úlohy

Úlohy disciplíny Vývoj se zaměřují na implementaci samotného řešení integrace dat, které je již konceptuálně navržené úlohami disciplíny Požadavky a popsané úlohami disciplíny Dokumentace. V rámci úloh disciplíny Vývoj je nejprve zvolena některý z integračních metod, na základě které se technická realizace integrace dat do Google Analytics odvíjí.

Na základě zvolené integrační metody je technicky navrženo a připraveno prostřední (podniková databáze, datový model GA či datová vrstva). Toto prostředí je pro každou integrační metodu a pro každý typ dat odlišné. Odlišnostmi a specifiky se zabývala tato diplomová práce v kapitole číslo 2. Jakmile je prostředí připravené, lze řešení implementovat (včetně funkčních požadavků) a zavést definované nefunkční požadavky.

Tabulka 5.4: Úlohy disciplíny vývoj

Úloha	Charakteristika
Volba integrační metody	Dle typu dat, technických možností a požadavků integrace je zvolena integrační metoda. Specifika metod viz kapitola 2.2.
Alokace prostoru v datovém modelu GA	Dle typu dat a zvolené integrační metody je alokován prostor v datovém modelu GA. Specifika modelu viz kapitola 2.1.
Návrh databáze podnikového systému	Pouze pro integrační metodu Measurement Protocol, pokud je požadavkem přiřazovat data konkrétnímu uživateli dle ID.
Návrh rozvoje datové vrstvy	Pouze pro integrační metodu Rozšířený měřící kód. Navržení datové vrstvy pro předávání dat mezi aplikací a kódem GA.
Rozvoj databáze podnikového systému	Pouze pro Measurement Protocol. Návrh rozvoje databáze je implementován. Databáze je tímto připravena k integraci.
Rozvoj datové vrstvy a implementace GTM	Pouze pro Rozšířený měřící kód. Navržená datová vrstva je implementována a volitelně může být implementováno GTM.
Implementace řešení	Probíhá implementace integrace dat (včetně funkčních požadavků) dle integrační metody v již připraveném prostředí.
Zavedení nefunkčních požadavků	Implementace integrace dat je rozšířena o vymezené nefunkční požadavky navržené v rámci úloh disciplíny Požadavky.

5.3.5 Disciplína Testování a její úlohy

Smyslem úloh této disciplíny je ověření implementované integrace, a to podle funkčních i nefunkčních požadavků vymezených úlohami disciplíny Požadavky. V rámci navržených a provedených testů mohou být odhaleny chyby různé závažnosti, které jsou buď vráceny roli Vývojář k opravě v rámci úlohy Implementace řešení v disciplíně Vývoj, nebo mohou být rolí Systémový integrátor akceptovány. V samotném testování hraje důležitou roli Webový analytik, který má hluboké znalosti o nástroji Google Analytics, a je tak schopen dle navržených testů rozklíčovat případné nedostatky implementované integrace, nebo potvrdit integraci jako úspěšně dokončenou s ohledem na funkční i nefunkční požadavky.

Tabulka 5.5: Úlohy disciplíny Testování

Úloha	Charakteristika
Návrh testů	Na základě funkčních a nefunkčních požadavků jsou navrženy testovací techniky, kompetence rolí v procesu a harmonogram.
Příprava testů	Během úlohy je vytvořen pracovní produkt Seznam testovacích nápadů a prostředí pro vykonání testů.
Provedení funkčních testů	Podle připraveného návrhu testů a testovacího prostředí jsou realizovány testovací případy zaměřené na funkční požadavky.
Provedení nefunkčních testů	Obdobným způsobem jako funkční testy jsou zrealizovány i testy zaměřené na nefunkční požadavky.

5.3.6 Disciplína Aplikace a její úlohy

Úlohy disciplíny Aplikace se zabývají reálnými případy užití, pro které je vyčleněn pracovní produkt Případy užití. Do úvodní fáze životního cyklu metodiky, kdy je vymezen rozsah integrace, je začleněna úloha Modelování případů užití, která si klade za cíl namodelovat konkrétní případy užití, jež díky konkrétní realizaci integrace dat do Google Analytics bude možné později aplikovat. Zbylé dvě úlohy této disciplíny jsou zařazeny do poslední fáze životního cyklu metodiky. V rámci těchto dvou úloh jsou nejprve další zainteresované role s proběhlou integrací seznámeny ve smyslu dopadů integrace na jejich pracovní úkony a následně jsou úvodem namodelované případy užití aplikovány přímo do praxe.

Tabulka 5.6: Úlohy disciplíny Aplikace

Úloha	Charakteristika
Modelování případů užití	Navržení možných případů vyplývajících z konkrétní integrace. Pomáhá odhadnout budoucí přínosy integrace.
Edukace dalších zainteresovaných rolí	Seznámení dalších zainteresovaných rolí s dopady realizované integrace dat do Google Analytics na jejich pracovní úkony.
Aplikování navržených případů užití	Aplikace namodelovaných případů užití v praxi. Typicky se jedná o datové analýzy nebo online marketingové cílení.

5.3.7 Disciplína Řízení změn a její úlohy

Smyslem úloh disciplíny Řízení změn je komunikace s rolemi, které s integrovanými daty z Google Analytics pracují v praxi. Od těchto rolí je sesbírána zpětná vazba, která je následně vyhodnocena. Smyslem této zpětné vazby je další odladění projektu formou návrhu na změnové řízení. Ze zkušenosti je totiž známo, že když je integrace implementována a aplikována do praxe, dostanou zainteresované role další nápady, jak ji zužitkovat pro další možné případy užití, které během modelování případů užití nevyšly najevo. Avšak některé z těchto nápadů vyžadují úpravy technických parametrů integrace, pro jejichž provedení slouží jako podklad návrh na změnové řízení, který musí schválit Systémový integrátor.

Tabulka 5.7: Úlohy disciplíny Řízení změn

Úloha	Charakteristika
Sběr zpětné vazby	Komunikace s rolemi, které s integrovanými daty pracují v praxi. Z této komunikace je sesbíraná věcná zpětná vazba.
Vyhodnocení zpětné vazby	Sesbíraná zpětná vazba je vyhodnocena a může se stát podkladem pro návrh na změnové řízení.
Návrh na změnové řízení	V rámci úlohy je sestaven pracovní produkt Seznam požadavků na změnu, který je podkladem pro další změny.

5.4 Pracovní produkty metodiky

Pracovní produkty metodiky lze chápat jako výstupy, které jsou v průběhu životního cyklu metodiky vytvářeny a modifikovány jednotlivými úlohami (60). Pracovní produkty mívají obvykle podobu dokumentu a slouží jako uchovatel a nositel informací pro komunikaci mezi zúčastněnými rolemi. Pracovní produkty jsou pro přehlednost v rámci této metodiky, obdobně jako úlohy, rozčleněny do disciplín, tedy množin zaměřených na určitou oblast zájmu (60). Metodika integrace dat do Google Analytics vymezuje celkem 15 pracovních produktů rozdělených do 7 disciplín. Každý z pracovních produktů je blíže představen a charakterizován v průběhu následujících kapitol.

5.4.1 Pracovní produkty disciplíny Požadavky

Boehm uvádí, že definice požadavků v úvodní fázi životního cyklu má z hlediska jakéhokoliv softwarového vývoje přímý dopad na snížení nákladů souvisejících s testováním a následnými opravami, spolehlivost, uživatelskou přívětivost a náročnost na údržbu (66). Ostatně požadavky jako takové jsou zároveň pevně zakotveny v úvodní fázi životního cyklus vývoje software dle standardu ISO/IEC/IEEE 12207 (67).

V souvislosti s velkým důrazem na jasné vymezení všech potřebných požadavků v úvodní fázi životního cyklu Metodiky integrace dat do Google Analytics byly navrženy 3 pracovní produkty, které slouží jako materiál sjednocující pohled zúčastněných rolí na projektu tak, aby všechny zúčastněné role měly konkrétní představu o výsledném výstupu. Těmito pracovními produkty jsou po vzoru MMSP Vize, Požadavky a Slovník pojmů, které jsou svým obsahem přizpůsobeny problematice integrace dat do Google Analytics.

Pracovní produkt Vize

Jedná se o jednoduchý, avšak výstižný textový dokument požadovaného cílového stavu. S pracovním produktem musí být seznámena každá role participující na projektu a každá role v průběhu účasti na projektu by měla konat v jeho souladu. Z hlediska řešené metodiky je obsahem pracovního produktu výstižná odpověď na tyto základní otázky: Jaký problém řešíme? Co bude naším cílem? Jak tento cíl naplníme? Pracovní produkt je sestaven rolí Systémového integrátora v rámci úlohy Sestavení záměru integrace.

Pracovní produkt Požadavky

Pracovní produkt Požadavky je chápán jako strukturovaný seznam funkčních i nefunkčních požadavků, jež by implementovaná integrace dat do Google Analytics měla splňovat. Tyto požadavky lze chápat jako popis očekávaných vlastností výstupu s uvedením akceptačního kritéria, tedy jakého stavu (či hodnoty) a za jakých podmínek musí daná vlastnost nabývat, aby byla akceptovatelná. Součástí každého z uvedených požadavků je i uvedení priority.

Za pracovní produkt Požadavky je zodpovědná role Systémového integrátora. Rozšiřovat jej však mohou i další role. O technické požadavky může pracovní produkt rozšířit role Vývojář a Webový analytik. O legislativní požadavky pak může pracovní produkt rozšířit Právník. Soulad dokončené integrace vůči pracovnímu produktu Požadavky je pak posuzován oproti výsledkům testování, případně vůči zpětné vazbě uživatelů.

Pracovní produkt Slovník pojmů

Od počátku projektu integrace dat do Google Analytics je budován pracovní produkt Slovník pojmů, který lze chápat jako abecední seznam definované terminologie, jež je užívána v průběhu projektu. Záměrem pracovního produktu je předejít možným komunikačním problémům ve smyslu nepřesného užití a výkladu termínů. Nepřesné užití či výklad konkrétního termínu může zcela pozměnit význam sdělení. Výsledným dopadem by pak mohlo být, že daná úloha neproběhne se správnými vstupními parametry. Do Slovníku pojmů přispívá primárně Systémový integrátor, který je za něj zároveň zodpovědný. Terminologii však mohou rozšířit či upřesnit i specializované role jako Webový analytik, Vývojář či Právník. Finální podoba pracovního produktu je schválena rolí Systémového integrátora v rámci úlohy Definice komunikace.

5.4.2 Pracovní produkty disciplíny Řízení projektu

Na integraci dat do Google Analytics lze hledět jako na projekt, který Rosenau popisuje jako *„organizované úsilí k dosažení určitého cíle“* (68). Metodika integrace dat do GA se na pracovní produkty Řízení projektu dívá ze dvou perspektiv. První je plánování projektu jako celku, kterou reprezentuje pracovní produkt Plán projektu. Druhou perspektivou je plánování jednotlivých iterací, které reprezentuje pracovní produkt Plán iterace. Obě perspektivy se potkávají v pracovním produktu Seznam pracovních položek, který je nejprve navržen pro projekt jako celek a následně používán pro plánování jednotlivých iterací.

