



Zavádění personalizace a cílení obsahu webové aplikace na základě zákaznického profilu

Bakalářská práce

Studijní program: B6209 – Systémové inženýrství a informatika

Studijní obor: 6209R021 – Manažerská informatika

Autor práce: **Antonín Doležal**

Vedoucí práce: Mgr. Tomáš Žižka



ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Antonín Doležal**
Osobní číslo: **E14000507**
Studijní program: **B6209 Systémové inženýrství a informatika**
Studijní obor: **Manažerská informatika**
Název tématu: **Zavádění personalizace a cílení obsahu webové aplikace
na základě zákaznického profilu**
Zadávací katedra: **Katedra informatiky**

Z á s a d y p r o v y p r a c o v á n í :

1. Základní pojmy vztahující se k personalizaci a cílení obsahu
2. Analýza současného stavu personalizačních možností a cílení obsahu ve vybrané B2C webové aplikaci Škoda Auto
3. Návrh principů zavedení personalizace a cílení obsahu ve vybrané webové aplikaci s ohledem na integraci s dalšími systémy
4. Časová, nákladová a riziková analýza návrhu
5. Zhodnocení navržených řešení

Rozsah grafických prací:

Rozsah pracovní zprávy: **30 normostran**

Forma zpracování bakalářské práce: **tištěná/elektronická**

Seznam odborné literatury:

GERMANAKOS, P. and M. BELK. Human - Centered Web Adaption and Personalization: From Theory to Practice. Berlin: Springer, 2016. ISBN 3319280503.

CASTELLANO, G. Web personalization in intelligent environments. Online-Ausg. Berlin: Springer, 2009. ISBN 9783642027932.

KRUG, S. Web design - nenuťte uživatele přemýšlet!. 2. vyd. Brno: Computer Press, 2006. ISBN 80-251-1291-8.

Elektronická databáze článků ProQuest (knihovna.tul.cz).

Vedoucí bakalářské práce:

Mgr. Tomáš Žižka

Katedra informatiky

Konzultant bakalářské práce:

Ing. Jana Svitalská

analytik, Trask solution a. s.

Datum zadání bakalářské práce:

31. října 2016

Termín odevzdání bakalářské práce:

31. května 2018



prof. Ing. Miroslav Žižka, Ph.D.
děkan



doc. Ing. Jan Skrbek, Dr.
vedoucí katedry

V Liberci dne 31. října 2016

Prohlášení

Byl jsem seznámen s tím, že na mou bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb., o právu autorském, zejména § 60 – školní dílo.

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

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

Bakalářskou práci jsem vypracoval samostatně s použitím uvedené literatury a na základě konzultací s vedoucím mé bakalářské práce a konzultantem.

Současně čestně prohlašuji, že tištěná verze práce se shoduje s elektronickou verzí, vloženou do IS STAG.

Datum: 04.05.2017

Podpis: Doležal

Anotace

Bakalářská práce se zabývá tématem zavádění personalizace a cílení obsahu webové aplikace na základě zákaznického profilu. V teoretické části práce jsou popsány základní pojmy vztahující se k personalizaci a cílení obsahu, vysvětleny metody sběru uživatelských dat, naznačena tvorba uživatelských modelů, rozděleny adaptační technologie a v závěru je upozorněno na lidské faktory vztahující se k personalizaci. V praktické části jsou na základě analýzy současného stavu personalizačních možností a cílení obsahu navrženy principy zavedení personalizace ve vybrané B2C webové aplikaci Škoda Auto s ohledem na integraci s dalšími systémy. Jsou zde navrženy jednotlivé business případy s možnými řešeními. Následují návrhy realizačních možností a zhodnocení jednotlivých návrhů pomocí časové, nákladové a rizikové analýzy.

Klíčová slova

Adaptační technologie, adaptivní webové aplikace, personalizace, uživatelské modely.

Annotation

Bachelor thesis deals with the implementation of personalization and context targeting of web applications based on the customer profile. The theoretical part describes the basic concepts related to personalization and content targeting, explains methods of collecting user data, indicates creation of user models, divides adaptation technologies and finally points out the human factors related to personalization. The practical part is based on an analysis of the current state of personalization options and content targeting are designed principles to introduce personalization in the selected B2C web application Škoda Auto with respect to integration with other systems. There are proposed various business cases with possible solutions. The following are suggestions of implementation options and evaluation of individual proposals through time, cost and risk analysis.

Key Words

Adaptation technologies, adaptive web applications, personalization, user models.

Obsah

Seznam zkratk	9
Seznam tabulek	10
Seznam obrázků	11
Úvod	12
1 Zhodnocení současného stavu problematiky z dostupné odborné literatury	13
1.1 Knihy	13
1.2 Odborné články	14
1.3 Kvalifikační práce	14
2 Personalizace	16
2.1 Dělení personalizace (kategorie)	16
2.1.1 Odkazová personalizace	17
2.1.2 Obsahová personalizace	18
2.1.3 Personalizované vyhledávání na webu	18
2.1.4 Kontextová personalizace.....	19
2.1.5 Autorizovaná personalizace.....	20
2.1.6 Humanizovaná personalizace	20
2.2 Metody sběru uživatelských dat	21
2.2.1 Explicitní metody sběru uživatelských dat.....	21
2.2.2 Implicitní metody sběru uživatelských dat.....	22
2.2.3 Shrnutí explicitních a implicitních metod	23
2.3 Vytváření uživatelského modelu	24
2.4 Shlukování (clustering)	24
2.5 Klasifikace	25
2.6 Asociační výsledky	25
2.7 Sekvenční vzory	27
2.8 Adaptační technologie	27
2.8.1 Uživatelské přizpůsobování.....	28
2.8.2 Filtrování založené na pravidlech.....	28
2.8.3 Obsahově založené filtrování	29
2.8.4 Kolaborativní filtrování	29
2.8.5 Web mining	31
2.8.6 Demograficky založené filtrování	33
2.8.7 Agent technologie.....	33

2.8.8	Shlukovací modely	33
2.9	Personalizace založená na historii vyhledávání	34
2.10	Lidské faktory v personalizaci	35
2.10.1	Vizuální vnímání	36
3	Analýza současného stavu personalizačních možností vybrané webové aplikace	38
3.1	Obsah, využití a cíle aplikace	38
3.2	SWOT z pohledu personalizačních možností	39
3.2.1	Silné stránky	39
3.2.2	Slabé stránky	42
3.2.3	Příležitosti	43
3.2.4	Hrozby	43
4	Návrh využití personalizace, možná řešení a jejich zhodnocení	45
4.1	Možnosti využití personalizace	45
4.2	Návrhy možných řešení	48
4.2.1	Vývoj vlastního řešení personalizace	48
4.2.2	Nákup hotového řešení personalizace	51
4.2.3	Časová, nákladová a riziková analýza variant řešení	52
4.3	Zhodnocení variant řešení	57
Závěr	60
Seznam použité literatury	61

Seznam zkratek

B2C	Business to customer
CMS	Content management system
COMS	Conference Management system
CSS	Cascading Style Sheets
CTA	Call To Action
GDPR	General Data Protection Regulation
GPS	Global positioning system
HCI	Human - Computer interaction
IP	Internet Protocol
MD	Man day
URL	Uniform Resource Locator
VW	Volkswagen

Seznam tabulek

Tabulka 1: Příklad kolaborativního filtrování	30
Tabulka 2: SWOT analýza webové aplikace www.skoda-auto.pl	39
Tabulka 3: Porovnání variant řešení	58

Seznam obrázků

Obrázek 1: Dělení personalizace	17
Obrázek 2: Proces clusteringu	25
Obrázek 3: Dělení web miningu.....	31
Obrázek 4: Příklad gestaltismu - Kanizsův trojúhelník.....	37
Obrázek 5: Ukázka novinky pro uživatele ze segmentu Motosport akce.....	40
Obrázek 6: Ukázka novinky pro uživatele ze segmentu vlastníci vozů	41
Obrázek 7: Ukázka novinky pro uživatele ze segmentu veteráni.....	41
Obrázek 8: Náhled webové aplikace www.skoda-auto.pl.....	47
Obrázek 9: Schéma procesu vlastního řešení	49
Obrázek 10: Hrubý odhad potřeby lidských zdrojů pro variantu vlastního řešení	54
Obrázek 11: Hrubý plán pro variantu vlastního řešení.....	54
Obrázek 12: Projektová analýza pro variantu vlastního řešení	55
Obrázek 13: Hrubý odhad potřeby lidských zdrojů pro variantu nákupu hotového řešení .	56
Obrázek 14: Hrubý plán pro variantu nákupu hotového řešení.....	56
Obrázek 15: Projektová analýza pro variantu nákupu hotového řešení	57

Úvod

Tato bakalářská práce se zabývá principy pro zavádění personalizace a cílení obsahu webové aplikace na základě zákaznického profilu a aktuálního kontextu uživatele.

Cílem celé práce je na základě analýzy současného stavu personalizačních možností a cílení obsahu ve vybrané B2C webové aplikaci Škoda Auto navrhnout principy zavedení personalizace a cílení obsahu s ohledem na integraci s dalšími systémy. Dílčími cíly práce jsou podrobení návrhů časové, nákladové a rizikové analýze a zhodnocení navržených řešení.

Práce se zabývá základními otázkami personalizace jako například: Jak personalizace obsahu vlastně probíhá? Jaké druhy personalizace existují? Jak získat od uživatelů, kterým bude poskytován personalizovaný obsah, potřebné informace o jejich zájmech a preferencích?

Práce je rozdělena na dvě části, teoretickou a praktickou část, které se dále dělí na další pod části. První, teoretická část, má za účel popsat základní pojmy vztahující se k personalizaci, jako je například rozdělení adaptačních technologií nebo popsání sběru dat.

Druhá pasáž je věnována části praktické, ve které je analýza současného stavu personalizačních možností a cílení obsahu. Na základě této analýzy jsou poté navrženy principy zavedení personalizace a cílení obsahu ve vybrané webové aplikaci. Následně je návrh podroben časové, nákladové a rizikové analýze. Závěrem praktické části je zhodnocení navržených řešení.

1 Zhodnocení současného stavu problematiky z dostupné odborné literatury

Obsahem této kapitoly je prozkoumání a následné popsání všech dostupných zdrojů, které souvisí s tématem bakalářské práce. Vzhledem k nedostatku související knižní publikace a odborných článků v českém jazyce je tato kapitola věnována především cizojazyčným publikacím.

1.1 Knihy

Knižní publikace zabývající se problematikou tohoto tématu byly rozděleny do tří skupin.

První skupinou jsou marketingově zaměřené publikace, které se orientují především na doporučující systémy. Jednou z těchto publikací je například kniha "Personalization techniques and recommender systems" od autorů G. Uchyigit a M. Y. Ma, kteří ve své publikaci popisují personalizační techniky a doporučující systémy s využitím mobilních přístupů. Další publikací je kniha pod názvem "Target Advertising Technologie in ICT Space" od autora Ch. Schlee, která se zaměřuje na cílení reklamních technologií a poskytuje přehled o trendech na trhu.

Druhou skupinou jsou technologicky orientované publikace. První takovou publikací je kniha "The adaptive web: Methods and Strategies of web personalization" od autorů P. Brusilovski, A. Kobsa, W. Nejdl, která se mimo jiného zabývá personalizačními technikami, modely personalizace a poukazuje na rozdíl mezi personalizací a přizpůsobením. Druhá publikace zabývající se technickým směrem personalizace je kniha s názvem "Web personalization in intelligent environments" od autorů G. Castellano a A. M. Fanelle, kteří ve své knize ukazují nové výzkumné směry webové personalizace a inteligentních prostředí. Třetí publikací je kniha "Evaluating the effectiveness of explanations for recommender systems" od autorů N. Tintarev a J. Masthoff zabývající se metrikami a efektivností personalizace. Ukazuje na nedostatečné a omezené metriky a zkoumá vliv personalizace na základě spokojenosti uživatele. Tato publikace částečně souvisí se třetí skupinou.

Třetí skupinou jsou publikace, které jsou zaměřené na uživatele, "na člověka". První publikací je kniha pod názvem "Human-Centered web adaption and personalization: from theory to practice" od autorů P. Germanakos a M. Belk, jedná se o jednu z nejaktuálnějších publikací, byla vydána v únoru 2016. Zabývá se návrhem a vývojem adaptivních interaktivních systémů zaměřených na uživatele. Hlavní myšlenkou této knihy je, že člověk, ve vztahu HCI (human-computer interaction), je nejdůležitějším článkem. Druhou publikací této skupiny je kniha "Privacy-Enhanced web personalization" od autora A. Kobsa. Tato kniha pojednává o právních úpravách v mnoha zemích a státech, zabývá se ochranou osobních údajů. Analyzuje napětí mezi personalizací a soukromím uživatelů na internetu a snaží se o jejich skloubení.

1.2 Odborné články

První článek, který se zabývá vybranou problematikou, nese název "User models for adaptive hypermedia and adaptive educational systems" od autorů P. Brusilovsky a E. Millán. Autoři v článku popisují uživatelské modely a přístupy k nim. Druhý článek "Model-driven development of context-aware web application" od autorů S. Ceri, F. Daniel, M. Matera, FM. Fecca se zaměřuje na konceptuální modelování a modelování kontextových vícekanálových aplikací, kde kontext neboli souvislost hraje největší roli.

Další článek pod názvem "Semantic personalization of web portal contents" od autorů C. Tziviskou a M. Brambilla poukazuje na konceptuální model pro extrakci personalizovaných doporučení na základě profilu uživatele.

