

Univerzita Palackého v Olomouci

Přírodovědecká fakulta

Katedra geoinformatiky

**WEBOVÁ ANALYTIKA A MOŽNOSTI JEJÍHO
PROSTOROVÉHO VYHODNOCENÍ**

Bakalářská práce

Marek POSPÍŠIL

Vedoucí práce Mgr. Rostislav Nėtek, Ph.D

Olomouc 2019

Geoinformatika a geografie / Geoinformatika

ANOTACE

Tato bakalářská práce se zabývá webovou analytikou a prostorovým vyhodnocením nashromážděných dat. Cílem je zhodnotit možnosti pokročilých analytických nástrojů podle předem definovaných aspektů a určit silné a slabé stránky těchto nástrojů. Teoretická část obsahuje obecné informace o webové analytice, metodiku sběru internetových dat a popis vybraných analytických nástrojů. Praktická část se zaměřuje na hodnocení nástrojů dle technických, uživatelských, kartografických a analytických aspektů a jejich závěrečné vyhodnocení prostřednictvím metody heuristické analýzy.

KLÍČOVÁ SLOVA

webová analytika; prostorové vyhodnocení

Počet stran práce: 53

Počet příloh: 2

ANOTATION

This bachelor thesis is focused on the web analytics and spatial evaluation of collected data. The aim is to evaluate the possibilities of advanced analytical tools according to predefined aspects and to identify the strengths and weaknesses of these tools. The theoretical part contains general information on web analytics, methodology for collecting Internet data, and description of selected analytical tools. The practical part focuses on evaluation of tools according to technical, user, cartographic and analytical aspects and their final evaluation through the heuristic analysis.

KEYWORDS

web analytics; spatial evaluation

Number of pages: 53

Number of appendixes: 2

Prohlašuji, že

- bakalářskou/diplomovou práci včetně příloh, jsem vypracoval(a) samostatně a uvedl(a) jsem všechny použité podklady a literaturu. *(Např. ve své programové aplikaci jsem použil modul pro transformaci vektorových dat mezi prostorovými referenčními systémy, vytvořený)*

- jsem si vědom(a), že na moji bakalářskou/diplomovou práci se plně vztahuje zákon č.121/2000 Sb. – autorský zákon, zejména § 35 – využití díla v rámci občanských a náboženských obřadů, v rámci školních představení a využití díla školního a § 60 – školní dílo,

- beru na vědomí, že Univerzita Palackého v Olomouci (dále UP Olomouc) má právo nevýdělečně, ke své vnitřní potřebě, bakalářskou/diplomovou práci užívat (§ 35 odst. 3),

- souhlasím, aby jeden výtisk bakalářské/diplomové práce byl uložen v Knihovně UP k prezenčnímu nahlédnutí,

- souhlasím, že údaje o mé bakalářské/diplomové práci budou zveřejněny ve Studijním informačním systému UP,

- v případě zájmu UP Olomouc uzavřu licenční smlouvu s oprávněním užit výsledky a výstupy mé bakalářské/diplomové práce v rozsahu § 12 odst. 4 autorského zákona,

- použít výsledky a výstupy mé bakalářské/diplomové práce nebo poskytnout licenci k jejímu využití mohu jen se souhlasem UP Olomouc, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly UP Olomouc na vytvoření díla vynaloženy (až do jejich skutečné výše).

V Olomouci dne

Děkuji vedoucímu práce Mgr. Rostislavu Nétkovi, Ph.D za podněty a připomínky při vypracování práce. Zároveň chci poděkovat rodině, která mě podporovala.

UNIVERZITA PALACKÉHO V OLOMOUCI

Přírodovědecká fakulta

Akademický rok: 2016/2017

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Marek POSPÍŠIL**

Osobní číslo: **R15483**

Studijní program: **B1301 Geografie**

Studijní obor: **Geoinformatika a geografie**

Název tématu: **Webová analytika a možnosti jejího prostorového vyhodnocení**

Zadávací katedra: **Katedra geoinformatiky**

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

Cílem bakalářské práce je zhodnotit možnosti pokročilých webových analytických nástrojů s ohledem na možnosti prostorového vyhodnocení dat. Student provede rešerši nástrojů poskytující možnosti pokročilé webové analytiky (GoogleAnalytics, Piwik, ParseLy, KISSmetrics, Clicky, Mint), na jejichž základě sestaví otevřenou databázi analytických nástrojů. Zásadní bude jednak analýza sběru a implementace prostorových dat, jednak analýza využití a vizualizace dat z analytických nástrojů v prostoru, resp. vizualizačních metod v mapě. Student zohlední vliv typu prostorové lokalizace (GPS, IP adresa) a možností vyplývajících z technologických možností (mobil/tablet vs. počítač). Student definuje a sestaví vhodná hodnotící kritéria, na jejichž základě provede vlastní hodnocení ověřených nástrojů na základě předem definovaných metrik (heuristik - z pohledu technického, kartografického, geoinformatického, uživatelského, apod). Student dále provede vyhodnocení dle objektivních i subjektivních parametrů, a na případových studiích definuje silné a slabé stránky prostorového vyhodnocení u jednotlivých nástrojů a vyvodí adekvátní závěry.

Student vyplní údaje o všech datových sadách, které vytvořil nebo získal v rámci práce, do Metainformačního systému katedry geoinformatiky a současně vytvoří zálohu údajů ve formě validovaného XML souboru. Celá práce (text, přílohy, výstupy, zdrojová a vytvořená data, XML soubor) se odevzdá v digitální podobě na CD (DVD) a text práce s vybranými přílohami bude odevzdán ve dvou svázaných výtiscích na sekretariát katedry. O diplomové práci student vytvoří webovou stránku v souladu s pravidly dostupnými na stránkách katedry. Práce bude zpracována podle zásad dle Voženílek (2002) a závazné šablony pro diplomové práce na KGI. Povinnou přílohou práce bude poster formátu A2.

Rozsah grafických prací: dle potřeby
Rozsah pracovní zprávy: max. 50 stran
Forma zpracování bakalářské práce: tištěná

Seznam odborné literatury:

Athanasiadis, E. I.Spyrou, G. M.: A Web-based database of studies with network-driven data exploration tools, Springer-Verlag Berlin Heidelberg 2004.
A Stewart Fotheringham, Peter A Rogerson: The SAGE Handbook of Spatial Analysis, 2008.

Clifton B. : Advanced Web Metrics with Google Analytics, 2008.

Kaushik A. Successful Analytics Gain Business Insights By Managing Google Analytics, 2015.

Luc Anselin, Yong Wook Kim, Ibnu Syabri: Web-based analytical tools for the exploration of spatial data, 2004.

Voženílek, Vít. Diplomové práce z geoinformatiky. Univerzita Palackého, 2002.

Vedoucí bakalářské práce: Mgr. Rostislav Néték, Ph.D.
Katedra geoinformatiky

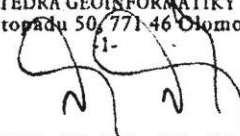
Datum zadání bakalářské práce: 10. května 2017

Termín odevzdání bakalářské práce: 10. května 2018

prof. RNDr. Ivo Frébort, CSc., Ph.D.
děkan

L.S.

UNIVERZITA PALACKÉHO V OLOMOUCI
PŘÍRODOVĚDECKÁ FAKULTA
KATEDRA GEOINFORMATIKY
17. listopadu 50, 771 46 Olomouc


prof. RNDr. Vít Voženílek, CSc.
vedoucí katedry

V Olomouci dne 15. května 2017

OBSAH

ÚVOD	7
1 CÍLE PRÁCE	10
2 METODY A POSTUPY ZPRACOVÁNÍ.....	11
2.1 Použitá data.....	11
2.2 Metodika sběru dat.....	11
2.2.1 Značkování stránek (Page-Tagging).....	11
2.2.2 Serverové logy (Web Server Logfiles).....	12
2.3 Sběr prostorových dat.....	12
2.4 Heuristická analýza	13
2.5 Použité programy	14
2.6 Postup zpracování.....	14
3 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY	15
3.1 Webová analytika.....	15
3.2 Analytické nástroje	16
3.3 Základní metriky webové analytiky	18
3.4 GDPR v souvislosti s webovou analytikou.....	20
4 DEFINICE ASPEKTŮ	21
4.1 Technické aspekty	21
4.2 Kartografické aspekty	21
4.3 Uživatelské a vizualizační aspekty.....	22
4.4 Analytické aspekty.....	22
5 ASPEKTY ANALYTICKÝCH NÁSTROJŮ	23
5.1 Google Analytics	23
5.2 Yandex Metrica	29
5.3 Woopra	32
5.4 Clicky Analytics	34
6 VÝSLEDKY HEURISTICKÉ ANALÝZY	38
6.1 Hodnocení Google Analytics	38
6.2 Hodnocení Yandex Metrica.....	38
6.3 Hodnocení Woopra.....	39
6.4 Hodnocení Clicky Analytics	39
7 PŘÍPADOVÉ STUDIE	40
8 VÝSLEDKY.....	43
8.1 Doporučení.....	43
8.2 Návrh vzorového řešení	46
9 DISKUZE	48
10 ZÁVĚR.....	49
POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE	8
PŘÍLOHY	11

SEZNAM POUŽITÝCH ZKRATEK

Zkratka	Význam
API	Application Programming Interface
CSV	Comma-separated values
EU	Evropská unie
GA	Google Analytics
GDPR	General Data Protection Regulation
GIS	Geografický informační systém
GPS	Global Positioning System
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
ID	Identifikátor
IP	IP adresa
ISP	Internet Service Provider
JSON	JavaScript Object Notation
LAU	Local Administrative Unit
NUTS	Nomenclature of Units for Territorial Statistics
OS	Operační systém
SDK	Software Development Kit
WGS 84	World Geodetic System 1984

ÚVOD

Obrovský počet webových stránek, stejně jako obecně vznik internetu v minulém století, předznamenal nové, doposud neznámé technologie pro analýzu obrovského množství internetových dat. Webová analytika je rozsáhlou problematikou s velice širokým polem působnosti, kterou využívají stovky milionů stránek po celém světě. V analytickém kontextu hovoříme o nástrojích, prostřednictvím kterých jsou shromažďována důležitá data o webových stránkách a jejich uživatelích. Webová analytika zkoumá tři obecné páky, které mají podniky k dispozici na internetu: marketing, merchandising a design webových stránek. Obchodníci často využívají těchto služeb k tomu, aby zjistili, jak dobře (nebo špatně) vynakládají své peníze (Peterson, 2005).

Pracovníci v oblasti marketingu využívají těchto služeb dennodenně, neboť jim pomáhají například zvyšovat zisk pomocí lepšího pochopení chování uživatelů v konkrétním webovém projektu. Mezi jednu z mnoha užitečných funkcí patří mimo jiné mapové výstupy shromážděných dat, jež pomáhají vizualizovat rozložení různých jevů v prostoru. Na základě těchto informací analytici zjišťují geografické oblasti, z nichž zákazníci nejvíce či nejméně navštěvují webové stránky, a zároveň předpovídají, které produkty budou zákazníci nejvíce a nejméně nakupovat v budoucnu. Například internetové obchody pak mohou pozorovat, zda splnily předem stanovené cíle, a pokud ne, mohou pomocí analytického nástroje najít důvod, proč se tomu tak nestalo.

Tato práce si klade za cíl provést analýzu jak využití statistických metod, tak hlavně analýzu využití prostorových dat a kartografických metod, čemuž se doposud nevěnovala žádná jiná odborná práce. Je důležité, aby tak účinný nástroj, jakým webová analytika bezesporu je, podával uživatelům nezkreslené prostorové informace. Nevhodně zvolené vyjadřovací prostředky, měřítko nebo dokonce kompozice totiž mohou negativně ovlivnit naměřená data, a tím pádem znehodnotit jejich celkovou interpretaci.

1 CÍLE PRÁCE

Hlavním cílem bakalářské práce je seznámení s problematikou webové analytiky, která je v současnosti stále více potřebná zejména k marketingovým účelům, a následné zhodnocení možností vybraných webových analytických nástrojů s důrazem na prostorové vyhodnocení jimi naměřených dat. Cíl je realizován pomocí několika dílčích částí.

Teoretická část zahrnuje rešerši odborné literatury a odborných článků týkajících se trendů ve webové analytice, zejména metodikou sběru dat a popisem jednotlivých metrik, které jsou v analytickém kontextu nezbytné pro porozumění webové analytiky jako celku. Teoretická část rovněž popisuje jednotlivé aspekty, na základě kterých jsou vybrané analytické služby detailně popsány – z hlediska technického, kartografického, uživatelského a vizualizačního a analytického.

Praktická část obsahuje aplikaci měřicích kódů na vybrané studie, díky čemuž je možné porovnávat weby s nízkou, střední a vysokou návštěvností, viz kapitola Studie. Dalším krokem je grafický návrh vzorového (ideálního) řešení, jež bude sestaveno na základě nabytých znalostí, předchozího hodnocení a definování silných a slabých stránek služeb, které byly využity pro účely bakalářské práce.

2 METODY A POSTUPY ZPRACOVÁNÍ

2.1 Použitá data

V rámci bakalářské práce byla shromažďována data prostřednictvím měřicích kódů jednotlivých analytických nástrojů, které byly nasazeny na studie (resp. webové stránky), kam byl autorovi práce udělen přístup. Konkrétně se jedná o analytická data webů Katedry Geoinformatiky, Pevnosti poznání a Univerzity Palackého. Tato data jsou sbírána a následně vizualizována ve formě grafů, tabulek či mapových výstupů přímo v prostředí analytických nástrojů.

2.2 Metodika sběru dat

Kapitola základních metodik sběru internetových dat pomáhá rozlišit, který způsob je vhodný pro konkrétní řešení. Obecně mluvíme o několika typech, z nichž nejpoužívanějšími jsou značkování stránek pomocí JavaScript kódu a sběr dat prostřednictvím serverových logů. Pro sběr všech použitelných informací je doporučeno tyto způsoby kombinovat, neboť se vzájemně doplňují.

2.2.1 Značkování stránek (Page-Tagging)

Metoda značkování, též tagování, je reprezentována fragmentem skriptovacího jazyku JavaScript, který je vložen na každou sledovanou stránku. Data o webu jsou pak sbírána pomocí webového prohlížeče návštěvníka. Základem jsou tzv. cookies – malé množství dat vytvořené webovým serverem a uložené do počítače prostřednictvím prohlížeče. Cookies se bude podrobněji zabírat kapitola 2.3.3 Cookies.

Obrovským přínosem metody tagování je flexibilita, umožňující sledovat různé metriky v libovolném časovém horizontu (včetně zpracování v reálném čase), snadná implementace kódu a menší požadavky na správu dat. Nevýhodou je náchylnost na software blokující reklamy či jiná rozšíření podobného charakteru. Metodě značkování předchází zavedení výše zmíněného měřicího kódu, jenž bývá zpravidla umístěn do hlavičky stránky. Analytický nástroj generuje pro každou webovou stránku unikátní identifikační číslo (žlutě zvýrazněná část kódu), příkladem budiž měřicí kód nástroje Google Analytics, viz obrázek níže. K vizualizaci dat dochází ve vybraném analytickém nástroji, a to v podobě tabulek, grafů či různých (také mapových) výstupů.

```
<!-- Global site tag (gtag.js) - Google Analytics -->
<script async src="https://www.googletagmanager.com/gtag/js?id=UA-69522209-1"></script>
<script>
  window.dataLayer = window.dataLayer || [];
  function gtag(){dataLayer.push(arguments);}
  gtag('js', new Date());

  gtag('config', 'UA-69522209-1');
</script>
```

Obr. 1 Příklad měřicího kódu služby Google Analytics

2.2.2 Serverové logy (Web Server Logfiles)

Podle článku Martina Čecha (2010) se surová data shromažďují do tzv. log souborů, jež představují souhrn záznamů o všech přístupech na webové stránky skrze http a https protokoly. Data (logy) jsou shromažďována bez ohledu na prohlížeč uživatele a primárně jsou koncipována pro sledování technických informací, jejich použití pro obchodní účely je tedy neefektivní.) Výhodou metody je zabezpečení, kdy protokolovací (log) server může být osamostatněn od webové stránky. Tuto možnost využívají zejména banky, pojišťovny a jiné korporace, které pracují s citlivými daty a potřebují chránit své zákazníky.

2.2.3 Cookies

Podkapitola 2.2.1 Značkování stránek nastínila, že mezi metody shromažďování uživatelských dat můžeme také zařadit soubory cookies. Pojem cookies je jednoduše charakterizován jako malé textové soubory rozlišující jednotlivé uživatele, kteří navštěvují webovou stránku. Při každé další návštěvě totožného serveru jsou pak tato data odesílána zpět na stranu serveru. Předpokladem cookies je dle serveru Adaptive (2018) rozlišit jednotlivého uživatele a uložit si o něm konkrétní informace, např. při další návštěvě vyplnit již dříve vyplněný přihlašovací formulář, nebo správně lokalizovat jazyk webových stránek na základě předešlých nastavení návštěvníka. Soubory cookies tedy představují pro každého uživatele naprosto unikátní identifikátor. Oproti identifikaci prostřednictvím IP adresy jsou jejich výhody téměř nezpochybnitelné – struktura IP adresy se zpravidla mění, případně nastává otázka, jak server rozliší dva unikátní uživatele připojené na stejný domácí router s pevnou IP adresou.

Mezi osobní data zaznamenaná v souborech cookies může patřit i geografická poloha uživatele. Geolokace zařízení je však závislá na nastavení parametrů v souborech cookies a nepatří tedy mezi primárně sbíraná data. Webové prohlížeče disponují možnostmi pro smazání, editaci či zobrazení cookies jednotlivých stránek. Nelze si nevšimnout obrovského rozdílu mezi počtem cookies na malém a velkém webu, zajímavým srovnáním jsou v tomto případě weby Katedry geoinformatiky a e-shopu Alza.cz. Tato data jsou volně dohledatelná ve vývojářské konzoli prohlížeče pro všechny webové stránky.

Z právního hlediska došlo v posledních letech k úpravě využívání cookies prostřednictvím evropské směrnice, jak uvádí server epravo.cz (2015): „směrnice zavázala členské státy EU, aby do svých právních rádek zavedly povinnost provozovatele webových stránek ukládat cookies v počítači či jiném zařízení pouze po předchozím souhlasu uživatele webových stránek. E-Privacy směrnice tedy vyžaduje zavedení tzv. opt-in režimu pro používání cookies, kdy uživatel webových stránek musí aktivně udělit souhlas s používáním cookies ještě před samotným uložením cookies v uživatelské počítači nebo jiném zařízení.“ Více v kapitole 3.4 GDPR v souvislosti s webovou analytikou.