Pracovní produkt Plán iterace

Iterace v rámci popisované metodiky probíhají v každé ze čtyř fází životního cyklu. Každá z iterací musí být naplánována a dle vytvořeného plánu následně řízena. Pracovním produktem, který iteraci detailně popisuje a uchovává harmonogram obsahující kladené nároky na výstupy dílčích úloh, je Plán iterace. Plán iterace je zřízen pro každou z iterací zvlášť, neboť každá z iterací má svůj specifický průběhu. Plán iterace obecně vychází z pracovního produktu Plán projektu. Výjimkou je však iterace v rámci první fáze životního cyklu, kdy na počátku této fáze pracovní produkt Plán projektu ještě neexistuje. Za pracovní produkt Plán iterace je zodpovědná role Systémového integrátora.

Pracovní produkt Plán projektu

Celá integrace dat do Google Analytics je pojata formou projektu. Aby byl projekt úspěšný, musí být určitým způsobem řízen. Základním podkladem pro řízení projektu jako celku je pracovní produkt Plán projektu, který má sám o sobe významnou oporu v této metodice, jelikož může využít navržených posloupností jednotlivých úloh, zúčastněných rolí či jednotlivých výstupů reprezentovaných pracovními produktu.

Z Plánu projektu v kontextu této metodiky lze vyčíst potřebné informace o průběhu projektu včetně jednotlivých pracovních položek, stanovených cílů a indikátorů, jež odhalují míru splnění daného cíle. Plán projektu současně slouží jako podklad pro monitoring projektu, kdy lze současný stav projektu vůči plánu porovnávat a vypočítat rozdíly očekávaného a současného stavu v čase, nákladech, popřípadě rozsahu. Za tento pracovní produkt je zodpovědná role Systémového integrátora.

5.4.3 Pracovní produkty disciplíny Dokumentace

Sommerville uvádí, že softwarová dokumentace je komunikační médium mezi vývojářem a uživatelem. Skrze dokumentaci je uživateli sděleno, jak daný produkt používat. Současně Sommerville popisuje, že dokumentace je stěžejním zdrojem informací o daném produktu, což hraje zásadní roli při návrhu jeho rozvoje a při rozvoji samotném (69).

Z obou uvedených důvodů jsou v projektu integrace dat do GA evidovány záznamy o datových modelech, technickém návrhu, datových tocích, či technických parametrech. Tyto záznamy jsou v rámci disciplíny Dokumentace rozděleny do dvou pracovních produktů, kterými jsou Datový model a Dokumentace. Přestože pracovní produkt Datový model by teoreticky mohl být součástí Dokumentace, vzhledem ke svému klíčovému významu byl vyčleněn záměrně. Význam těchto produktů je blíže popsán v následujících odstavcích.

Pracovní produkt Datový model

Obsahem pracovního produktu je popis dvou datových modelů. Prvním z nich je datový model podnikového (tedy zdrojového) systému. Druhým je pak aktuální datový model Google Analytics. Poznání obou datových modelů hraje klíčovou roli pro návrh pomyslného datového toku, tedy určení odkud z datového modelu zdrojového systému kam do datového modelu Google Analytics potečou data. Tento pracovní produkt zároveň usnadňuje zúčastněným rolím orientaci v obou integrovaných systémech pro účely vývoje, testování či budoucí modifikace. Na pracovní produkt Datový model se dále odkazuje pracovní produkt Dokumentace, který na datových modelech názorně popisuje vytvořené datové toky, technický návrh a užité technické parametry integrace.

Popis datového modelu podnikového systému je obvykle ve své aktuální verzi součástí dokumentace podnikového software. Zatímco popis datového Modelu Google Analytics je součástí této diplomové práce, dostupný na obrázku 2.1. Protože datový model GA může v průběhu času dostát různých změn, byla do první fáze životního cyklu metodiky začleněna úloha Ověření platnosti datového modelu GA, kdy role Webového analytika ověří, zda přiložený datový model je stále platný a pokud ne, příslušným způsobem jej modifikuje.

Pracovní produkt Dokumentace

Pracovní produkt je nositelem řady informací, které v průběhu realizace projektu vyplývají najevo. Prostřednictvím Dokumentace jsou zároveň sdíleny technické náležitosti aktuálního stavu realizace mezi jednotlivými rolemi. Dokumentace je textový dokument rozdělený do několika sekcí. Těmito sekcemi jsou charakteristika datových toků, technický návrh integrace dat a dokumentace technické realizace.

Dokument je přístupný všem zúčastněným rolí ve své aktuální podobě. Zodpovědnost za pracovní produkt však nese role Systémového integrátora, který zároveň pracovní produkt reviduje a nejasnosti upřesňuje s konkrétními rolemi. Do dokumentace v průběhu životního cyklu přispívají role Datový steward, Webový analytik a Vývojář.

5.4.4 Pracovní produkty disciplíny Vývoj

Pokud uvažujeme tradiční rigorózní přístup k vývoji, pak Buchalcevová uvádí, že lze vývoj rozpadnout na několik dílčích prvků, které lze podle tradičního přístupu jasně definovat. Jedná se prvky: problém, řešení problému, vývojář (nositel řešení) a prostředí (34).

V kontextu navržené metodiky rozumí sestavení technického návrhu realizace a jeho implementace do vymezeného prostředí. Vymezené prostředí je definováno produkty disciplín Požadavky a Dokumentace. Disciplína Vývoj na základech vymezeného prostředí staví pracovní produkt Návrh, který je podkladem pro pracovní produkt Implementace.

Pracovní produkt Návrh

Smyslem pracovního produktu je technický popis řešení, který vychází z požadavků vymezených pracovním produktem Požadavky v rámci první fáze životního cyklu. Pracovní produkt Návrh je sestaven ve druhé fázi životního cyklu metodiky a obsahuje rozhodnutí o volbě integrační metody, popis alokovaného prostoru v datových modelech a případná specifika týkajících se konkrétních integračních metod, kterým je například v případě Measurement Protocolu návrh rozvoje databáze podnikového systému nebo v případě Rozšířeného měřícího kódu přes GTM návrh rozvoje datové vrstvy. Pracovní produkt Návrh je podkladem pro roli Vývojáře, která na základě něj provádí implementaci. Za pracovní produkt zodpovídá role Webového analytika.

Pracovní produkt Implementace

Pracovní produkt Implementace je samotná funkcionalita integrace, kterou vyvinula primárně role Vývojáře na základě pracovního produktu Návrh. V případě integrační metody Measurement Protocol zde může participovat i role Databázového specialisty. Pro potřeby úprav na straně Google Analytics může do pracovního produktu přispívat i role Webového analytika. Tento pracovní produkt tak lze chápat jako zdrojový kód zasazený do vymezeného prostředí podnikového systému a současně soupis užitých parametrů na straně podnikového systému i na straně služby Google Analytics. Vysvětlení ke zdrojovému kódu a užitým parametrům doplňuje role Vývojáře, popřípadě Webového analytika do pracovního produktu Dokumentace. Za pracovní produkt Implementace zodpovídá role Vývojáře.

5.4.5 Pracovní produkty disciplíny Testování

Bureš vnímá testování jako prostředek k dosažení určité úrovně kvality, která do určité míry předchází defektům, případně tyto defekty včas odhaluje (70). Pokud by se kritický defekt objevil na produkční verzi, může to mít z hlediska funkčnosti integrace i dodatečných nákladů fatální dopady. Testování tak můžeme chápat jako nástroj, který pomáhá snižovat případné neočekávané náklady a současně pomáhá dodat výstup v očekávané kvalitě.

Metodika integrace dat do Google Analytics klade na testování vysoký důraz a to z výše uvedených důvodů. V rámci disciplíny Testování pracuje metodika se čtyřmi vzájemně provázanými pracovními produkty, kterými jsou Seznam testovacích nápadů, Plán testů, Testovací případ a Záznam výsledků testů. V Metodice integrace dat do GA nevystupuje role Testera, která je zde považována za nadbytečnou, jelikož se nejedná o klasický software, nýbrž pouze o datové propojení existujících komponent. Testování tak realizují role Webového analytika a Vývojáře, které by neměly být reprezentovány stejnými osobami, které prováděli implementaci řešení.

Pracovní produkt Seznam testovacích nápadů

Jedná se o strukturovaný seznamu konkrétních oblastí implementovaného řešení, které potenciálně mohou vykazovat určitou chybovost s návrhem na jejich otestování. Tyto oblasti je žádoucí nejprve identifikovat. Identifikace vychází z vymezených požadavků, ze sestavené dokumentace a z rozhovorů se zainteresovanými rolemi. K takto identifikovaným oblastem jsou vygenerovány možné způsoby jejich otestování, tedy samotné testovací nápady. Jelikož každá integrace dat do GA je svým způsobem specifická, pak i testovací nápady mohou být různého charakteru v závislosti na užité integrační metodě, způsobu implementace či dostupného prostředí pro realizaci testů. Za pracovní produkt Seznam testovacích nápadů je zodpovědná role Systémového integrátora.

Pracovní produkt Plán testů

Plán testů je souvislý text, který popisuje samotný proces testování vycházející ze Seznamu testovacích nápadů. Správně zpracovaný Plán testů je předpokladem úspěšného procesu testování implementovaného řešení, jelikož odpovídá na otázky co, jak, kdy a kým bude testováno. Součástí plánu je tedy vymezení testovaných oblastí s popisem způsobu testování, časový harmonogram vymezující celý proces testování a přiřazení testovacích aktivit příslušným osobám. Za pracovní produkt odpovídá Systémový integrátor.

Pracovní produkt Testovací případ

Pracovní produkt Testovací případ je kolekcí stručných záznamů průběhu o konkrétních testovacích aktivit. Záznamy evidují informaci o tom, jakým způsobem bylo připraveno prostředí pro vykonání testovací aktivity, jakými specifiky bylo testovací prostředí charakteristické (např. prohlížeč, typ zařízení, deaktivovaný JavaScript a jiné) a jak testovací aktivita probíhala. Tyto záznamy slouží především k tomu, aby bylo možné rozlišit a otestovat co nejširší spektrum potenciálních scénářů. Za každý záznam v tomto pracovním produktu zodpovídá osoba, která danou testovací aktivitu vykonala.

Pracovní produkt Záznam výsledků testů

Z každého záznamu v pracovním produktu Testovací případ vyplývá určitý závěr, respektive výsledek testu. Výsledek testu je souhrnem nálezů, které v průběhu daného testovacího případu byly zjištěny. Pracovní produkt Záznam výsledků testů shromažďuje veškeré souhrny nálezů každého testovacího případu. Záznam výsledků testů tak lze chápat jako klíčový zdroj informací pro případné opravy nebo jako důkaz, že výsledná integrace dat nabývá vymezených funkčních i nefunkčních požadavků, jež byly testovány. Za každý souhrn nálezů zodpovídá osoba, která danou testovací aktivitu vykonala.