1.3 Kvalifikační práce

Diplomová práce s názvem "Adaptivita webových aplikací a modelování uživatele", studenta Martina Podborského z Mendelovy univerzity v Brně, se zabývá problematikou adaptivity webových aplikací a modelování uživatele. V praktické části jsou navržena a implementována obecná komponenta pro podporu adaptivity a personalizace webových systémů zahrnující modelování uživatele.

Disertační práce pod názvem "Designing Personalization in Technology-Based Services", studenta Min Kyung Lee z univerzity jménem Carnegie Mellon University Pittsburgh, rozšiřuje koncept personalizace, navrhuje a hodnotí adaptivní personalizační strategii, která se přizpůsobuje změnám lidských preferencí, a zachycuje uživatelsky orientovaný proces tvorby designu.

2 Personalizace

V souvislosti s webem či internetovou sítí personalizace znamená dodání dynamického obsahu, jako jsou textové prvky, odkazy, reklamy, nabídky produktů a podobně, který je přizpůsoben potřebám a zájmům konkrétního uživatele nebo segmentu uživatelů. Rozlišujeme "automatickou personalizaci" a tzv. "přizpůsobení", jak bývá personalizace také označována. Oba pojmy, jak personalizace, tak přizpůsobení, se váží k poskytování obsahu na míru pro konkrétního uživatele. Tím hlavním, co odděluje tyto dva pojmy, je to, kdo ovládá tvorbu uživatelských profilů stejně jako prezentaci prvků uživatelského rozhraní pro uživatele. Bavíme-li se o přizpůsobení, uživatelé jsou pod kontrolou (často manuální) s uvedením svých preferencí a požadavků, na jejichž základě jsou pak vytvořeny prvky uživatelského rozhraní. Automatická personalizace na druhé straně znamená to, že uživatelské profily jsou tvořeny a případně aktualizovány automaticky pomocí systému, který vyžaduje minimální kontrolu ze strany uživatele (Brusilovsky, 2007).

Obecně platí, že proces personalizace se skládá ze dvou fází. Jde o fázi sběru dat, ve které jsou získávány informace týkající se zájmů uživatele, a fázi učení, v níž jsou vytvořeny uživatelské profily na základě získaných informací. Učení se z dat může být rozděleno mezi paměťově založené učení (označováno také jako líné), nebo modelově založené učení (tzv. toužící), které závisí na tom, jak učení probíhá. Může probíhat online, mezitím co systém provádí personalizační úkoly, nebo offline pomocí využití tréninkových dat (Brusilovsky, 2007).

2.1 Dělení personalizace (kategorie)

Personalizaci můžeme rozdělit do šesti kategorií (viz obrázek 1): Odkazová personalizace, obsahová personalizace, personalizované vyhledávání, kontextová personalizace, autorizovaná personalizace a humanizovaná personalizace. Každá z těchto kategorií bude popsána v následujících podkapitolách (Germanakos, 2016).



Obrázek 1: Dělení personalizace
Zdroj: vlastní

2.1.1 Odkazová personalizace

Odkazová personalizace se věnuje přizpůsobení struktury webu a prezentaci hypertextových odkazů v interaktivním systému. Toho je dosaženo pomocí výběru takových odkazů, které jsou nejvíce relevantní pro uživatele (např. na základě jeho zájmu či preferencí), změnou původního navigačního prostoru, díky zmenšení či zvětšení jednotlivých prvků, a připojením prezentovaných odkazů.

Odkazová personalizace je využívána především v aplikacích elektronické komerce (E-Komerce) jako doporučující prostředek příslušných produktů na základě zákaznickovy historie nákupů a jeho hodnocení jednotlivých produktů nebo kategorií (Germanakos, 2016).

2.1.2 Obsahová personalizace

Obsahová personalizace se zaměřuje na přizpůsobení a personalizaci obsahu uživatelského rozhraní. Obsahovou personalizaci můžeme rozdělit do dvou kategorií:

Strukturovaná personalizace obsahuje filtrování obsahu, které je relevantní pro uživatele, či na základě zájmu daných *uživatelů*. Uživatelé mohou přímo uvést své preference, nebo mohou být odvozeny z jejich statistik. Například aplikace se sportovními novinkami může zobrazovat lokalizované informace na základě umístění GPS.

1. Obsahová personalizace je jemnější než strukturovaná personalizace. Jde o přizpůsobení informací o stejném prvku různým uživatelům. Například systém E-Komerce nabízí stejný produkt s různými slevami různým uživatelům na základě jejich historie nákupů (Germanakos, 2016).

2.1.3 Personalizované vyhledávání na webu

Personalizované vyhledávání na webu se zabývá přizpůsobením a personalizací výsledků vyhledávání s ohledem na zájmy jednotlivce. Personalizované vyhledávání na webu může být realizováno buď na straně serveru jako součást metod vyhledávače, nebo na straně klienta, tj. na počítači uživatele například jako plugin webového prohlížeče.

V rámci personalizovaného vyhledávání systém vytváří uživatelské modely charakteristik, zájmů a preferencí na základě specifických koncepčních kategorií. Uživatelé modely mohou obsahovat:

- Demografické informace jako je věk, pohlaví, jazyk, vzdělání
- Geolokační údaje jako je země, adresa
- Zájmy a preference
- Historii vyhledávání jako jsou předchozí dotazy uživatele, navštívené stránky, stažené soubory
- Chování jako je kliknutí myši, pohyb myši, posouvání obrazu
- Uživatelé akce jako je vytváření záložek, nastavení webové stránky jako oblíbené

Podle Wena a jeho společníků lze modelování uživatelských údajů dosáhnout pomocí následujících technik:

- Personalizované vyhledávání na základě obsahové analýzy, ve které systém porovnává a kontroluje obsahovou podobnost mezi webovou stránkou a uživatelským modelem.
- Personalizované vyhledávání založené na analýze odkazů, ve kterém systém vypočítá individuální důležitost webových dokumentů pro každého uživatele.
- Personalizované vyhledávání na základě spolupracujících přístupů, kterými systém prezentuje podobné výsledky vyhledávání pro uživatele s podobnými uživatelskými modely (Germanakos, 2016).

2.1.4 Kontextová personalizace

Kontextová personalizace se zakládá na přizpůsobení informací, které jsou použité v různých kontextech. Kontextová personalizace může vycházet z polohy uživatele, interakčního zařízení, fyzického prostředí nebo sociálního kontextu. Například personalizované systémy založené na umístění doporučují nová zajímavá místa pro uživatele na základě jejich zájmů a preferencí, s využitím obsahových nebo spolupracujících doporučujících technik.

Herder a jeho společníci identifikovali čtyři proudy výzkumu, které se zaměřují na služby založené na poloze.

- Vzory pohybu lidí (počet návštěv určitého místa atd.).
- Předpovídání dalšího místa, které využívají dolovací techniky, jako jsou Markovovy modely a Bayesovy modely, pro predikci uživatelova příštího umístění.
- Umístění a sociální média, která se snaží analyzovat pohyb lidí. Nicméně výzkum v této oblasti není příliš účinný, protože uživatelé jen zřídka sdílejí jejich přesnou polohu.
- Služby založené na umístění, které poskytují především doporučení na základě umístění uživatele a kontextové výsledky vyhledávání.

Dalším příkladem kontextové personalizace může být využití kontextových parametrů. Pomocí těchto parametrů jsou uživatelům doporučovány blízké restaurace na základě polohy a aktuálního času (například obědy, večere), (Germanakos, 2016).

2.1.5 Autorizovaná personalizace

Autorizovaná personalizace se využívá v interaktivních systémech, které poskytují odlišné přístupy k informacím a možnostem akcí pro uživatele s různými rolmi v systému. Nejrozšířenější přístup k autorizované personalizaci je řízení přístupu na bázi rolí, ve kterých jsou přístupová práva jednotlivých částí systému rozdělena do kategorií podle názvu role. V tomto přístupu platí vztah M:N mezi uživateli, rolmi a oprávněními, ve kterém každý uživatel patří do konkrétních rolí a každá role se vztahuje ke konkrétním oprávněním. Například v Conference Management systému (COMS) mají uživatelé různé přístupy a možnosti akcí závislé na jejich rolích (autoři, recenzenti, organizátoři). Autoři mohou pouze předložit a nahrát svůj vlastní papír, zatímco recenzent může zobrazit a spravovat více dokumentů od různých autorů. Řízení přístupů na základě rolí je dobře škálovatelné a hierarchie rolí umožňuje snadná přiřazení oprávnění (Germanakos, 2016).

2.1.6 Humanizovaná personalizace

Humanizovaná personalizace se zaměřuje na vytváření personalizovaných uživatelských rozhraní založených na vnitřních lidských faktorech, jako jsou emocionální faktory (úzkost, stress), osobnostní rysy, kognitivní styly, schopnost učení se, vizuální pozornost atd. Vzhledem k velmi komplexnímu a multi-dimenzionálnímu charakteru těchto faktorů je personalizace obsahu a funkčnost interaktivních systémů založených na lidských faktorech stále v ještě v plenkách. Kromě toho tu je hodně praktických omezení vztahujících se k těmto faktorům.

Proces modelování vyžaduje explicitní metody sběru uživatelských dat, v nichž uživatel absolvuje řadu psychometrických testů, odpovídá na speciálně navržené dotazníky nebo se podílí na kontrolovaných laboratorních studiích. Tyto studie využívají externí hardwarové zařízení, jako je například zařízení pro sledování očí, pomocí něhož je zjišťována vizuální

pozornost a vizuální vyhledávání uživatelů, fyziologické senzory na sledování krevního tlaku, srdeční frekvence atd.

Nicméně existující výzkum naznačuje, že jednotlivci se v těchto lidských faktorech liší stejně jako empirické poznatky, které ukazují dopad personalizace a funkčnosti interaktivních systémů založených na těchto vnitřních faktorech. Je potřeba dalšího zkoumání účinků a dopadů při uplatňování těchto přízpůsobení v personalizovaných systémech ve prospěch uživatele (Germanakos, 2016).

2.2 Metody sběru uživatelských dat

Všechny přístupy k personalizaci, zejména pak personalizace založená na dolování dat, vyžadují sbírání dat, která přesně odrážejí zájmy uživatelů, a jejich interakce s aplikacemi a předměty. Personalizační systémy se neliší pouze v algoritmu, který slouží ke generování doporučení nebo provedení predikce, ale také ve způsobu použití podkladových dat pro tvorbu uživatelských profilů. Metody sběru uživatelských dat k tvorbě uživatelských profilů se dělí do dvou skupin: Explicitní metody a implicitní metody (Brusilovsky, 2007).

2.2.1 Explicitní metody sběru uživatelských dat

Explicitní metody sběru uživatelských dat spoléhají na osobní informace poskytnuté samotným uživatelem, většinou pomocí registračních formulářů. Shromážděná data nejčastěji obsahují: demografické informace (tj. věk, pohlaví, povolání), zájmy nebo preference. Běžné techniky pro získání explicitní informace zahrnují použití zaškrtačích tlačítek, rozvírací seznamy nebo textové pole, kde uživatel může vyjádřit svůj názor. Všechny tyto techniky mají výhodu v tom, že formáty odpovědí jsou standardizovány.

Hlavní nevýhodou je uživatelské vědomí o tom, že systém tyto informace ukládá, proto může nastat několik případů:

- Uživatel nebude chtít poskytnout potřebné informace
- Uživatel nebude mít k systému dostatek důvěry
- Uživatel nebude mít dostatek času

Na základě toho pak může být proces přerušen. Výsledky z těchto technik jsou také náchylné k lidským chybám, protože v případě, že otázky nejsou pečlivě navrženy, pak mohou být nepřesné, nejednoznačné nebo v nejhorším případě zavádějící.

Přístupy založené na explicitním sběru uživatelských dat jsou běžně využívány pro přizpůsobení uživatelského rozhraní. Nasbírané informace jsou použity k vytvoření uživatelského modelu, podle něhož jsou přizpůsobené poskytované služby. Například Google se uživatelů přímo ptá na jejich osobní informace, které jsou poté využity k tvorbě uživatelských modelů. Obsah webové stránky je poté dynamicky organizován na základě preferencí uživatele.

Důležitou nevýhodou těchto přístupů je, že uživatel nemusí přesně nebo úplně nahlásit své preference a charakteristiky. Dalším problémem je, že většina interaktivních systémů, které využívají těchto přístupů, stěží donutí uživatele, aby tyto informace aktualizovali a zřídka kdy, mají natolik inteligentní mechanismy, aby poznaly, že se něco v preferencích uživatele změnilo. Uživatelské preference se v průběhu času mohou měnit a to může mít za následek, že se uživatelský model může stát velmi nepřesným (Germanakos, 2016).

2.2.2 Implicitní metody sběru uživatelských dat

Uživatelské modely mohou být také dynamicky generovány na základě implicitních informací, jako je například navigační chování uživatelů. Mechanismy používané ke konstrukci těchto modelů jsou uživatelsky transparentní a nevyžadují žádné další úsilí.