2.3 Sběr prostorových dat

Data s prostorovou složkou jsou stěžejním předmětem vizualizace naměřených dat do formy mapových výstupů. Shromážděné informace nabývají prostorové informace pomocí tří základních metod, kterými jsou IP adresa, GPS a triangulace sítě. Podmínkou pro lokalizaci zařízení je pak jeho dosažitelnost, tedy například komunikace s jakoukoliv sítí (mobilní, internetovou). Jestliže je zařízení vypnuto, nejsme schopni určit jeho aktuální polohu.

Lokalizace zařízení může probíhat pomocí internetového protokolu (IP). IP adresou disponují všechna zařízení, která prostřednictvím tohoto protokolu komunikují, například tedy počítače. Ačkoliv se jedná o poměrně rychlou metodu geolokace, její přesnost je zpravidla omezena na polohu poskytovatele internetu. V případě GPS je poloha určována pomocí známých poloh vysílačů (BTS stanic). Čím blíže je tedy třeba mobilní telefon bližší těmto stanicím, tím přesnější je určení jeho polohy. Obdobným způsobem jsme schopni lokalizovat zařízení na základě okolních Wi-Fi sítí, které mají rovněž prostorové souřadnice.

Pakliže není zařízení připojeno k GPS ani k internetu, je mobilní operátor schopen určit jeho přibližnou polohu pomocí triangulace sítě na základě výše zmiňovaných BTS stanic pomocí vysílaného mobilního signálu. Tato data jsou však přístupná pouze operátorům a je nutné zmínit, že metoda triangulace je nejméně přesná. V rámci webové analytiky jsou tak použitelnými faktory pro geolokační proces pouze IP adresa a GPS. I přes to ale hovoříme o přesnosti v rámci jednoho města, podrobnější poloha je bez zásahu poskytovatele internetu (ISP) téměř nemožná.

2.4 Heuristická analýza

Heuristická analýza použitelnosti představuje kvalitativní zhodnocení kompletního webu, aplikace, nebo systému, díky kterému získáme přesnou představu o silných a slabých místech testovaného objektu (Nielsen, 2004). Přestože se jedná o subjektivní metodu, hodnocení probíhá na základě předem stanovených kritérií, které vychází z komplexního vzorku zkoumaných řešení. Díky tomu je subjektivní vliv eliminován na co nejnižší možnou míru. Heuristická analýza použitelnosti vychází z 10 principů použitelnosti dle Nielsen (2004), (Inflow 2015):

- Viditelnost stavu systému,
- Propojení systému a reálného světa,
- Uživatelská kontrola a svoboda
- Standardizace a konzistence
- Prevence chyb
- Rozpoznání namísto vzpomínání
- Flexibilní a efektivní použití
- Estetický a minimalistický
- Schopnost uživatelů pochopit, poznat a vzpamatovat se z chyb
- Náповěda a návody

Bylo vytvořeno celkem 170 hodnotících otázek, které byly rozděleny do 10 logických kategorií odpovídajících základním aspektům: Obecné, GIS, Karto, Použitelnost, Administrace, Bezpečnost, Obsah, Grafika, Vyhledávání, Technologické. Heuristická analýza použitelnosti je reprezentována XLSX souborem pro využití v aplikaci Microsoft Excel, přičemž výše zmiňovaných deset kategorií je obsaženo na jednotlivých listech s tematicky příbuznými otázkami, např.: „K dispozici je grafické měřítko“. Každé z otázek je přiřazena odpověď na základě předem definované množiny hodnot: -1 = nesplňuje, 0 = splňuje částečně, 1 = splňuje, prázdné, pokud otázka není relevantní. Otázka je uvedena v poli B, přičemž hodnotitel zadává odpověď do zeleně zvýrazněného pole D. V závislosti na hodnotě koeficientu, který může, ale nemusí být zadán podle důležitosti otázky, je vypočítáno skóre konkrétní otázky v poli E. Celkové výsledky jsou zpracovány v listu „Výsledky“, kde jsou souhrnně zobrazeny podstatné statistiky (počet otázek, odpovědí, % zodpovězených), výsledky za jednotlivé listy i celkový výsledek. Tyto výsledky lze odečítat v absolutních hodnotách (ve formě bodů) i procentuálním přepočtem (ve formě skóre).

Jak uvádí Čapková (2010): „heuristiky nejsou pevný seznam definitivních faktů, jedná se o jistá vodítka, která může hodnotitel použitelnosti použít jako startovní čáru, od které začne aplikovat svoji expertizu“. Na základě tohoto tvrzení mohou být otázky upravovány dle potřeb, hodnotitel může doplňovat další důležité otázky dle konkrétní aplikace. Pro potřeby bakalářské práce proto byly parametry heuristické analýzy použitelnosti upraveny. Některé méně podstatné listy byly zcela odstraněny, protože se pro danou problematiku nepovažují za důležité, jiné (např. GIS a Karto) sloučeny dohromady. Výsledkem je tedy 5 logických kategorií čítajících 10–27 otázek, přičemž konečný počet otázek je 85.

2.5 Použité programy

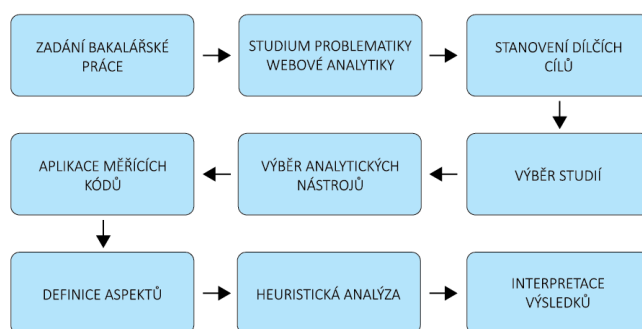
Většina práce s naměřenými daty probíhala v prohlížeči. Následné vyhodnocování analytických nástrojů pomocí heuristické analýzy bylo provedeno v aplikaci Microsoft Excel. Pro grafické výstupy byl využit open-source vektorový nástroj Inkscape. Finalizace textu bakalářské práce probíhala v programech balíčku Microsoft Office 2016.

2.6 Postup zpracování

Prvotním krokem práce bylo seznámení s problematikou webové analytiky a výběr vhodných analytických nástrojů, které budou vyhovovat požadavkům práce. Následovalo zorientování se v jednotlivých nástrojích a aplikace měřících kódů na webové stránky pro sběr dat. Následující postup zpracování bakalářské práce byl rozdělen na několik teoretických a praktických částí.

Důležitým krokem bylo definovat aspekty a jejich dílčí části, na základě kterých budou vybrané webové analytické nástroje hodnoceny. Počet aspektů byl důsledkem odborného šetření snižen, a to kvůli tematické podobnosti některých z nich. Ve finální verzi se hodnotí tyto aspekty: technický, kartografický, uživatelský a vizualizační a analytický.

Po definování předchozí části byly srovnány silné a slabé stránky vybraných analytických nástrojů, a to na základě heuristické analýzy, jejíž parametry byly autorem práce upraveny pro účely bakalářské práce. Na základě zjištěných plusů a minusů bylo navrženo grafické prostředí ideální analytické služby (s důrazem na prostorové vyhodnocení dat). V neposlední řadě byla sepsána textová část práce, vytvořeny webové stránky a poster.



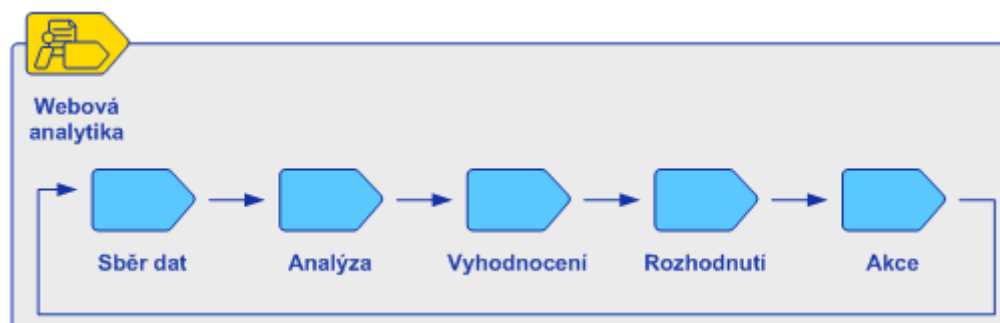
Obr. 2 Vývojový diagram postupu práce

3 SOUČASNÝ STAV ŘEŠENÉ PROBLEMATIKY

3.1 Webová analytika

Vývoj webové analytiky jako disciplíny je těsně spjat s vývojem webu a jeho využitím pro marketingové účely. Definice webové analytiky se mění v čase podle způsobů získávání dat, zdrojů informací, způsobu zpracování, podle vývoje marketingových analýz, potřeb marketérů a možností prezentace výsledků. Definice pojmu webová analytika není ustálená a také vzhledem k vývoji není možná jednoznačná definice (Wikidot, 2014).

Mnoho zdrojů však uvádí definici zdůrazňující všechny důležité faktory: webová analytika je měření, sběr, sledování a vyhodnocování internetových dat za účelem porozumění a optimalizace webových stránek. Jednoduše řečeno, využitím analytických nástrojů lépe porozumíme chování uživatelů na webu. Obrovský dopad mají tyto nástroje zejména v oblasti marketingu, kde správné nastavení a interpretace dat zpravidla zvyšuje konverzi, tj. akci, kterou má návštěvník vykonat během doby strávené na webu – např. stažení aplikace, vstup do nákupního košíku aj. (Jan Petryl, 2017). Problematika webové analytiky mimo jiné zasahuje do designu webu, zásadní může být kupříkladu špatné umístění určitého tlačítka.



Obr. 3 K čemu slouží webová analytika?
zdroj: www.krcmic.cz

Webová analytika se ale diametrálně odlišuje od základních statistik webových stránek. Zatímco základní statistiky poukazují na jednoduché metriky, například kolik lidí navštívilo web za minulý týden, webová analytika hledá odpovědi na sofistikovanější otázky: *Proč se tomu tak děje? Co uživatelé na daném webu dělají (anebo nedělají)? Proč nechodí na určité stránky, jaké možné překážky asi musí řešit uživatel během návštěvy webu (tj. co je třeba zlepšit)? Odkud uživatelé na web přišli? Přinášejí financované kampaně výsledky? Pokud ano, jsou efektivní? Kde z webu nejčastěji odcházejí (rozlišujeme přímou návštěvnost, odkazující stránky, přirozenou návštěvnost atp.)?*

Mezi těmito dvěma statistickými metodami samozřejmě existuje nějaký přesah, použijeme-li automobilovou analogii: pokud jsou základní webové statistiky tachometrem vozu, webová analytika je jeho palubní deskou (Logaholic, 2018). Analytika ve webovém kontextu zkrátka představuje nikdy nekončící koloběh shromažďování, analýzy a interpretace dat za účelem co nejlepší optimalizace webu.

3.2 Analytické nástroje

Seznam nástrojů pro webovou analýzu je dlouhý, každý se navíc liší pokročilostí (ve smyslu propracovanosti metrik) a výše zmiňovanou metodikou sběru internetových dat. Obecně tedy není jednoduché určit, které řešení bude pro daný projekt vhodné. Je důležité podrobně zhodnotit požadavky projektu/firmy, která chce tyto nástroje využívat, a z dlouhého výčtu analytických služeb vybrat tu nejvhodnější. Nutno podotknout, že nasazení některé z dostupných služeb nelze brát jako podmínku úspěchu, ale jako příležitost efektivně získat cenné informace. Podobně jako metody sběru dat, je i kombinace více nástrojů přínosem. Optimální řešení dle Martina Čecha (2010) tedy zohledňuje:

- Potřeby firmy;
- Metodiku měření dat;
- Náklady (implementaci, provoz, zaškolení pracovníků);
- Způsob sběru dat a následnou práci nástroje s daty;
- Objemy dat, které je nástroj schopen zpracovat;
- Bezpečnost a ochranu dat včetně způsobu uchování dat;
- Grafické možnosti výstupů, statistik;
- Míru automatického předzpracování/vyhodnocení;
- Možnosti přizpůsobení nástroje dle potřeb analytika/firmy;
- Nároky na uživatele při práci s nástrojem;
- Integraci, tj. propojení s ostatními aplikacemi a nástroji, databázemi, firemními systémy
- Možnosti efektivního exportu, importu dat;
- Náročnost implementace a nastavení;
- Producentickou podporu včetně jazykové lokace

Nejen v České republice, ale také v celosvětovém měřítku, má největší podíl bezplatná služba Google Analytics (viz níže). Z placených nástrojů jsou pak oblíbenými alternativami programy SiteCatalyst od americké společnosti Adobe, Webtrends, nebo IBM Customer Analytics (*dříve Coremetrics*). Stručný popis některých nejvíce využívaných služeb slouží k jejich prvotnímu srovnání, autor se bude vybranými bezplatnými možnostmi podrobněji zabývat v jiné části práce.

Google Analytics

Služba Google Analytics (dále jen GA) je produktem společnosti Google a podle mnohých zdrojů představuje nejpoužívanější analytický nástroj současnosti. Z průzkumů vyplývá, že nástroj GA je nejčastěji nasazován u méně rozsáhlých webů – tedy čím je web rozsáhlejší, tím více jsou využívány složitější a propracovanější analytické nástroje a řešení. Služba je dostupná ve dvou variantách: základní bezplatné verzi Standard, která se může volně využívat pro komerční účely, a placené verzi Premium 360 s cenovkou 150 000 dolarů ročně. Rozdíl je především v limitech a požadavcích při dotazování na server.

Yandex Metrica

Analytický nástroj Yandex Metrica patří pod ruskou nadnárodní korporaci Yandex, jež se zaměřuje na internetové služby (*podobně jako v České republice Seznam.cz*). Největší rozmach zažívá zejména v oblasti Ruska, kde je dokonce používanější než Google Analytics. Existuje pouze v bezplatné verzi, není však limitován žádnými faktory.

Společnost Yandex na oficiálních stránkách uvádí, že službu využívá více než osm milionů webů; podle některých zdrojů tedy zabírá druhou příčku v tomto sektoru služeb.

Adobe SiteCatalyst (dříve Omniture)

Webový analytický nástroj Omniture vznikl v roce 1996, následně v roce 2009 proběhla akvizice společností Adobe Systems. Dále mluvíme o projektu SiteCatalyst. Cenová politika Adobe však příliš nenahrává nástroji do karet: v roce 2012 byl měsíční paušál za středně velký web 770 € (*v přepočtu téměř 20 tisíc korun*). SiteCatalyst proto kvůli výše zmíněným důvodům zabírá na trhu zhruba jen dvouprocentní podíl.

Matomo

Matomo, do začátku roku 2018 Piwik, představuje open-source aplikaci pro měření a optimalizaci webových stránek. Jako jedna z mála funguje prostřednictvím PHP/MySQL serveru, a tudíž nejsou shromážděná data sdílena s žádnou službou třetí strany. K tomuto roku je nasazena na více než milionu stránek, zabírá 1,3 % v segmentu analytických nástrojů. Služba Matomo je poskytována zcela zdarma.

Clicky Analytics

Službu Clicky v současnosti využívá více než milion webových stránek (Clicky, 2018), číslo se však dennodenně zvyšuje. Základní tarif pro web s denní návštěvností do 3000 uživatelů je bezplatný, další varianty se pak v závislosti na počtu spravovaných projektů a návštěvnosti pohybují od 80 do 160 dolarů ročně. Clicky je na poli analytických nástrojů přes deset let.

Woopra

Nástroj Woopra se převážně zaměřuje na co nejpřesnější zpracování dat v reálném čase. Je nabízen v několika variantách, kdy se cena odvíjí podle počtu akcí, které uživatelé vykonají (*tj. počet návštěveních stránek, počet stažení, události...*), a také na velikosti objednatele. Menší firmy může stát roční provoz webu až šest tisíc dolarů, v přepočtu takřka 124 tisíc korun. Tarif Enterprise je podle oficiálních webových stránek Woopra (2018) zpravidla určen pro větší podniky a je řešen individuálně, přičemž oproti předchozí variantě disponuje integrací prémiových služeb (*MySQL, Adwords, Salesforce apod.*). Projekt Woopra byl v této oblasti služeb mnohokrát vyzdvihován různými internetovými magazíny.

Závěrem je nutné zmínit, že ačkoliv některé nástroje využívají stejnou metodiku sběru dat, nelze naměřená data mezi sebou srovnávat. Důvodem je odlišný technický základ každého z nástrojů, výsledná data jsou také jinak prezentována. Argumentů pro nesrovnalosti v datech je podle Martina Čecha (2010) hned několik:

- Odlišný postup a frekvence pravidelného zpracování a sběru dat: data mohou být při porovnání rozdílná, neboť každý systém může mít jinou frekvenci zpracování (1 den, 1 hodina apod.);
- Filtry a specifické nastavení: různé systémy mohou mít rozdílné způsoby přiřazení filtrů a mohou být specificky nastavené;
- Negativní transakce elektronického obchodu: schopnost systému operovat zpětně s daty při negativní transakci (vrácení produktu);

- Problém bezpečnosti kódu u stránkových značek: při zjištění kódu může být zneužit a vložen i na jiný web, čímž dojde ke zkreslení dat (např. při konkurenčním boji);
- Doba trvání souborů cookies: každý systém může mít nastavenou jinou dobu pro trvání cookies.;
- Návštěvník či Návštěva: použití jiných algoritmů pro potřeby rozlišování návštěvy od návštěvníka vedou k rozdílným počtům;
- Kompletnost označování webu a nutnost úpravy značek pro měření: každý systém nemusí zvládnout měřit určité aktivity. Pro tento účel je vyžadována úprava původního měřicího kódu, při níž však může dojít k syntaktickým chybám. Nejen při těchto dodatečných úpravách, ale obecně se nesmí zapomenout, že každá stránka webu by měla být opatřena měřicím kódem daného analytického nástroje/služby;
- Umístění měřicích (stránkových) značek: různé způsoby umístění měřicích značek v rámci zdrojového kódu mohou znamenat rozdílnost výsledných naměřených dat;
- Vlastní versus cizí soubory cookies: nízká míra souvztažnosti mezi oběma druhy cookies, způsobená vyšší mírou blokování cizích souborů cookies.

3.3 Základní metriky webové analytiky

Webová analytika, jak bylo zmiňováno v kapitole 3.1, je měření, sběr, sledování a vyhodnocování internetových dat za účelem porozumění chování uživatelů a optimalizace webových stránek. Analytické nástroje sledují obrovský počet informací, jež je nutné správně interpretovat.