5.4.6 Pracovní produkty disciplíny Aplikace

Hohpe uvádí, že existuje celá řada důvodů, proč integraci na datové úrovni realizovat. Jedním z nich je podpora business procesů sdílením dat napříč aplikacemi za účelem větší efektivnosti, spolehlivosti a bezpečnosti (71). Tento důvod je současně hlavní motivací pro integraci dat z podnikových systémů do Google Analytics, kde jsou integrovaná data využita pro pokročilé datové analýzy, optimalizaci webu či online marketingové aktivity.

Pracovní produkt Případy užití

Pracovní produkt Případy užití popisuje konkrétní scénáře, jakým způsobem budou integrovaná data zužitkována z praxi. Návrhy případů užití dané integrace dat jsou navrženy již v první fázi životního cyklu metodiky. Návrhy případů užití dávají jasnější a konkrétnější představu o výsledném výstupu a o projektu integrace jako celku. V poslední fázi jsou pak navržené případy užití aplikovány do praxe. Za návrh každého případu užití je zodpovědná osoba, která bude jejím hlavním aktérem v praxi, typicky se jedná o role Webového analytika nebo Online marketingového specialisty.

5.4.7 Pracovní produkty disciplíny Řízení změn

Procesy řízení změn jsou dle ITIL důležitou součástí správy informačních systémů. Ostatně ITIL definuje pro řízení změn celý životní cyklus od identifikace změn, přes jejich plánování, posouzení, implementaci, zhodnocení až po uzavření změnového řízení (72). Metodika integrace dat do Google Analytics se na proces řízení změn dívá zjednodušeně, kdy je sbírána zpětná vazba od zainteresovaných stran, která je následně posouzena a případně předána k implementaci formou změnového řízení.

Pracovní produkt Seznam požadavků na změnu

Každého případu užití se týkají různí aktéři, kteří integrovaná data určitým způsobem využívají. Od těchto aktérů je do Seznamu požadavků na změnu sbírána zpětná vazba, která je následně posouzena a vyhodnocena. V případě, že byla zpětná vazba posouzena jako předmětná, pak může být podkladem pro změnové řízení, v rámci kterého bude změna implementována. Za pracovní produkt Seznam požadavků na změnu je zodpovědná role Systémového integrátora, která zpětnou vazbu eviduje, vyhodnocuje a připravuje podklady pro případná změnová řízení. V případě potřeby si role Systémového integrátora může pomoci pro vyhodnocení zpětné vazby specialistou, například Webovým analytikem.

6 Ukázka užití metodiky

Ukázka aplikace navržené metodiky je provedena postupným průchodem úloh jednotlivých fází životního cyklu. Ke každé z úloh je doplněn stručný komentář nesoucí informaci o tom, jak byla daná úloha na daném případu užití řešena a jakým způsobem daná úloha ovlivnila pracovní produkty metodiky. Před aplikací navržené metodiky je čtenář seznámen s konkrétní situací, na které je integrace dat do Google Analytics realizována.

6.1 Popis případu užití

Pro ověření užití metodického postupu je uvažován fiktivní maloobchodní řetězec prodejen potravin označený obchodním názvem RychloMarket, který disponuje celkem deseti prodejnami napříč různými městy v České republice. Zákazníci RychloMarketu mohou získat věrnostní kartu, která jim umožňuje sbírat věrnostní body. Tyto body lze uplatnit formou slevy na další nákup. O věrnostní kartu může zákazník zažádat skrze formulář na webových stránkách obchodního řetězce. V průběhu žádosti o věrnostní kartu má zákazník možnost zřídit si uživatelský účet na webových stránkách, prostřednictvím kterého bude informován o nových slevových akcích prodejny. V uživatelském účtu je zároveň evidována informace o zákazníkově věrnostní kartě a informace s aktuálním stavem věrnostních bodů.

Marketingové oddělení RychloMarketu se rozhodlo spustit online marketingovou kampaň, která si klade za cíl informovat zákazníky o výhodách věrnostní karty. Za tímto účelem byly sestaveny 3 uživatelské segmenty. Každému ze segmentů je určené jiné sdělení. Tyto segmenty a jim komunikovaná sdělení jsou popsány v tabulce 6.1.

Tabulka 6.1: Uživatelské segmenty online marketingové kampaně

Segment	Popis segmentu	Komunikované sdělení
A	Nemají věrnostní kartu	RychloMarket nabízí věrnostní kartu
B	Mají věrnostní kartu a za posledních 30 dní neudělali žádný nákup	Výhody plynoucí z aktivního používání věrnostní karty
C	Mají věrnostní kartu a za posledních 30 dní udělali alespoň 1 nákup	Žádné sdělení o věrnostní kartě

RychloMarket používá na svých webových stránkách nástroj Google Analytics, ve kterém by rád segmenty definoval a naimportoval do Google Ads, kde bude daná kampaň spuštěna. K tomu, aby mohly být segmenty v Google Analytics definovány, je nutné do GA předávat informace o tom, zda daný uživatel má věrnostní kartu a kdy naposledy nakoupil. Procesem předávání informací (též integrace dat) provede Metodika integrace dat do Google Analytics navržená touto diplomovou prací.

6.2 Aplikace metodiky

Popsaný případ užití bude dle Metodiky integrace dat do Google Analytics realizován ve čtyřech fázích: Zahájení, Rozpracování, Konstrukce a Zavedení. Každá z fází životního cyklu metodiky disponuje řadou úloh, které různým způsobem modifikují pracovní produkty metodiky. Aplikace metodiky na popsaném případu užití proběhne formou stručných komentářů ke každé z úloh. Obsahem komentáře je vysvětlení, jak je daná úloha vykonána na popsaném případu užití a jaké konkrétní dopady má na pracovní produkty metodiky. Samotný obsah pracovních produktů součástí aplikace metodiky na tomto ukázkovém případu užití není. Součástí navržené metodiky jsou dále role. Doporučené role jsou popsány v kapitole 5.2. Avšak nutno podotknout, že v praxi menších projektů může jedna osoba zastupovat více rolí, proto v aplikaci metodiky na popsaném případu užití bude na role poukázáno v rámci zmíněného komentáře jednotlivých úloh. V opačném případě platí, že role vykonává takové úlohy, které jsou popsány v kapitole 5.2.

6.2.1 Fáze Zahájení

Předmětem úloh fáze Zahájení na uvedeném případu užití je zhodnocení, zda je daný projekt technicky realizovatelný. V kladném případě bude výstupem úvodní fáze schválený projektový plán. Pokud by se zjistilo, že projekt technicky realizovat nelze, pak by bylo nutné upravit zadání, nebo projekt ukončit.

Úloha Plánování iterace – Na základě Metodiky integrace dat do GA jsou zvoleny všechny úlohy první fáze, které jsou uspořádány do časového harmonogramu. Úlohám jsou přiřazeny odpovědností konkrétních osob a očekávané výstupy. Výsledným výstupem úlohy je pracovní produkt Plán iterace, který je současně řídícím dokumentem dané iterace.

Úloha Poznání stávajících dat – Datový steward prozkoumal databázové struktury podniku a zjistil, že podnik disponuje jednou centrální databází, která je provozována na databázovém serveru Oracle. Tato centrální databáze je současně zdrojem dat většiny podnikových systémů včetně CRM. Vedle toho provozuje podnik webovou aplikaci napojenou na MySQL databázi, která je nastavena tak, že dokáže komunikovat a vyměňovat data s centrální databází Oracle. Dále podnik disponuje pokladním systémem obsluhující kasy na jednotlivých prodejnách. Pokladní systém odesílá data do centrální Oracle databáze skrze lokální dočasné úložiště dat.

Úloha Identifikace dat k integraci do GA – Byly identifikovány dva typy dat, které je nutno integrovat do Google Analytics. Prvním je informace, zda uživatel věrnostní kartu registroval. Vzhledem k tomu, že se věrnostní karta registruje přes webovou aplikace, pak je informace uložena v instanci MySQL databáze, kde lze využít integrační metodu Rozšířený měřící kód. Tato informace je díky komunikaci mezi MySQL databází a Oracle databází současně propisována do Oracle databáze.

Druhým typem dat je informace, zda zákazník svou věrnostní kartu použil na prodejně skrze pokladní systém. Pokladní systém tuto informaci předává skrze dočasné úložiště dat do centrální Oracle databáze, kde lze pracovat s integrační metodou Measurement Protocol. Veškeré zjištěné poznatky jsou zaznamenány do pracovního produktu Datový model.

Úloha Ověření dostupnosti dat – Datový steward prověřil, že na identifikovaná data se nevztahují žádné interní restrikce, které by neumožňovaly jejich integraci do GA. Tato data jsou zároveň v databázích uložena tak, že k nim lze relativně snadno přistoupit skrze zdrojový kód. Identifikovaná data tak byla shledána jako dostupná k integraci do GA.

Úloha Ověření platnosti datového modelu GA – Datový model Google Analytics, jež je součástí metodiky v rámci pracovního produktu Datový model, byl na základě dostupné specifikace shledán jako aktuální a platný bez nutnosti jakékoliv další modifikace.

Úloha Sestavení záměru integrace – Byl sestaven cíl projektu, kterým je integrace dat do Google Analytics. Integrovanými daty jsou informace, zda zákazník registroval věrnostní kartu a pokud ano, kdy ji naposledy využil k nákupu v prodejně. Účelem je vytvoření segmentů pro potřeby online marketingové kampaně. Cíl a účel byly zaznamenány do pracovního produktu Vize. Dále byly vytvořeny produkty Požadavky a Slovník pojmu.

Úloha Zhodnocení technické proveditelnosti – Vývojář na základě dostupných podkladů posoudil, že je možné projekt technicky zrealizovat. Avšak po konzultaci s Webovým analytikem zjistil, že je potřeba rozšířit centrální Oracle databázi i MySQL databázi o nový atribut, kterým je uživatelův identifikátor vygenerovaný službou Google Analytics. Informace byly zaznamenány do pracovního produktu Požadavky a zároveň byl rozšířen pracovní produkt Slovník požadavků za účelem jednoznačné komunikace.

Úloha Zhodnocení legislativních aspektů – Právník uvedl, že se musíme vyvarovat zasílání identifikátorů věrnostních karet do Google Analytics, jelikož se jedná o údaj umožňující identifikaci konkrétních osob, což je v rozporu s podmínkami Google Analytics. Tento identifikátor tedy předáván nebude. Do GA budou integrována pouze informace zda uživatel registroval věrnostní kartu a kdy ji naposledy použil na prodejně, což je podle Právníka z hlediska legislativních aspektů v pořádku.