Nejběžnějším zdrojem implicitních informací o uživatelích je jejich historie prohlížení, ze které se extrahují zájmy uživatelů. Historie prohlížení obsahuje URL adresu navštívené stránky, datum a čas návštěvy. Z tohoto by tedy mohly být získány smysluplné informace o počtu návštěv jednotlivých stránek a čas strávený na těchto stránkách. Informace o historii prohlížení mohou být shromažďovány dvěma způsoby: Uživatelé sdílí jejich prohlížečím mezipaměti (cache) v pravidelných intervalech nebo uživatelé mají nainstalovaný proxy server, který funguje jako brána k internetu, čímž se zachytí všechny webové činnosti generované uživatelem. První technika využívá systémovou mezipaměť webového prohlížeče, která ukládá historii procházení webu uživatele. Tato technika nevyžaduje žádnou instalaci specifického softwaru. Nicméně za účelem získávání smysluplných informací ze shromážděných údajů musí uživatel ukládat mezipaměť pravidelně. V druhé technice proxy servery umožňují snadné zachycení informace bez jakéhokoliv zatížení pro uživatele, protože vyžadují pouze počáteční nastavení a nevyžadují žádný software, který by musel být instalován a později aktualizován na pracovní ploše uživatele. Důležitou nevýhodou však je, že žádný uživatelský model nemůže být vytvořen bez povolení k proxy serveru od uživatele.

Další přístup ke shromažďování implicitních informací, zatímco uživatel prochází interaktivní systém, je pomocí využití agentů (například plugin webového prohlížeče). Prohlížečím agenti mohou být nainstalováni na počítači uživatele a jsou schopni zachytit všechny činnosti uživatele během prohlížení. Na rozdíl od sbírání uživatelské historie prohlížení (tj. navštívené URL adresy) prohlížečím agenti přesně shromažďují informace o akcích provedených uživatelem na dané webové stránce, jako je vytváření záložek a stahování souborů. Na základě této dodatečné informace o prohlížečím činnosti uživatele může agent navrhnout zajímavé odkazy na aktuální stránce. Důležitou nevýhodou tohoto přístupu je, že vyžaduje nainstalování specializovaného softwaru (Germanakos, 2016).

2.2.3 Shrnutí explicitních a implicitních metod

Explicitní sběr dat obvykle vyžaduje aktivní účast uživatele (on-line průzkumy, registrační údaje, osobní či finanční údaje při nákupu). Nicméně mnoho personalizačních systémů využívá implicitní sběr dat, který používá clickstream nebo jiný typ dat s chováním

uživatele. V implicitním sběru dat se systém pokouší změřit zájem jednotlivců nebo skupin o objekt na základě heuristických ukazatelů (například čas strávený prohlížením položky).

Výhodou využití implicitního sběru pro tvorbu uživatelských profilů je to, že odstraňuje zátěž pro uživatele spojenou s poskytováním osobních informací. Systém sám shromažďuje relevantní údaje na základě pozorovaného chování uživatelů a vyvozuje informace specifické pro uživatele. Implicitní profilování znamená, že systém musí být schopen sledovat a monitorovat chování uživatelů za účelem identifikace při procházení nebo nákupních zvyklotech.

Implicitní údaje mohou být shromažďovány jak na straně klienta, tak na straně serveru. Klíčová odlišnost mezi těmito přístupy je šíře dat, která jsou k dispozici pro personalizaci systému. Na straně klienta jsou k dispozici pouze údaje o jednotlivých uživateli, a tím je i přístup možný pouze individuálně. Na straně serveru má podnik schopnost shromažďovat údaje o všech svých návštěvnicích, tím pádem mohou být použity jak individuální, tak spolupracující přístupy (Brusilovsky, 2007).

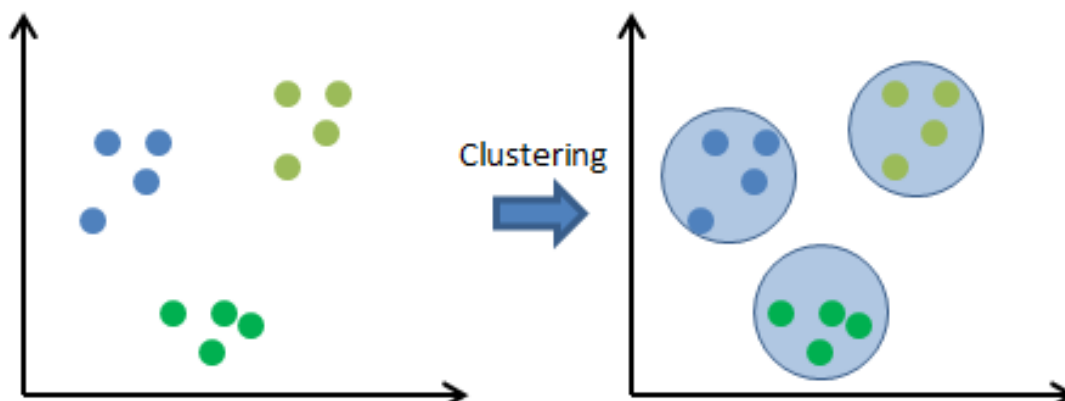
2.3 Vytváření uživatelského modelu

Nejjednodušším přístupem k vytváření uživatelského modelu je, že informace shromážděné od uživatele se používají tak, jak jsou v nezměněném stavu. Například uživatelé mohou explicitně vyjádřit svůj zájem o specifická témata týkající se zpravodajského publikačního systému, který bude pomocí jednoduchých mechanismů založených na pravidlech zobrazovat vybraná témata v horní části rozhraní uživatele. Inteligentnější přístupy k vytváření uživatelských modelů zahrnují případy, kdy je zachycena historie procházení uživatele, rozpoznání zákonitostí v uživatelských cestách a následné integrování do uživatelského modelu (Germanakos, 2016).

2.4 Shlukování (clustering)

Shlukování je bez dozorový proces, který spojuje uživatele do shluků (clusterů) na základě jejich podobné charakteristiky či navigačního chování (viz obrázek 2) a poté odvodí

specifické modely uživatelů z každého shluku. Algoritmy založené na shlukování obvykle začínají s počátečním souborem segmentů, které často obsahují jednoho náhodně vybraného uživatele. Algoritmy vytvoří vektory segmentů, podle nichž jsou uživatelé s největší podobností klasifikováni do jednotlivých segmentů (Germanakos, 2016).



Obrázek 2: Proces clusteringu
Zdroj: www.codeproject.com

2.5 Klasifikace

Klasifikace je kontrolovaný proces učení, který mapuje informace o uživateli (například interakční data) do jedné z několika předem určených tříd, které obvykle představují různé uživatelské modely. Hlavním cílem je pochopit stávající data a předpovídat, jak by se mohly chovat další případy. Pomocí klasifikace lze modelovat chování uživatele založené na předdefinovaných skupinách uživatelů. Mezi nejčastější klasifikační metody patří rozhodovací strom, Bayesovy klasifikátory a umělé neuronové sítě (Germanakos, 2016).

2.6 Asociační výsledky

Asociační pravidla vyhledávacích technik, jako je například Apriori algoritmus, byla zprvu určena ke sběru dat v supermarketových košících, ale od té doby byla použita v nejrůznějších doménách včetně sbírání internetových dat. Pravidlo asociace vzniká na

základě výsledků skupin předmětů stránek, které jsou běžně nakupované nebo procházené dohromady. Toto umožňuje webovým stránkám efektivněji uspořádat obsah webu nebo poskytovat efektivnější doporučení křížového prodeje produktů.

Taková pravidla mohou být také použita k optimalizaci struktury webu. Například pokud web neposkytuje přímou vazbu mezi dvěma stránkami A a B, získání pravidla A – je závislé na B by naznačovalo, že poskytuje přímý hypertextový odkaz, pomocí něhož by mohli uživatelé najít zamýšlené informace.

Jedním z problémů pro doporučující systémy na základě asociačních pravidel je, že systém nemůže poskytnout doporučení, pokud je datový soubor příliš řídký (což bývá častý případ Web usage miningu a kolaborativních filtrovacích aplikací), i větší položky často nesplňují minimální podporu omezení. Tímto problémem se zabýval Fu a jeho kolegové a navrhli dvě možná řešení tohoto problému. Prvním řešením je zařadit všechna zjištěná pravidla vypočtená podle míry průniku mezi levou stranu pravidel a uživatelskou aktivní relací a poté vygenerovat nejlepších K doporučení. Druhým řešením je využívání kolaborativního filtrování: systém najde „blízké sousedy“, kteří mají podobné zájmy jako cílový uživatel, a vytváří doporučení na základě historie blízkých sousedů.

V kolaborativním doporučujícím systému bylo představeno využití asociačních pravidel. Navrhovaný algoritmus pro získávání dat najde odpovídající počet pravidel pro každého cílového uživatele pomocí automaticky vybrané minimální podpory. Doporučující systém generuje asociační pravidla pro každého uživatele jak mezi uživateli, tak mezi předměty. Pokud má uživatel větší minimální podporu, než je prahová hodnota, systém generuje doporučení založené na uživatelské asociaci, jinak využívá asociaci mezi předměty. Vzhledem k tomu, že je obtížné najít odpovídající pravidlo předchůdce s plným uživatelským profilem (například celá relace ve Web usage miningu), asociačně založené doporučující systémy obvykle používají posuvné okno během aktivního profilu nebo relace cílového uživatele. Velikost tohoto okna se opakovaně snižuje až do té doby, než není nalezena přesná shoda s předchůdcem pravidla (Brusilovsky, 2007).

2.7 Sekvenční vzory

Dolování sekvenčních vzorů se zaměřuje na nalézání relevantních vzorů mezi datovými položkami, které jsou dodávány v pořadí. Dolování sekvenčních vzorů lze rozdělit do dvou široce zkoumaných metod:

- Deterministické metody, které se zaměřují na sledování navigačního chování uživatelů
- Stochastické metody, které představují přechod mezi webovými stránkami s cílem předvídat další návštěvu.

Nejpoužívanější stochastickou metodou je Markovův model, který zobrazuje přechody uživatelů na webových stránkách. Markovův model čerpá informace z aktuálního umístění a předchozího navigačního chování uživatele a na základě toho určuje skupiny uživatelů se stejnými nebo podobnými cestami (Germanakos, 2016).

2.8 Adaptační technologie

Hlavní schopností personalizačních systémů je přizpůsobení obsahu a doporučení takových položek, které nejpravděpodobněji uživatel hledá, to tedy znamená, že systém musí být schopen odvodit, co uživatel vyžaduje na základě předchozí nebo současné interakce s daným uživatelem a případně s dalšími uživateli. Úkol personalizace může být tedy vnímán jako problém predikce: systém se musí pokusit předpovědět zájem uživatele o danou položku, její užitečnost, specifické obsahové kategorie, stránky nebo předměty a všechny tyto předpovědi ohodnotit podle jejich predikované hodnoty (Brusilovsky, 2007).

Adaptační a personalizační systémy využívají specifické algoritmy, které rozhodují o tom, jaký druh přizpůsobení bude použit s ohledem na obsah funkcionalitu. V průběhu let byly navrženy různé technologie. Mezi nejvíce rozšířené a vyhledávané patří: uživatelské přizpůsobování, filtrování založené na pravidlech, obsahově založené filtrování, kolaborativní filtrování, Web mining, demograficky založené filtrování, agent technologie a shlukovací modely (Germanakos, 2016).

2.8.1 Uživatelské přizpůsobování

Uživatelské přizpůsobování nabízí mechanismus, který umožňuje uživatelům vytvořit vlastní reprezentaci rozhraní založené na jejich preferencích. Poté, co uživatel zadá potřebné informace, proces párování, jak je označován, vyhledá položky, které splňují zadaná kritéria a zobrazí je uživateli. V tomto případě není systém považován za adaptivní, ale adaptabilní, protože to, jak bude přizpůsoben obsah a funkcionality, je nakonfigurováno výhradně uživatelem.

Většina dnešních velkých poskytovatelů služeb nabízí několik uživatelských přizpůsobení jako součást jejich služeb. Například Gmail od společnosti Google umožňuje uživatelům nastavit hustotu zobrazení emailů (tj. „komfortní“ a „útulný,“ zobrazení na větších obrazovkách, a „kompaktní“, zobrazení na menších displejích). Gmail také nabízí mechanismus pro konfiguraci a přeskupení emailových kategorií (pomocí funkce drag-and-drop) nebo možnost vybrat si z předdefinovaných vizuálních témat (Germanakos, 2016).

2.8.2 Filtrování založené na pravidlech

Mechanismus založený na pravidlech je proces produkování vysoko úrovněových informací na základě nízko úrovněových metrik, které se týkají jak statických, tak dynamických kontextových informací uživatele.

Vzhledem k tomu, že dynamická část kontextového datového modelu může být aktualizována v reálném čase, je zřejmé, že podpora schopnosti rozhodování poskytuje přidanou hodnotu, která pomáhá uživatelům v různých úkolech. Taková pravidla pak mohou vyvolat systém automatických akcí nebo porovnat prediktivní model interakce s aktuálními daty o interakci uživatele shromážděných v reálném čase. Na základě toho tak systém poskytuje cenné poznatky o aktuálních cílech uživatele a zvyšuje efektivnost interakcí.

Například online bankovníctví může obsahovat pravidlo, že pokud uživatel zadá dvakrát špatně heslo, zobrazí se možnost pro kontaktování zákaznické podpory (Germanakos, 2016).

2.8.3 Obsahově založené filtrování

Obsahově založené filtrování navrhuje označování odkazů na základě analýzy obsahu stránky. Typický obsahově založený systém obsahuje následující kroky:


























- Předběžné načítání obsahu (prefetch) pod odkazy na aktuální prohlížené stránce
- Analyzování předběžně načtených stránek k vytvoření váženého vektoru klíčových slov ke každé stránce
- Porovnání vážených vektorů klíčových slov jednotlivých stránek s preferencemi uživatele, které jsou obvykle také popsány, využitím váženého vektoru klíčových slov
- Navrhnutí stránek, které mají stejné vektory jako preference uživatele

Tato technika je primárně charakterizovaná dvěma slabinami, obsahovým omezením a nadměrnou specializací. Obsahové omezení spočívá v metodách vyhledávání informací, které mohou být použity pouze na několik druhů obsahu, jako je například text a obrázky. Na druhé straně obsahově založené doporučující systémy poskytují doporučení pouze na základě uživatelských modelů, tudíž uživatelé nemají možnost objevování nových věcí, které nejsou podobné těm položkám, které mají ve svých modelech, a to může vést k nadměrné specializaci (Germanakos, 2016).