Přestože jsou následující ukazatele považovány za zcela základní a jsou zastoupeny v téměř všech analytických nástrojích, může se lišit metodika jejich výpočtu. Pro správné pochopení metrik je tedy vhodné přečíst příslušnou dokumentaci konkrétního nástroje. Během psaní bakalářské práce bylo nicméně zjištěno, že ve většině případech jsou tyto definice nedohledatelné. Logicky z toho vyplývá, že nasadí-li uživatel na jednu webovou stránku více analytických nástrojů, jejich shromážděné statistiky se budou s největší pravděpodobností lišit, někdy velmi výrazně. Autor však nedoporučuje využívat měřicí kódy několika služeb na jednom webovém projektu, neboť může docházet ke vzájemným „kolizím“.

Návštěvy (Visits)

Návštěvou se rozumí taková situace, kdy někdo navštívil web, nějakou dobu procházel stránkami a pak web opustil. Návštěva je také charakterizována jako časový úsek (relace, anglicky session) mezi prvním a posledním požadavkem na webový server. Různé nástroje však tuto metriku definují jinak, například služba Google Analytics ji definuje následovně:

„Jeden uživatel může provést více návštěv. Ty mohou proběhnout ve stejný den, ale i během několika dnů, týdnů, či měsíců. Jakmile jedna návštěva skončí, může začít nová. Návštěva se ukončuje dvěma možnými způsoby:

- Časové vypršení platnosti:
 - o po 30 minutách nečinnosti návštěvníka,
 - o o půlnoci.
- Změna kampaně
 - o pokud uživatel na web přijde prostřednictvím jedné kampaně, odejde a vrátí se prostřednictvím jiné kampaně.“

Návštěvníci (Visitors)

Často používaným ukazatelem je podle mediálního slovníku Mediaguru (2018) počet takzvaných unikátních návštěvníků (unique visitors), kteří představují počet unikátních návštěv webové stránky v určitém časovém období. Při jejich další návštěvě se návštěva již nezapočítává. Metrika tedy udává, kolik reálných lidí přišlo na web za toto období. Údaje se sbírají prostřednictvím souborů cookies, díky čemuž je nutno brát tento ukazatel za mírně zkreslený – na jednom počítači může být více uživatelů, případně cookies mohou být zakázány.

Čas strávený na stránce a čas strávený na webu (Time on page, Time on site)

Tento ukazatel udává, kolik času strávil uživatel na konkrétní stránce, než ji opustil nebo než odešel na jinou stránku. Čas strávený na webu naopak představuje součet časů, které uživatel strávil na těchto stránkách. Problémem je určit čas opuštění poslední stránky, tedy samotného webu, protože tento údaj zkrátka není dostupný. Řešením může být například implementace tzv. onbeforeunload – kódu, který se při odchodu uživatele z webu zeptá, zda chce opravdu opustit web, čímž získá chybějící časový údaj opuštění webu (Web-analytics, 2014).

Míra okamžitého opuštění (Bounce Rate)

Míra okamžitého opuštění představuje procento návštěv, které opustily web po shlédnutí jediné stránky. V závislosti na účelu webu nemusí vysoké číslo nutně znamenat problém, což dokazují například weby tvořené jednou stránkou. Pokud je ale vysoká míra okamžitého opuštění zastoupena na webech, od kterých se to neočekává, problémem může být design webu (pak by se měl zvážit jeho redesign), nebo chybná implementace měřicího kódu (nemusí být vložen na všech stránkách).

Míra odchodů (Exit Rate)

Tato metrika udává, kolik procent návštěvníků opustilo web z konkrétní stránky (ze stránky, u které odchod není očekáván, např. z domovské stránky). Míra odchodu představuje procentuální podíl případů, kdy byla daná stránka poslední navštívenou stránkou, vzhledem k celkovému počtu zobrazení stránky (Google, 2018).

Míra zájmu (Engagement)

Míra zájmu prostřednictvím ostatních, výše zmiňovaných metrik, měří zájem návštěvníka o konkrétní webovou stránku (Wikidot, 2014). Lze ji zjistit různými metodami, např. dotazníkovým šetřením při odchodu z webu, přímým marketingovým šetřením trhu či pravidelnou analýzou zákazníků (např. pomocí emailové kampaně).

Míra konverze (Conversion Rate)

Míra konverze je dle webu JakMeritWeb.cz (2018) údaj nejčastěji udávaný jako procentuální poměr mezi počtem uskutečněných cílů (konverzí) a počtem návštěv. Například 5 konverzí ze 100 návštěv odpovídá konverznímu poměru 5 %. Za konverzi považujeme např. objednávku zboží na e-shopu, stažení souboru, kliknutí na odkaz a podobně.

3.4 GDPR v souvislosti s webovou analytikou

Jak bylo letmo zmíněno v předcházejících kapitolách, v roce 2018 vstoupilo v platnost nařízení GDPR – dlouho připravovaná legislativa Evropské unie, jejímž hlavním úkolem je chránit a modernizovat osobní údaje evropských občanů. GDPR znamená Obecné nařízení o ochraně osobních údajů, v angličtině General Data Protection Regulation, a nahrazuje původní a dnes již zastaralou směrnici z roku 1995 (Info.cz, 2018). Nová směrnice se dotkne zejména subjektů, které shromažďují anějakým způsobem zpracovávají osobní údaje občanů EU. V rámci GDPR jsou za osobní údaje považovány např. jméno, věk, datum narození, ale také technické parametry jako poloha, email, IP adresa, číslo kreditní karty či výše zmiňované cookies. Jak uvádí server Info.cz, tyto subjekty budou muset vést záznamy o činnostech, které s osobními daty provádějí, a ohlásit všechny případné úniky osobních údajů. V případě nedodržení pravidel legislativy se mohou sankce dle závažnosti případu vyšplhat až na 20 milionů korun.

Za reálnou integraci nařízení GDPR na webových stránkách považujeme povinnost informovat návštěvníky webu o tom, jaká data web zpracovává. Mluvíme o takzvané cookie liště, která (v závislosti na účelu zpracování) návštěvníky o těchto náležitostech informuje, nebo vyžaduje jejich aktivní souhlas k použití údajů pro další procesy (zejména tedy pro marketingové účely), (Miroslav Pecka, 2018). Rovněž se dá předpokládat, že v budoucnu budou moci uživatelé požádat o poskytnutí kompletního výpisu všech svých údajů, jejich změnu, smazání, anebo zastavení jejich dalšího zpracování na konkrétní webové stránce.

Pro účely tohoto nařízení se dle oficiálních definic rozumí: "osobními údaji" veškeré informace o identifikované nebo identifikovatelné fyzické osobě (dále jen "subjekt údajů"); identifikovatelnou fyzickou osobou je fyzická osoba, kterou lze přímo či nepřímo identifikovat, zejména odkazem na určitý identifikátor, například jméno, identifikační číslo, lokační údaje, síťový identifikátor nebo na jeden či více zvláštních prvků fyzické, fyziologické, genetické, psychické, ekonomické, kulturní nebo společenské identity této fyzické osoby (Privacy Regulation, 2018).

Závěrem kapitoly o GDPR je potřeba zmínit, že nová evropská směrnice dalším způsobem neovlivnila problematiku webové analytiky. Zejména v rámci vizualizace prostorových dat nedošlo k signifikantním změnám, co se přesnosti dat či využitých kartografických metod týče. Některé z webových analytických nástrojů (např. Google Analytics) se v souvislosti s ochranou osobních údajů vydaly cestou maskování IP adres. Na základě tohoto opatření nemají subjekty shromažďující data přístup k celé IP adrese, ale pouze k její části. Nutno však podotknout, že i tato metoda nijak neovlivňuje geolokaci návštěvníků webových stránek, ale spíše anonymizuje osobní údaje dle požadavků evropské legislativy.

4 DEFINICE ASPEKTŮ

Na základě předešlé teoretické části a stručného popisu používaných analytických nástrojů byly vybrány čtyři nástroje a definována kritéria, která jsou v rámci těchto nástrojů stěžejní. Konkrétně byly na weby nasazeny měřicí kódy bezplatných služeb Google Analytics, Yandex.Metrica, Clicky Analytics a Woopra. Jiné analytické služby nebyly uvažovány kvůli jejich finanční dostupnosti, nebo kvůli jejich malé uživatelské základně.

Prvotním plánem bylo srovnat tyto možnosti v rámci několika předem určených hledisek – kartografického, geoinformatického, technického, uživatelského, bezpečnostního... Na základě odborného šetření se však dospělo k závěru, že některé z aspektů jsou si tematicky podobné, a tak by měly být sloučeny. Tyto služby tedy budou podrobně definovány pomocí následujících upravených hledisek: technického, kartografického, uživatelského a vizualizačního a analytického. Stěžejním tématem je práce s prostorovými daty, tj. odkud je služby berou, jak je vizualizují atd. Každá tato heuristika obsahuje mnoho podmnožin, které je nutné definovat a určit jejich důležitost.

4.1 Technické aspekty

Technické parametry analytických nástrojů zahrnují sběr prvotních dat, tzn. jaká metodika sběru dat byla u konkrétního nástroje zvolena a jakým způsobem data nabyta jejich prostorovou informací (prostřednictvím IP adresy, GPS, HTML5...), případně zda se do těchto elementárních informací může vůbec nahlížet, nebo je to v rozporu s ochranou osobních údajů konkrétní služby. Důležitými faktory jsou zpracování dat v reálném čase (real-time processing) či šifrování služby skrze zabezpečený https protokol. Z technického hlediska jsou klíčovými a nezbytnými funkcemi export dat (důležitým aspektem je výčet formátů) a podpora pluginů, extenzí a API pro pokročilejší analýzy. Výhodou je přítomnost aplikace pro nejzastoupenější mobilní platformy Android či iOS. V rámci technických aspektů popisujeme:

- metodiku sběru dat
- možnost zjištění IP adresy uživatelů (ochrana osobních údajů)
- systém řízení báze dat
- zpracování v reálném čase (real-time processing)
- zabezpečení analytického nástroje
- export shromážděných dat
- popis pluginů, extenzí a API
- mobilní aplikace

4.2 Kartografické aspekty

Aspekty analytických nástrojů z kartografického hlediska se primárně zaměřují na naměřená data, konkrétněji na jejich vizualizaci do mapových výstupů, nicméně obsahují také popis využitých geoinformatických technologií. U jednotlivých analytických nástrojů popisují samotné mapové pole, tzn. s jakými územními jednotkami je v rámci analýzy možné pracovat, legendu (většinou v podobě barevné stupnice), či měřítko. Obecně mluvíme o mapové kognici, tzn. jak mapy na uživatele působí, zda jim návštěvník rozumí, zda obsahují potřebné kompoziční prvky a podobně. Mapové výstupy jednotlivých služeb budou subjektivně hodnoceny.

V rámci mapového výstupu je dále posuzována použitá kartografická metoda, především tedy její korektnost vzhledem k nashromážděným datům. Důležitou otázkou je charakter dat, tedy zda se jedná o absolutní či relativní hodnoty. Z hlediska geoinformatického pak popisujeme použité mapové online řešení (např. Leaflet, Mapbox, Google Maps API...).

- Vizualizace prostorových dat zahrnující hodnocení mapového pole, jazyk mapy, kompoziční prvky a kartografické metody
- Mapová kognice (subjektivní vnímání a hodnocení mapového výstupu)
- Interaktivní prvky mapových výstupů
- Použití online mapové řešení (Google Maps API, Leaflet, Mapbox, OpenStreetMap apod.)

4.3 Uživatelské a vizualizační aspekty

Požadavky uživatelů na analytické nástroje se mohou lišit. Kategorie uživatelských aspektů vnímá rozhraní analytického nástroje jako celek. V potaz se bere například celková estetika prostředí služeb, tedy jak na uživatele konkrétní nástroj působí, zahrnující například použitou barevnou symbologii a její celkovou vhodnost. Z hlediska webového designu se kapitola uživatelských aspektů zaměřuje na použité prvky webové kompozice – zda jsou důležité informace, například počet návštěv, dostupné na hlavní stránce, případně v záložkách, nebo se k nim uživatel dostane až prostřednictvím hledání v hlavní nabídce služby. V tomto kontextu můžeme posuzovat, zda je orientace v prostředí snadná či zbytečně složitá, a subjektivně tak upozornit na silné a slabé stránky každé ze služeb ve smyslu kritických chyb nebo drobných nedostatků. Důležitým faktorem je využití responzivního designu, díky kterému je prostředí nástroje včetně obsahu pomocí HTML kódu optimalizováno pro komfortní zobrazení na více typech zařízení, a to zcela nezávisle na velikosti úhlopříčky displeje (telefon, tablet, notebook...). Podstatný vliv na porozumění naměřených dat mají použité metody vizualizace, v tomto případě přehlednost tabulek a různé typy grafů (výšečový, liniový, sloupcový, grid a podobně).

- Estetika rozhraní analytických služeb včetně použité barevné symbologie, kompozice důležitých prvků a jejich dostupnosti
- Responzivní design
- Metody vizualizace dat (např. grafy)

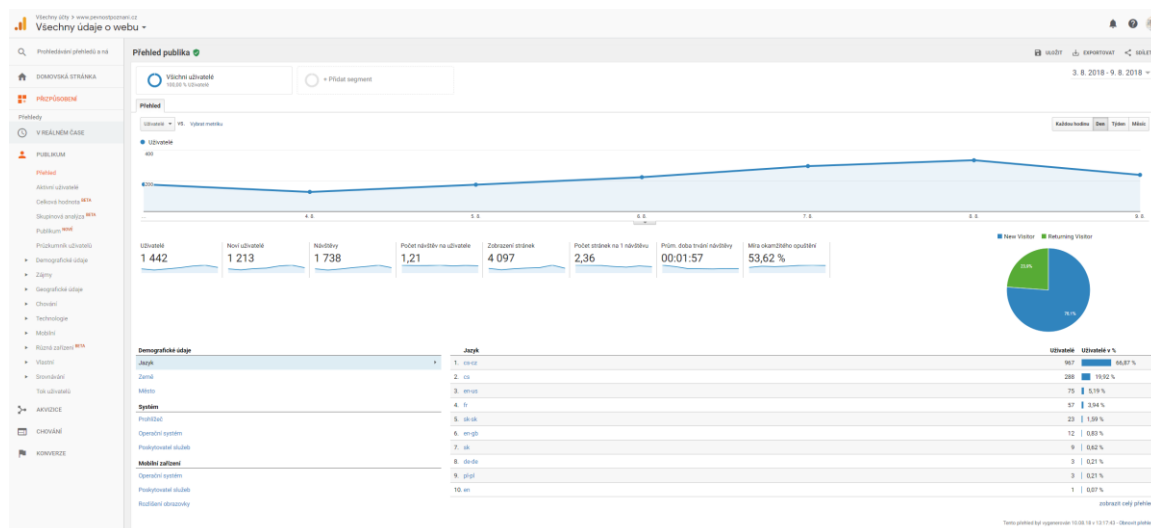
4.4 Analytické aspekty

Mezi analytické aspekty patří možnosti interpretace naměřených dat a schopnosti jednotlivých analytických nástrojů, tzn. jaké nástroje patří k nejpokročilejším a co všechno umí, respektive jaká data zpracovávají. V obecném kontextu hovoříme o měření základních a složitějších metrik. K jistým rozdílům mezi analytickými nástroji dochází především kvůli odlišnému způsobu měření těchto metrik. Některé z analytických nástrojů dokážou zjistit nejen zcela elementární údaje, ale na základě shromážděných údajů také definovat pokročilejší chování uživatelů, např. porovnat nové a vracející se uživatele, dobu trvání návštěvy, počet stránek na návštěvu, zdroje návštěvnosti a další. Z pohledu hlubších analytických parametrů hovoříme o přesnosti zpracování v reálném čase, případně možnost provázání analytické služby s některou ze současných sociálních sítí jako jsou Facebook, Instagram, Twitter a další.

5 ASPEKTY ANALYTICKÝCH NÁSTROJŮ

5.1 Google Analytics

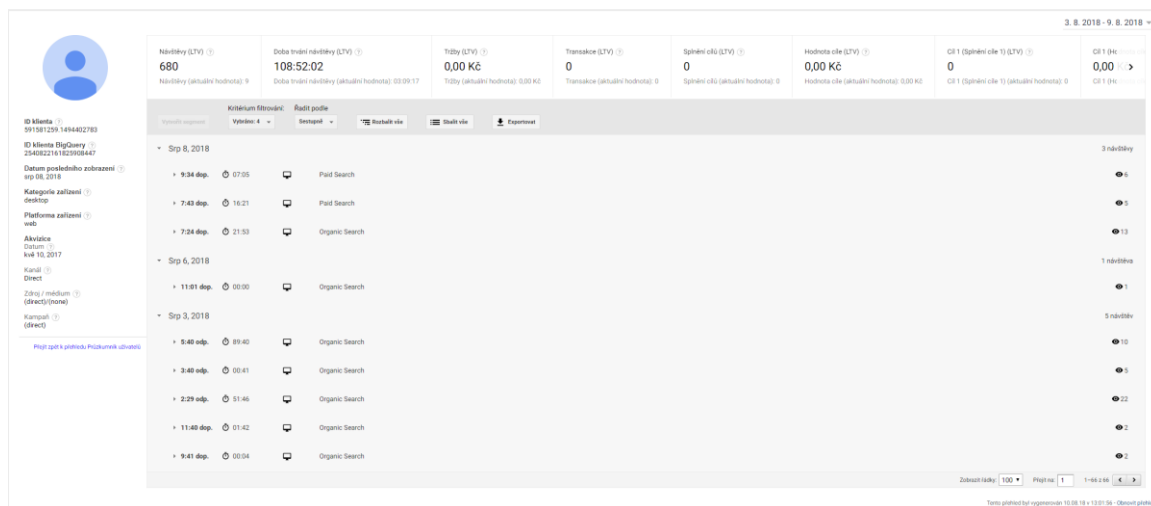
Nástroj Google Analytics (dále GA) je vůbec nepoužívanějším nástrojem současnosti. Je to způsobeno zejména přesností měření, do karet mu ale nahrává také jednoduchá implementace měřicího kódu, velké množství pluginů (hlavně pro weby postavené na platformě WordPress) a uživatelsky přívětivé webové prostředí. Mnoho marketingových pracovníků si pochvaluje možnost vlastního přizpůsobení panelů a přehledů, což práci s nástrojem výrazně usnadňuje. Prostřednictvím jednoho GA účtu je možné zdarma spravovat až sto webových projektů a libovolně se mezi nimi přepínat.