Úloha Modelování případů užití – Online marketingový specialista identifikoval hlavní případ užití, a sice vytvoření online marketingových segmentů dle nových dat, které budou zužitkovány v inzertním systému Google Ads. Webový analytik poté zjistil, že tato data lze zohlednit i při datových analýzách a reportingu. Zjištěné poznatky byly zaznamenány do pracovního produktu Případy užití.

Úloha Vymezení rozsahu integrace – Na základě předchozích úloh zrevidoval Systémový integrátor pracovní produkty Požadavky, Slovník pojmů a Vize. Zároveň bylo jednoznačně vymezeno, že budou do GA integrována následující data:

1. Věrnostní karta (ano/ne), zdroj: MySQL databáze.
2. Datum nákupu s věrnostní kartou, zdroj: Oracle databáze.

U data nákupu bude přiřazen a odesílán i GA identifikátor z cookie _ga (viz kapitola 2.2.3).

Úloha Plán činností – Byl sestaven pracovní produkt Plán projektu, který přímo vychází z jednotlivých úloh této metodiky. Dále bylo součtem jednotlivých úloh odhadnuto, že časová náročnost projektu bude celkem 120 hodin a náklady dosáhnou 100 000 Kč. Do Plánu projektu byly dále zakomponovány požadavky na kvalitu projektu (respektive metriky a indikátory funkcí integrace), které vycházejí z pracovního produktu Požadavky.

Úloha Alokace kapacit pracovních rolí – K jednotlivým úlohám byly přiřazeny konkrétní osoby a dle jejich časových kapacit byl upřesněn časový odhad projektu. Projekt bude realizován po dobu 30 pracovních dnů, v rámci kterých spotřebují jednotlivé zúčastněné role počet hodin dle tabulky 6.2. V součtu odhadovaných 120 hodin. Celková vydaná částka (hodinová sazba x časová alokace) odpovídá odhadovaným 100 000 Kč.

Tabulka 6.2: Časová alokace jednotlivých rolí

Role	Časová alokace	Hodinová sazba
Systémový integrátor	35 hodin	800 Kč
Datový steward	12 hodin	700 Kč
Webový analytik	20 hodin	750 Kč
Právník	3 hodiny	1800 Kč
Online marketingový specialista	12 hodin	600 Kč
Databázový specialista	10 hodin	800 Kč
Vývojář	28 hodin	1000 Kč

Úloha Identifikace a návrh řízení rizik – Byla identifikována následující dvě rizika: nedostatečná kvalita dat v podnikových databázích a zásah do nastavení GA nebo GTM dalšími nezainteresovanými uživateli. Pokud se zjistí, že data jsou v nedostatečné kvalitě, pak budou aplikována pravidla, která odešlou jen ta data, kde je záruka dostatečné kvality. U ostatních dat bude hledáno řešení, jak kvalitu vylepšit. Riziko týkající se zásahu dalšími nezainteresovanými uživateli bude ošetřeno tak, že všichni uživatelé s přístupem do GA nebo GTM budou o projektu informováni. Rizika byla zaznamenána do Plánu projektu.

Úloha Definice komunikace – Za účelem komunikace mezi projektovým týmem byla zvolena webová aplikace Basecamp, kam každá zúčastněná osoba obdržela své přístupy.

Úloha Kompletace projektového plánu – Systémový integrátor zrevidoval pracovní produkt Plán projektu a doplnil jej o rezervy, které činí 20% odhadnutého času a rozpočtu. Následně jej nechal schválit zadavatelem projektu, který projekt schválil.

Úloha Řízení iterace – Úlohy této fáze probíhaly podle Plánu iterace, tedy bez výrazných odchylek, které by byly nad rámec povolené tolerance, která činí 10%. Průběžné výstupy jednotlivých úloh byly dodávány včas a v odpovídající kvalitě.

Úloha Zhodnocení výsledků iterace – První fáze byla zhodnocena jako úspěšně dokončená, protože bylo dosaženo stanoveného milníku, kterým je schválený projektový plán zadavatelem projektu. V rámci úlohy bylo současně vypočteno, že se během první fáze spotřebovalo celkem 21 hodin a vynaložila částka 15 000 Kč. Systémový integrátor zhodnotil, že je možné přejít do druhé fáze projekt podle Metodiky integrace dat do GA.

6.2.2 Fáze Rozpracování

Předchozí fáze zhodnotila popsaný případ užití jako realizovatelný a dále navrhla plán realizace celého projektu integrace dat do GA. Fáze Rozpracování pak sestavuje návrh technických náležitostí a zaznamenává je do příslušných částí pracovního produktu Dokumentace. Dokumentace je následně podkladem pro samotnou implementaci.

Úloha Plánování iterace – V souladu s pracovním produktem Plán projektu byla vytvořena nová instance pracovního produktu Plán iterace, která vymezila úlohy fáze Rozpracování na základě Metodiky integrace dat do GA. Tato instance Plánu iterace je hlavním řídícím dokumentem pro vykonání jednotlivých úloh této fáze.

Úloha Volba integrační metody – Vzhledem ke specifikům integrovaných dat bylo posouzeno, že budou zvoleny následující integrační metody:

1. Věrnostní karta (ano/ne), zdroj: MySQL databáze – Rozšířený měřící kód.
2. Datum nákupu s věrnostní kartou, zdroj: Oracle databáze – Measurement Protocol.

Úloha Alokace prostoru v datovém modelu GA – Pro integraci dat tohoto případu užití bylo posouzeno, že je nutné definovat nové vlastní dimenze v Google Analytics:

1. Věrnostní karta (ano/ne) – Název dimenze: *Věrnostní karta*, rozsah dimenze: uživatel. Rozsah uživatel byl zvolen, protože informace nenáleží návštěvě nebo odeslanému požadavku na server, ale uživateli samotnému.
2. Datum nákupu s věrnostní kartou – Název dimenze: *Datum nákupu s kartou*, rozsah dimenze: požadavek na server. Tento rozsah byl zvolen, protože se jedná o jednorázový úkon, což odpovídá dimenzi požadavek na server.

Úloha Návrh rozvoje databáze podnikového systému – Vzhledem k tomu, že skrze integrační metodu Measurement Protocol bude předáván identifikátor uživatele z cookie _ga, je potřeba tento identifikátor uložit do podnikové databáze. V databázi bude uchována vždy pouze aktuální hodnota identifikátoru, která bude v rámci požadavku odesílána skrze Measurement Protocol. Tím dojde ke správnému připárování dat ke konkrétnímu uživateli v GA. Databázový specialista proto sestavil návrh, jak rozšířit stávající podnikovou databázi o nový atribut, kam bude identifikátor ukládán. Návrh rozvoje databáze podnikového systému byl Databázovým specialistou zaznamenán do pracovního produktu Návrh.

Úloha Návrh rozvoje datové vrstvy – Informace o tom, zda uživatel vlastní věrnostní kartu je známa již ve webové aplikaci, jelikož o ni lze žádat pouze přes webový formulář. V případě, že by uživatel promazal mezipaměť prohlížeče nebo použil nové zařízení, pak webová aplikace zjistí, že je vlastníkem věrnostní karty po přihlášení do uživatelského účtu. Webová aplikace bude tato data uchovávat ve své MySQL databázi a předávat skrze Rozšířený měřící kód, který bude implementovaný v GTM. Jako komunikační rozhraní pro předávání dat mezi webovou aplikací a GTM poslouží datová vrstva. Webový analytik v rámci této úlohy sestavil implementační manuál pro příslušnou implementaci datové vrstvy. Tento implementační manuál je součástí pracovního produktu Návrh.

Úloha Vymezení nefunkčních požadavků – Kromě funkčních požadavků vymezil Systémový integrátor i požadavky nefunkční. Vymezené nefunkční požadavky se týkají integrační metody Measurement Protocol. Jedná se o odeslání dat do GA nejpozději do 30 vteřin od jejich zjištění, zaznamenávání stavových hlášení odeslaných HTTP požadavků a v případě chybového stavového hlášení odesílat data každých 5 minut dokud nebudou úspěšně odeslána. Tyto požadavky byly zaznamenány do pracovního produktu Požadavky.

Úloha Návrh testů – V rámci této úlohy bylo rámcově vymezeno, jaké role se budou účastnit testování, jak budou vymezeny kompetence rolí v procesu testování a dále byl upřesněn odhad pracnosti testů. Testování se budou účastnit role Vývojáře a Webového analytika. K účasti role Vývojáře byla stanovena podmínka, že se nesmí jednat o tutéž osobu, která implementuje řešení. Kompetence testů jsou vymezeny ze dvou různých úhlů. Z úhlu zdrojového kódu, kde bude role Vývojáře testovat za pomoci debuggeru a zároveň bude kontrolovat, zda lze zdrojový kód zoptimalizovat s ohledem na výpočetní složitost. A dále z úhlu Google Analytics, kde bude role Webového analytika vyhodnocovat zpracovaná data a hledat v nich případné anomálie indikující nesrovnalosti. V průběhu úlohy byl vytvořen pracovní produkt Plán testů, kam Systémový integrátor zaznamenal vše výše uvedené.

Úloha Dokumentace datových toků – Vzhledem k tomu, že v tuto chvíli je zřejmé, jaká data podnikového datového modelu budou integrována do jaké části datového modelu GA, je v této úloze rolemi Datového stewarda a Webového analytika každý z datových toků řádné zdokumentován do pracovního produktu Dokumentace.

Úloha Dokumentace technického návrhu – V průběhu této fáze byl navržen rozvoj podnikové databáze, datová vrstva a zvoleny integrační metody, které mají své specifické nároky na implementaci. Všechny tyto informace byly zaznamenány do sekce týkající se technického návrhu v pracovním produktu Dokumentace.

Úloha Předání dokumentace vývojářům – V pracovním produktu Dokumentace je v současné chvíli popsán kompletní technický návrh realizace včetně implementačních manuálů a současně jsou zdokumentovány budoucí datové toky. Tento pracovní produkt je hlavním podkladem pro samotnou implementaci řešení. Pracovní produkt je proto zrevidován Systémovým integrátorem a předán Vývojářům.

Úloha Řízení iterace – Úlohy této fáze proběhly v pořádku podle pracovního produktu Plán iterace. Výstupy úlohy byly dle Systémového integrátora v odpovídající kvalitě.

Úloha Řízení odchylek vůči plánu – Projekt jako celek probíhá podle pracovního produktu Plán projektu. Výraznější odchylky zaznamenány nebyly. Pouze bylo zaznamenáno několik porušení pravidel komunikace, kdy zúčastněné role komunikovaly mimo definované prostředí v aplikaci Basecamp. Role byly ponaučeny.

