

Vizuální zpracování dat

Gorazd Ratulovský

Bakalářská práce
2019



Univerzita Tomáše Bati ve Zlíně
Fakulta multimediálních komunikací

Univerzita Tomáše Bati ve Zlíně
Fakulta multimediálních komunikací
Ateliér Digitální design
akademický rok: 2018/2019

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: Gorazd Ratulovský
Osobní číslo: K15007
Studijní program: B8206 Výtvarná umění
Studijní obor: Multimédia a design – Digitální design
Forma studia: prezenční

Téma práce: Vizualní zpracování dat

Zásady pro vypracování:

1. Rešerše
2. Analýza
3. Vypracování
4. Zhodnocení

- a) teoretická část v rozsahu 25 – 30 normostran textu
- b) prototyp nebo funkční model nebo fyzický model v měřítku 1:1, 1:2, 1:3, 1:5, 1:10 podle charakteru projektu a konzultace s vedoucím práce
- c) grafická prezentace v rozsahu minimálně 2,8 m²

Rozsah bakalářské práce: viz. Zásady pro vypracování
Rozsah příloh: viz. Zásady pro vypracování
Forma zpracování bakalářské práce: tištěná/elektronická

Seznam odborné literatury:

WEINSCHENK, Susan. 100 věcí, které by měl každý designér vědět o lidech. Brno: Computer Press, 2012. ISBN 9788025136492.

NORMAN, Donald A. Design pro každý den. Praha: Dokořán, 2010. ISBN 978-80-7363-314-1.

STEELE, Julie. a Noah P. N. **ILIINSKY.** Beautiful visualization: [looking at data through the eyes of experts]. Sebastopol, CA: O'Reilly, c2010. ISBN 1449379869.

<https://www.smashingmagazine.com/>

<http://www.creativeapplications.net/>

Vedoucí bakalářské práce: **MgA. Bohuslav Stránský, Ph.D.**
Ateliér Digitální design
Datum zadání bakalářské práce: **3. prosince 2018**
Termín odevzdání bakalářské práce: **10. května 2019**

Ve Zlíně dne 3. prosince 2018

doc. Mgr. Irena Armutidisová
děkanka



MgA. Bohuslav Stránský, Ph.D.
vedoucí ateliéru

PROHLÁŠENÍ AUTORA BAKALÁŘSKÉ / DIPLOMOVÉ PRÁCE

Beru na vědomí, že

- bakalářská/diplomová práce bude uložena v elektronické podobě v univerzitním informačním systému a bude dostupná k nahlédnutí;
- na moji bakalářskou/diplomovou práci se plně vztahuje zákon č. 121/2000 Sb. o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon) ve znění pozdějších právních předpisů, zejm. § 35 odst. 3;
- podle § 60 odst. 1 autorského zákona má UTB ve Zlíně právo na uzavření licenční smlouvy o užití školního díla v rozsahu § 12 odst. 4 autorského zákona;
- podle § 60 odst. 2 a 3 mohu užit své dílo – bakalářskou/diplomovou práci - nebo poskytnout licenci k jejímu využití jen s předchozím písemným souhlasem Univerzity Tomáše Bati ve Zlíně, která je oprávněna v takovém případě ode mne požadovat přiměřený příspěvek na úhradu nákladů, které byly Univerzitou Tomáše Bati ve Zlíně na vytvoření díla vynaloženy (až do jejich skutečné výše);
- pokud bylo k vypracování bakalářské/diplomové práce využito softwaru poskytnutého Univerzitou Tomáše Bati ve Zlíně nebo jinými subjekty pouze ke studijním a výzkumným účelům (tj. k nekomerčnímu využití), nelze výsledky bakalářské/diplomové práce využít ke komerčním účelům;
- pokud je výstupem bakalářské/diplomové práce jakýkoliv softwarový produkt, považují se za součást práce rovněž i zdrojové kódy, popř. soubory, ze kterých se projekt skládá. Neodevzdání této součásti může být důvodem k neobhájení práce.