Obr. 4 Webové rozhraní služby Google Analytics

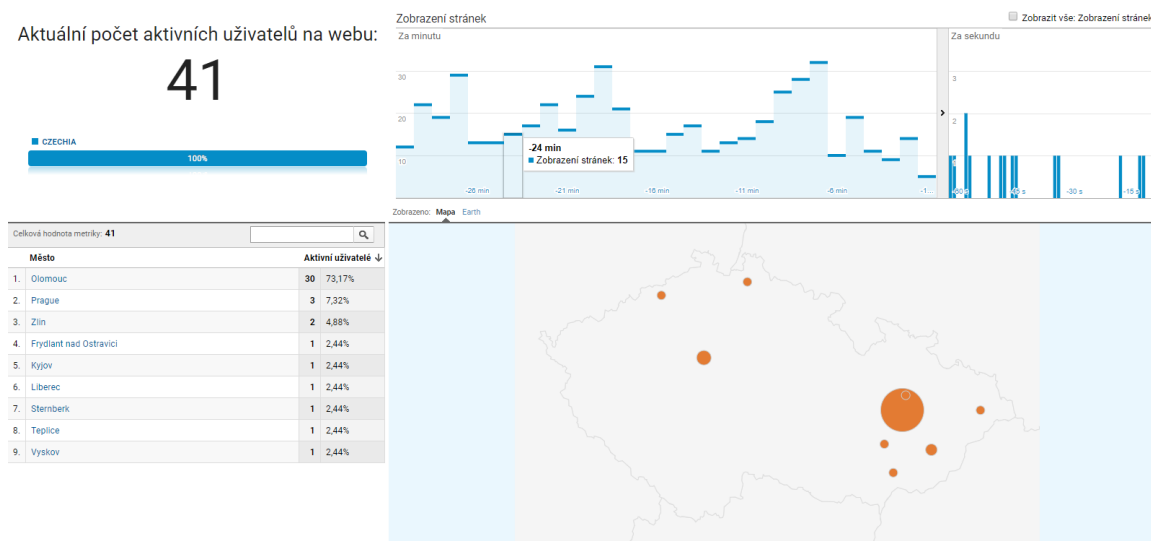
Technické aspekty

Služba Google Analytics obecně pro sběr informací využívá metodu značkování stránek pomocí JavaScript kódu, který je vložen na každou stránku, jíž chce uživatel sledovat. Prostorová data nástroje GA jsou získávána pomocí IP adres a ISP (Internet Service Provider – poskytovatel internetového připojení). Přístup na úroveň jednoho uživatele, tedy jedné IP adresy, je v rozporu s ochranou osobních údajů společnosti Google, nicméně pomocí filtračních nástrojů je lze zpřístupnit. Google v rámci bezpečnosti využívá metodu známou jako „maskování IP“, díky které mohou webové projekty využívající Google Analytics využívat pro geolokaci pouze část IP adresy, nikoliv celou adresu. Co se přesnosti lokalizace týče, GA zaokrouhluje zeměpisnou šířku a délku na 4 číslice, čímž poskytuje maximální možnou přesnost zhruba 11,1 metru (Hamel, 2016).



Obr. 5 Přístup na úroveň jednoho uživatele (Google Analytics)

Zpracování v reálném čase funguje se zpožděním několika vteřin, přičemž nabízí přesnost v rámci měst. Ukazuje aktuální počet aktivních uživatelů a zároveň kolik a jaké stránky jednotlivý uživatel navštívil. Zpřesnění zpracování v reálném čase pro účely podrobnějších informací lze docílit implementací Real-Time Reporting API, viz níže.



Obr. 6 Zpracování dat v reálném čase (Google Analytics)

Obrovský počet shromážděných dat je zpracováván ve vysoce optimalizovaném cloudovém databázovém systému BigTable. Stejná databáze řídí například Vyhledávání Google, Gmail, Google Maps a další Google služby s extrémními objemy dat. Projekt BigTable představuje NoSQL sloupcovou (wide-column) databázi zvládající mohutný počet požadavků s neměnnou nízkou latencí. Tato naměřená data je možné jednoduše filtrovat v závislosti na čase a exportovat do formátů PDF, CSV, XLSX a Tabulek Google, případně komukoliv sdílet e-mailem v některém ze zmiňovaných formátů. Výhodou je velký počet pluginů pro weby postavené na platformě WordPress (např. Google Analytics Dashboard for WP nebo Google Analytics for WordPress by MonsterInsights), jež poskytují základní údaje přímo v rozhraní tohoto redakčního systému. Rovněž implementace měřicího kódu a optimalizace služby je díky pluginům vhodná i pro laickou veřejnost. Webové rozhraní GA je zabezpečeno https protokolem. Mobilní aplikace má zastoupení na platformách Android a iOS, uživatelé mohou využít také

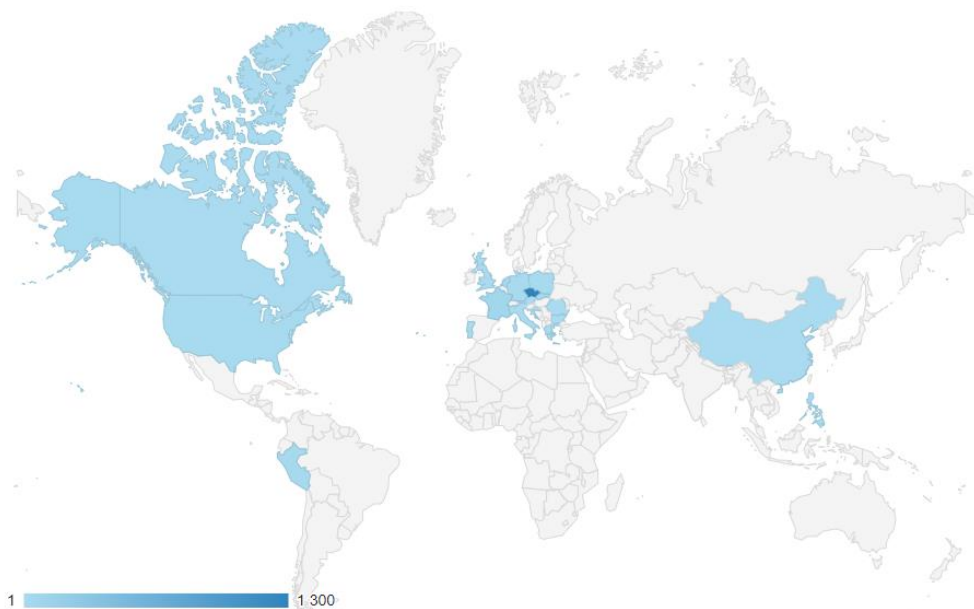
některé rozšíření pro prohlížeč Google Chrome. Google Analytics rovněž disponuje množstvím API pro hlášení a konfiguraci určených pro vývojářské účely:

- **Core Reporting API** – umožňuje dotazování na dimenze a metriky při vytváření přizpůsobených přehledů;
- **Embed API** – možnost snadno vytvářet a vkládat „dashboard“ (tj. panel s vybranými údaji) na webové stránky třetích stran
- **Multi-Channel Funnels Reporting API** – ověřuje zdroje návštěvnosti, které vedou ke konverzi
- **Real Time Reporting API** – zaznamenává aktuálně probíhající činnost na projektech
- **Metadata API** – přístup k seznamu dimenzí a metrik včetně jejich atributů
- **Management API** – zobrazení a správa účtů, vlastností, zobrazení, filtrů, nahrávek, oprávnění atd.

Kartografické aspekty

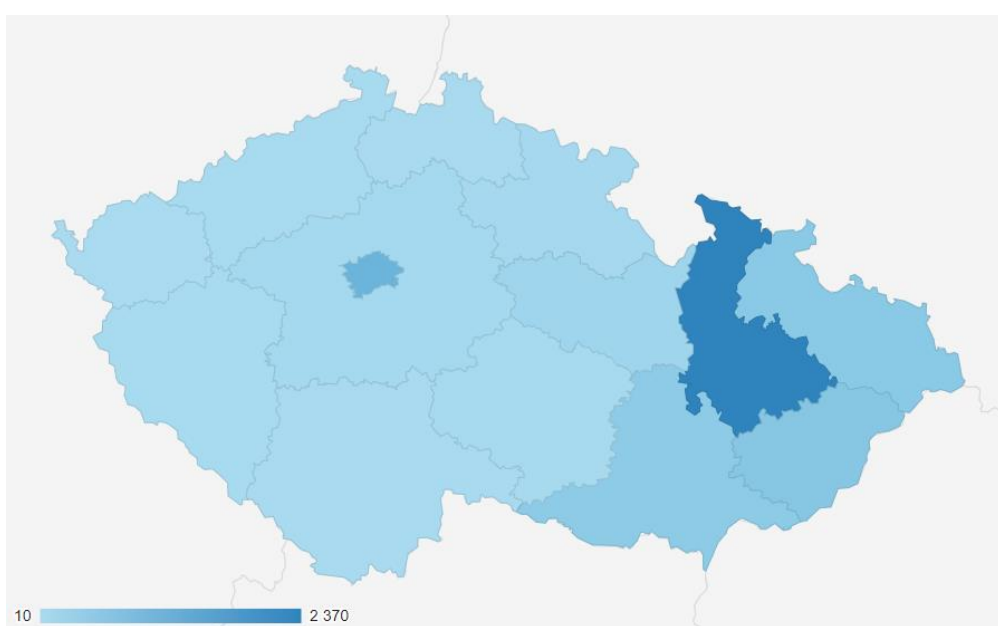
Služba Google Analytics disponuje mapovým výstupem v sekci Lokalita s možností výběru územní jednotky na úrovních Země, Kontinent, Subkontinent a Město. Na prvních třech zmiňovaných úrovních jsou kvantitativní data vizualizována prostřednictvím areálové metody, použitá kartografická metoda je nicméně zvolena nevhodně. Příčinou jsou data v podobě absolutních hodnot, které nejsou přepočteny na jednotku plochy, a tak se nejedná ani o kartogram, ani o nepravý kartogram (takzvaný pseudokartogram). Adekvátním řešením by bylo použití atraktivní metody kartografické anamorfózy, což je dle Voženílka (2011) metoda založená na geometrické přeměně vybraného parametru jevu (např. plochy území) podle určitých pravidel při zachování neměnného tematického prvku mapy (např. tvar území) pro jeho zvýraznění. Topologické atributy mapy (délky, plochy, úhly, tvary apod.) jsou účelně deformovány podle zvoleného matematického pravidla, nebo kartogramu.

Tato mapa světa je zobrazena v Mercatorově válcovém zobrazení a vyšší mírou podléhá generalizaci, i tak ale uživatelům poskytuje rychlý náhled na oblast celého světa. Použitým souřadnicovým systémem je celosvětově uznávaný systém WGS 84. Interaktivita výstupů je omezena na minimum, protože neobsahuje funkci přiblížení/oddálení, dokonce ani žádné měřítko. Jediným interaktivním prvkem je vyskakovací okno s atributy názvu území a hodnotou jevu v jevem zastoupených oblastech. Tento prvek najdeme na všech mapových výstupech. Mapa slouží obecně pro informativní účely, kartografická přesnost tedy není primárním účelem.



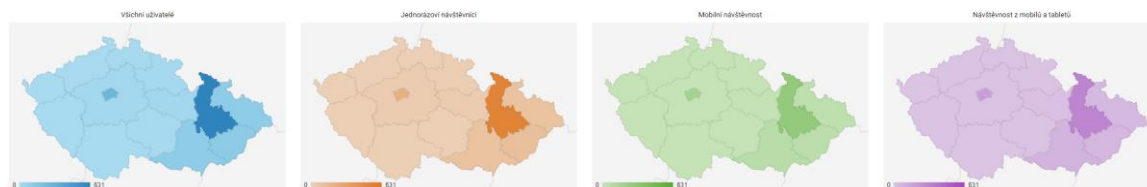
Obr. 7 Mapový výstup počtu návštěv Pevnosti poznání ve světě (Google Analytics)

Při kliknutí na příslušnou oblast dojde k přiblížení konkrétní oblasti na ještě menší územní celky (v ČR až na úroveň NUTS 3 a LAU 2). Česká republika zde má novodobý název Czechia. Tato úroveň interaktivity, tedy možnost přiblížení na úroveň států, respektive na ještě menší jednotky, rovněž podléhá částečné generalizaci. I zde je k vizualizaci prostorových dat použita chybná kartografická metoda, která je barevnou stupnicí rozdělena na čtyři třídy. U měst je zastoupena metoda kartodiagramu. Hustota jevu je znázorněna barevnou stupnicí, kde nejtmaší odstín představuje jeho největší hustotu. Naopak územní jednotka s nulovou hodnotou jevu je vyobrazena světle šedou barvou. Samotná legenda barevné stupnice, respektive klasifikace hodnot, je nevhodná a zkrslující, neboť zde hodnoty nejsou rozděleny do intervalů, ale podle každé hodnoty zvlášť.



Obr. 8 Mapový výstup počtu návštěv Pevnosti poznání v krajích ČR (Google Analytics)

V případě porovnávání více dimenzí, tj. například všech uživatelů a vracejících se návštěvníků, dochází u dalších výstupů ke změně barevné stupnice. Nejvíce můžeme porovnávat až čtyři dimenze zároveň, přičemž použitými barevnými stupnicemi jsou modrá, oranžová, zelená a fialová. Tyto barvy nejsou jakkoliv rušivé, přechod mezi třídami je dobře viditelný. Dimenze mohou být využity na všech zmiňovaných úrovních, nicméně v rámci celého světa a měst se zobrazí specifický prvek lupy, který zlepšuje přehlednost dané lokality. Mapa na domovské stránce Google Analytics je v závislosti na její velikosti nejvíce generalizovaná. Estetickým prvkem je přítomnost vlajek při najetí myši na konkrétní oblast.



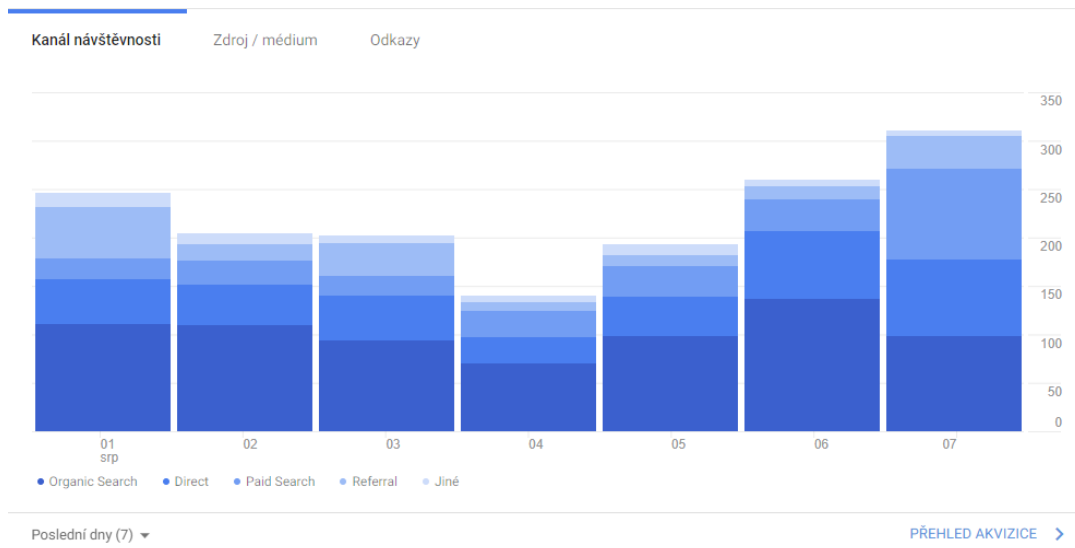
Obr. 9 Porovnání čtyř dimenzí (Google Analytics)

Google postavil mapové výstupy na svém vlastním řešení Google Maps API, což je rozhraní pro vkládání výstupů Google Maps na jiné webové stránky. Projekt byl spuštěn v roce 2005 a pokud uživatel splňuje podmínky, tj. web, na němž se bude API nacházet, je veřejně přístupný, neúčtuje si žádné poplatky a nepřekročí 25 000 přístupů za den, je zdarma pro komerční využití (Nétek, 2018). Nástroj Google Analytics však disponuje odlišnou verzí, co se technického, funkčního a vizualizačního hlediska týče. Například je obsažena upravená a zjednodušená symbologie (šedý podklad, bílé hranice atp.) Jakákoliv úprava mapy v rámci Google Analytics již není možná. GA využívá Google Maps API ve verzi 3, která již vyžaduje unikátní API klíč. Ze zdrojového kódu lze vyčíst, že Google Maps API prostřednictvím parametru „library=visualization“ dokonce umožňuje vkládat heat mapy.

Uživatelské a vizualizační aspekty

Z hlediska uživatelských aspektů patří služba Google Analytics mezi velice povedené, co se uživatelsky přívětivého, respektive intuitivního prostředí a estetičnosti týče. K bílo-šedému pozadí byla zvolena kontrastní modrá barva, a tak vystupují do popředí všechny důležité informace již na domovské stránce. Přechody mezi jednotlivými odstíny modré jsou navíc dobře vidět, Google zároveň použil svoje písmo Roboto v několika řezech pro dobrou čitelnost údajů. Web disponuje rozbalovací hlavní nabídkou na levé straně, přičemž zde jsou jednotlivé sekce i podsekce zvýrazněny oranžovou barvou. Přepínání mezi účty je umístěno v levém horním rohu.

Téměř všechna základní data na hlavní stránce GA jsou rozdělena do takzvaných karet (respektive widgetů), které interaktivní formou záložek nabízejí možnost přepínání sledovaných metrik včetně výběru časového rozmezí a sdružují informace podobného charakteru. Tyto informace jsou dále vizualizovány různými metodami, kdy z větší části převažuje vizualizace pomocí grafů (zastoupeny jsou liniový, výsečový a sloupcový typ). Některé údaje byly zpracovány formou „gridu“, například metrika *Počet uživatelů podle denní doby*, prostorové informace tvoří mapový výstup velkého měřítka, viz výše. Nevyhovuje-li uživateli výchozí rozložení domovské stránky, může si v sekci *Prizpůsobení* vytvořit vlastní přehledy a rozmístit widgety dle svého uvážení. Právě tato možnost úpravy dle vlastních požadavků je mezi analytickými pracovníky považována za velmi pokročilou a oblíbenou.



Obr. 10 Obrazovka Jak získáváte uživatele? (Google Analytics)

Google Analytics umožňuje vytvoření až 100 účtů a 50 služeb, kdy jeden účet představuje jeden webový projekt a služba zpravidla jeho subdoménu. Více služeb stejného účtu (webového projektu) je tvořeno stejným UA kódem a liší se číslem za pomlčkou. Tyto webové projekty mohou být samozřejmě spravovány více uživateli, kterým jsou v sekci Správce udělována práva k úpravě, spolupráci, čtení a analýze, a kompletní administrátorská práva. Správa uživatelů je řízena prostřednictvím účtu Google, respektive e-mailové adresy Google.

Analytické aspekty

Nástroj Google Analytics je možné provázat s dalšími Google službami, jež jsou zaměřeny převážně na zvýšení efektivity při práci s naměřenými daty. Mezi tyto služby patří AdSense, AdWords a Ad Exchange, jejichž prostřednictvím může uživatel svůj webový projekt zpeněžit na základě relevantních reklam a inzercí. Následující služby mohou díky propojení s účtem Google Analytics pomoci analytikům v mnoha obchodních i analytických faktorech.