2.8.4 Kolaborativní filtrování

Kolaborativní filtrování využívá doporučení lidí, kteří něco zažili (například čtení knihy, sledování filmu atd.), k doporučení jiným lidem. Systém kolaborativního filtrování je založen na předpokladu, že pokud uživatelé X a Y ohodnotí podobně n předmětů nebo mají podobné chování při různých aktivitách (například nakupování, sledování něčeho), budou mít společné zájmy (Germanakos, 2016).

Tabulka 1: Příklad kolaborativního filtrování

Zdroj: vlastní zpracování podle gerardnico.com

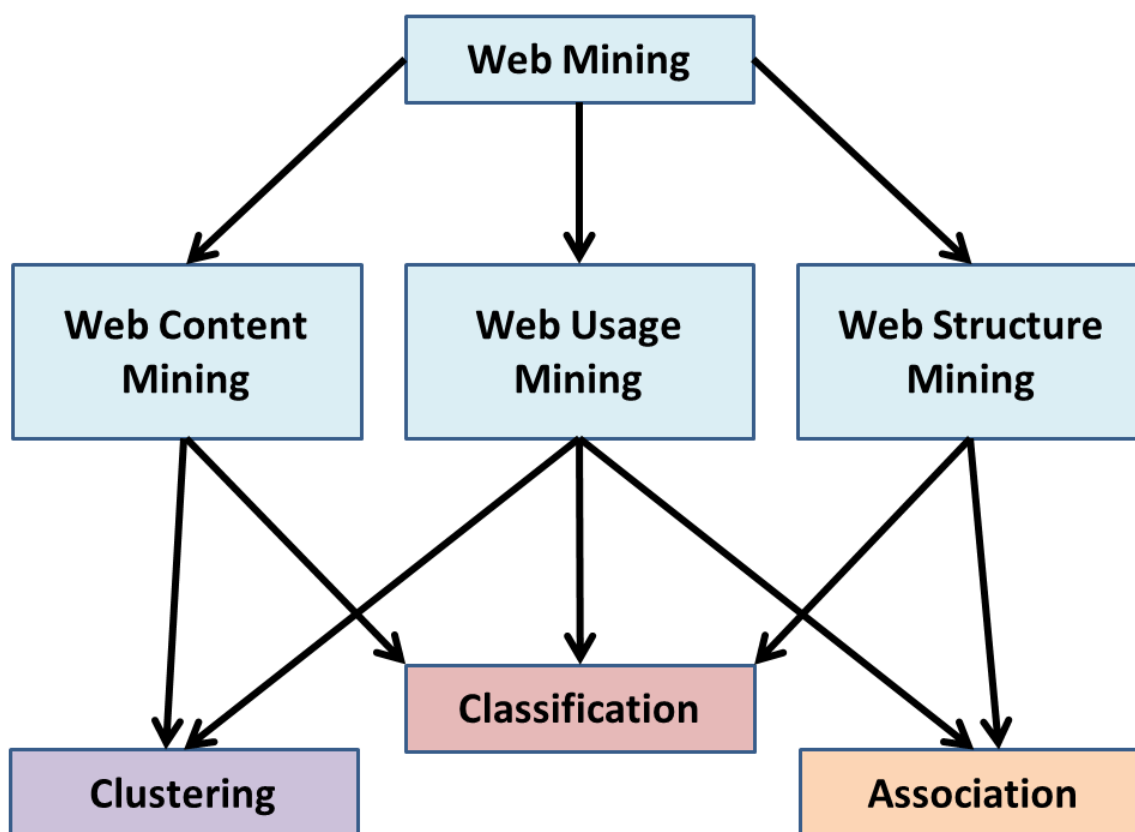
Adaptační a personalizační systémy využívají kolaborativního filtrování zejména k doporučování odkazů, o které by mohl mít uživatel zájem na základě dřívějších hodnocení, nebo podle chování podobných uživatelů (viz tabulka 1). Překonává nevýhody obsahově založeného filtrování, protože k doporučení obvykle nepoužívá samotný obsah položky. Nicméně mezi nevýhody kolaborativního filtrování patří:

- Algoritmus je nepraktický na velké datové soubory (může snížit kvalitu doporučení)
- Uživatelský vstup může být zkreslený
- Explicitní hodnocení od uživatelů nemusí být k dispozici
- Modely mohou být statické, a díky tomu se mohou stát rychle zastaralé
- Nemožnost doporučit nové položky, které dosud nebyly hodnoceny jinými uživateli
- Nefunguje příliš dobře s uživatelem, který není dostatečně podobný s jinými existujícími uživateli (Germanakos, 2016).

2.8.5 Web mining

Web mining zahrnuje techniky dolování dat s cílem identifikovat modely z webových systémů. Web mining je rozdělen do tří hlavních kategorií (viz obrázek 3):

- Obsahový web mining, ve kterém jde o extrakci a integraci dat a znalostí z obsahu webové stránky
- Strukturní web mining, který se zaměřuje na analýzu uzlů a spojovací struktury webové stránky
- Web usage mining, který se zabývá sbíráním užitečných informací o interakci uživatelů ze serverových protokolů (například zjistit, co uživatel na dané webové stránce hledá), (Germanakos, 2016).



Obrázek 3: Dělení web miningu
Zdroj: commons.wikimedia.org

Web usage mining primárně souvisí s adaptací a personalizací. Tento proces používá statistické a dolovací techniky na data ze serverových protokolů, na základě čehož potom vznikají soubory užitečných modelů, které ukazují navigační chování uživatelů.

Celkový proces webové personalizace založené na využití Web usage miningu se skládá ze tří fází: příprava dat a transformace, rozpoznání vzoru a doporučení. Z nich pouze druhá fáze probíhá v reálném čase (Brusilovsky, 2007).

Fáze přípravy dat transformuje "syrové" záznamy webových souborů do uživatelského profilu nebo do webové datové transakce, která může být zpracována pomocí data mining tasků. Tato fáze zahrnuje také integraci dat z různých zdrojů, jako jsou backend databáze, aplikační servery nebo obsah webových stránek (Brusilovsky, 2007).

Ve fázi rozpoznání vzoru mohou být použité různé techniky data miningu, jako je clustering, získávání společných pravidel, sekvenční rozpoznávání vzoru a pravděpodobnostní modelování. Výsledky z dolovací fáze jsou transformovány do celkových uživatelských modelů vhodných pro použití ve fázi doporučení. Poté doporučující systém poskytne aktivnímu uživatelskému profilu personalizovaný obsah na základě rozpoznávaných vzorů (Brusilovsky, 2007).

Hlavní výhody web usage miningu jsou shrnuty v následujících bodech:

- Modely jsou dynamicky získávány z uživatelských vzorů, tím nedochází ke snižování hodnoty systému v průběhu času.
- Použití samotného obsahu jako způsobu získávání agregačních modelů může mít za následek chybějící důležité vztahy mezi webovými objekty založenými na jejich použití. Tím pádem web usage mining sníží potřebu pro získávání subjektivních hodnocení od uživatelů nebo zjišťování preferencí při registraci uživatele.
- Modely jsou založeny na objektivních informacích (jak uživatelé skutečně využili webovou stránku).
- Není potřeba explicitní uživatelské hodnocení nebo interakce s uživateli (šetří čas a další komplikace).
- Podporuje zachování soukromí pomocí efektivního využití anonymních dat.

- Jsou zachycena data, která obsahově založené přístupy vynechávají.
- Může zvýšit účinnost kolaborativních a obsahově založených filtrovacích technik.

Nicméně pokud je k dispozici málo údajů týkajících se používání některých objektů, může být využití v personalizaci problematické (Germanakos, 2016).

2.8.6 Demograficky založené filtrování

Demograficky založené filtrování doplňuje jiné adaptační technologie, jako je filtrování na základě pravidel a kolaborativní filtrování s cílem vylepšit výsledek personalizace. Určité demografické údaje o uživateli, jako jsou například věk, pohlaví, profese atd., mohou být použity k odvození zájmu uživatele a poté k doporučení konkrétních předmětů. Tato metoda využívá demografické údaje k identifikaci typů uživatelů, kteří preferují určitý předmět, a k identifikaci jednoho shluku z několika již existujících, ke kterému uživatel patří. Cílem je přizpůsobit doporučení na základě informací o ostatních v tomto shluku (Germanakos, 2016).

2.8.7 Agent technologie

Agenti jsou procesy, které se zaměřují na plnění úkolů pro své uživatele, hrají roli osobních asistentů. Agenti obvykle řeší běžné problémy, které mohou uživatele na internetu zažít, jako je historie, zkratky nebo sledování webových stránek. Některé z hlavních charakteristik agentů můžeme rozlišit podle jejich použitých schopností nebo podle úkolů, které vykonávají. K prvně jmenovaným patří charakteristiky jako inteligence, samospráva, sociální kapacity a mobilita, zatímco druhý typ agentů lze kvalifikovat jako informační agenti, průzkumovní agenti, doporučující agenti, agenti pro elektronický trh nebo agenti pro správu sítě (Germanakos, 2016).

2.8.8 Shlukovací modely

Shlukovací modely jsou využívány především v oblasti elektronické komerce (E-Komerce) a mohou být charakterizovány jako doporučující E-Komerce algoritmy. Slouží k hledání

zákazníku, kteří jsou stejní s uživatelem, shlukovací metody rozdělí zákazníky do mnoha segmentů. Cílem algoritmu je přiřadit uživatele do segmentu, který obsahuje co nejvíce podobné zákazníky. Na základě nákupů a hodnocení zákazníků z toho segmentu pak systém generuje doporučení. Segmenty jsou obvykle vytvořeny pomocí shlukování nebo jiných bez dohledových algoritmů, ačkoli některé aplikace používají ručně určené segmenty (Germanakos, 2016).

Shlukování si klade za cíl rozdělit datovou sadu do skupin nebo seskupení, v nichž jsou podobnosti uvnitř shluku maximalizovány a zároveň podobnosti mezi jednotlivými shluky minimalizovány. Metody shlukování lze rozdělit do tří kategorií (Brusilovsky, 2007).

- Metody rozdělení, které vyvářejí K oddílů dané datové sady, kde každý oddíl představuje shluk. Nejrozšířenější metodou je k-means algoritmus.
- Hierarchické metody využívají buď přístup shora dolů (tzv. dělicí), nebo přístup zdola nahoru (tzv. aglomerativní) k vytvoření hierarchie shluků. Dělicí metody vycházejí z celého datového souboru jako jednoho shluku a rekurzivně rozdělují tato data. Aglomerativní metody vycházejí z jednotlivých předmětů jako shluků a opakovaně spojují menší shluky.
- Metody založené na modelu, který se snaží najít nejlepší shodu mezi datovými body a matematickým modelem, bývají specifikované jako rozdělení pravděpodobnosti.

Například algoritmus nazvaný PageGather byl použit k objevení významné skupiny stránek založené na uživatelských přístupových vzorech (Brusilovsky, 2007).

2.9 Personalizace založená na historii vyhledávání

Vyhledávací dotazy uživatelů jsou nepochybně důležitým zdrojem při rozpoznávání informačních potřeb a přizpůsobení interakce mezi člověkem a počítačem (HCI). Vyhledávač je schopen veškeré tyto informace zpracovávat neinvazivním způsobem, to znamená bez instalace externích proxy serverů nebo desktopových agentů, a proto lze přizpůsobit výsledky dotazu na základě předchozích žádostí a zájmů. Pro identifikaci jednotlivých uživatelů mohou být použity jednoduché přihlašovací formuláře nebo cookies a relativní clickstreamová data místo komplexu heuristických metod založených na

IP adresách, časů posledních přístupů nebo dat z uživatelských agentů, které stejně nelze pokládat jako zcela přesné (Brusilovsky, 2007).

Přístupy založené na historii vyhledávání mohou být rozděleny do dvou skupin. Offline přístupy, které využívají informace v před zpracovatelským kroku, obvykle analyzují vztahy mezi dotazy a navštívenými dokumenty. Online přístupy zachycují tato data, jakmile jsou k dispozici, ovlivňují uživatelské modely a poskytují personalizační výsledky s ohledem na poslední interakce uživatele. Přestože online přístupy poskytují aktuální návrhy, offline přístupy mohou zpracovávat daleko složitější algoritmy, protože v offline přístupech je obvykle méně naléhavé časové omezení než v online přístupech (Brusilovsky, 2007).

2.10 Lidské faktory v personalizaci

Snaha zavést lidské individuální rozdíly do různých návrhů adaptačních a personalizačních systémů pomocí vytváření uživatelských modelů a adaptačních mechanismů, které budou schopny regulovat zpracovávání informačních faktorů ve prospěch jednotlivého uživatele, je náročným směrem vzhledem k uživatelsky orientovaným rozhraním. Takováto snaha je brzděna především skutečností, že existuje pouze omezené množství zkušeností s tím, které vlastnosti jsou ve webové interakci nejdůležitější.