Prohlašuji, že:

- jsem na bakalářské/diplomové práci pracoval samostatně a použitou literaturu jsem citoval. V případě publikace výsledků budu uveden jako spoluautor.

Ve Zlíně dne: 30. 4. 2019

Jméno a příjmení studenta: GORAZD RATULOVSKÝ

podpis studenta

ABSTRAKT

Tato bakalářská práce se zabývá vizualizací osobních dat autora získaných ze sociální sítě Facebook. První část práce se zabývá zkoumáním teorie dat, nakládáním s nimi ve velkých technologických společnostech, jako je například Facebook nebo Google, na to navazuje část, jak data prakticky vizualizovat. Druhá část práce se opírá o získané dovednosti z teorie a věnuje se následnému praktickému zpracování dat. Obsahuje získávání osobních dat z platformy Facebook, jejich následné zpracování a výstupem této části je finální vizuální koncept a jeho prezentace prostřednictvím webové stránky.

Klíčová slova: vizualizace dat, sociální sítě, data, big data, facebook

ABSTRACT

This bachelor thesis is dedicated to the visualization of author's personal data retrieved from the Facebook platform. The first part of the thesis deals with the data theory, its usage within large Tech companies such as Facebook or Google. After that, the next chapter how to visualize data in practice interlocks to the first part of the thesis. The second part of the thesis is based on gained knowledge from the theoretical part and is consequently dedicated to the practical data processing. It contains personal data retrieved from the Facebook platform, its further processing and delivers output of its final visual concept presented through the web site.

Keywords: data visualisation, social networks, Facebook, data, big data

Pod'akovanie

Rád by som využil tento priestor na pod'akovanie môjmu vedúcemu práce MgA. Bohuslavovi Stránskemu, Ph.D., ktorý mi veľmi pomohol svojimi cennými radami, skúsenosťami na konzultáciách a dokázal ma nasmerovať k finálnemu konceptu. Ďalej by som chcel pod'akovať mojej priateľke za to, že mi vedela vždy povedať konštruktívny názor a bola mi počas posledných mesiacov veľkou oporou a tiež priateľom a rodine, ktorí mi akokoľvek poradili alebo dali spätnú väzbu.

Prehlásenie

Prehlasujem, že odovzdaná verzia bakalárskej/diplomovej práce a verzia elektronická nahraná do IS/STAG sú totožné.

OBSAH

ÚVOD	8
I. TEORETICKÁ ČASŤ	9
1 DÁTA	10
1.1 BIG DATA	10
1.2 DÁTA AKO NOVÁ KOMODITA V 21. STOROČÍ.....	12
1.3 SOCIÁLNE MÉDIÁ A OCHRANA OSOBNÝCH ÚDAJOV	13
1.3.1 Dvaja veľikáni.....	14
1.4 AKÉ DÁTA O NÁS PLATFORMA FACEBOOK ZBIERA	15
1.5 AKÁ JE HODNOTA ĽUDÍ V DIGITÁLNO M PRIESTORE.....	15
1.5.1 Priemerný príjem na užívateľa.....	16
1.5.2 Ktoré dáta ovplyvňujú našu hodnotu.....	17
1.5.3 Abstraktné poňatie.....	18
2 VIZUALIZÁCIA DÁT	19
2.1.1 Informačný design a historická nadväznosť.....	19
2.1.2 Infografika.....	21
2.1.3 Vizualizácia dát.....	21
2.1.4 Informačná architektúra.....	22
2.2 UPLATŇOVANÉ ZÁSADY PRI TVORBE	23
2.2.1 Informatívnosť.....	23
2.2.2 Efektívnosť.....	23
2.2.3 Estetika	24
2.2.4 Nadčasovosť	25
2.2.5 Animácie.....	26
2.2.6 Zhrnutie	26
2.3 AKO POSTUPOVAŤ PRI EFEKTÍVNEJ VIZUALIZÁCII.....	26
2.3.1 Proces tvorby pri vizualizácii dát.....	27
2.3.2 Vizúálna interpretácia dát.....	29
2.3.3 Vizualizácia dát pomocou grafov.....	34
2.3.4 Technológie a software na spracovanie a vizualizáciu dát.....	36
II. PROJEKTOVÁ ČASŤ	39
3 CIEĽ A KONCEPT VÝSTUPNEJ PRÁCE.....	40
3.1 CIEĽ PRÁCE	40
3.2 KONCEPT	40
4 PROCES TVORBY	42