- **BigQuery** – nástroj pro vývojáře, provádějící dotazy na rozsáhlé datové sady
- **DoubleClick Bid Manager (DBM)** – řešení umožňující velkým zákazníkům nakupovat na RTB burzách (prodej bannerových reklam v aukci)
- **DoubleClick Campaign Manager** – řešení pro správu a zobrazování reklam
- **DoubleClick Search** – nástroj poskytující reklamní technologie k vytváření, fungování a správě digitální inzerce
- **Volání postback** – funkce umožňující informovat jiné reklamní sítě než Google v případě, kdy je zjištěna konverze v podobě instalace nebo přímého odkazu
- **Search Console** – služba umožňující zjistit, jak uživatelé nacházejí webový projekt pomocí Vyhledávání Google

aplikace, zahrnující desítky inzertních služeb a real-time zpracování s podrobnějšími technickými informacemi. V rámci pokročilejších analýz existuje možnost implementace Metrica API, jež dovoluje uživateli ještě hlubší správu metrik a nastavení přístupových práv bez použití webového rozhraní. Pro přístup je potřeba vygenerovat unikátní Yandex.OAuth klíč. Zmiňované rozhraní API se dle oficiální dokumentace Yandex (2018) skládá z následujících prvků:

- **Management API** – správa počítadel, cílů, filtrů ve smyslu jejich vytváření a úpravy parametrů, udělování přístupových práv mimo webové rozhraní apod.
- **Reports API** – porovnávání segmentů, metrik
- **API compatible with the Google Analytics Core Reporting API** – API pro podporu GA Core Reporting API s omezeným výběrem metrik
- **Logs API** – správa a prohlížení surových dat (log file soubory)

Kartografické aspekty

Mapový výstup služby Metrica je umístěn v sekci Reports, Geography a je určen primárně pro oblast Ruska a sousedních států Ukrajina, Bělorusko a Turecko. Právě pro tyto státy je dostupná možnost podrobnějšího prohlížení prostřednictvím tlačítek pod mapou, díky čemuž jsou data vizualizována také v nižších správních celcích. Ostatní oblasti jsou zobrazeny pouze v mapě celého světa a postup na nižší úroveň již není možný. Mapa je zobrazena v Mercatorově zobrazení, přičemž vyšší mírou podléhá generalizaci ve všech měřítcích a ve všech územních jednotkách. Použitým souřadnicovým systémem je standardní WGS84. Přítomnost interaktivních prvků mapy se projevuje už při najetí myši na konkrétní oblast zastoupenou jevem, a to zobrazením vyskakovacího okna s hodnotou daného jevu (např. počet návštěvníků za určitou časovou jednotku). Mapový výstup je rovněž obohacen o posuvné měřítko, kterým se uživatel může přibližovat či oddalovat, a také o prvek určený pro posouvání v mapovém poli.



Obr. 12 Mapový výstup počtu návštěv Pevnosti poznání (Yandex Metrica)



Obr. 13 Mapový výstup počtu návštěv Pevnosti poznání v oblasti Ruska (Yandex Metrica)

Stejně jako Google Analytics vizualizuje nástroj Metrica kvantitativní data pomocí areálové metody, i zde je však kartografická metoda zvolena zcela nevhodně v závislosti na charakteru prostorových dat, která nejsou přepočtena na plošnou jednotku. Mapa má spíše informativní účel. Doporučeným řešením je již výše zmiňovaná kartografická

anamorfóza, nebo kartogram. Zavádějící je rovněž vizualizace dat ve zmiňovaných podrobnějších oblastech Ruska, Ukrajiny, Běloruska a Turecka, která je v kontextu kartografických metod také chybně zvolena, jelikož se jedná o absolutní hodnoty. Z hlediska barev je v mapě použita barevná stupnice, kde zelená barva představuje oblasti s největším zastoupením daného jevu a červená barva s jeho nejmenším zastoupením. Území s nulovou hodnotou jsou zobrazena odstínem šedé barvy. Přechody mezi třídami jsou dobře viditelné, a to i přes zvolenou legendu barevné stupnice s plynulým přechodem. Klasifikace hodnot legendy je však, podobně jako u služby Google Analytics, velmi zkreslující.

Mapový výstup analytického nástroje Metrica je postaven na Yandex Maps API, tedy na vlastním alternativním řešení ruské korporace Yandex. Podkladem pro mapu jsou volně dostupné vrstvy projektu OpenStreetMap, jenž je formou copyrightu zmíněn v tiráži mapového pole. Těmito vrstvami jsou státy celého světa, v případě čtyř vybraných států pak nižší správní celky. API je v případě dodržení smluvních podmínek poskytováno zdarma, přičemž těmito podmínkami se rozumí:

- Umístění dobře viditelného loga Yandex do mapového pole,
- Nemožnost spravovat či uchovávat data,
- Mapa může být vložena na zdarma dostupný web bez nutnosti registrace či pozvání,
- API je zakázáno využívat pro sledování zaměstnanců a vozidel
- Méně než 25 000 požadavků denně v případě JavaScript API a Geocoder HTTP API

V případě zvolení placeného tarifu se cena odvíjí od denního počtu požadavků na mapový server: v nejlevnějším případě uživatel za 1 000 požadavků denně zaplatí 120 000 rublů ročně, v přepočtu přibližně 43 000 Kč. Je-li počet požadavků na velkých webech větší než 100 000 denně, uživatel ročně zaplatí až 360 000 Kč. Pakliže uživatel chce pro vlastní potřeby shromážděná data ukládat a pracovat s nimi, cena se navýší o dalších 500 000 rublů za rok (Yandex, 2018).

Uživatelské a vizualizační aspekty

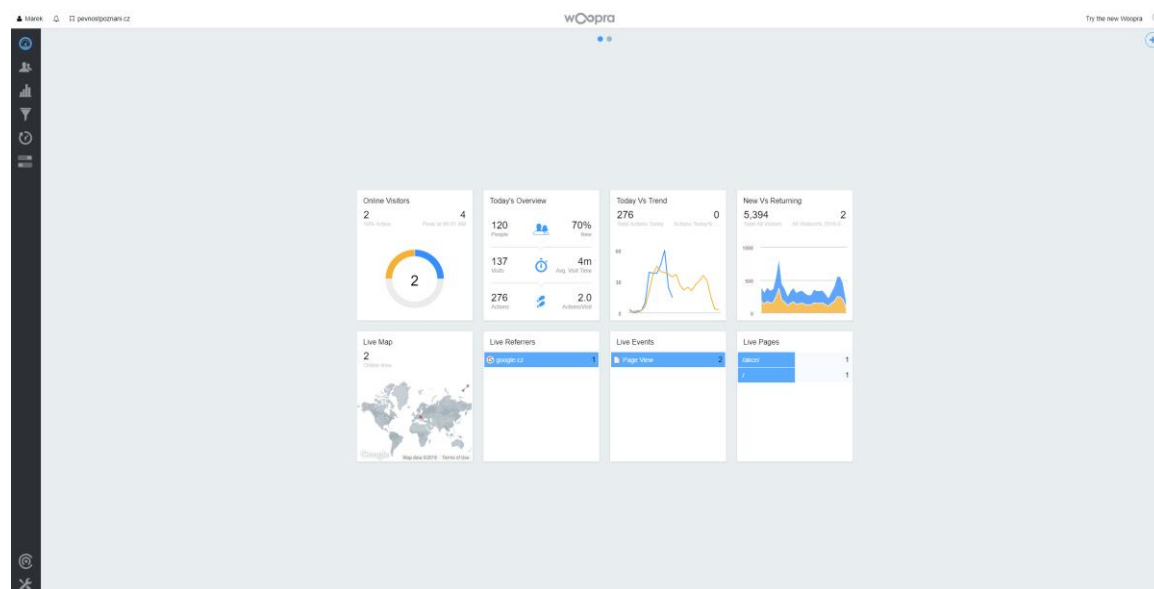
Nástroj Metrica disponuje velice pohodlným pracovním prostředím s logicky uspořádanými přehledy a vhodně zvolenými vizualizačními metodami naměřených dat. Přestože Yandex Metrica nemá zcela definovaný vizuální styl, barevná symbologie není nijak rušivá, naopak je příjemná pro používání. Levá strana webového rozhraní je obsazena rozbalovací hlavní nabídkou, kde jsou umístěny nejdůležitější prvky nástroje rozdělené do menších podsekcí. Kontrastní barvou je oranžová. Přepínání mezi vícero účty lze docílit kliknutím na logo Yandex Metrica umístěné v levém horním rohu.

Naměřená data jsou z velké části vizualizovaná formou výsečových, sloupcových a líniových grafů, přičemž každý z grafů disponuje interaktivním vyskakovacím oknem při najetí na libovolnou hodnotu. Hlavní stránka je dále doplněna o možnost vytvoření vlastních widgetů, které si uživatel prostřednictvím knihovny může jakkoliv přemísťovat, přidávat či odstraňovat dle svých představ. Kliknutím na widget je uživatel přesměrován na jeho stránku, kde jej může důkladně prohlížet. Pro tyto účely jsou k dispozici tabulky v dolní části webového prostředí, jež jsou naplněny veškerými potřebnými informacemi, a pro potřeby uživatele je dostupná funkce vypnutí některé z hodnot pro lepší srovnání.

Konkrétní počet webových projektů není ve službě Metrica specifikován. Správa více uživatelů, kteří mohou analyzovat data konkrétního projektu, se provádí v administrativní sekci. Tito uživatelé jsou přidáni na základě existujícího Yandex účtu, a mohou jim být přidělována práva k prohlížení a úpravě.

5.3 Woopra

Analytický nástroj Woopra se primárně zabývá analýzou jednotlivých návštěvníků webových stránek, čímž poskytuje velice přesné zpracování dat v reálném čase pro účely pochopení chování uživatele za účelem optimalizace webu. Přestože má služba Woopra poměrně malý tržní podíl, v minulosti byla mnohokrát vyzdvižována různými internetovými portály, a to právě kvůli propracovanému real-time faktoru, kterým Woopra vyplňuje značnou mezeru na trhu.



Obr. 14 Webové rozhraní služby Woopra

Technické aspekty

Woopra sleduje a měří webové projekty prostřednictvím metody značkování stránek, tedy využitím JavaScript kódu. Prostorová data jsou získávána skrze IP adresu, a na rozdíl od ostatních služeb nabízí Woopra možnost nahlédnutí na celou strukturu IP adresy u jednotlivých uživatelů. V rámci získaných informací je pak administrátor schopen pozorovat, jak se konkrétní návštěvník na webu choval, tj. jaké stránky navštívil, na jaké stránce web opustil, z jakého zařízení na web přistupoval a podobně. Díky souborům cookies nástroj ukazuje i dny, kdy návštěvník na web přišel, na základě čehož jednoduše rozlišuje zcela nové a vracející se návštěvníky.

Naměřená data lze exportovat ve zvoleném časovém období do formátů CSV, CSV (raw), PDF a HTML. Ačkoliv existuje mnoho funkčních řešení, vývojáři služby Woopra navrhli vlastní, sloupcově orientovaný databázový systém, založený na konkrétní potřebě okamžitě analyzovat právě naměřené údaje v rozmezí 100–200 milisekund. Klíčovou roli sehrálo navržnutí vlastního formátu pro způsob zápisu dat, Flat Map (FlaM), jenž je znatelně rychlejší než doposud známé alternativní formáty jako JSON nebo MessagePack (Jasani, 2017). Woopra je velice otevřeným analytickým nástrojem, který poskytuje značné množství API pro pokročilou konfiguraci specifických požadavků správců webových projektů:

- **Server and Mobile Tracking SDK's** – SDK jsou malé balíčky s funkcemi napsanými v různých programovacích jazycích (PHP, Node.js, Java, Django, Python, Ruby atp.), v nichž je zahrnuto jádro pro měření údajů
- **HTTP tracking API** – rozhraní pro sledování událostí a návštěvníků a jejich dalších atributů
- **HTTP API**
- **HTTP Reporting API** – rozhraní pro dotazování naměřených údajů
- **HTTP Woopra Files API** – rozhraní pro správu dat (včetně schémat, přehledů, štítků atd.)
- **HTTP Import/Export API** – rozhraní pro konfiguraci import a export funkcí
- **Profiles/Search API**

Webové projekty založené na platformě WordPress mohou využít plugin pro zobrazení užitečných informací na hlavní stránce redakčního systému (například Woopra Analytics Plugin). Kromě toho je služba Woopra dostupná ve formě mobilní aplikace pro operační systémy Android a iOS. Webová stránka nástroje Woopra je zabezpečena skrze https protokol.

Kartografické aspekty

Mapový výstup služby Woopra vizualizuje návštěvníky v reálném čase bodovou metodou na mapě celého světa, oproti ostatním analytickým nástrojům ale nenabízí kvantitativní vizualizační metody pro srovnání jednotlivých metrik, ani srovnání nasbíraných dat ve zvoleném časovém období. Živá mapa je dostupná jako widget na domovské stránce, přičemž po jejím zvětšení na celou obrazovku ukazuje všechny současné návštěvníky jako červený bod v mapovém poli. Interaktivita mapového výstupu je omezena na minimum – při kliknutí na červený bod nedojde k zobrazení vyskakovacího okna s podrobnějšími údaji, uživatel se smí po mapě pouze pohybovat, nebo přibližovat a oddalovat pomocí tlačítka na myši. Podkladová mapa využívá odstíny šedé barvy pro znázornění výškopisu, kdy tmavší barva znázorňuje oblasti s vyšší nadmořskou výškou. Zajímavostí je přítomnost názvů států pouze pro území Severní a Jižní Ameriky a také pro jižní část Asie, nicméně ostatní země byly ponechány bez popisu. Jádro mapy použité ve službě Woopra představuje Google Maps s pozměněným vzhledem, jak bylo zmiňováno výše.



Obr. 15 Mapový výstup právě aktivních uživatelů (Woopra)

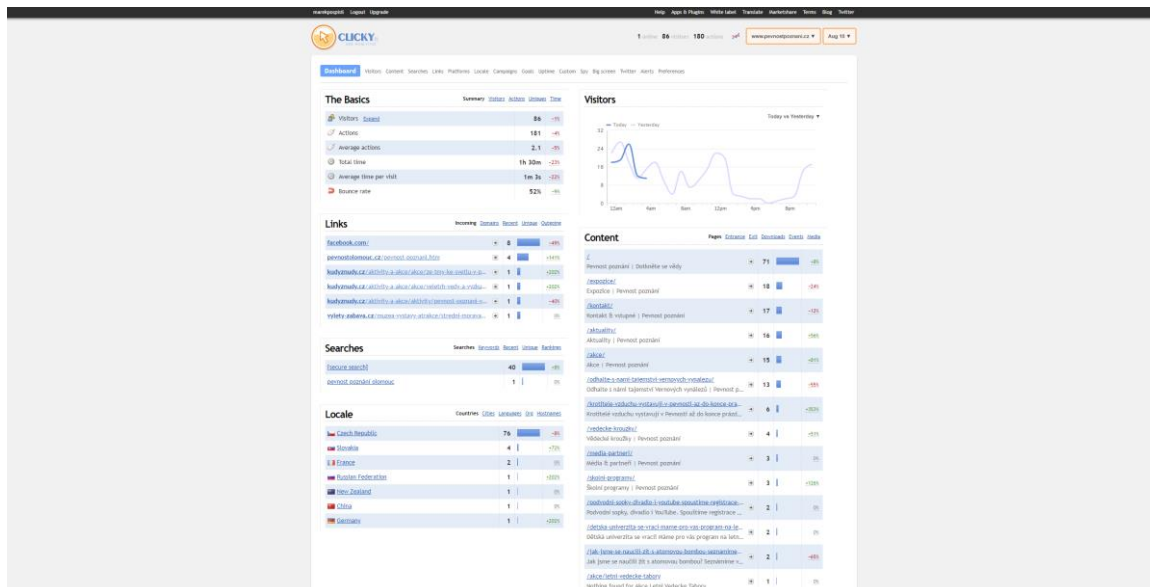
Uživatelské a vizualizační aspekty

Služba Woopra v nedávné době představila nové webové rozhraní, které je stále ve fázi veřejného beta-testování. Klasické prostředí služby je velmi snadné na orientaci, čemuž napomáhá možnost přidávat, odstraňovat a přemísťovat všechny widgety na hlavní stránce podle vlastních potřeb. Pokud je překročen počet osmi widgetů na pracovní ploše, vytvoří se další oddíl, jenž je dostupný skrze posuvný prvek umístěný pod logem Woopra. Prostor je tvořeno nabídkou umístěnou na levé straně, po kliknutí na konkrétní sekci se zobrazí podrobnější možnosti. Z hlediska designu byly zvoleny kontrastní barvy oranžová a modrá. Služba Woopra vizualizuje naměřená data metodou grafických výstupů, konkrétně liniovými a výsečovými grafy. Najede-li uživatel myší na libovolný graf, zobrazí se vyskakovací okno s hodnotou, na niž se uživatel nachází. Každá z metrik je pak detailně popsána atributy a hodnotami v tabulce ve spodní části webového rozhraní.

Přidělování práv dalším uživatelům se provádí v administrativní sekci projektu, a to na základě e-mailové adresy jednotlivce. Tito uživatelé mohou být zařazeni do určité skupiny dle svého zaměření (např. analytik, vývojář), přičemž tyto skupiny jsou zpravidla spravovány hlavním administrátorem, který má mimo jiné možnost určit, jaká data mohou a nemohou ostatní uživatelé vidět.

5.4 Clicky Analytics

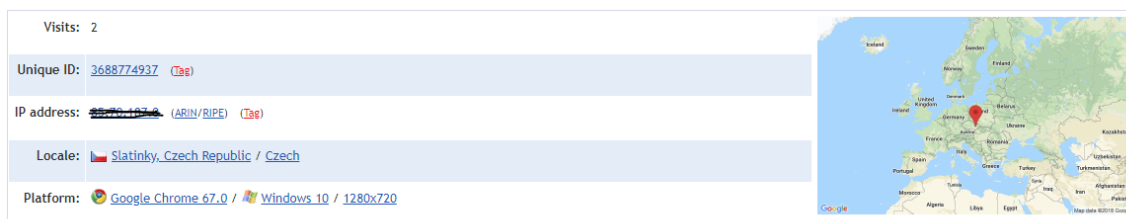
Služba Clicky patří mezi nejstarší, v této oblasti služeb se pohybuje již přes deset let. V současnosti využívá tento nástroj ke sběru a analýze dat více než milion webových projektů, přičemž toto číslo každým měsícem pravidelně roste. Přestože se v současnosti analytici zaměřují spíše na modernější řešení, je Clicky stále mocným nástrojem ve sféře analytických nástrojů.