Zkoumání emocí je daleko složitější a nepředvídatelné zejména v důsledku obtížné identifikace, extrakce a interpretace mnoha faktorů a událostí. Kromě toho je nedostatek shody mezi definicí emoce a různorodosti reakcí jednotlivců v konkrétních situacích, tím pádem je téměř nemožné generalizovat, či vytvořit více deterministické přístupy výzkumu. Chování a reakce určité osoby na určitý stimul může být řízena psychofyzilogickými změnami, osobnostními rysy, subjektivními zkušenosti a podobně, které vedou k různým vyjádřením a široké škále těžko formulovatelných souvisejících aktivit.

Interaktivní systémy vyžadují vizuální i slovní (čtení) zpracovávání informací, které ovlivňují vysoko úroňové kognitivní faktory lidského mozku, jako jsou kognitivní styly způsobu, jakým si jednotlivci zapamatovávají informace. Lidé s omezenými schopnostmi v základních kognitivních procesech, jako je například provozní kapacita mozku, mohou

mít s těmito systémy potíže. Například generování nadměrného obsahu během interakcí nebo zvýšení hladiny emočního vzrušení může vést k nerovnovážným situacím (Germanakos, 2016).

2.10.1 Vizualní vnímání

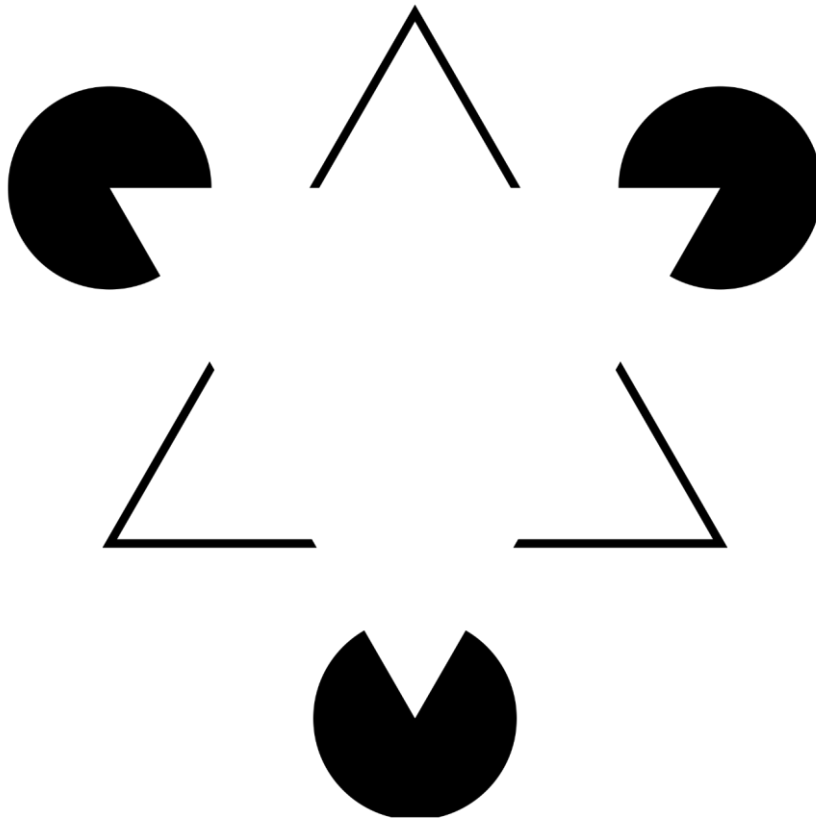
V každodenním životě je vizuální vnímání velmi důležité, neboť poskytuje jednotlivcům mnoho možností: volně se pohybovat, rozeznat ostatní lidi či předměty, číst knihy, identifikovat hloubku a vzdálenost nebo sledovat videa a filmy. Vizuální vnímání, stejně jako pojmy vizuální vyhledávání nebo vizuální pozornost, jsou jedněmi z nejdůležitějších faktorů, které ovlivňují oblast HCI (Human-computer interaction), neboli interakce člověka s počítačem.

Jak je známo, nejčastějším komunikačním médiem mezi člověkem a počítačem je obrazovka zařízení, kde uživatel optickým kanálem vnímá informace, které mu jsou skrze zařízení prezentovány. Použití grafiky na obrazovce počítače umožňuje reprezentaci obrazů, které připomínají skutečné objekty a vyvolávají stejné rozpoznávací procesy jako v reálném světě (Germanakos, 2016).

Gestaltismus (Gestalt psychologie), neboli tvarová psychologie, se mimo jiného zabývá tím, jak jednotlivci vnímají vizuální komponenty a jak je organizují. Gestaltova teorie přispívá především ke konceptu vnímání seskupení, pokouší se rozumět schopnostem jednotlivců a snaží se získat a udržet smysluplné vnímání. Gestaltovy principy jsou shrnuty následovně:

- Princip vzdálenosti – objekty, které jsou k sobě blíže, vytváří skupiny
- Princip podobnosti – objekty, které jsou podobné tvarem nebo barvou patří do stejné skupiny
- Princip uzavření – objekty, které nejsou úplné, mají tendenci být vnímány jako úplné (naše mysl vyplní vizuální mezery)
- Princip kontinuity – prvky objektů, které jsou zarovnané uvnitř objektu, mají vlastnost být seskupeny a vnímány jako integrovaný celek

- Princip společného osudu – prvky, které směřují stejným směrem nebo stejnou rychlostí jsou vnímány jako prvky společné skupiny.
- Princip symetrie – prvky objektů mezi symetrickými hranicemi vytváří souvislé tvary (Germanakos, 2016).



Obrázek 4: Příklad gestaltismu - Kanizsův trojúhelník
Zdroj:coopervision.cz

Tyto principy mají přímé použití v HCI a návrhu interaktivních hyper mediálních systémů, protože jejich použitelnost a přijatelnost ze strany uživatelů je závislá značnou měrou na tom, zda: usnadňují seskupení podobných objektů či tvarů, v případě chybějících informací spojují detaily tvarů nebo obsah na základě rozumných způsobů, poskytují smysluplné popisy o aktivitách, používají metafory, stejné termíny na podněty, nabídky, udržují stejnou posloupnost akcí v podobné situaci atd. (Germanakos, 2016).

3 Analýza současného stavu personalizačních možností vybrané webové aplikace

Tato kapitola je rozdělena na dvě podkapitoly. V první podkapitole je popsána obsahová stránka vybrané aplikace spolu s využívanými technologiemi. Druhá část je pak věnována SWOT analýze z pohledu personalizačních možností.

3.1 Obsah, využití a cíle aplikace

Pro analýzu současného stavu personalizačních možností a cílení obsahu byla vybrána B2C webová aplikace www.skoda-auto.pl. Jedná se o moderní web obsahující modelovou prezentaci nabídky Škoda Auto a odkazující se na specializované nástroje pro zákazníky (např. aplikace pro sestavení vlastní konfigurace vozu, vyhledávání prodejců, kalkulačka pro financování vozů apod.).

Hlavním cílem této webové aplikace je získání a udržení zájmu zákazníka o značku Škoda Auto a zejména získání kontaktů na nové zákazníky (především formou odeslání formulářů, např. žádost o testovací jízdu).

Sekundárním cílem je získání informací o chování návštěvníků webu pro další analýzu a tvorbu marketingové strategie. K tomuto účelu je do webu integrována webová analytika (Google Analytics) a vlastní poskytovatel anonymních identifikátorů uživatelů na webových a mobilních aplikacích Škoda Auto. Sběr těchto dat o návštěvnících je podmíněn jejich souhlasem s využíváním cookies.

Webová aplikace, včetně systému pro správu obsahu, je realizována vlastním řešením na straně vývojového týmu Škoda Auto za využití technologií .NET (na straně CMS), jQuery, Azure a Bootstrap CSS (na straně prezentační vrstvy).

3.2 SWOT z pohledu personalizačních možností

Pro lepší představu a přehlednost jsem sestrojil SWOT analýzu vybrané webové aplikace.

Tabulka 2: SWOT analýza webové aplikace *www.skoda-auto.pl*

	Pozitivní	Negativní
Interní	Silné stránky	Slabé stránky
	<ul style="list-style-type: none">- Rychlost načítání- Responzivní design- Unikátní obsah pro určité uživatele- Customer's journey	<ul style="list-style-type: none">- Jeden obsah pro všechny uživatele- Nevyužívání získaných informací
Externí	Příležitosti	Hrozby
	<ul style="list-style-type: none">- Cílení marketingových kampaní- Zlepšení zkušenosti zákazníka	<ul style="list-style-type: none">- Ztráta rychlosti načítání- Konkurenční weby- Poškození značky- Legislativní změny

Zdroj: vlastní

3.2.1 Silné stránky

Rychlost načítání stránek (i přes načítání velkého objemu dat)

Z pohledu personalizace tedy existuje prostor pro další logiku bez omezení rychlosti webu.

Responsivní design, tj. přizpůsobení prezentace jednoho obsahu pro různá rozlišení.

Z pohledu personalizace to přináší tu výhodu, že bude nutné spravovat jen jeden obsah ke konkrétním segmentům zákazníků.

Unikátní obsah určený pro určité segmenty uživatelů

Na webu jsou prezentovány např. novinky a promo akce cílené na určité skupiny uživatelů (např. Motorsport akce, promo akce "výměna letních pneu" pro vlastníky vozů apod.).



Zimowa runda WRC

Rajdowe Mistrzostwa Świata, a wraz z nim zespół SKODA Motorsport, w lutym przenoszą się na mroźne i zaśnieżone odcinki specjalne Rajdu Szwecji. Po fantastycznym początku sezonu i zwycięstwie w klasyfikacji WRC 2 za sprawą Andreeasa Mikkelsena w Rajdzie Monte-Carlo, ekipa z bazą w Młade Bolesław powalczy o kolejne cenne punkty w drodze do obrony tytułu mistrzów świata.

[Dowiedz się więcej...](#)

Obrázek 5: Ukázka novinky pro uživatele ze segmentu Motorsport akce

Zdroj: www.skoda-auto.pl



Obrázek 6: Ukázka novinky pro uživatele ze segmentu vlastníci vozů
Zdroj: www.skoda-auto.pl



60 lat modelu 450. Narodziny ikony.

Wprowadzenie przez SKODĘ na rynek w 1955 roku modelu 440, zwanego potocznie Spartakiem, okazało się ogromnym sukcesem. W Czechosłowacji po samochod ustawiły się gigantyczne kolejki, a ilość rynków eksportowych rosła w tempie wykładniczym. Miaz przedwojennej techniki Populara z modnymi kształtami nadwozia i wnętrza oraz solidnym montażem, okazał się prawdziwym strzałem w dziesiątkę.

[Dowiedz się więcej...](#)

Obrázek 7: Ukázka novinky pro uživatele ze segmentu veteráni
Zdroj: www.skoda-auto.pl

Obsah pro každou část "zákaznické cesty" (customer's journey)

Informační architektura webu pokrývá obsah (či na něj naviguje do jiné aplikace) pro téměř všechny fáze zákaznického procesu:

- Ve fázi výběru vozu ke koupi či pronájmu má zákazník k dispozici graficky bohatou modelovou prezentaci, ceníky základních konfigurací, navigaci do aplikace pro sestavení vozu či výpočet financování vozu.
- Ve fázi těsně před koupí zákazník většinou navštíví konkrétního prodejce, příp. využije kontaktní formulář na webu.
- Ve fázi po koupi vozu zákazník může na webu najít obsah či navigování na obsah vhodný pro vlastníky vozů (servisní informace, palubní literatura apod.).

3.2.2 Slabé stránky

Jeden obsah pro všechny uživatele

Na webu sice existuje obsah určený pro určité skupiny uživatelů, ale uživatel musí tento obsah sám najít, což se nemusí vždy za dobu jeho návštěvy podařit.

Nevyužitý potenciál nasbíraných dat o návštěvnících webu

O návštěvnících webu se v současném stavu sbírají data, na základě kterých se vytvářejí pouze statistiky používání stránek a jejich jednotlivých částí (zpětná analytika). Identifikace anonymního uživatele je však sledována napříč více aplikacemi Škoda Auto a tato data nejsou aktuálně systematicky využívána pro personalizaci (pouze jako jednorázové cílení obsahu emailových a reklamních kampaní, které však vychází z velice jednoduchých statistik).

3.2.3 Příležitosti

Cílení marketingových kampaní v rámci webu

Pod touto příležitostí se skrývá především vytvoření si znalosti o zákaznících a jejich preferencích na základě nasbíraných dat nejen z vybrané webové aplikace, ale i z jiných systémů. S využitím těchto znalostí se poté snažit o vytváření cílených marketingových kampaní ke spokojenosti zákazníků.

Zlepšení uživatelské zkušenosti zákazníka

K naplnění této příležitosti by měla směřovat celá podstata personalizace. Uživatel by měl rychle najít to, co hledá, o co má zájem. Na základě toho by se měla snížit pravděpodobnost, že zákazník odejde nespokojený.

3.2.4 Hrozby

Ztráta rychlosti načítání stránek

S personalizací vyvstává i otázka, zda neovlivní rychlost načítání webové stránky. Velká část procesu webové personalizace probíhá online, a jestliže by se obsah webové stránky načítal příliš dlouho, mohlo by to vést ke ztrátě potenciálních zákazníků.

Rychlejší konkurence

Personalizace se v současné době stává velmi populární, a pokud by došlo k promeškání zásadní inovace webu, mohla by nastat situace, kdy konkurence získá nové, potřebné znalosti, mohli bychom říci "know-how" o zákaznících, a to by mohlo vést nejen ke ztrátě potenciálně nových zákazníků, ale i těch stávajících.