Úloha Zhodnocení výsledků iterace – Druhá fáze životního cyklu byla úspěšně dokončena, jelikož bylo dosaženo milníku Technický návrh realizace. Systémový integrátor i Vývojář posoudili, že zpracovaný technický návrh je pro potřeby implementace zcela dostačující. Je proto možné přejít do třetí fáze. Během druhé fáze byly spotřebovány prostředky ve výši 14 000 Kč, kdy role strávily na těchto úlohách celkem 22 hodin.

6.2.3 Fáze Konstrukce

Samotné řešení integrace dat je v současné chvíli navrženo a zdokumentováno. Vývojář obdržel potřebné podklady a může řešení implementovat. Implementované řešení bude následně testováno z hlediska funkčních i nefunkčních požadavků. Implementace řešení a následné testování jsou hlavními aktivitami fáze Konstrukce.

Úloha Plánování iterace – V souladu s pracovním produktem Plán projektu byla vytvořena nová instance pracovního produktu Plán iterace, která vymezila úlohy fáze Konstrukce na základě Metodiky integrace dat do GA. Nová instance Plánu iterace je hlavním řídícím dokumentem pro vykonání jednotlivých úloh této fáze.

Úloha Rozvoj databáze podnikového systému – Podniková databáze byla rozšířena dle pracovních produktů Návrh a Dokumentace. Nově je tak podniková Oracle databáze i MySQL databáze schopna k entitě popisující zákazníka evidovat i jeho identifikátor z cookie _ga. Přidaný zdrojový kód byl zaznamenán do pracovního produktu Implementace.

Úloha Rozvoj datové vrstvy – Pro potřeby předávání dat skrze integrační metodu Rozšířený měřící kód byla dle implementačního manuálu implementována datová vrstva, která předává informace do GTM, kde je nasazen měřící kód nástroje GA. Zároveň byla zkontrolována implementace GTM. Nově přidaná část zdrojového kódu byla zaznamenána do pracovního produktu Implementace.

Úloha Implementace řešení – Dle pracovních produktů Dokumentace a Návrh implementoval Vývojář řešení, respektive implementoval integrační metodu Measurement Protocol a datovou vrstvu pro GTM. V GTM pak příslušná nastavení ve značce GA provedla role Webového analytika. Zdrojový kód nových prvků a popis nastavení GTM byl zaznamenán do pracovního produktu Implementace.

Úloha Zavedení nefunkčních požadavků – Ve druhé fázi životního cyklu byly vymezeny následující nefunkční požadavky pro integrační metodu Measurement Protocol:

1. Odeslání dat do GA nejpozději do 30 vteřin od jejich zjištění.
2. Zaznamenávání stavových hlášení odeslaných HTTP požadavků.
3. V případě chybového stavového hlášení odesílat data každých 5 minut dokud nebudou úspěšně odeslána.

Odeslání požadavku na server skrze Measurement Protocol je automaticky spuštěno bezprostředně po uložení dat z pokladního systému do Oracle databáze, čímž je vyhověno prvnímu požadavku. Stavová hlášení odeslaných HTTP požadavků jsou automaticky sledována a v případě výskytu chybového hlášení je tentýž požadavek odeslán s pětiminutovou prodlevou znovu. Veškerá stavová hlášení jsou evidována v separátním textovém souboru. Úpravy zdrojového kódu byly zaznamenány do produktu Implementace.

Úloha Příprava testů – Bylo vytvořeno několik testovacích nápadů, kterými jsou otestování extrémních scénářů (například atypická struktura dat, chybějící údaj či více požadavků najednou), porovnání integrovaných dat z rozhraní GA vůči podnikové databázi a vytvoření segmentů v GA, kde bude sledován trend v počtu zaznamenaných uživatelů. Vše bylo zaznamenáno do nově vytvořeného pracovního produktu Seznam testovacích nápadů.

Úloha Provedení funkčních testů – Na základě pracovního produktu Seznam testovacích nápadů proběhl test funkčních požadavků. Z testu vyplynulo, že přibližně 10% nově registrovaných věrnostních karet skrze webový formulář nemají v Oracle databázi uložený identifikátor z cookie _ga. Bylo zjištěno, že se jedná o uživatele s vypnutým JS nebo uživatele používající adblockery blokující měřící kód GA. Systémový integrátor posoudil, že se jedná o zjištění, které lze akceptovat bez nutnosti dalších úprav. V testech zaměřených na integrační metodu Rozšířený měřící kód žádný nedostatek nalezen nebyl. Výsledky byly zaznamenány do pracovního produktu Záznam výsledků testů.

Úloha Provedení nefunkčních testů – Nefunkční požadavky byly vztaženy pouze na integrační metodu Measurement Protocol, kde byly testovány extrémní scénáře (více požadavků najednou, nestandardní data, chybějící údaj). Ve všech případech byly obdrženy očekávaná stavová hlášení. Testování tak nezjistilo žádné nedostatky. Výsledky byly zaznamenány do pracovního produktu Záznam výsledků testů.

Úloha Dokumentace technické realizace – Během implementace řešení vytvořil Vývojář novou část zdrojového kódu, kterou zasadit do existujícího prostředí CRM systému a webové aplikace. Popis zasazení nové části zdrojového kódu byl popsán do sekce technické realizace pracovního produktu Dokumentace. Součástí téže sekce jsou dále vysvětlivky k jednotlivým částem zdrojového kódu. Význam zmíněné sekce Dokumentace je v tom, aby případné další osoby byly schopny vytvořenému zdrojovému kódu snadno porozumět.

Úloha Revize dokumentace – Systémový integrátor novou část pracovního produktu Dokumentace zrevidoval a požádal o doplnění Webového analytika o svůj úhel pohledu, především o pohled z hlediska nastavení parametrů značky GA v GTM. Webový analytik na tento podnět Dokumentaci o svou část rozšířil. Následně Systémový integrátor posoudil, že Dokumentace je obsahově dostatečná a lze ji považovat za kompletní.

Úloha Řízení iterace – Fáze Konstrukce skončila o 3 dny později než předpokládal Plán iterace. Důvodem byla zaneprázdněnost Vývojáře, který musel věnovat svou časovou kapacitu požadavku s vyšší prioritou z jiného projektu. Tato prodleva vznikla čekáním na vykonání úlohy Implementace řešení. Ostatní úlohy proběhly přesně dle Plánu iterace.

Úloha Řízení odchylek vůči plánu – Prodleva čekáním na úlohu Implementace řešení nemá z hlediska projektu jako celku zásadní dopad, jelikož prodlevu vykryla stanovená rezerva pracovního produktu Plán projektu. Součástí Plánu projektu byla v úvodní fázi vymezena dvě rizika. Tato rizika se dosud neprojevila, přesto budou monitorována i nadále, a to až do ukončení poslední fáze životního cyklu.

Úloha Zhodnocení výsledků iterace – Fáze Konstrukce je považována za úspěšně dokončenou, jelikož bylo dosaženo milníku stanoveného metodikou, kterým je Validní datový tok. V případě tohoto projektu se dokonce jedná o dva validní datové toky prostřednictvím dvou různých integračních metod. Že jsou toky validní lze usuzovat na základě skutečnosti, že úlohy aktivity Testování neodhalily žádné závažné nedostatky. Systémový integrátor proto zhodnotil, že je možné přejít do závěrečné fáze. V průběhu fáze Konstrukce bylo spotřebováno celkem 52 000 Kč, což odpovídalo 48 hodinám.

6.2.4 Fáze Zavedení

Data jsou v tuto chvíli do Google Analytics integrována a proto je možné přejít k samotnému využití v praxi. V případě RychloMarketu se tímto užitím rozumí sestavení uvedených segmentů a jejich zužitkování pro online marketingové cílení. Dále se fáze Zavedení zabývá edukací dalších zainteresovaných rolí a sběrem zpětné vazby. Diskuze vyplývající ze edukace i zpětná vazba jsou cennými zdroji inspirace pro dosud neodhalené případy užití a pro případná vylepšení stávajícího řešení. Fáze závěrem projekt vyhodnocuje a ukončuje.

Úloha Plánování iterace – Obdobným způsobem jako v případě předchozích fází byla Systémovým integrátorem vytvořena nová instance pracovního produktu Plán iterace, která je řídícím dokumentem pro úlohy fáze Zavedení.

Úloha Edukace dalších zainteresovaných rolí – S novou integrací dat do GA seznámil Webový analytik v rámci pravidelné porady další zainteresované role, kde nad tématem proběhla krátká diskuze. Z diskuze bylo patrné, že zainteresované role vidí další případy použití, či mají návrhy na drobná vylepšení a rozšíření. Rolím účastním se diskuze bylo oznámeno, že v rámci úlohy Sběr zpětné vazby budou požádány o zpětnou vazbu. Zároveň byla stanovena role Webového analytika jako role kontaktní pro ty, kteří se chtějí o nové integraci dovědět více nad rámec probíraného rozsahu na poradě.

Úloha Aplikování navržených případů užití – Webový analytik sestavil v GA úvodem vymezené segmenty uživatelů, které následně naimportoval do inzertního systému Google Ads. V inzertním systému Google Ads pak role Online marketingového specialisty sestavila kampaně se stanoveným sdělením cíleným na nové segmenty. Tímto krokem byl dosažen úvodem vymezený cíl případu užití řetězce RychloMarket. Kromě toho bylo v rámci první fáze navrženo, že nová data lze používat pro datové analýzy a reporting. Webový analytik proto sestavil nový report v Google Analytics, který je přístupný vedení RychloMarketu a kde lze zároveň pozorovat případné odchylky v datech.

Úloha Sběr zpětné vazby – Vazba je sbírána ze dvou zdrojů. Prvním je role Online marketingového specialisty, který na vytvořené segmenty cílí inzerci. Online marketingový specialista má pouze jednu připomínku, a sice že vytvořené segmenty mají v tuto chvíli relativně málo uživatelů. Druhým zdrojem jsou role, které se na poradě zapojily do diskuze. Zde padl návrh o vizualizaci nových dat v přehledných grafech prostřednictvím nástroje Google Data Studio, který má vestavěný konektor na Google Analytics.

Úloha Vyhodnocení zpětné vazby – Připomínka o malém počtu uživatelů v nových segmentech byla vyhodnocena jako předmětná, avšak jediným řešením je vyčkat, než se segmenty díky nově integrovanými datům zaplní. Druhá připomínka týkající se vizualizace dat v Google Data Studio byla také vyhodnocena, tato připomínka bude zrealizována.

Úloha Návrh na změnové řízení – Nebyla shledána žádná připomínka, která by byla podkladem na návrh změnového řízení, jež by v důsledku znamenala úpravu implementovaného řešení. Jako jediná předmětná připomínka byla vyhodnocena vizualizace dat prostřednictvím Google Data Studio, ta však žádné změny v současně implementovaném řešení nevyžaduje. Pro sestavení vizualizace dat v Google Data Studio pověřil Systémový integrátor roli Webového analytika.