Obr. 16 Webové rozhraní služby Clicky Analytics

Technické aspekty

Nástroj Clicky Analytics využívá rovněž implementaci měřicího JavaScript kódu na každé stránce. Mimo jiné je obohacen o funkci nahlížení do elementárních údajů, správci webového projektu tedy mají přístup k celé struktuře IP adresy jednotlivých uživatelů včetně jejich přibližné polohy, počtu akcí, operačního systému, času stráveného na webové stránce atd. Každý uživatel pak disponuje unikátním identifikátorem, na základě čehož je nástroj Clicky s pomocí cookies schopen rozpoznat nové a vracející se návštěvníky. Tyto hodnoty ID lze značkovat dle potřeb, například zakázat jejich sledování kvůli eliminaci sledování návštěv administrátorů a dalších interních pracovníků, nebo je přejmenovat.



Visits by this user in the last 7 days

Select a date or date range to view all visits from a different or larger period.

Aug 10 2018 4:32am	85.70.187.0	3 actions	1m 7s	google.cz [secure search]	●
Aug 8 2018 2:06pm	85.70.187.0	10 actions	4m 49s	google.cz [secure search]	

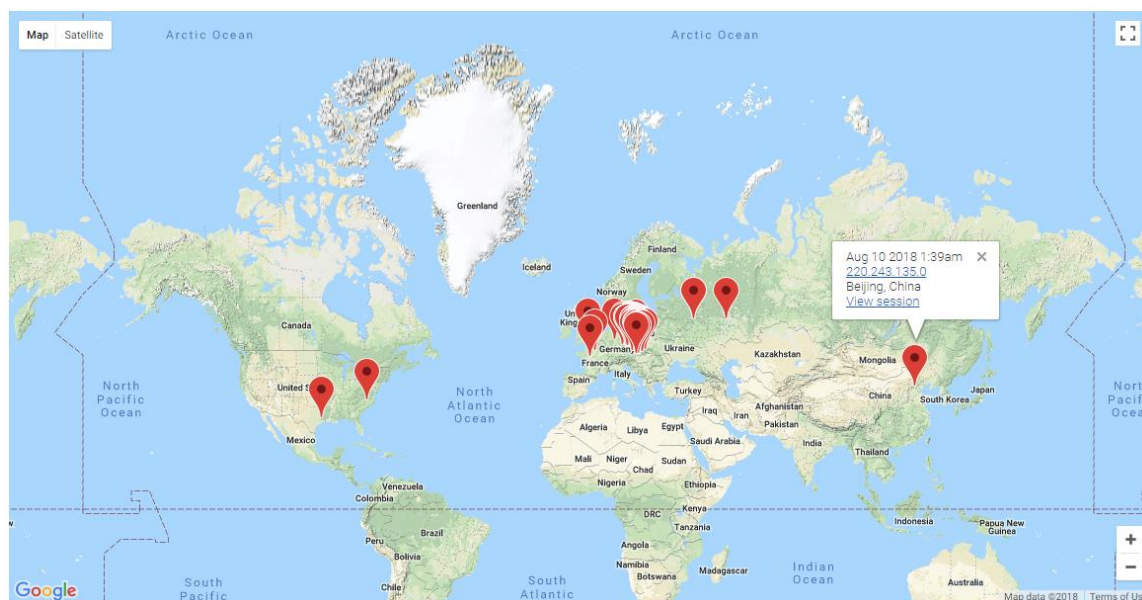
Obr. 17 Nahlížení na úroveň jednoho uživatele (IP adresa je přeškrtnuta z důvodu ochrany osobních údajů)

Seznam uživatelů obsahuje prvek zpracování dat v reálném čase, a to formou zelené tečky s významem „online“ u právě aktivního návštěvníka. Počet právě aktivních uživatelů lze vidět v horní části webové stránky. Clicky Analytics na rozdíl od ostatních služeb blíže nespécifikuje svůj systém řízení báze dat. Shromážděná data mohou být exportována prostřednictvím tlačítka do souborů ve formátech CSV, XML a PDF, přičemž některé informace nabízí podporu pouze pro soubor PDF.

Vývojářská sekce nástroje popisuje rozhraní Analytics API určené ke konfiguraci a extrahování naměřených dat do několika běžných formátů v rámci potřeb vlastní aplikace. API přijímá požadavky prostřednictvím standardní HTTP GET metody a může reagovat pomocí formátů XML, JSON, CSV a serializovaného PHP. Služba Clicky je dostupná ve formě pluginů pro celou řadu webových redakčních systémů (WordPress, Drupal, Joomla atd.) a také jako mobilní aplikace pro platformy Android, iOS a Windows Phone. Kromě toho ji mohou uživatelé využívat na desktopových operačních systémech Windows a Mac OS, nebo jako rozšíření v prohlížeči Google Chrome. Webové rozhraní je zabezpečeno https protokolem.

Kartografické aspekty

Mapový výstup naměřených dat je umístěn v sekci Locale. V průběhu bakalářské práce byl výstup nahrazen jiným, přičemž ten původní se nápadně podobal výstupu z Google Analytics. V současnosti Clicky Analytics vizualizuje data prostřednictvím bodové metody (poslední 100, 500 a 1000 uživatelů). Jednotliví uživatelé představují červený bod v mapě, kliknutím na něj se zobrazí vyskakovací okno s doplňujícími informacemi (je zahrnuta i IP adresa). V levém horním rohu Clicky disponuje tlačítkem pro změnu podkladové mapy s možností vypnutí výškopisu a popisu. Mapa je rovněž obohacena o prvek pro přiblížení a oddálení v pravém dolním rohu, případně zvětšení na celou obrazovku. Tato prostorová data ale nelze vizualizovat v libovolném časovém období, možnost zobrazení více map pro srovnávání metrik také zcela chybí. Clicky Analytics vizualizuje prostorová data prostřednictvím Google Maps API.



Obr. 18 Mapový výstup služby Clicky Analytics

Uživatelské a vizualizační aspekty

Webové rozhraní služby Clicky Analytics je funkční a snadné na orientaci, nicméně vzhledem k současným trendům ve webovém designu se řadí mezi poměrně zastaralé. Z hlediska použitých vizualizačních metod shromážděných dat nabízí hlavní stránka vše potřebné – zcela základní údaje (tj. návštěvy, počet akcí apod.), nejvíce navštěvované stránky, zdroje návštěvnosti a také srovnání počtu návštěvníků (graf závislosti počtu návštěvníků na čase), a to formou porovnání aktuálního dne s jiným datem (včera, před týdnem, týdenní průměr atd.). Najetím myši na část křivky se otevře vyskakovací okno s konkrétní hodnotou počtu návštěvníků v daný čas. Tyto metriky jsou rozděleny do skupin podle podobnosti, přičemž disponují možností zobrazení podrobnějších údajů napravo od názvu každé skupiny (například skupina Obsah dále ukazuje údaje o stránkách, vstupních stránkách, výstupních stránkách, počtu stažení, událostech atd.). Jednotlivé hodnoty jsou znázorněny pomocí sloupcových grafů, které ukazují jak hodnotu, tak procentuální nárůst či pokles oproti jinému časovému období. Clicky Analytics neumožňuje posouvat, přidávat ani odstraňovat widgety na hlavní stránce. Mezi kontrastní barvy webového rozhraní patří modrá a oranžová. Nabídka je umístěna nahoře, nikoliv nalevo jako u výše zmiňovaných služeb, nicméně dostupnost důležitých informací není jakkoliv narušena. Stránka využívá responzivní design.

Základní verze Clicky Analytics umožňuje sledovat jeden webový projekt jedním uživatelem. V závislosti na konkrétních požadavcích si pak zákazník vybere jeden z nabízených tarifů, které se liší počtem spravovaných webů, počtem denních návštěv a samozřejmě cenou. Využíváním některého z těchto tarifů se zákazníkovi odemknou prémiové funkce, kam patří například heat mapy, možnost exportu vybraných dat za zvolené období, odstranění reklam a další.

6 VÝSLEDKY HEURISTICKÉ ANALÝZY

Podrobný popis vybraných analytických nástrojů z technického, kartografického, uživatelského a vizualizačního a analytického hlediska sloužil k jejich prvotnímu srovnání. Na základě předchozí kapitoly byly všechny tyto nástroje hodnoceny prostřednictvím heuristické analýzy použitelnosti.

6.1 Hodnocení Google Analytics

Nástroj Google Analytics dosáhl výsledného skóre 85 %. Znatelným problémem je podle hodnocení kategorie Karto & GIS, kde nástroj obdržel poměrně slabé skóre 55 %. To je zapříčiněno především nevhodně zvolenou kartografickou metodou a absencí jakýchkoliv kompozičních prvků, případně pokročilejších funkcí. Vzhledem k charakteru analytických nástrojů, který je spíše informativní, se ale podobné parametry ani nemohou očekávat. Z hlediska jiných logických kategorií (Obecné, Použitelnost, Grafické, Technologické) dosáhla služba Google Analytics dobrých výsledků, viz obrázky níže.

Výsledky								
Skupina aspektů	Bodů	Bodů s koef.	Otázek	Odpovědí	Zodpovězeno	Skóre	Skóre s koef.	Koef součet
Obecné	9	11	9	9	100%	100%	100%	11
Karto & GIS	2	3.5	21	21	100%	55%	57%	23.5
Použitelnost	22	25	24	24	100%	96%	96%	27
Grafické	16	17	16	16	100%	100%	100%	17
Technologické	9	9.5	13	13	100%	85%	83%	14.5
Celkem	58	66	83	83	100%	85%	85%	93

Obr. 19 Výsledek hodnocení Google Analytics

6.2 Hodnocení Yandex Metrica

Analytická služba Yandex Metrica oproti konkurenčním řešení ztrácí na absenci žádaného zpracování dat v reálném čase. Výsledné skóre 82 % je pak primárně způsobeno velkými mezerami v kategorii Karto & GIS, kde Metrica dosáhla pouhých 55 %, a to z podobných důvodů jako výše uvedený nástroj Google Analytics – tedy kvůli nevhodně zvolené kartografické metodě či nepřítomnosti kompozičních prvků v mapovém poli. V rámci použitelnosti a grafických aspektů dosáhla služba dobrých výsledků.

Výsledky								
Skupina aspektů	Bodů	Bodů s koef.	Otázek	Odpovědí	Zodpovězeno	Skóre	Skóre s koef.	Koef součet
Obecné	9	11	9	9	100%	100%	100%	11
Karto & GIS	-8	-5.5	21	21	100%	31%	38%	23.5
Použitelnost	22	25	24	24	100%	96%	96%	27
Grafické	16	17	16	16	100%	100%	100%	17
Technologické	11	12.5	13	13	100%	92%	93%	14.5
Celkem	50	60	83	83	100%	80%	82%	93

Obr. 20 Výsledek hodnocení Yandex Metrica

6.3 Hodnocení Woopra

Nástroj Woopra dosáhl celkového skóre 82 %. Největším nedostatkem byla kategorie Karto & GIS, kde kvůli nepřítomnosti důležitých prvků obdržel jen 50 %. Oproti ostatním nástrojům však využívá vhodně zvolenou kartografickou metodu, nicméně vizualizuje data pouze v reálném čase, a tak neumožňuje srovnávat data v libovolném časovém období. V rámci ostatních kategorií vyčnívá Woopra v možnosti nahlížení do elementárních dat, tj. na kompletní strukturu IP adresy jednotlivých uživatelů.

Výsledky								
Skupina aspektů	Bodů	Bodů s koef.	Otázek	Odpovědí	Zodpovězeno	Skóre	Skóre s koef.	Koef součet
Obecné	9	11	9	9	100%	100%	100%	11
Karto & GIS	-8	-5.5	21	21	100%	31%	38%	23.5
Použitelnost	22	25	24	24	100%	96%	96%	27
Grafické	16	17	16	16	100%	100%	100%	17
Technologické	11	12.5	13	13	100%	92%	93%	14.5
Celkem	50	60	83	83	100%	80%	82%	93

Obr. 21 Výsledek hodnocení Woopra

6.4 Hodnocení Clicky Analytics

Přestože Clicky Analytics patří, co se estetičnosti týče, mezi zastaralé řešení, je velice funkční. Služba dosáhla celkového skóre 80 %, kdy stejně jako předchozí nástroje získala nelichotivé hodnocení v kategorii kartografických a GIS parametrů (52 %). Od ostatních služeb ji značně odlišuje pokročilá možnost sledování konkrétních uživatelů pro pochopení jejich chování na webových stránkách, a to skrze kompletní strukturu IP adresy.

Výsledky								
Skupina aspektů	Bodů	Bodů s koef.	Otázek	Odpovědí	Zodpovězeno	Skóre	Skóre s koef.	Koef součet
Obecné	7	8	9	9	100%	89%	86%	11
Karto & GIS	-4	-3.5	21	21	100%	40%	43%	23.5
Použitelnost	21	24	24	24	100%	94%	94%	27
Grafické	15	15.5	16	16	100%	97%	96%	17
Technologické	11	12.5	13	13	100%	92%	93%	14.5
Celkem	50	56.5	83	83	100%	80%	80%	93

Obr. 22 Výsledek hodnocení Clicky Analytics

V rámci hodnocení analytických nástrojů pomocí heuristické analýzy použitelnosti bylo vyzorováno, že všechny srovnávané nástroje disponují podobnými nedostatky v kategorii mapových výstupů. Těmi jsou obecně nevhodně zvolená kartografická metoda, nepřítomnost měřítka (grafického, číselného) či jiných kompozičních prvků (posouvání, přiblížení/oddálení, vyhledávání, přepínání podkladových map atd.), které by mohly zvýšit efektivitu při analýze shromážděných dat v mapě. Rovněž je zajímavé pozorovat využívání známých mapových řešení v těchto nástrojích, autoři se nepouštějí do experimentů a dávají přednost zaběhnutým službám.

Z hlediska ostatních logických kategorií nedocházelo k velkým odchylkám mezi nástroji, rozdílovým parametrem budiž zpracování dat v reálném čase, případně možnost nahlížení do elementárních údajů (IP adresy návštěvníků) či celková estetika webového prostředí služeb, jež je však vnímána velmi subjektivně. Nejlépe hodnoceným nástrojem je podle použité metody hodnocení Google Analytics, jenž dosáhl výsledného skóre 85 %.

7 PŘÍPADOVÉ STUDIE

Pro pochopení funkčnosti analytických nástrojů byly jejich měřicí kód nasazeny do zdrojového kódu některých webových stránek, ke kterým byl autorovi udělen přístup. Funkčností se rozumí pochopení základních měřených metrik, viz kapitola 3.4 Základní metriky webové analytiky, a také seznámení s ovládním a webovým prostředím každého analytického nástroje, aby bylo možné je v následujících krocích popisovat dle předem stanovených aspektů a následně hodnotit na základě heuristické analýzy použitelnosti.

Autorovi byl pod smlouvou o mlčenlivosti povolen přístup k oficiálním datům Univerzity Palackého, respektive webu www.upol.cz (v práci budou proto čísla uvedena jen zaokrouhleně). Dalšími weby, jejichž data byla sledována, jsou univerzitní projekt www.pevnostpoznani.cz a web Katedry Geoinformatiky www.geoinformatics.upol.cz. Tyto tři weby byly vybrány v závislosti na jejich návštěvnosti, aby mohly být objektivně posouzeny vizualizační metody jednotlivých analytických nástrojů na webech s menší (desítky až stovky návštěv denně) a velkou návštěvností (tisíce až desetitisíce návštěv denně).

Katedra Geoinformatiky (www.geoinformatics.upol.cz)

Webová stránka www.geoinformatics.upol.cz prezentuje Katedru geoinformatiky Univerzity Palackého v Olomouci. Je určena jak pro uchazeče o budoucí studium na tomto oboru, tak pro současné studenty a zaměstnance katedry, kteří na ni mohou dohledat všechny potřebné informace nutné ke studiu, zahrnující například vizuální dokumenty pro tvorbu textových prací či odkaz na e-learningovou platformu Moodle. Web funguje na open-source publikačním systému WordPress (PHP, MySQL).

V rámci statistik navštíví web Katedry Geoinformatiky až několik tisíc lidí měsíčně, přičemž jsou to především vracející se uživatelé (můžeme říci, že současní studenti a pracovníci katedry). Stránky fungují na produktu Windows Server 2008 R2 s PHP ve verzi 7.0.4, v současnosti stále nejsou zabezpečeny https protokolem. Následující tabulka ukazuje základní statistiky za určité časové období. Roční data jsou nedostupná z důvodu výpadku měřicího kódu během této doby, a proto byl zvolen 9 měsíční interval.

	ZA 1 MĚSÍC (1.–31.12. 2018)	ZA 6 MĚSÍCŮ (1.10.2018–1.4. 2019)	ZA 9 MĚSÍCŮ (1.7.2018–1.4. 2019)
Počet uživatelů:	813	6 002	7 204
Počet nových uživ:	594	5 946	7 205
Počet návštěv:	1292	13 628	15 896
Počet návštěv na uživ.:	2,29	2,20	2,21
Zobrazení stránek:	2958	29 934	35 478
Prům. doba návštěvy:	00:01:53	00:01:58	00:01:58
Bounce rate:	51,01 %	52,99 %	52,42 %

Tab. 1 Statistika KGI

Pevnost poznání (www.pevnostpoznani.cz)

Pevnost poznání je dle www.pevnostolomouc.cz „první centrum popularizace vědy a výzkumu na střední Moravě, realizované Univerzitou Palackého v Olomouci. Vedle čtyř stálých interaktivních expozičních z oblasti přírodních a humanitních věd nabízí laboratoře, dílny, ateliéry, studia a multifunkční Laudonův sál.“ Webové stránky www.pevnostpoznani.cz tedy informují všechny zájemce o aktuálním i budoucím dění na Pevnosti poznání, o otevírací době, vstupném a dalších nezbytných údajích. Web funguje na publikačním systému WordPress.

Web Pevnost poznání navštěvuje několik tisícovek lidí měsíčně, primárně se jedná o uživatele z České republiky (konkrétněji z města Olomouc). Dle dlouhodobého šetření trvá průměrná doba návštěvy kolem dvou minut, přičemž uživatelé nejčastěji na web přistupují z desktopového systému Windows, dále následují mobilní platformy Android a iOS. Měřené statistiky z různých období vypadají následovně:

	ZA 1 MĚSÍC (1.–31.12. 2018)	ZA 6 MĚSÍCŮ (1.10.2018–1.4. 2019)	ZA 12 MĚSÍCŮ (1.4.2018–1.4. 2019)
Počet uživatelů:	4 110	25 486	54 442
Počet nových uživ.:	3 592	24 387	53 370
Počet návštěv:	5 265	35 934	80 351
Počet návštěv na uživ.:	1,28	1,41	1,48
Zobrazení stránek:	14 067	96 075	210 198
Prům. doba návštěvy:	00:02:08	00:02:16	00:02:14
Bounce rate:	47,75 %	48,05 %	48,35 %

Tab. 2 Statistiky Pevnosti poznání

Univerzita Palackého (www.upol.cz)

Web www.upol.cz představuje studii s řádově největší návštěvností (tisíce návštěvníků denně), přičemž prezentuje samotnou Univerzitu Palackého. Stránky jsou určeny jak pro uchazeče o studium, současně studenty, tak i zaměstnance z akademických řad. Obsahuje potřebné informace zahrnující univerzitní novinky, informace o jednotlivých fakultách, vizuální styl univerzity apod. Mimo jiné odkazuje na všem potřebný STAG (systém pro administraci studijní agendy).