Legislativní změny

Možnou hrozbou mohou být i legislativní změny, jako je například Obecné nařízení o ochraně osobních údajů pod zkratkou GDPR (General Data Protection Regulation). GDPR je nová legislativa Evropské unie, která byla přijata již minulý rok v dubnu 2016,

ale v platnost přichází až 25. května 2018. Měla by přinést zatím největší revoluci v ochraně osobních údajů a pomoci proti neoprávněnému zacházení s daty a osobními informacemi občanů EU (GDPR, 2016).

Poškození značky

K poškození značky může dojít vlivem neadekvátního personalizovaného obsahu. Personalizace je vnímána jako velký pokrok v tvorbě webového obsahu, neznamena to ale, že je potřeba personalizaci přehánět a snažit se o personalizaci všeho. Pokud by takováto situace nastala, došlo by tím k přílišné personalizaci, což bychom mohli nazvat jako "přepersonalizování", pro uživatele by to pak bylo zavádějící.

4 Návrh využití personalizace, možná řešení a jejich zhodnocení

Tato kapitole je rozdělena na tři podkapitoly, v první podkapitole jsou stručně popsány jednotlivé byznys případy možností využití personalizace. Druhá podkapitola je věnována možným řešením a závěr této kapitoly je zhodnocením daných řešení.

4.1 Možnosti využití personalizace

Při prvním zobrazení webové aplikace se přímo nabízí personalizace úvodního obrázku. Obrázek by se mohl měnit na základě informací o uživateli a jeho preferencích. Například pokud by uživatel v minulosti již stránky navštívil a v aplikaci pro sestavení vozu by vybral své preference, byl by na úvodním obrázku zobrazen vůz podobný tomu, který byl nakonfigurován.

Dalším příkladem by mohlo být zařazení uživatele do segmentu na základě jeho historie vyhledávání. Uživatel je ve věkové kategorii 18- 22 let, předpokládáme tedy, že se jedná o začínajícího řidiče, který bude mít nejpravděpodobněji zájem o model Fabia nebo Citygo.

Pokud se posuneme stránkou o kousek níž, narazíme zde na panel novinek. V personalizaci se můžeme řídit i pomocí ročních období. Blíží se jaro, a pokud uživatel již vlastní automobil, bude zajisté potřebovat přezout zimní pneumatiky na letní.

Dalším příkladem využití personalizace může být propagování např. střešních nosičů kol uživateli, který v předchozích interakcích projevoval zájem o jízdní kola (např. si prohlížel stránku o cyklistickém závodě sponzorovaném Škoda Auto).

Poslední částí úvodní obrazovky je seznam plánovaných akcí. Časové řazení akcí je jednou možností, ale s personalizací to nemá nic společného. Jako možnost se tedy nabízí zobrazit nadcházející akce například podle lokace uživatele. Není potřeba se snažit uživatele nalákat na akci, která se koná 200km od jeho bydliště, ale spíše ho upozornit na akce v jeho okolí.

Další možností personalizace by mohlo být zobrazování určitých částí stránky jen v případě splnění určitých podmínek. Například na základě zjištěných údajů předpokládáme, že uživatel žije v nižší příjmové třídě, měli bychom mu tedy promovat nabídku s možnostmi financování. Podobně by to mohly být například před vyplnění finanční atributy ve finanční kalkulačce, či zobrazování nabídky s aktuálními vozy na skladě podle preferencí uživatele.

Jednou z posledních možností je řazení odkazů v menu podle pravděpodobného zájmu uživatele nebo podle jeho nejčastějších navštívených záložek. Poslední možností by mohly být vyskakovací formuláře při opuštění stránky, či formuláře se speciálními nabídkami určitých modelů nebo doplňkových služeb.












Modele ŠKODY

Wybierz samochód, który będzie jak członek Twojej rodziny i który najlepiej odpowiada Twoim potrzebom.

Przejdź do modelu ŠKODY

 CITIGO Doskonale do miasta, idealny na co dzień.	 FABIA Wytenczość i styl dla rodziny i praktyczny na co dzień.	 RAPID Wytenczość i styl dla rodziny i praktyczny na co dzień.	 OCTAVIA Lider rynku, idealny do użytku rodzinnego i w firmie.
 YETI TOUR Kompaktowy SUV o wielu zastosowaniach.	 SUPERB Lśniący wizerunek, elegancki w klasie biznes.	 KODIAQ Duży SUV luboju wyposażenia. Pomieści nawet 7 pasażerów.	

 NOWY PORTAL ROWEROWY WELOVECYCLING.PL		 NOWA ŠKODA KODIAQ
 WIOSENNE WYZWANIE - OPON WYMIENIANIE	 NOWA ŠKODA CITIGO	 INFORMACJA DLA KLIENTÓW
 SAMOCHOZY UŻYWANE	 NOWA ŠKODA RAPID	
 ELEKTRYCZNA ŠKODA VISION E		 NOWA ŠKODA RAPID

OSTATNIE NEWSY



Auto Shanghai 2017: światowa premiera pierwszego w pełni elektrycznego samochodu ŠKODY i prezentacja strategii elektromobilności

Strategia elektromobilności i elektryfikacji węgla samochodowej w Šanghaju. 19-23 kwietnia w Šangaju światowa premiera pierwszego w pełni elektrycznego samochodu ŠKODY i prezentacja strategii elektromobilności.

[Czytaj więcej...](#)

Jak zmieniły się wymagania dotyczące obowiązkowych systemów bezpieczeństwa stosowanych w samochodach?

Temat bezpieczeństwa nie sposób jest duży, szczególnie w kontekście nowych przepisów dotyczących bezpieczeństwa pojazdów. Wymagania dotyczące obowiązkowych systemów bezpieczeństwa stosowanych w samochodach. Jak zmieniły się wymagania dotyczące obowiązkowych systemów bezpieczeństwa stosowanych w samochodach?

[Czytaj więcej...](#)

[Wszystkie newsy](#)

POBIERZ APLIKACJĘ ŠKODY



 ZAPYTAJ O OFERTĘ Zgłoś się po więcej informacji o interesującym modelu.	 JAZDA PRÓBNA Próbny jazda, spotkanie z ekspertem.	 ZNAJDŹ DEALERA Znajdź, gdzie jest najbliższy Salon ŠKODA.	 KONFIGURATOR Sprawdź, jak może wyglądać ŠKODA.
Więcej	Więcej	Więcej	Więcej

Informacje prasowa Praktyczny kontakt Regulamin serwisu Polityka prywatności	ŠKODA dla firm Oferta dla firm	Tvoja ŠKODA Twój pojazd Jaką próbę konfiguracji Sprawdź ofertę Newsletter	Usługi i serwis Opcje finansowania Umowa wzięcia w serwisie Złoty Akcja przynależności Technologia obsługi	Pozostaw w komentarzu do ŠKODY Kontakt z ekspertem Kontakt z ekspertem
--	--	--	--	---

Obrázek 8: Náhled webové aplikace www.skoda-auto.pl
 Zdroj: www.skoda-auto.pl

4.2 Návrhy možných řešení

V této kapitole jsou popsány dva základní přístupy k řešení zavedení personalizace a cílení obsahu na webu Škoda Auto:

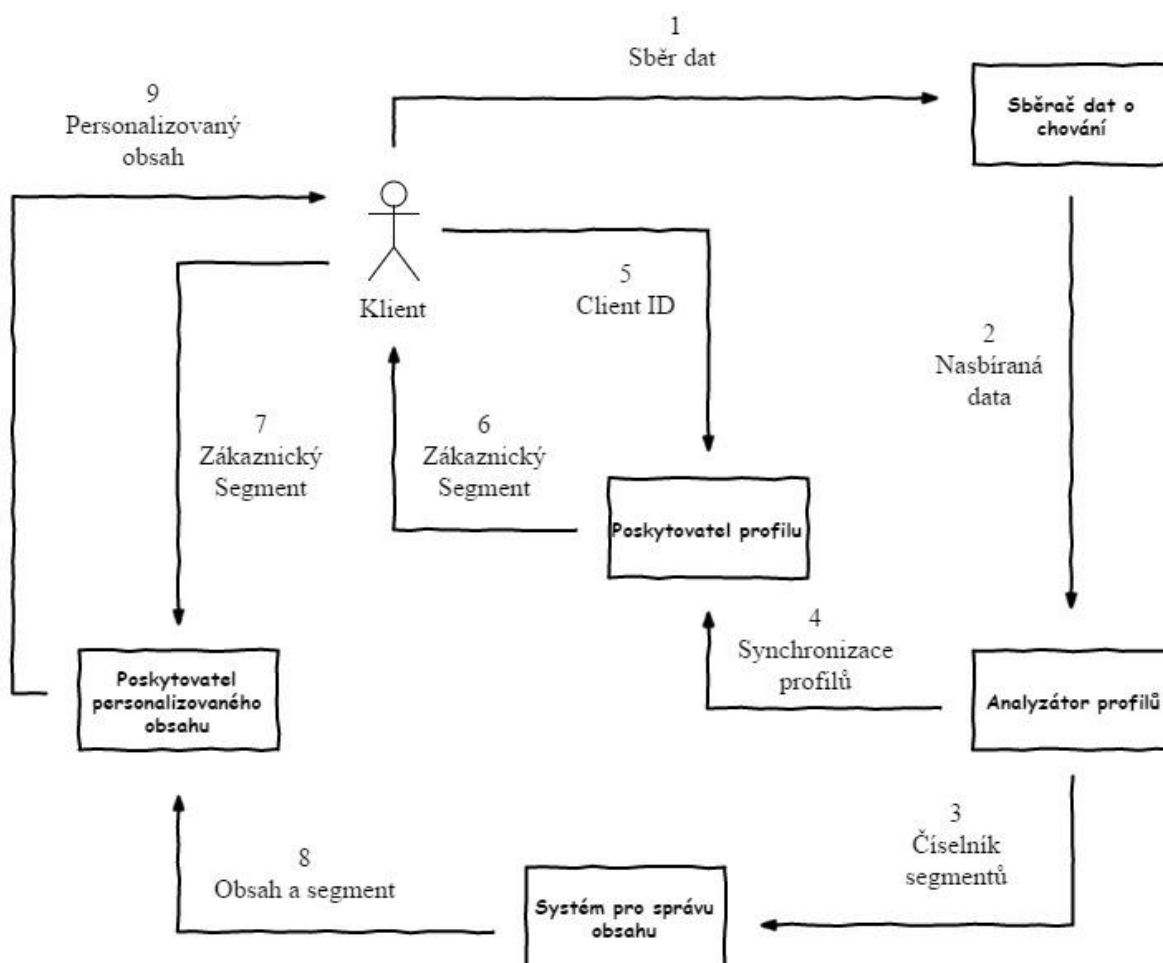
- Vývoj vlastního řešení, jehož předností je možnost vytvořit řešení "šité na míru" potřebám prostředí Škoda Auto a možnost integrace s dalšími interními systémy Škoda Auto, které mohou přinášet vstupy do personalizačního mechanismu (např. vstupy ze CRM), nebo mohou využívat výstupy tohoto řešení (např. napojení personalizačních mechanismů pro emailové kampaně apod.). Přináší tak možnost unifikovaného přístupu k řešení personalizace napříč systémy Škoda Auto.
- Nákup hotového řešení, jehož předností je využití "best practices" v oblasti personalizace a rychlé zavedení řešení do webu Škoda Auto.

Cílem této kapitoly je diskutovat o základních výhodách a nevýhodách těchto přístupů, porovnat je a zhodnotit. Výstup této kapitoly je pak možné použít jako vstup pro další analýzu a rozhodování.

4.2.1 Vývoj vlastního řešení personalizace

Jedná se o vlastní řešení s využitím některých komponent od třetích stran. Vyvíjení veškerých komponent vlastní cestou by bylo neefektivní, časově i nákladově velmi náročné, a proto jsou využívány některé komponenty od třetích stran. Mluvíme tu například o Google Analytics pro sběr dat, Azure Stream Analytics pro analýzu dat v reálném čase nebo o nejrůznějších technologiích strojového učení (machine learning).

Pro lepší představu celého návrhu bylo vytvořeno schéma vlastního řešení personalizace.



Obrázek 9: Schéma procesu vlastního řešení

Zdroj: vlastní

Schéma je velmi zjednodušeno a obsahuje pouze hlavní logické komponenty procesu, které jsou popsány v následujících podkapitolách. Šipky mezi jednotlivými komponenty představují datové toky, číslování je pouze pro lepší orientaci, neznamená pořadí datových toků.

Přijímač dat o chování

Celý proces začíná právě sběrem potřebných anonymních dat (viz datový tok 1) o uživateli (klientovi) zmiňovaném v teoretické části této práce. Můžeme si to představit například takto: uživatel pod ID 465 klikl na vůz Octavie.

Možným nástrojem pro sběr uživatelských dat by mohla být například webová analytika Google Analytics, která je v současné době na webu integrována. Všechna tato nasbíraná data jsou následně stažena analyzátozem profilů (viz datový tok 2), kde proběhne analýza dat.

Analyzátor profilů

Je jedním z nejdůležitějších prvků celého procesu. Stahuje data z přijímače dat, analyzuje je a následně z nich odvozuje jednotlivé segmenty klientů. Tyto segmenty jsou přiřazeny klientům, a tak vznikají zákaznické profily.

Přiřazování klientů do segmentů probíhá vyhodnocováním vybraných charakteristik zákazníka na základě pravidel, která jsou určena marketingovým plánem společnosti, případně pomocí systému využívajícího technologie strojového učení. Například systém dokáže sám pochopit, že se uživatel nachází ve stavu "před koupí" na základě jeho chování. Navíc o uživateli ví, že se jedná o klienta s rodinou a pomocí těchto informací zařadí uživatele do segmentu.