Úloha Řízení iterace – Úlohy fáze Zavedení proběhly v pořádku v souladu s vymezenou instancí pracovního produktu Plán iterace. Veškeré úlohy proběhly s povolenou tolerancí 10% z hlediska rozpočtu i spotřebovaných hodin. Výstupy jednotlivých úloh zhodnotil Systémový integrátor jako zcela dostačující.

Úloha Zhodnocení výsledků iterace – Fáze Zavedení dosáhla svého milníku, kterým je Aplikace integrovaných dat v praxi. Milníkem v kontextu tohoto případu užití se rozumí spuštění online marketingové kampaně na nové segmenty, které mohly být vymezeny díky implementované integraci dat do GA. Závěrem byla sesbírána zpětná vazba, která identifikovala příležitost ve vizualizaci nových dat prostřednictvím nástroje Google Data Studio. Fázi Zavedení tak lze považovat za úspěšně dokončenou. V průběhu této fáze byla spotřebována částka 13 000 Kč, která odpovídala 19 hodinám práce jednotlivých rolí.

Úloha Zhodnocení realizace projektu jako celku – Před ukončením projektu jako celku je dle navržené metodiky žádoucí projekt vyhodnotit. Projekt jako celek lze považovat za úspěšně zdařilý, jelikož dosáhl cíle popsaného případu užití. Kromě dosažení cíle se projekt vešel do hranic vymezených počtem hodin jednotlivých rolí a stanoveného rozpočtu na spotřebovaný čas pracovními rolemi. V první fázi projektu bylo odhadnuto, že bude spotřebováno 120 hodin a vydaná částka bude 100 000 Kč. Nyní lze vyhodnotit, že reálně bylo spotřebováno 110 hodin a vydaná částka byla 94 000 Kč. Oba údaje vychází z úloh Zhodnocení výsledků iterace, které uzavíraly každou z jednotlivých fází.

6.3 Shrnutí aplikace metodiky

Na ukázce fiktivního maloobchodního řetězce prodejen potravin RychloMarket byla demonstrována aplikace Metodiky integrace dat do Google Analytics, kde marketingové oddělení RychloMarketu chtělo vytvořit specifické segmenty pro personalizovanější sdělení prostřednictvím inzertního nástroje Google Ads. Pro potřeby sestavení vymezených segmentů bylo nutno integrovat data z podnikových databází do Google Analytics, kde byly posléze segmenty vytvořeny a skrze vestavěné propojení naimportovány do Google Ads.

Procesem změny původního stavu, kdy Google Analytics nedisponoval potřebnými daty o věrnostních kartách až do cílového stavu, kdy data o věrnostních kartách byla do GA integrována, na jejichž základě mohly být vymezeny segmenty, provedla Metodika integrace dat formou stručného komentáře ke každé z úloh metodiky. Čtenář byl seznámen s tím, jak se jednotlivé úlohy mohou snadno přizpůsobit konkrétnímu případu užití a jak lze celý proces integrace dat do Google Analytics díky navržené metodice výrazně usnadnit.

Význam šesté kapitoly této diplomové práce byl ukryt nejen v názorné demonstraci metodiky na konkrétním případu užití řetězce potravinových prodejen, ale zároveň v ověření platnosti metodiky. S ohledem na skutečnost, že díky aplikaci navržené metodiky na popsaném případu užití bylo dosaženo cíle popsaného případu užití, lze v tuto chvíli považovat Metodiku integrace dat do Google Analytics za validní.

Závěr

V průběhu diplomové práce byl čtenář seznámen s rozsahem problematiky integrace dat do nástroje Google Analytics, a to především v kontextu integrace dat s podnikovými systému. Přestože nástroj Google Analytics je určen primárně pro vykreslení charakteristik webového uživatele a analýzu jeho chování v rámci měřené webové stránky, lze tato data patřičným způsobem obohatit o data z dalších systémů a tím značně rozšířit možnosti využití Google Analytics z hlediska datových analýz nebo využití pro online marketing.

Problematika integrace dat do Google Analytics byla dosud popsána pouze v kusých materiálech popisující konkrétní integrační metodu nebo konkrétní případ užití. Avšak dosud nebylo na tuto integraci hleděno v širším kontextu pokrývající záměr integrace, návrh integrace, implementaci řešení a aplikování integrovaných dat na praktických případech užití. Právě integrace dat do Google Analytics v širších souvislostech byla předmětem této práce, přičemž celý projekt integrace dat včetně zmíněných souvislostí byl zmapován formou metodiky pojmenované jako Metodika integrace dat do Google Analytics.

Před navržením samotné metodiky byl čtenář seznámen s potřebným teoretickým základem charakterizující stěžejní komponenty nástroje Google Analytics pro potřeby integrace. Jako stěžejní komponenty byly identifikovány integrační metody, datový model a legislativní regulace. Na problematiku bylo dále nahlíženo optikou širšího kontextu charakterizující stav současného řešení, která byla formalizovaná dimenzemi metodiky MMDIS. Po vykreslení důležitých souvislostí a nastavení potřebné hladiny znalostí bylo možné přistoupit k samotnému návrhu metodiky. Navržená metodika je adaptací metodiky MMSP přizpůsobující jednotlivé prvky metodiky, kterými jsou životní cyklus, fáze, iterace, aktivity, úlohy, disciplíny, role a pracovní produkty, řešené problematice integrace dat do Google Analytics. Platnost navržené metodiky byla závěrem ověřena na ukázkovém případu užití.

Navržená metodika pokrývá celý životní cyklus integrace dat do Google Analytics po vzoru životního cyklu metodiky MMSP. Platnost navržené metodiky byla ověřena na ukázkovém případu užití, kde bylo názornou aplikací dokázáno, že navržená metodika je pro svůj účel validní. Úvodem stanovený cíl, kterým je návrh metodiky integrace dat do Google Analytics, lze tímto považovat za naplněný.

Osobně věřím, že navržená metodika nalezne své uplatnění v praxi a to nejen díky tomu, že se jedná o ucelený materiál popisující problematiku v celé své šíři, ale i díky zpracování metodiky do podoby webové verze, která je přílohou této práce (viz příloha B). Webová verze metodiky totiž umožňuje za pomoci hypertextových odkazů rychlou a přehlednou orientaci a zároveň názorněji vykresluje vzájemné závislosti konkrétních prvků navržené metodiky.

Vzhledem k tomu, že prostředí nástroje Google Analytics se postupem času vyvíjí a stejně tak se vyvíjejí možnosti integrace dat do Google Analytics, přikládám k této diplomové práci současně zdrojový kód webové verze navržené metodiky (viz příloha C), který umožňuje s ohledem na případné změny obsah metodiky v budoucnu libovolně modifikovat. Přiložený zdrojový kód lze otevřít a dále modifikovat za pomoci nástroje EPF Composer.

Použitá literatura

1. **GÁLA, Libor, POUR, Jan a ŠEDIVÁ, Zuzana.** *Podniková informatika. 2., přeprac. a aktualiz. vyd.* Praha : Expert (Grada), 2009. ISBN 978-80-247-2615-1.

2. **GOOLSBEE, Austan D. a KLENOW, Peter J.** *Internet Rising, Prices Falling: Measuring Inflation in a World of E-Commerce.* [Dostupné z: http://www.klenow.com/internet-rising-prices-falling_Goolsbee_Klenow.pdf] Pittsburgh, PA : AEA Papers and Proceedings vol. 108, 2018.

3. **PHILLIPS, Judah.** *Ecommerce analytics: analyze and improve the impact of your digital strategy.* Old Tappan, New Jersey : Pearson Education, 2016. ISBN 978-0-13-417728-1.

4. **W3Techs.** Usage statistics and market share of Google Analytics for websites. *Web Technology Surveys.* [Online] Q-Success, Leden 2019. [Citace: 12. Leden 2019.] Dostupné z: https://w3techs.com/technologies/details/ta-googleanalytics/all/all.

5. **TURBAN, Efraim, a další.** *Electronic commerce: a managerial and social networks perspective.* Cham : Springer, 2015. ISBN 978-3-319-10091-3.

6. **DE RIDDER, Jente.** How accurate are Google Analytics Demographic Reports. *humix.* [Online] 17. Březen 2014. [Citace: 12. Leden 2019.] Dostupné z: https://www.humix.be/blog/accurate-google-analytics-demographic-reports/.

7. **GÁLA, Libor, POUR, Jan a TOMAN, Prokop.** *Podniková informatika: počítačové aplikace v podnikové a mezipodnikové praxi, technologie informačních systémů, řízení a rozvoj podnikové informatiky.* Praha : Grada, 2006. ISBN 80-247-1278-4.

8. **BRUNEC, Jan.** *Google Analytics.* Praha : Grada Publishing, 2017. ISBN 978-80-271-0338-6.

9. **Google.** Integrating CRM Data with Google Analytics to create Google Ads Remarketing Audiences. *Google Developers.* [Online] 30. Červenec 2018. [Citace: 2. Březen 2019.] Dostupné z: https://developers.google.com/analytics/solutions/crm-integration.

10. **Google**. Vlastní dimenze a metriky: Přidejte do přehledů i nestandardní údaje. *Nápověda Analytics.* [Online] 2019. [Citace: 13. Leden 2019.] Dostupné z: https://support.google.com/analytics/answer/2709828?hl=cs.

11. **Google**. Jak je v Analytics definována návštěva webu. *Nápověda Analytics.* [Online] 2019. [Citace: 13. Leden 2019.] Dostupné z: https://support.google.com/analytics/answer/2731565?hl=cs.

12. **CLIFTON, Brian.** *Advanced Web metrics with Google Analytics. 2nd ed.* Indianapolis : Wiley Pub., 2010. ISBN 978-0-470-56231-4.

13. **Google.** Measurement Protocol Parameter Reference. *Google Developers*. [Online] Červenec 2018. [Citace: 15. Leden 2019.] Dostupné z: https://developers.google.com/analytics/devguides/collection/protocol/v1/parameters.

14. **Google**. Track Google Analytics Events. *Google Developers*. [Online] 30. Červenec 2018. [Citace: 9. Únor 2019.] Dostupné z: https://developers.google.com/analytics/devguides/collection/gtagjs/events.

15. **Google**. Import údajů: Získejte jednotný přehled o svých údajích. *Nápověda Analytics*. [Online] [Citace: 9. Únor 2019.] Dostupné z: https://support.google.com/analytics/answer/3191589?hl=cs.

16. **W3Techs.** Usage statistics and market share of Google Tag Manager for websites. *Web Technology Surveys*. [Online] Q-Success. [Citace: 23. Únor 2019.] Dostupné z: https://w3techs.com/technologies/details/tg-google/all/all.

17. **WEBER, Jonathan a WAISBERG, Daniel.** *Practical Google Analytics and Google Tag Manager for developers*. New York, NY : Apress, 2015. ISBN 978-1-4842-0266-1.

18. **WAISBERG, Daniel.** *Google analytics integrations*. Indianapolis, IN : Wiley, 2015. ISBN 978-1-119-05306-4.