Webové stránky fungují na redakčním open-source systému TYPO3 a měsíčně je navštíví desítky tisíc uživatelů, jimž trvá návštěva kolem dvou a půl minut. Logicky lze vypočítat, že návštěvnost přes letní měsíce oproti akademickému školnímu roku znatelně klesá. Měřené statistiky z různých období vypadají následovně:

	ZA 1 MĚSÍC	ZA 6 MĚSÍCŮ	ZA 12 MĚSÍCŮ
Počet uživatelů:	+ 30 000	+ 200 000	+ 430 000
Počet nových uživ.:	+ 20 000	+ 200 000	+ 400 000
Počet návštěv:	+ 80 000	+ 800 000	+ 1 700 000
Počet návštěv na uživ.:	+ 2	+ 3	+ 4
Zobrazení stránek:	+ 150 000	+ 1 500 000	+ 3 500 000
Prům. doba návštěvy:	+ 2,5 min	+ 2,5 min	+ 2,5 min
Bounce rate:	+ 50 %	+ 50 %	+ 50 %

Tab. 3 Statistiky UP (hodnoty jsou zaokrouhlené)

Srovnání tří různě velkých webových stránek, kdy web Katedry geoinformatiky považujeme za nejméně navštěvovaný, Pevnost poznání za středně navštěvovaný a web UP za nejvíce navštěvovaný, se logicky odlišuje v počtu návštěv za vybrané časové období. Další statistiky, jako třeba poloha návštěvníků, se liší v závislosti na účelu webové stránky. Například web KGI je nejčastěji navštěvován z města Olomouc z univerzitní internetové sítě, tudíž lze předpokládat, že jeho hlavními uživateli mohou být současní studenti či učitelé, kteří využívají jeho služeb v průběhu studia nebo během výuky (přístup na Moodle atp.). Totéž můžeme uvažovat pro univerzitní web Univerzity Palackého, odkud je přístup do informačního systému studijní agendy STAG.

8 VÝSLEDKY

Tato kapitola aplikuje dosavadní nabyté znalosti z předchozího studia problematiky webové analytiky. Výsledky bakalářské práce s názvem *Webová analytika a možnosti jejího prostorového vyhodnocení* představují splnění všech dílčích cílů, které si autor v počáteční části práce určil. Hlavním výsledkem práce je definice aspektů, stanovení hodnotících metrik a následně jejich reálné ověření prostřednictvím heuristické analýzy s ohledem na prostorovou vizualizaci dat. Autor také považoval za důležité navrhnout grafické řešení, které na základě dříve nabytých informací eliminuje slabé stránky analytických nástrojů. Na základě získaných výsledků, které byly z kartografického hlediska zpravidla chybné, navrhl autor vzorové řešení, viz další kapitola Návrh vzorového řešení.

Definice hodnotících aspektů zpočátku počítala s mnohem větším počtem kritérií, šetření však ukázalo podobnost několika aspektů. Počet těchto aspektů se ustálil na čtyřech, konkrétně na technickém, kartografickém, uživatelském a vizualizačním a analytickém aspektu. Pomocí zmiňovaných kritérií byly jednotlivé analytické služby detailně popsány (viz kapitola 5 Aspekty analytických nástrojů).

Prvním typem výstupů jsou výsledky hodnocení jednotlivých analytických nástrojů prostřednictvím metody heuristické analýzy. Po konzultaci s vedoucím práce Mgr. Rostislavem Nétkem, Ph.D. byly parametry heuristické analýzy upraveny pro potřeby práce. Výsledky ukázaly silné a slabé stránky těchto nástrojů. Z hlediska tématu bakalářské práce se projevily silné nedostatky v oblasti mapových výstupů, respektive ve vizualizaci prostorových informací. Podle výsledku heuristické analýzy a celosvětového využití považuje autor za nejvhodnější službu Google Analytics.

8.1 Doporučení

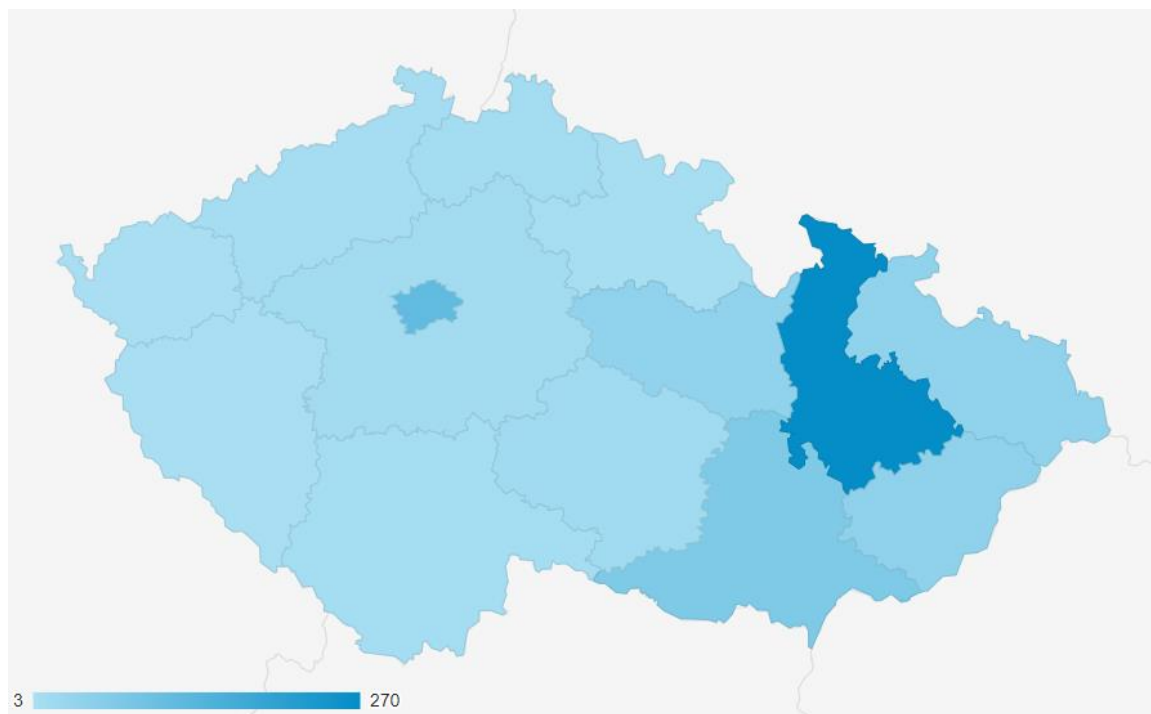
Práce nastiňuje důležitost definování patřičných aspektů, na jejichž základě je doporučeno analytické nástroje popisovat či následně hodnotit. Výsledkem práce je tedy kromě definice těchto aspektů také stanovení hodnotících metrik a jejich reálné ověření metodou heuristické analýzy. Vzhledem ke všem nedostatkům, které byly v mapových výstupech analytických nástrojů zjištěny, existuje celá řada doporučení pro jejich vylepšení.

Důležitým faktorem je bezesporu správná interpretace naměřených dat. V kontextu mapové tvorby tedy hovoříme o výběru vhodné kartografické metody (kartogram, pseudokartogram, bodová metoda...) a přítomnosti nezbytných kompozičních prvků – v několika případech totiž zcela chybí legenda či grafické měřítko, které představují nedílnou součást mapových výstupů, podobně jako korektně použitá legenda naměřených hodnot. Některé z kompozičních prvků, např. absence přepínání podkladových vrstev, představují velmi subjektivní pohled jednotlivců na danou problematiku. Přestože autorům analytických služeb nelze nepřítomnost možnosti vyčítat, v mnohých případech by mohla minimálně zlepšit orientaci v prostředí, nebo lépe vizualizovat shromážděná data.

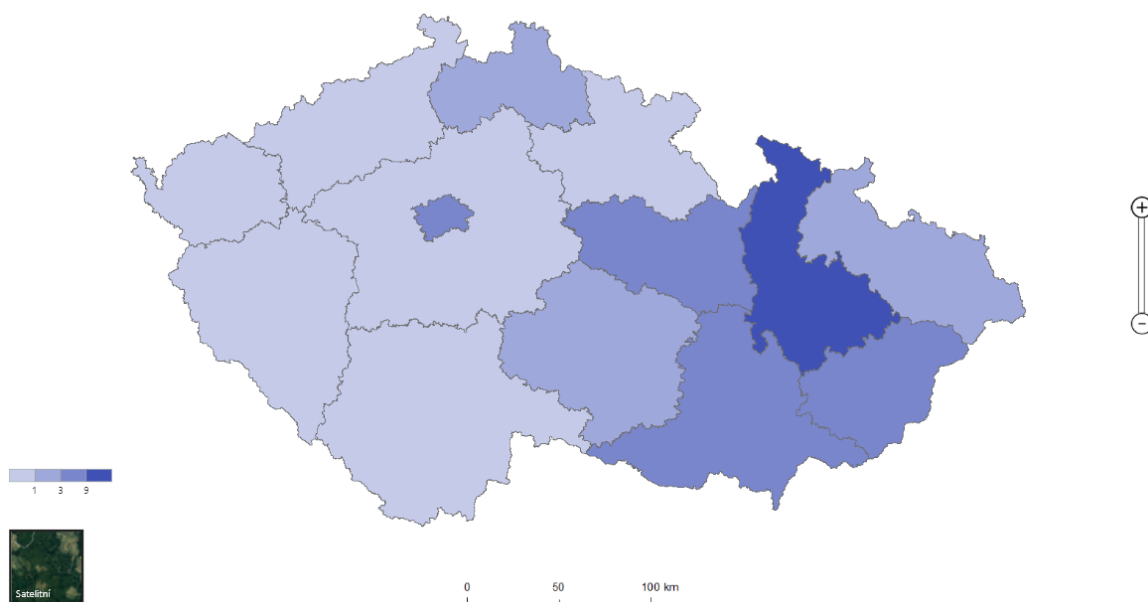
Za stěžejní problém analytických nástrojů lze částečně (v některých případech) považovat práci s absolutními daty, která jsou pro tvorbu mapových výstupů (zejména u metrik, pro jejichž vizualizaci jsou příhodná relativní data) nevhodná. Ačkoliv nejsou mapy primárním účelem těchto nástrojů, jednoduchý vzorec či využití relativních hodnot vztažených na jednotku plochy by mapové výstupy (u některých ukazatelů) bezesporu obohatil.

Bakalářská práce přináší rovněž doporučení grafického rozložení ideálního analytického nástroje (konkrétně se zaměřením na mapové pole), které bylo vypracováno na základě praktických znalostí z rozhraní Google Analytics, Yandex Metrica, Woopra a Clicky Analytics. Tyto znalosti vedly ke zjištění nedostatků a doporučení, od nichž se autor práce odrážel při návrhu koncepčního řešení. Konkrétním příkladem budiž volba vhodného „titulu“ mapového pole, tedy název vizualizované metriky ve vybrané oblasti ve zvoleném časovém období. Uživatelsky přívětivé možnosti byly autorem v konceptu ponechány. Například předdefinované časové období, exportování shromážděných dat do nejpoužívanějších formátů (PDF, XLSX, CSV), nebo přepínání konkrétního webového projektu z každé obrazovky.

Následující dva mapové výstupy ukazují rozdílnost použitých kartografických metod. Oba tyto výstupy zobrazují počet uživatelů webu Katedry geoinformatiky za měsíc prosinec. Výstup nahoře patří službě Google Analytics, zatímco výstup dole ukazuje autorem přepracovaný pseudokartogram počtu uživatelů v krajích ČR vztahený na 1000 obyvatel. Můžeme si všimnout, že ačkoliv výstupy vizualizují data stejného časového období, hodnoty (resp. barevné odstíny) u některých krajů se liší. Vlastní mapový výstup pak disponuje grafickým měřítkem a intervalovým rozdělením legendy, což je dle autora mnohem přehlednější varianta, co se zjišťování hodnot v jednotlivých územních jednotkách týče.



Obr. 23 Mapový výstup – počet uživatelů KGI v krajích ČR za měsíc prosinec podle služby Google Analytics

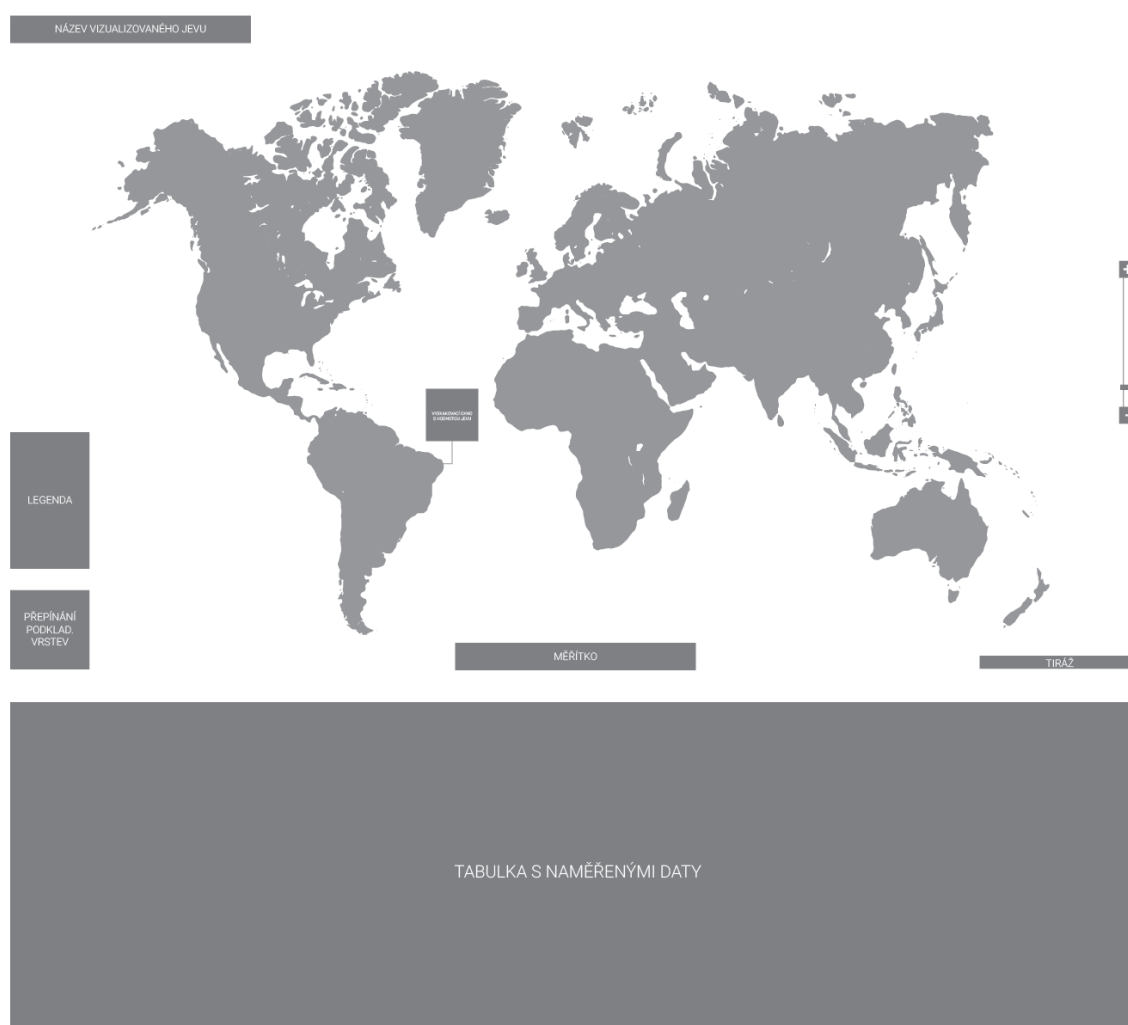


Obr. 24 Mapový výstup – počet uživatelů KGI za měsíc prosinec, přepracovaný pseudokartogram počtu uživatelů v krajích ČR na 1000 obyvatel

8.2 Návrh vzorového řešení

Návrh vzorového řešení vznikl na základě znalostí, které byly nabyty po čas psaní bakalářské práce. V předchozích kapitolách bylo mnohokrát zmíněno, že samotné mapové pole je ve většině případech nevhodné. Vzorové řešení proto klade důraz zejména na část mapového pole se všemi nezbytnými prvky mapové kompozice, jež ve vybraných analytických nástrojích chyběly. Vzorové řešení rovněž zahrnuje mapový výstup s vhodně zvolenou kartografickou metodou.

První návrh prezentuje optimální kompozici nezbytných prvků – vlevo nahoře název vizualizovaného jevu, dále legendu a přepínání podkladových vrstev, které by rozhodně pomohlo při vizualizaci některých jevů. Vzorové řešení rovněž disponuje vyskakovacím oknem při kliknutí na konkrétní oblast, v němž se zobrazí hodnota vizualizovaného jevu, dále grafickým měřítkem, tiráží s nezbytnými informacemi o použitých datových sadách a v neposlední řadě prvek pro přibližování a oddalování.