Systém pro správu obsahu

Do tohoto systému je editorem vkládán obsah. Pokud je tento obsah vhodný pro určitou skupinu lidí (segment), tak ho na základě rozhodnutí editora systém označí, tj. přiřadí k němu číslo z číselníku segmentů, který sem byl poskytnut z analyzátoru profilů. Takto označený obsah (vazba: obsah + segment) je posléze stažen poskytovatelem personalizovaného obsahu (viz datový tok 8).

Poskytovatel profilu

Poskytovatel profilu by mohl být charakterizován jako "prostředník" mezi analyzátozem profilů a klientem. Poskytovatel profilu pravidelně, např. jednou denně, provede synchronizaci s analyzátozem profilů, tj. stáhne nové profily s informací o segmentech, či zaktualizuje ty stávající (viz datový tok 4). Klient se poté poskytovatele profilu zeptá, k jakému segmentu patří (viz datový tok 5), a poskytovatel mu pošle zpět zákaznický

segment (viz datový tok 6), který získal předem z analyzátoru profilů a se kterým uživatel pokračuje k poskytovateli personalizovaného obsahu (viz datový tok 7).

Poskytovatel personalizovaného obsahu

Stahuje segmentovaný obsah ze systému pro správu obsahu (viz datový tok 8). Tento obsah je prezentován klientovi na základě zákaznického segmentu poskytnutého od klienta (viz datový tok 7). Zjednodušeně řečeno - systém obdrží zákaznický segment a vrátí specifický obsah klientovi (viz datový tok 9).

Klient

Pracuje s webem a vědomě či nevědomě poskytuje osobní informace přijímači dat (viz datový tok 1). Na základě zákaznického segmentu (profilu) mu je následně poskytován personalizovaný obsah (viz datový tok 9).

4.2.2 Nákup hotového řešení personalizace

Na trhu softwarových produktů jsou nabízena nejrůznější hotová řešení pro personalizaci a cílení marketingových kampaní (např. Sophus, IgnitionOne nebo Evergage).

Speciálně Sophus nabízí komplexní řešení personalizace webových aplikací především pro automobilové společnosti. Mezi nejznámější klienty patří například BMW, Mercedes, VW, Renault, Mazda a mnoho dalších známých automobilových společností. Služby Sophusu využívají nejmodernější data z prostředí automobilového průmyslu, která kombinují spolu s chytrým dynamickým algoritmem k získání co nejvíce potenciálních zákazníků.

Tato hotová řešení jsou založena na analýze dat, která vede k tvorbě adaptivního samo-učícího se algoritmu. Na základě interakce uživatele s webovou aplikací pak systém vyhodnocuje pozici uživatele k zákaznické cestě a nabídne relevantní cílený obsah, který ho většinou má motivovat k vykonání nějaké akce (tzv. CTA prvek = call to action), a tak se zvyšuje pravděpodobnost konverze uživatele v zákazníka (kontakt).

Např. uživateli se zobrazí vyskakovací okno ve chvíli, kdy začne hledat cestu, jak opustit webovou stránku. Takové okno většinou obsahuje nějaký CTA prvek a snaží se o to, aby uživatel na stránce zůstal a zkusil si například nakonfigurovat své vlastní auto.

Jiným příkladem může být možnost živé diskuze (live chat), kde systém předem vyhodnotí, zda by aktuální uživatel mohl mít zájem o takovouto komunikaci, či nikoliv. Na základě tohoto vyhodnocení se následně uživateli zobrazí, či nezobrazí tato možnost.

4.2.3 Časová, nákladová a riziková analýza variant řešení

V této kapitole je uvedena prvotní analýza variant řešení navrhovaných výše z pohledu jejich časové a nákladové náročnosti a rizik, která s sebou přináší. Cílem je stanovit hrubý projektový rámec každé varianty, který by mohl být podkladem pro další rozhodování v tématu personalizace a který by měl být dále rozpracován ve studii proveditelnosti. Z pohledu bakalářské práce se jedná o nutný vstup pro celkové zhodnocení variant uvedený v další kapitole.

- Z pohledu analýzy časové a nákladové náročnosti jsou definovány:
- klíčové aktivity, které daná varianta obnáší a musí být zohledněny v projektovém plánu
- klíčové zdroje, které musí být v projektu zajištěny

Dále uvedené odhady časové náročnosti jsou pouze orientační a je potřeba je v případě další analýzy a práce na projektu revidovat. Slouží pouze pro možnost indikativního odhadu náročnosti řešení.

Časová analýza je hodnocena u obou variant z pohledu standardního rozdělení projektu do etap, které následují jedna po druhé:

- Globální analýza a návrh
- Detailní analýza a návrh
- Implementace řešení
- Zavedení řešení

Nákladová analýza je založena pouze na vydefinování časové náročnosti (v MD = man days) a rolí, kterých se tato náročnost týká. Není proveden převod na finanční částku, protože sazby společnosti za jednotlivé role nemohou být uváděny. Předmětem této analýzy není ani stanovení pořizovacích a provozních nákladů spojených s nákupem nových strojů a licencí, neboť nelze bez detailní analýzy a návrhu řešení toto nyní ani hrubě určit a bylo by pro porovnání spíše zavádějící.

Pro účely zhodnocení časové a nákladové náročnosti obou řešení jsou tato hlediska dostatečným vstupem pro porovnání v závěrečné kapitole.

Z pohledu analýzy rizik jsou pak definována hlavní projektová rizika s pravděpodobností jejich výskytu, dopadu a možná preventivní opatření pro zamezení jejich vzniku.

Vlastní vývoj řešení

Vlastní vývoj řešení klade velké nároky na kompetence a zapojení interních zaměstnanců z byznysu i IT oddělení. Ve fázi analýzy a návrhu řešení bude nutné sladit potřeby personalizace s celkovou marketingovou a informační strategií společnosti za větší účasti architektů a vlastníka projektu. Řešení musí být navrženo s dostatečnou vizí do budoucna, protože se jedná o strategický systém digitální komunikace, musí být rozděleno na logické komponenty, které musí respektovat technologická a bezpečnostní omezení. Samotná implementace bude zahrnovat pravděpodobně orchestraci několika vývojových týmů, které budou jednotlivé komponenty vyvíjet a testovat. Součástí testů by měly být i integrační a výkonnostní testy, které prokáží robustnost řešení. Ve fázi zavedení systému pak je potřeba počítat se spoluprací provozního oddělení, kterému řešení bude předáváno.

V tabulkách níže je pro představu uvedena hrubá odhadovaná časová náročnost jednotlivých fází dle potřebných rolí a jejich alokací, z toho je odvozen hrubý projektový plán a dále jsou uvedena projektová rizika.

	GAN		DAN		IM		ZA		Celková náročnost (MD)
	Alokace	Náročnost (MD)	Alokace	Náročnost (MD)	Alokace	Náročnost (MD)	Alokace	Náročnost (MD)	
Byznys analytik	300%	60	100%	20	50%	10	20%	4	94
Byznys architekt	100%	30	20%	4	0%	0	0%	0	34
Byznys vlastník	20%	4	20%	4	20%	4	20%	4	16
Dokumentarista	0%	0	0%	0	50%	10	200%	40	50
IT analytik	100%	20	300%	100	150%	30	50%	10	160
IT architekt	50%	10	100%	30	20%	4	0%	0	44
IT vedoucí	20%	4	20%	4	100%	40	20%	4	52
Projektový manažer	50%	10	50%	10	20%	4	20%	4	28
Test analytik	20%	4	100%	40	100%	20	50%	10	74
Test manažer	20%	4	20%	4	100%	20	20%	4	32
Tester	0%	0	0%	0	600%	120	200%	40	160
Vývojář	0%	0	0%	0	1000%	400	300%	60	460
Celková náročnost (MD)		146		216		662		180	1204

Obrázek 10: Hrubý odhad potřeby lidských zdrojů pro variantu vlastního řešení
Zdroj: vlastní

	t+1M	t+2M	t+3M	t+4M	t+5M	t+6M	t+7M	t+8M	t+9M	t+10M
Globální analýza a návrh (GAN)										
Definice hlavních funkcí a mimofunkčních požadavků										
Vytvoření aplikační architektury										
Definice systémů a subsystémů										
Sběrač dat o chování										
Analýzátor profilů										
Poskytovatel profilu										
Systém pro správu obsahu										
Poskytovatel personalizovaného obsahu										
Definice rozhraní mezi systémy a okolím										
Schvalování a akceptace analýzy a návrhu										
Detailní analýza a návrh (DAN)										
Vytvoření technologické architektury										
Podrobná funkční a datová analýza										
Podrobná analýza rozhraní mezi systémy										
Test analýza										
Schvalování a akceptace analýzy a návrhu										
Implementace řešení (IM)										
Vývoj nových komponent řešení										
Úprava existujících komponent řešení										
Testování										
Funkční testy										
Výkonnostní testy										
Integrační testy										
Akceptační testy										
Tvorba dokumentace										
Školení uživatelů										
Zavedení řešení (ZA)										
Instalace a konfigurace technického vybavení (produkční prostředí)										
Instalace a konfigurace databázového a aplikačního SW										
Zkušební provoz										
Vyhodnocení zkušebního provozu										
Předání řešení do rutinního provozu										

Obrázek 11: Hrubý plán pro variantu vlastního řešení
Zdroj: vlastní

Rizikový faktor	Míra výskytu	Dopad rizika (stupeň)	Dopad rizika (popis)	Ohodnocení rizika	Preventivní opatření
Nedostatečná znalost interních lidí v problematice personalizace	2	3	Návrh řešení, které bude náchylné k chybám a neustálým změnám.	6	Zajištění externích odborných konzultantů, příp. odborné vzdělávání interních zaměstnanců v problematice personalizace.
Složitost schvalovacího procesu	2	2	Zpoždění fáze analýzy a návrhu.	4	Zajištění pravidelných kontrolních dnů se všemi schvalovateli v průběhu analýzy.
Nedostatečně naplánované lidské zdroje	2	3	Nedodržení termínů z důvodu nedostatečné kapacity pracovníků.	6	Průběžná revize kapacit lidských zdrojů
Technické problémy	2	2		4	Zajištění dostatečné časové rezervy v implementační a zaváděcí fázi.
Nedbalost pracovníků	2	2	Neefektivnost práce lidí na projektu, zátěž na financování projektu.	4	Pečlivý výběr lidí na projektu, zajištění motivace týmu a neustálý dohled nad plněním povinností.
Koordinace jednotlivých týmů a oddělení	2	2	Zpoždění celého projektu (např. jeden tým bude záviset na výstupech jiného týmu).	4	Obsazení seniorního projektového manažera, který zná dobře vazby jednotlivých týmů a oddělení, a který bude pravidelně sledovat a reportovat závislosti mezi týmy a odděleními.

Obrázek 12: Projektová analýza pro variantu vlastního řešení

Zdroj: vlastní

Nákup hotového řešení

Tato varianta zahrnuje hlavně aktivity spojené s pečlivým výběrem hotového řešení, který by měl být zahájen zmapováním trhu, stanovením kritérií a vah pro rozhodování a následně by měl být následován vypsáním výběrového řízení a provedením alespoň dvou kol výběru, než dojde k rozhodnutí. V detailní analýze a návrhu bude pak nutné vydefinovat nutné úpravy spojené s integrací hotového řešení, a to jak na straně samotného vybraného produktu (přizpůsobení řešení speciálním potřebám společnosti), tak i na straně již existujících systémů, na které integrace může mít dopad.

V oblasti lidských zdrojů bude zejména zapotřebí oproti předchozí variantě role externího konzultanta, resp. zástupce hotového řešení, který pomůže s integrací řešení. Mělo by tedy být zde uplatněno know-how zvenčí, a nároky na interní zaměstnance proto nebudou tak vysoké.

V tabulkách níže je pro představu uvedena hrubá odhadovaná časová náročnost jednotlivých fází dle potřebných rolí a jejich alokací, z toho je odvozen hrubý projektový plán a dále jsou uvedena projektová rizika.