19. **Ministerstvo vnitra České republiky.** Co je GDPR. *Orientace v GDPR*. [Online] [Citace: 24. Únor 2019.] Dostupné z: https://www.mvcr.cz/gdpr/clanek/co-je-gdpr.aspx.

20. **VAŠÍČEK, Libor.** *To podstatné z GDPR | OsobniData.cz*. Praha : LegalPartners, 2018.

21. **KMOŠKOVÁ, Jitka a KORNEL, Martin.** Co je, co není a co bude osobní údaj podle GDPR. *Frank Bold Advokáti*. [Online] Frank Bold, 10. Únor 2018. [Citace: 24. Únor 2019.] Dostupné z: https://www.fbadvokati.cz/cs/clanky/541-co-je-co-neni-a-co-bude-osobni-udaj-podle-gdpr.

22. **Úřad pro ochranu osobních údajů.** GDPR stručně. *GDPR (obecné nařízení)*. [Online] Úřad pro ochranu osobních údajů, 2018. [Citace: 24. Únor 2019.] Dostupné z: https://www.uoou.cz/gdpr-strucne/ds-4843.

23. **KINDLOVÁ, Ilona a ŠVEC, Michal.** Práva a povinnosti související s právní úpravou tzv. „cookies“. *epravo.cz*. [Online] KŠD LEGAL advokátní kancelář s.r.o., 4. Červen 2014. [Citace: 24. Únor 2019.] Dostupné z: https://www.epravo.cz/top/clanky/prava-a-povinnosti-souvisejici-s-pravni-upravou-tzv-cookies-94488.html.

24. **Česká republika.** Listina základních práv a svobod. [Online] usnesení předsednictva České národní rady, 16. Prosinec 1992. Dostupné z: https://www.psp.cz/docs/laws/listina.html.

25. **Česká republika**. Trestní zákoník (40/2009 Sb.). *Část druhá, Hlava II, Díl 2, § 180*. [Online] 19. Září 2015. Dostupné z: http://zakony.centrum.cz/trestni-zakonik/cast-2-hlava-2-dil-2-paragraf-180.

26. **European Commission.** EU-US Privacy Shield. *Data protection*. [Online] Justice and Consumers Department, 2016. Dostupné z: https://ec.europa.eu/info/law/law-topic/data-protection/data-transfers-outside-eu/eu-us-data-transfers_en.

27. **VEČERKOVÁ, Eva.** *Nekalá soutěž a reklama: (vybrané kapitoly)*. Brno : Masarykova univerzita, 2005. ISBN 80-210-3607-9.

28. **Sdružení agentur pro výzkum trhu a veřejného mínění.** Kodex v sedmi bodech. *Mezinárodní Kodex ICC/ESOMAR*. [Online] [Citace: 24. Únor 2019.] Dostupné z: http://simar.cz/standardy/kvalitativni-standardy-icc/esomar/kodex-v-sedmi-bodech.html.

29. **Association of National Advertisers.** Guidelines for Ethical Business Practice. *Data Marketing & Analytics*. [Online] [Citace: 24. Únor 2019.] Dostupné z: https://thedma.org/accountability/ethics-and-compliance/dma-ethical-guidelines/.

30. **Digital Analytics Association.** The Web Analyst's Code of Ethics. [Online] [Citace: 24. Únor 2019.] Dostupné z: https://www.digitalanalyticsassociation.org/codeofethics.

31. **Google.** Measurement Protocol, SDK, and User ID Feature Policy. *Google Developers*. [Online] 15. Říjen 2018. [Citace: 24. Únor 2019.] Dostupné z: https://developers.google.com/analytics/devguides/collection/protocol/policy.

32. **Google**. Anonymizace IP adres ve službě Analytics. *Nápověda Analytics*. [Online] [Citace: 24. Únor 2019.] Dostupné z: https://support.google.com/analytics/answer/2763052?hl=cs.

33. **BRUCKNER, Tomáš.** *Tvorba informačních systémů: principy, metodiky, architektury*. Praha : Grada, 2012. ISBN 978-80-247-4153-6.

34. **BUCHALCEVOVÁ, Alena.** *Metodiky vývoje a údržby informačních systémů: kategorizace, agilní metodiky, vzory pro návrh metodiky*. Praha : Grada, 2005. ISBN 80-247-1075-7.

35. **Google.** Dimensions & Metrics Explorer. *Google Developers*. [Online] [Citace: 15. Leden 2019.] Dostupné z: https://developers.google.com/analytics/devguides/reporting/core/dimsmets.

36. **DRKUŠIC, Emil.** Website Analytics: A Google Analytics Data Model. *Vertabelo*. [Online] 4. Říjen 2017. [Citace: 3. Březen 2019.] Dostupné z: https://www.vertabelo.com/blog/technical-articles/website-analytics-a-google-analytics-data-model.

37. **GoodData Corporation.** Analyzing Website Traffic Using Google Analytics. *GoodData Help*. [Online] 2019. [Citace: 3. Březen 2019.] Dostupné z: https://help.gooddata.com/display/doc/Analyzing+Website+Traffic+Using+Google+Analytics.

38. **SELIG, Jeff.** Ultimate Google Analytics Dimensions and Metrics List. *Overdrive Interactive*. [Online] 3. Únor 2018. [Citace: 2. Březen 2019.] Dostupné z: https://www.ovrdrv.com/ultimate-google-analytics-dimensions-and-metrics-list/.

39. **HALL, Doug.** Mixing Google Analytics Dimensions and Metrics. *Digital Debrief: Data Related Thoughts*. [Online] 17. Květen 2016. [Citace: 2. Březen 2019.] Dostupné z: http://www.kristaseiden.com/mixing-google-analytics-dimensions-and-metrics/.

40. **ROSCHE, Emily.** Understanding Scope In Google Analytics Reporting. *Bounteous: Creating Big-Picture Digital Solution*. [Online] 30. Listopad 2016. [Citace: 2. Březen 2019.] Dostupné z: https://www.bounteous.com/insights/2016/11/30/understanding-scope-google-analytics-reporting/?ns=l.

41. **SHARMA, Himanshu.** Google Analytics GDPR Checklist. Become GDPR compliant using GA. *Optimize Smart - Analytics Consulting & Conversion Optimization*. [Online] 28. Září 2018. [Citace: 2. Březen 2019.] Dostupné z: https://www.optimizesmart.com/checklist-to-become-gdpr-compliant-with-google-analytics/.

42. **KOKS, Paul.** How to Deal With Personally Identifiable Information (PII) in Google Analytics. *Online Metrics | Google Analytics Courses and Consulting*. [Online] 3. Leden 2017. [Citace: 2. Březen 2019.] Dostupné z: https://online-metrics.com/pii-google-analytics/.

43. **SHERMAN, Rick.** *Business intelligence guidebook: from data integration to analytics*. Amsterdam : Elsevier, 2015. ISBN 978-0-12-411461-6.

44. **DAYAL, Umeshwar, a další.** *Data Integration Flows for Business Intelligence*. Saint Petersburg : EDBT 2009, 2009. ISBN 978-1-60558-422-5.

45. **ELMELEEFY, Hazem, a další.** *Overview of Turn Data Management Platform for Digital*. Trento : VLDB Endowment, 2013.

46. **VAN DER VURST, Bart.** Can Google Analytics be my DMP? *element61 | Thought-leading Business Analytics & Performance Management*. [Online] 28. Srpen 2018. [Citace: 3. Březen 2019.] Dostupné z: https://www.element61.be/en/resource/can-google-analytics-be-my-dmp.

47. **HÄRDER, Theo.** *DBMS Architecture—Still an Open Problem*. Kaiserslautern : BTW vol.65, 2005.

48. **ROUSE, Margaret.** Data stewardship definition. *Building an effective data governance framework*. [Online] TechTarget, Červen 2013. [Citace: 3. Březen 2019.] Dostupné z: https://searchdatamanagement.techtarget.com/definition/data-stewardship.

49. **Gartner, Inc.** SI (system integrator). *IT Glossary*. [Online] [Citace: 3. Březen 2019.] Dostupné z: https://www.gartner.com/it-glossary/si-system-integrator.

50. **Microsoft.** What is middleware? *Microsoft Azure*. [Online] [Citace: 3. Březen 2019.] Dostupné z: https://azure.microsoft.com/en-us/overview/what-is-middleware/.

51. **BARNES, Samantha.** Using Google Analytics Measurement Protocol With Webhooks. *Bounteous: Creating Big-Picture Digital Solutions*. [Online] 9. Červen 2016. [Citace: 3. Březen 2019.] Dostupné z: https://www.bounteous.com/insights/2016/06/09/using-google-analytics-measurement-protocol-webhooks/.

52. **Lantronix, Inc.** Lantronix Demonstrates Solution to Connect Internet of Things Applications to Google Analytics. *Industrial Internet of Things (IIoT) Connectivity Solutions | IoT Devices*. [Online] 25. Únor 2014. [Citace: 3. Březen 2019.] Dostupné z: https://www.lantronix.com/newsroom/press-releases/lantronix-demonstrates-solution-connect-internet-things-applications-google-analytics/.

53. **MANGOLD, Benjamin.** Co potřebujete vědět o službě Google Analytics a soukromí. *Loves Data - Learn Google Analytics, Google Ads & More*. [Online] 2016. [Citace: 3. Březen 2019.] Dostupné z: https://www.lovesdata.com/blog/2016/what-you-need-to-know-about-google-analytics-and-privacy.

54. **ALLEN, Matt.** Top Data Security Concerns Around Data Integration. *MarkLogic | Data integration and NoSQL Databases for Your Business*. [Online] 16. Říjen 2017. [Citace: 3. Březen 2019.] Dostupné z: https://www.marklogic.com/blog/top-data-security-issues-integrating-data/.

55. **HOLMAN, Victor.** The Data Integration Lifecycle. *Business Marketing & Consulting Expert Florida | Victor Holman*. [Online] 7. Listopad 2013. [Citace: 3. Březen 2019.] Dostupné z: http://www.victorholman.com/2013/11/07/the-data-integration-lifecycle/.

56. **VAHALÍK, Tomáš.** Odhadujete pracnost projektu? *Systémový integrátor*. [Online] 2013. [Citace: 3. Březen 2019.] Dostupné z: http://www.komix-inxmail.cz/upload/noviny2004c.pdf.

57. **GAURAW, Kumar.** Data Integration Techniques And Their Challenges. *Data Integration Ninjas*. [Online] 27. Únor 2015. [Citace: 3. Březen 2019.] Dostupné z: http://www.dataintegration.ninja/data-integration-techniques-and-its-challenges/.

58. **LEVY, Alon Y.** Logic-Based Techniques in Data Integration. [autor knihy] Jack MINKER. *Logic-Based Artificial Intelligence*. Springer, Boston, MA : The Springer International Series in Engineering and Computer Science, 2000.