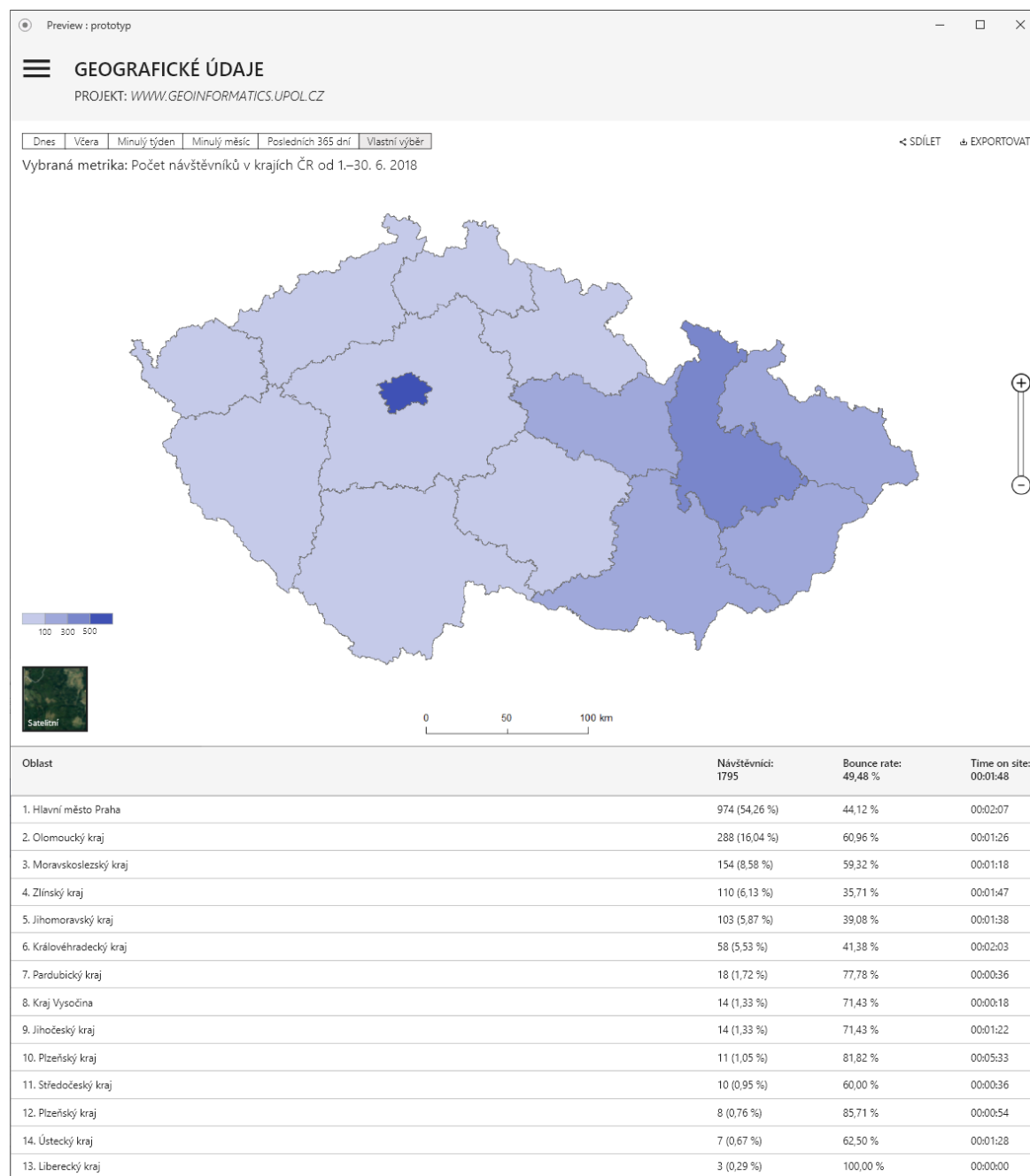


Obr. 25 Wireframe vzorového řešení

Druhý návrh představuje komplexnější aplikaci prvků z prvního návrhu, přičemž již obsahuje vhodně zvolenou kartografickou metodu, a sice pseudokartogram počtu uživatelů v krajích České republiky vztážený na 1000 obyvatel. V případě měřítka bylo zvoleno grafické, protože analytické nástroje jsou kompatibilní se všemi prohlížeči, které umožňují zvětšování či zmenšování obrazovky, případně zobrazují obsah na různě

velkých monitorech a zařízeních, proto nemá smysl využívat číselné měřítko. Modrá barevná stupnice byla zvolena zejména kvůli dobré odlišitelnosti jednotlivých odstínů mezi sebou a rovněž kvůli dobrému kontrastu na monitorech a jiných zařízeních.

Prototyp byl zhotoven ve volně dostupném programu Adobe Xd a je částečně funkční – při kliknutí na určité oblasti mapového výstupu disponuje interaktivními prvky, jako například vyskakovacími okny, možnost přepnutí na satelitní mapu, porovnání jednotlivých časových období atd. Prototyp je dostupný na webových stránkách bakalářské práce.



Obr. 26 Komplexní návrh vzorového řešení

9 DISKUZE

Bakalářská práce se zabývá seznámením a následným hodnocením analytických nástrojů, které byly autorem práce vybrány na základě dostupnosti a celkového využívání na světovém trhu. Prvotním problémem však bylo tyto nástroje vybrat, a to zejména kvůli jejich finanční náročnosti – většina řešení totiž disponuje bezplatnou verzí, která je ale limitována počtem dotazů na server. U zpoplatněných služeb bývají částky zpravidla astronomické a vyplatí se pouze velkým společnostem s finančními prostředky. Kvůli těmto omezením byly tedy vybrány studie, jež splňují limity těchto bezplatných řešení. Při několikaměsíčním sledování studií pak docházelo k mnohým problémům, kdy například z důvodu havárie serveru zcela zmizely měřicí kódy z některých univerzitních projektů (v případě bakalářské práce došlo opakovaně k resetování naměřených dat na webech Katedry Geoinformatiky a Pevnosti poznání).

Přestože se bakalářská práce zabývá prostorovým vyhodnocením dat vybraných analytických nástrojů, mezi primární účely těchto nástrojů patří odlišné výstupy a statistiky. V rámci jejich hodnocení dle heuristické analýzy nepřesáhla ani jedna služba hranici 60 % ve skupině kartografických a geoinformatických aspektů, což je zapříčiněno zcela rozdílným zaměřením analytických služeb. Víceméně všechny mapové výstupy se dají považovat pouze za informativní zpestření naměřených dat, nikoliv za směrodatný výstup splňující veškerá kartografická pravidla. Vzhledem k důležitějším informacím, které tyto nástroje v analytickém kontextu primárně poskytují, je tedy nutné brát tyto výstupy s jistou rezervou. Z různých hledisek je ale nepřijatelné poskytovat uživatelům chybně interpretované mapy, které jsou ve většině případů zkesleny nevhodně zvolenou kartografickou metodou a jinými nežádoucími faktory. Autor se během celé práce zaměřil na eliminaci těchto vlivů.

Webová analytika je v současné době velkým trendem s příznivými vyhlídkami do budoucna. Zejména v oblasti marketingu hovoříme o pojmech jako SEO (Search Engine Optimization) či PPC (Pay Per Click), které značně napomáhají pozici hledané webové stránky ve výsledcích vyhledávání. Mimo jiné tomuto faktu nasvědčuje nedávný vznik obrovského projektu Google Marketing Platform, jenž zařazuje dva doposud velice využívané nástroje Google Analytics 360 a DoubleClick Digital Marketing. Dle slov Brada Bendersa (2018) „poskytne relevantnější a efektivnější marketingový nástroj, přičemž zajistí respektování soukromí zákazníků a jejich kontrolu nad vlastními daty.“ Konkrétně kontrolou nad vlastními daty se zabývá často probírané obecné nařízení na ochranu osobních údajů (známe spíše pod zkratkou GDPR). Oblast webové analytiky je velmi dynamická a každým dnem přináší novinky, nebo naopak omezení.

Autor doporučuje mít nasazený měřicí kód, i když webová stránka primárně žádná data nesbírá. V případě potřeby pak mají správci stránek z čeho čerpat. Během psaní bakalářské práce několikrát došlo k výpadku měřicích kódů, jednou kvůli havárii serveru.

10 ZÁVĚR

Hlavním cílem bakalářské práce bylo zhodnotit prostorové vyhodnocení dat u pokročilých webových analytických nástrojů, které byly vybrány s důrazem na jejich dostupnost a limity pro autorem vybrané studie. V rámci rešeršní části byly vysvětleny obecné informace týkající se problematiky webové analytiky, rovněž byly nastíněny základní metriky, které tyto nástroje měří. V neposlední řadě autor definoval několik aspektů, na jejichž základě byly zjišťovány silné a slabé stránky čtyř vybraných analytických služeb. Ty byly následně vyhodnoceny prostřednictvím heuristické analýzy, jež poukázala na silné nedostatky těchto nástrojů zejména v kontextu prostorového vyhodnocování shromážděných dat, kdy všechny použité nástroje získaly poměrně malý počet bodů vzhledem k absenci důležitých kartografických prvků v mapových výstupech. Rovněž některé kartografické metody byly zvoleny zcela nevhodně, nicméně po konzultaci byla navržena alternativní řešení.

Správnou vizualizaci prostorových dat včetně ideálního prostředí a nezbytných kompozičních prvků mapového pole ukazuje kapitola 8.1 Návrh vzorového řešení. Návrh tohoto vzorového řešení vychází z porovnání, analýz a informací získaných během teoretického i praktického osvojení problematiky v rámci zpracování bakalářské práce i studia geoinformatiky obecně napříč předměty, tudíž kombinuje objektivní i subjektivní poznatky autora.

Diskuze prezentuje dynamický vývoj webové analytiky ve světě. Rovněž upozorňuje na zcela jiný účel těchto nástrojů, které nejsou primárně určeny pro tvorbu mapových výstupů či podobnou vizualizaci prostorových dat. Jak bylo mnohokrát zmíněno, dostupné mapové výstupy slouží k atraktivnímu či alternativnímu znázornění shromážděných informací (podobně jako například tabulky a grafy).

Nelze s jistotou předpovědět, zda se autoři analytických nástrojů budou vůbec někdy zaměřovat na bezchybnost těchto prostorových výstupů, nebo je to z hlediska účelu těchto nástrojů zcela nedůležité. Je však potřeba dbát na správnou interpretaci naměřených dat, v případě bakalářské práce na bezchybnost poskytovaných map. Analytické nástroje by měly uživatelům poskytovat data, která nebudou ovlivněna žádnými faktory – například nevhodně zvolenou kartografickou metodou, jež znehodnocuje celý vzorek naměřených dat. Autor práce se svými doporučeními snažil eliminovat nežádoucí vlivy, vznikající v analytických nástrojích při vizualizaci prostorových dat.

Na základě heuristické analýzy získala největší počet bodů služba Google Analytics, což se ovšem vzhledem k jejímu celosvětovému zastoupení dalo očekávat. Je nicméně zarážející, kolik nedostatků a hrubých chyb (zejména ve vizualizaci prostorových dat) Google Analytics obsahuje. A to zejména s přihlédnutím k tomu, že Google zaštiťuje nejvyužívanější mapovou aplikaci Google Maps. Ostatní analytické nástroje nenabízejí tolik možností a nemají zdaleka takové zastoupení, přesto se v oblasti webové analytiky dají považovat za konkurenční.

Hlavní cíl práce byl splněn realizací všech dílčích částí, které byly definovány v její prvotní fázi. Bakalářská práce jako první přináší zhodnocení analytických nástrojů z různých předem definovaných hledisek, a nastiňuje problémy týkající se využívání nevhodných vyjadřovacích prostředků v mapových výstupech. Tvůrci analytických nástrojů se mohou poučit z častých chyb, nebo načerpat inspiraci z návrhu vzorového řešení v podobě wireframu.

SEZNAM ILUSTRACÍ

- Obr. 1 Příklad měřicího kódu služby Google Analytics
- Obr. 2 Vývojový diagram postupu práce
- Obr. 3 K čemu slouží webová analytika?
- Obr. 4 Webové rozhraní služby Google Analytics
- Obr. 5 Přístup na úroveň jednoho uživatele
- Obr. 6 Zpracování dat v reálném čase
- Obr. 7 Mapový výstup počtu návštěv Pevnosti poznání ve světě
- Obr. 8 Mapový výstup počtu návštěv Pevnosti poznání v krajích ČR
- Obr. 9 Porovnání čtyř dimenzí
- Obr. 10 Obrazovka Jak získáváte uživatele?
- Obr. 11 Webové prostředí služby Yandex Metrica
- Obr. 12 Mapový výstup počtu návštěv Pevnosti poznání
- Obr. 13 Mapový výstup počtu návštěv Pevnosti poznání v oblasti Ruska
- Obr. 14 Webové rozhraní služby Woopra
- Obr. 15 Mapový výstup právě aktivních uživatelů
- Obr. 16 Webové rozhraní služby Clicky Analytics
- Obr. 17 Nahlížení na úroveň jednoho uživatele
- Obr. 18 Mapový výstup služby Clicky Analytics
- Obr. 19 Výsledek hodnocení Google Analytics
- Obr. 20 Výsledek hodnocení Yandex Metrica
- Obr. 21 Výsledek hodnocení Woopra
- Obr. 22 Výsledek hodnocení Clicky Analytics
- Obr. 23 Mapový výstup – počet uživatelů KGI v krajích ČR za měsíc prosinec podle služby Google Analytics
- Obr. 24 Mapový výstup – počet uživatelů KGI za měsíc prosinec, přepracovaný pseudokartogram počtu uživatelů v krajích ČR na 1000 obyvatel
- Obr. 25 Wireframe vzorového řešení
- Obr. 26 Komplexní návrh vzorového řešení

SEZNAM TABULEK

- Tab. 4 Statistiky KGI
- Tab. 5 Statistiky Pevnosti poznání
- Tab. 6 Statistiky UP

POUŽITÁ LITERATURA A INFORMAČNÍ ZDROJE

PETRTYL, Jan. Webová analytika přehledně. Naučte se ji využít!. *MarketingMind* [online]. 18. 2. 2017 [cit. 2018-08-06]. Dostupné z: <https://www.marketingmind.cz/webova-analytika/>

ČECH, Martin. Nástroje webové analytiky. *InFlow* [online]. 6. 6. 2010 [cit. 2018-08-06]. Dostupné z: <http://www.inflow.cz/nastroje-webove-analytiky>

Clicky: Real Time Web Analytics [online]. [cit. 2018-08-06]. Dostupné z: <https://clicky.com/>

Woopra: Pricing [online]. [cit. 2018-08-06]. Dostupné z: <https://www.woopra.com/pricing>
Mediaguru: Mediální slovník [online]. [cit. 2018-08-06]. Dostupné z: <https://www.mediaguru.cz/slovník-a-mediatypy/slovník/klicova-slova/unikatni-navstevnik-unique-user/>

Základní metriky. *Web-Analytics* [online]. 14. 4. 2014 [cit. 2018-08-06]. Dostupné z: <http://web-analytics.wikidot.com/printer--friendly//zakladni-metriky>

Míra odchodu vs. míra okamžitého opuštění [online]. © 2018 Google [cit. 2018-08-06]. Dostupné z: <https://support.google.com/analytics/answer/2525491?hl=cs>

Yandex: Metrica API [online]. [cit. 2018-08-06]. Dostupné z: <https://tech.yandex.com/metrika/doc/api2/concept/about-docpage/>

KARTNER, JUDr. Martin a Mgr. Jiří PROUZA. *Epravo: Právní regulace cookies v České republice* [online]. [cit. 2018-08-06]. Dostupné z: <https://www.epravo.cz/top/clanky/pravni-regulace-cookies-v-ceske-republice-98406.html>

Začíná platit GDPR: Jaké změny přinese? 10 věcí, které byste měli o tajemné zkratce vědět [online]. [cit. 2019-04-14]. Dostupné z: <https://www.info.cz/evropska-unie/zacina-platit-gdpr-jake-zmeny-prinese-10-veci-ktere-byste-meli-o-tajemne-zkratce-vedet-30760.html>

PECKA, Miroslav. *GDPR cookie lišta – vše v jednom* [online]. [cit. 2019-04-14]. Dostupné z: <https://miroslavpecka.cz/blog/gdpr-cookie-lista/>

Článek 4 EU obecné nařízení o ochraně osobních údajů "Definice" [online]. [cit. 2019-04-14]. Dostupné z: <http://www.privacy-regulation.eu/cs/4.htm>

BENDER, Brad. Introducing Google Marketing Platform. *Google* [online]. 27. 6. 2018 [cit. 2018-08-07]. Dostupné z: <https://www.blog.google/products/marketingplatform/360/introducing-google-marketing-platform/>

Čapková, V. (2010) Testování a hodnocení použitelnosti portálu VS, Univerzita Pardubice

Nielsen. J (1994). Heuristic evaluation. New York, USA, John Wiley & Sons

SNOZOVÁ, Martina. Heuristická analýza. *Inflow* [online]. 2013 [cit. 2018-08-08]. Dostupné z: <http://www.inflow.cz/heuristicka-analyza>

Voženílek, V., Kaňok, J., a kol. (2011): Metody tematické kartografie – Vizualizace prostorových jevů. Univerzita Palackého v Olomouci, 216s.9788024427904

Nétek, R., Burian, T. (2018): Free and open source v geoinformatice. Univerzita Palackého v Olomouci. ISBN 978-80-244-5291-3.

JASANI, Hiral. A Sneak Peak into the Real-time Engine of Woopra. *Woopra* [online]. 20. 11. 2017 [cit. 2018-08-08]. Dostupné z: <https://blog.woopra.com/part-2-a-sneak-peak-into-the-real-time-engine-of-woopra-986d223580d1>

O webové analytice. *Wikidot* [online]. 8. 4. 2014 [cit. 2018-08-08]. Dostupné z: <http://web-analytics.wikidot.com/o-webove-analytice>

What can Web Analytics do for me?. *Logaholic* [online]. 2018 [cit. 2018-08-08]. Dostupné z: <http://www.logaholic.com/manual/references/what-is-web-analytics/>

PETERSON, Eric T. What Is Web Analytics. *XML* [online]. 12. 10. 2005 [cit. 2018-08-08]. Dostupné z: <https://www.xml.com/pub/a/2005/10/12/what-is-web-analytics.html>

Usage of traffic analysis tools for websites. *W3techs* [online]. 2018 [cit. 2018-08-08]. Dostupné z: https://w3techs.com/technologies/overview/traffic_analysis/all

Pricing and terms. *Yandex* [online]. 2018 [cit. 2018-08-08]. Dostupné z: <https://tech.yandex.com/maps/commercial/>

HAMEL, Stéphane. Case Study: Accuracy & Precision of Google Analytics Geolocation. *Radical-analytics* [online]. 29. 3. 2016 [cit. 2018-08-08]. Dostupné z: <https://radical-analytics.com/case-study-accuracy-precision-of-google-analytics-geolocation-4264510612c0>

Cookies. *Adaptic* [online]. 2018 [cit. 2018-08-08]. Dostupné z: <http://www.adaptic.cz/znalosti/slovnicek/cookies/>

Clickhouse. *Fedora* [online]. 2018 [cit. 2018-08-08]. Dostupné z:
<https://copr-fe.cloud.fedoraproject.org/coprs/tartare/clickhouse/>

PŘÍLOHY

SEZNAM PŘÍLOH

Popis struktury DVD

Adresáře:

/BP_Pospisil.pdf – text práce

/BP_Pospisil_poster.pdf – poster

- **heuristicka_analyza**
 - heuristicka_analyza_clicky.xlsx
 - heuristicka_analyza_GA.xlsx
 - heuristicka_analyza_woopra.xlsx
 - heuristicka_analyza_yandex.xlsx
- **web**