	GAN		DAN		IM		ZA		Celková náročnost (MD)
	Alokace	Náročnost (MD)	Alokace	Náročnost (MD)	Alokace	Náročnost (MD)	Alokace	Náročnost (MD)	
Byznys analytik	200%	40	100%	20	50%	10	20%	4	74
Byznys architekt	100%	20	20%	4	0%	0	0%	0	24
Byznys vlastník	20%	4	20%	4	20%	4	20%	4	16
Dokumentarista	0%	0	0%	0	50%	10	100%	20	30
Externí konzultant	0%	0	100%	20	100%	20	50%	10	50
IT analytik	100%	20	200%	40	50%	10	20%	4	74
IT architekt	50%	10	50%	10	10%	2	0%	0	22
IT vedoucí	20%	4	20%	4	50%	10	10%	2	20
Projektový manažer	50%	10	50%	10	20%	4	20%	4	28
Test analytik	10%	2	50%	10	50%	10	20%	4	26
Test manažer	20%	4	20%	4	100%	20	20%	4	32
Tester	0%	0	0%	0	200%	40	50%	10	50
Vývojář	0%	0	0%	0	200%	40	50%	10	50
Celková náročnost (MD)		114		126		180		76	496

Obrázek 13: Hrubý odhad potřeby lidských zdrojů pro variantu nákupu hotového řešení
Zdroj: vlastní

	t + 1M	t + 2M	t + 3M	t + 4M	t + 5M
Globální analýza a návrh (GAN)					
Definice funkčních a mimofunkčních požadavků					
Zmapování trhu hotových řešení pro personalizaci					
Stanovení kritérií a vah pro výběrové řízení					
Hrubý výběr hotových řešení					
Hodnocení hrubého výběru hotových řešení					
Detailní analýza a návrh (DAN)					
Jemný výběr hotových řešení					
Podrobná analýza nutných přizpůsobení hotového řešení					
Podrobná analýza integrace hotového řešení do současné architektury					
Test analýza					
Odsouhlasení si výběru hotového řešení					
Domluvení SLA					
Implementace řešení (IM)					
Provedení vývojových a konfiguračních úprav					
Testování					
Tvorba dokumentace					
Školení uživatelů					
Akceptace řešení, podpis SLA, zajištění nákupu					
Zavedení řešení (ZA)					
Instalace a konfigurace produkčního prostředí					
Zkušební provoz					
Vyhodnocení zkušebního provozu					
Předání řešení do rutinního provozu					

Obrázek 14: Hrubý plán pro variantu nákupu hotového řešení
Zdroj: vlastní

Rizikový faktor	Míra výskytu	Dopad rizika (stupeň)	Dopad rizika (popis)	Ohodnocení rizika	Preventivní opatření
Nespolehlivost dodavatele řešení	1	3	Nedodržení termínu projektu.	3	Mít připravené záložní dodavatele.
Nevhodnost hotového řešení	2	2	Dodané řešení, které nebude splňovat všechny požadavky.	4	Jasná definice požadavků, rozpracovaných do detailů.
Integrace s vnitřními systémy Škoda Auto	3	3	Nemožnost integrace řešení s určitými systémy z důvodu ochrany interních (firemních) dat	9	Detailní analýza technického řešení, zapojení oddělení pro infrastrukturu a bezpečnost.
Problémy s komunikací	2	2	Špatná nebo žádná komunikace, která vede ke špatnému definování požadavků k řešení.	4	Domlouvat jak osobní, tak elektronickou komunikaci v pravidelných intervalech.
Prozrazení firemních údajů dodavateli	1	3	Poškození značky, možné právní problémy v budoucnu.	3	Dostatečné proškolení interních zaměstnanců v jednání s dodavatelem.

Obrázek 15: Projektová analýza pro variantu nákupu hotového řešení

Zdroj: vlastní

4.3 Zhodnocení variant řešení

Z analýzy uvedené v praktické části vyplynulo, že každý přístup k řešení má své výhody a nevýhody a záleží jen na managementu společnosti, jakou cestu zvolí:

- Vývoj řešení personalizace vlastními silami s sebou přináší na jedné straně výhody v budování vlastního unikátního know-how, kterým se může společnost odlišovat od konkurence, ale na straně druhé se jedná o časově náročné řešení (z analýzy vyplynul min. hrubý časový rámec jednoho roku).
- Nákup hotového řešení personalizace znamená nespornou výhodu ve využití již vyzkoušených metod a sdílených nákladů s jinými odběrateli, na druhé straně se může jednat o řešení, které bude vyžadovat změnu v procesech, strategii a funkcích stávajícího digitálního marketingu a které nemusí být žádoucí pro vedení společnosti.

Pro kvalifikované rozhodnutí bude ještě potřeba další analýza a diskuze mezi zúčastněnými stranami za účelem odvození detailní potřeby a požadavků na personalizaci ve vazbě na firemní strategii a cíle.

V tabulce č. 3 níže jsou uvedena hlediska pro porovnání obou variant řešení z předchozích kapitol s uvedením výhod a nevýhod.

Tabulka 3: Porovnání variant řešení

Porovnávací hledisko	Vlastní vývoj řešení	Nákup hotového řešení
Sběr dat	(+) Sběr dat pod vlastní kontrolou, možno uzpůsobit logiku sběru dat každému trhu v souladu s legislativou (-) Sběr dat probíhá pouze z vlastních webů.	(+) Sběr dat o chování uživatelů i z jiných (i konkurenčních) webů s podobným zaměřením (-) Poskytování nasbíraných dat třetí straně.
Analýza dat	(+) Algoritmy jsou šité na míru, přesně odpovídají potřebám podniku (-) Nutno implementovat vlastní algoritmy pro vyhodnocení dat, příp. najít vhodnou technologii na bázi strojového učení. (-) Absence vyzkoušených a fungujících technik.	(+) Již hotové a osvědčené algoritmy pro vyhodnocování dat (best practice) (-) Algoritmy jsou know-how dodavatele, jsou do jisté míry přizpůsobitelné potřebám zákazníka, ale není možné zcela přesně definovat logiku vyhodnocování.
Nasazení do dalších B2C aplikací	(+) Možnost bezproblémového rozšíření do dalších aplikací.	(+) Nezávislá implementace do dalších aplikací časově méně náročná (-) Integrace jednotlivých aplikací by mohla být problematická.
Vyhodnocování segmentů	(+) Možnost vyhodnocování segmentů podle vlastních kritérií (-) Chybějící znalosti, absence ověřených a funkčních technik.	(+) Segmenty jsou určeny dynamicky na základě předchozích výsledků a jejich vyhodnocování je neustále zlepšováno (-) Vyhodnocování lze v určitých mezích měnit, ale ne tak přesně jako u vlastního řešení.
Integrace s interními systémy (CRM apod.)	(+) Možnost integrace s jakýmkoliv interním systémem.	(-) Omezená integrace z důvodu ochrany firemních údajů
Podpora řešení	(+) Možnost okamžitého řešení.	(-) Závislost na dodavateli při řešení problémů (-) Možnost zániku dodavatele.
Time to market	(-) Dlouhá dodací doba (vytvoření koncepce, zajištění vlastního vývoje, testování), viz Obrázek 11 a Obrázek 14 (min. dvojnásobná časová náročnost oproti hotovému řešení).	(+) Řešení je možné dostat do provozu v řádu několika měsíců.
Náklady	(-) Vysoké náklady na analýzu, návrh, vývoj i údržbu (viz O).	(+) Nižší náklady než při vlastním řešení (využití úspory z rozsahu – hotové řešení používá mnoho klientů a dělí se o náklady).
Konkurence	(+) Vlastní řešení, které nemůže použít konkurence a může tak přinést nové příležitosti	(-) Řešení využívá i konkurence, a tak nemusí přinášet samotné řešení novou příležitost.
Bezpečnost dat	(+) Návrh uložení a práce s daty je plně v kompetenci společnosti, a odpovídá tak standardům přístupu práce s údaji.	(-) Riziko úniku informací.

Zdroj: vlastní

V rámci fáze globálního návrhu by mohly být zvažovány i tyto další možné varianty:

- kombinace obou přístupů (využití výhod z obou variant), tj. nákup vybraných komponent řešení a vývoj vlastních komponent, které by mohly přinést specifickou konkurenční výhodu
- vytvoření prototypu vlastního řešení za účelem analýzy a ověření vybraných vlastností a funkcí systému personalizace
- zkušební provoz hotového řešení, nabízí-li jej dodavatel, a jeho následné zhodnocení

Závěr

Předkládaná práce se zabývala tématem zavádění personalizace a cílení obsahu webové aplikace na základě zákaznického profilu. Práce měla za cíl popsání základních a pokročilejších pojmů z oblasti personalizace a adaptivity webových aplikací, analýzu současného stavu personalizačních možností a cílení obsahu ve vybrané B2C webové aplikaci Škoda Auto, spolu s návrhem jednotlivých možností řešení. Návrhy měly být následně podrobeny časové, nákladové a rizikové analýze. Závěrem práce mělo být zhodnocení navržených řešení.

Jednotlivých cílů bylo v průběhu práce dosaženo. V teoretické části práce byly popsány základní pojmy vztahující se k personalizaci a adaptivitě webových aplikací. Následně byla provedena analýza vybrané B2C webové aplikace Škoda Auto, která popisuje obsahovou stránku, využívané technologie a hlavní cíle této aplikace. Na základě této analýzy, byly navrženy dvě odlišné možnosti řešení, které by mohly být námětem diplomové práce.

U jednotlivých možností byl stanoven hrubý projektový rámec, který by mohl být podkladem při dalším rozhodování v tématu personalizace. Byly provedeny odhady časové náročnosti, nákladová analýza a byla definována hlavní projektová rizika s pravděpodobností jejich výskytu, dopadu a možných preventivních opatření. Následně bylo provedeno porovnání jednotlivých variant a vytvořeno zhodnocení návrhů.

Hlavním přínosem práce je pomoc při rozhodování společnosti v otázkách personalizace a cílení obsahu ve vybrané webové aplikaci. V praktické části bylo na základě analýzy zjištěno, že každá možnost řešení má své výhody a nevýhody. Vývoj řešení vlastními silami nabízí možnost tvorby vlastního know-how v oblasti personalizace, na druhou stranu jde o časově i nákladově velmi náročné řešení.

Nákup hotového řešení je naopak časově příznivější variantou, navíc je využíváno již vyzkoušených metod a sdílených nákladů s jinými odběrateli. Tyto metody ale nemusí být shodné s cíly společnosti a proto bude nutné dalších analýz a diskuzí mezi zúčastněnými stranami, než bude stanoveno kvalifikované rozhodnutí.

Seznam použité literatury

- BRUSILOVSKY, Peter., Alfred. KOBZA a W. NEJDL. The adaptive web: methods and strategies of web personalization. New York: Springer, c2007. Lecture notes in computer science, 4321. ISBN 35-407-2078-2.
- CASTELLANO, Giovanna a LAKHMI C. JAIN AND ANNA MARIA FANELLI (EDS.). Web personalization in intelligent environments. Online-Ausg. Berlin: Springer, 2009. ISBN 978-364-2027-949.
- CERI, Stefano, Florian DANIEL, Maristella MATERA a Federico M. FACCA. Model-driven development of context-aware Web applications. ACM Transactions on Internet Technology. 2007, 7(1), 2-es. DOI: 10.1145/1189740.1189742. ISSN 15335399. Dostupné také z: <http://portal.acm.org/citation.cfm?doid=1189740.1189742>
- Clustering using K-Means Algorithm [online]. Canada: Samir Kunwar, 2013 [cit. 2017-04-26]. Dostupné z: <https://www.codeproject.com/Articles/439890/Text-Documents-Clustering-using-K-Means-Algorithm>
- Collaborative filtering [online]. Netherland: Nicolas Gerard, 2016 [cit. 2017-04-27]. Dostupné z: https://gerardnico.com/wiki/data_mining/collaborative_filtering
- GDPR. Obecné nařízení o ochraně osobních údajů [online]. Praha: Mgr. Eva Škorníčková, 2016 [cit. 2017-04-23]. Dostupné z: <https://www.gdpr.cz/gdpr/>
- GERMANAKOS, Panagiotis. a Merios BELK. Human-centred web adaptation and personalization: from theory to practice. Švýcarsko: Springer, 2016. ISBN 978-3-319-28048-6.
- Optické klamy [online]. Česká Republika: Cooper Companies, 2015 [cit. 2017-04-27]. Dostupné z: <https://coopervision.cz/blog/optick%C3%A9-klamy>
- PODBORSKÝ, Martin. Adaptivita webových aplikací a modelování uživatele. Brno, 2013. Mendelova univerzita v Brně. Vedoucí práce Bc. Ing. Roman Malá, Ph.D.
- SCHLEE, Christian. A Use Case Driven Analysis of Targeted Advertising Technologies in the ICT Space. Wiesbaden: Vieweg, 2012. ISBN 978-383-4823-953.

SWOT analýza: jak a hlavně proč ji sestavit [online]. Praha 5: Magdalena Čevelová, 2011 [cit. 2017-04-28]. Dostupné z: <https://www.cevelova.cz/proc-swot-analyza/>

The general relationship between the categories of Web Mining and objectives of Data Mining [online]. Russia, 2013 [cit. 2017-04-27]. Dostupné z: https://commons.wikimedia.org/wiki/File:The_general_relationship_between_the_categories_of_Web_Mining_and_objectives_of_Data_Mining.png

TINTAREV, Nava a Judith MASTHOFF. Evaluating the effectiveness of explanations for recommender systems. *User Modeling and User-Adapted Interaction*. 2012, 22(4-5), 399-439. DOI: 10.1007/s11257-011-9117-5. ISSN 0924-1868. Dostupné také z: <http://link.springer.com/10.1007/s11257-011-9117-5>

TSIHRINTZIS, George A. *New directions in intelligent interactive multimedia*. Berlin: Springer, c2008. ISBN 978-3-540-68126-7.

TZIVISKOU, Christina a Marco BRAMBILLA. Semantic personalization of web portal contents. *Proceedings of the 16th international conference on World Wide Web - WWW '07*. New York, New York, USA: ACM Press, 2007, , 1245-. DOI: 10.1145/1242572.1242788. ISBN 9781595936547. Dostupné také z: <http://portal.acm.org/citation.cfm?doid=1242572.1242788>

UCHYIGIT, G. a Matthew Y. MA. *Personalization techniques and recommender systems*. Hackensack, NJ: World Scientific, c2008. ISBN 978-981-279-701-8.

VÝBĚR A IMPLEMENTACE INFORMAČNÍHO SYSTÉMU PRO FIRMU. Brno, 2014. Bakalářská práce. VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ. Vedoucí práce Doc. Ing. MILOŠ KOCH, CSc.

Web Škoda Auto [online]. Polsko: ŠKODA Polska Volkswagen Group Polska Sp. z o.o., 2015 [cit. 2017-04-23]. Dostupné z: <http://www.skoda-auto.pl/>