59. **SCHWALBE, Kathy.** *Řízení projektů v IT*. Brno : Computer Press, 2007. ISBN 978-80-251-1526-8.

60. **REJNKOVÁ, Petra.** Lokalizace a přizpůsobení metodiky OpenUP. *Diplomová práce*. Praha : Vysoká škola ekonomická v Praze, 2011.

61. **BUCHALCEVOVÁ, Alena a STANOVSKÁ, Iva.** *Příklady modelů analýzy a návrhu aplikace v UML*. Praha : Oeconomica, 2013. ISBN 978-80-245-1922-7.

62. **BUCHALCEVOVÁ, Alena.** *Metodiky budování informačních systémů*. Praha : Oeconomica, 2009. ISBN 978-80-245-1540-3.

63. **The Eclipse Foundation.** Eclipse Process Framework (EPF) Composer 1.0 Architecture Overview. *Enabling Open Innovation & Collaboration | The Eclipse Foundation.* [Online] [Citace: 9. 3 2019.] Dostupné z: https://www.eclipse.org/epf/composer_architecture/.

64. **ŠILHAVÝ, Radek, a další.** *Vybrané aspekty návrhu webových informačních systemů.* Vsetín : Šilhavý, s.r.o., 2013. ISBN 978-80-904741-3-0.

65. **The Eclipse Foundation.** Eclipse Process Framework Project (EPF). *Enabling Open Innovation & Collaboration | The Eclipse Foundation.* [Online] 5. Červen 2018. [Citace: 16. Březen 2019.] Dostupné z: https://www.eclipse.org/epf/.

66. **BOEHM, Barry W.** Verifying and Validating Software Requirements and Design Specifications. *IEEE Software.* 1, 1984, Sv. 1.

67. **Organisation Internationale de Normalisation.** ISO/IEC/IEEE 29148:2018(en). *Online Browsing Platform (OBP).* [Online] 2018. [Citace: 22. Březen 2019.] Dostupné z: https://www.iso.org/obp/ui/#iso:std:iso-iec-ieee:29148:ed-2:v1:en.

68. **ROSENAU, Milton D.** *Řízení projektů.* Praha : Computer Press, 2000. ISBN 80-7226-218-1.

69. **SOMMERVILLE, Ian.** Software Documentation. [Online] Pearson Education Ltd., 7. Listopad 2001. [Citace: 24. Březen 2019.] http://www.hostemostel.com/software/73.pdf.

70. **BUREŠ, Miroslav, a další.** *Efektivní testování softwaru: klíčové otázky pro efektivitu testovacího procesu.* Praha : Grada, 2016. ISBN 978-80-247-5594-6.

71. **HOHPE, Gregor a WOOLF, Bobby.** *Enterprise integration patterns: designing, building, and deploying messaging solutions.* Boston : Addison-Wesley, 2004. ISBN 0-321-20068-3.

72. **CABINET OFFICE.** *ITIL service transition.* London : TSO, 2011. ISBN 0113313063.

Přílohy

Příloha A: Vazby mezi prvky navržené metodiky

Obsahem přílohy A jsou tabulky (A.1 až A.4) reprezentující vzájemné vazby mezi prvky v jednotlivých fázích životního cyklu navržené metodiky.

Tabulka A.1: Vazby mezi prvky navržené metodiky ve fázi Zahájení (zdroj: vlastní zpracování)

Aktivita	Úloha	Role	Disciplína	Pracovní produkt
Řízení a plánování iterace	Plánování iterace	Systémový integrátor	Řízení projektu	Seznam pracovních položek / Plán iterace
	Řízení iterace	Systémový integrátor	Řízení projektu	Seznam pracovních položek / Plán iterace
	Zhodnocení výsledků iterace	Systémový integrátor	Řízení projektu	Plán iterace
Analýza současného stavu	Poznání stávajících dat	Datový steward	Dokumentace	Datový model
	Identifikace dat k integraci do GA	Datový steward / Webový analytik	Dokumentace	Datový model
	Ověření dostupnosti dat	Datový steward	Dokumentace	Datový model
	Ověření platnosti datového modelu GA	Webový analytik	Dokumentace	Datový model
Sestavení a zhodnocení záměru	Sestavení záměru integrace	Systémový integrátor	Požadavky	Slovník pojmů / Požadavky / Vize
	Zhodnocení technické proveditelnosti	Vývojář / Webový analytik	Požadavky	Slovník pojmů / Požadavky
	Zhodnocení legislativních aspektů	Právník	Požadavky	Slovník pojmů / Požadavky
	Modelování případů užití	Webový analytik / OM specialista	Aplikace	Případy užití
	Vymezení rozsahu integrace	Systémový integrátor	Požadavky	Slovník pojmů / Požadavky / Vize
Příprava plánu realizace	Plánování činností	Systémový integrátor	Řízení projektu	Seznam pracovních položek / Plán projektu
	Alokace kapacit pracovních rolí	Systémový integrátor	Řízení projektu	Seznam pracovních položek / Plán projektu
	Identifikace a návrh řízení rizik	Systémový integrátor	Řízení projektu	Plán projektu
	Definice komunikace	Systémový integrátor	Řízení projektu	Slovník pojmů / Plán projektu
	Kompletace projektového plánu	Systémový integrátor	Řízení projektu	Plán projektu
Milník: Schválený projektový plán				

Tabulka A.2: Vazby mezi prvky navržené metodiky ve fázi Rozpracování (zdroj: vlastní zpracování)

Aktivita	Úloha	Role	Disciplína	Pracovní produkt
Řízení a plánování iterace	Plánování iterace	Systémový integrátor	Řízení projektu	Seznam pracovních položek / Plán iterace / Plán projektu
	Řízení iterace	Systémový integrátor	Řízení projektu	Seznam pracovních položek / Plán iterace
	Řízení odchylek vůči plánu	Systémový integrátor	Řízení projektu	Seznam pracovních položek / Plán iterace / Plán projektu
	Zhodnocení výsledků iterace	Systémový integrátor	Řízení projektu	Seznam pracovních položek / Plán iterace
Příprava realizace	Volba integrační metody	Vývojář / Webový analytik	Vývoj	Návrh
	Alokace prostoru v datovém modelu GA	Webový analytik	Vývoj	Návrh
	Návrh rozvoje db podnikového systému	Databázový specialista	Vývoj	Návrh
	Návrh rozvoje datové vrstvy	Webový analytik	Vývoj	Návrh
	Návrh testů	Systémový integrátor	Testování	Seznam testovacích nápadů / Plán testů
	Vymezení nefunkčních požadavků	Systémový integrátor	Požadavky	Požadavky
Vedení dokumentace	Dokumentace datových toků	Datový steward / Webový analytik	Dokumentace	Dokumentace
	Dokumentace technického návrhu	Datový steward / Webový analytik	Dokumentace	Dokumentace
	Předání dokumentace vývojářům	Systémový integrátor	Dokumentace	Dokumentace
Milník: Technický návrh realizace				

Tabulka A.3: Vazby mezi prvky navržené metodiky ve fázi Konstrukce (zdroj: vlastní zpracování)

Aktivita	Úloha	Role	Disciplína	Pracovní produkt
Řízení a plánování iterace	Plánování iterace	Systémový integrátor	Řízení projektu	Seznam pracovních položek / Plán iterace / Plán projektu
	Řízení iterace	Systémový integrátor	Řízení projektu	Seznam pracovních položek / Plán iterace
	Řízení odchylek vůči plánu	Systémový integrátor	Řízení projektu	Seznam pracovních položek / Plán iterace / Plán projektu
	Zhodnocení výsledků iterace	Systémový integrátor	Řízení projektu	Seznam pracovních položek / Plán iterace
Vývoj	Rozvoj databáze podnikového systému	Databázový specialista	Vývoj	Implementace
	Rozvoj datové vrstvy a implementace GTM	Vývojář	Vývoj	Implementace
	Implementace řešení	Vývojář	Vývoj	Implementace
	Zavedení nefunkčních požadavků	Vývojář	Vývoj	Implementace
Testování řešení	Příprava testů	Vývojář / Systémový integrátor	Testování	Seznam testovacích nápadů / Plán testů
	Provedení funkčních testů	Vývojář / Webový analytik	Testování	Testovací případ / Záznam výsledků testů
	Provedení nefunkčních testů	Vývojář / Webový analytik	Testování	Testovací případ / Záznam výsledků testů
Vedení dokumentace	Dokumentace technické realizace	Vývojář	Dokumentace	Dokumentace
	Revize dokumentace	Systémový integrátor	Dokumentace	Dokumentace
Milník: Validní datový tok				

Tabulka A.4: Vazby mezi prvky navržené metodiky ve fázi Zavedení (zdroj: vlastní zpracování)

Aktivita	Úloha	Role	Disciplína	Pracovní produkt
Řízení a plánování iterace	Plánování iterace	Systémový integrátor	Řízení projektu	Seznam pracovních položek / Plán iterace / Plán projektu
	Řízení iterace	Systémový integrátor	Řízení projektu	Seznam pracovních položek / Plán iterace
	Řízení odchylek vůči plánu	Systémový integrátor	Řízení projektu	Seznam pracovních položek / Plán iterace / Plán projektu
	Zhodnocení výsledků iterace	Systémový integrátor	Řízení projektu	Seznam pracovních položek / Plán iterace
	Zhodnocení realizace projektu jako celku	Systémový integrátor	Řízení projektu	Požadavky / Plán projektu
Aplikování řešení	Edukace dalších zainteresovaných rolí	Webový analytik	Aplikace	Případy užití / Seznam pracovních rolí
	Aplikování navržených případů užití	Webový analytik / OM specialista	Aplikace	Případy užití
Vyhodnocení zpětné vazby	Sběr zpětné vazby	Systémový integrátor	Řízení změn a konfigurací	Seznam požadavků na změnu
	Vyhodnocení zpětné vazby	Systémový integrátor / Webový analytik	Řízení změn a konfigurací	Seznam požadavků na změnu
	Návrhy na změnové řízení	Systémový integrátor	Řízení změn a konfigurací	Seznam požadavků na změnu
Milník: Aplikace integrovaných dat v praxi				

Příloha B: Webová verze metodiky

Společně s touto diplomovou prací byla odevzdána webová verze Metodiky integrace dat do Google Analytics sestavená nástrojem EPF Composer. Webová verze metodiky je dostupná po boku této diplomové práce ve studijním informačním systému VŠE v Praze. K otevření webové verze metodiky se doporučuje prohlížeč Internet Explorer nebo Mozilla Firefox.

Příloha C: Zdrojové soubory metodiky

Přílohou diplomové práce jsou zdrojové soubory navržené metodiky, které byly společně s prací uloženy do studijního informačního systému VŠE v Praze. Zdrojové soubory metodiky byly vytvořeny nástrojem EPF Composer ve verzi 1.5.2. Stejným nástrojem je lze otevřít, modifikovat a sestavit aktualizovanou či rozšířenou verzi navržené metodiky.