4.1	ZBER DÁT.....	42
4.2	ROZDELENIE DÁT.....	42
4.3	DEFINOVANIE CIELOVÉHO PUBLIKA A VYTvoreNIE PRÍBEHU.....	43
4.4	VÝBER VÝSTUPNÉHO MÉDIA.....	44
4.5	ŠTRUKTÚRA WEBU.....	44
4.6	TEXTOVÁ ČASŤ WEBU.....	45
4.7	VIZUÁLNE SPRACOVANIE.....	48
5	GRAFY A VIZUALIZÁCIA DÁT.....	50
5.1	INTERPRETÁCIA POMERU A MNOŽSTVA.....	50
5.2	INTERPRETÁCIA LOKALITY.....	50
5.3	INTERPRETÁCIA ČASU (MNOŽSTVA V ČASE).....	50
5.4	INTERPRETÁCIA OSTATNÝCH DÁT.....	51
6	VIZUÁLNE SPRACOVANIE KONCEPTU A VIZUÁLNE PRINCÍPY.....	52
6.1	FARBY.....	52
6.2	TYPOGRAFIA.....	52
	ZÁVER.....	54
	ZOZNAM POUŽITEJ LITERATÚRY.....	55
	ZOZNAM POUŽITÝCH SYMBOLOV A SKRATIEK.....	59
	ZOZNAM OBRÁZKOV.....	60
	ZOZNAM TABULIEK.....	62
	ZOZNAM GRAFOV.....	63

ÚVOD

Dáta, ich ochrana a zneužívanie v digitálnom priestore sa v tejto dobe stáva čoraz väčšou témou. Medzi najhodnotnejšími spoločnosťami na svete sú technologické spoločnosti, ktoré zhromažďujú obrovské množstvo osobných dát svojich používateľov. Tieto informácie im poskytujeme dobrovoľne a s dôverou v to, že sú v bezpečí. Osobne verím, že tomu tak naozaj je, ale čoraz častejšie škandály o únikoch, hackerských útokoch alebo predaji týchto dát tretím stranám vzbudzujú mnohé otázky a pochybnosti. A na tie som sa snažil nájsť odpovede v projektovej časti práce.

Úvodná časť práce sa bude venovať získaniu teoretických poznatkov k vybranej problematike práce, ktorá pozostávala z teórie popisujúcej význam dát ako takých a ich jednotlivé charakteristiky. Okrem problematiky dát teoretická časť práce bude popisovať jedného z významných hráčov na trhu, ktorý zhromažďuje najväčšie množstvá užívateľských dát. Následne svoju pozornosť upriamim na zisťovanie informácií o samotnej vizualizácii dát, ako je dobre postupovať, aké formy vizualizácie poznáme, a ako môžeme dáta získavať. V závere tejto časti práce bude zhrnutý prístup ako dáta efektívne vizualizovať.

Druhá, veľmi významná časť práce, bude venovaná projektovej časti bakalárskej práce. V úvodnej časti bude charakterizovaný význam samotnej práce a jej cieľ. Následne na to naviaže proces tvorby samotnej vizualizácie od zberu dát, cez ich triedenie, obsahovú a informačnú štruktúru ich prezentácie, až po samotné vizuálne spracovanie, ktoré predstavuje primárny zámer tejto práce.

Čo ma však viedlo k výberu samotnej témy tejto práce? Prvotným cieľom bolo zdokonaľiť sa vo vybranej oblasti a prehĺbiť doposiaľ získané poznatky o práci s dátami a ich interpretáciou. Cieľom práce bude vizuálne interpretovať osobné užívateľské dáta tak, aby sám človek, ktorého sa tieto dáta týkajú mal prehľad o tom, aké všetky informácie sú o ňom zistiteľné z kľúčových zdrojov dát, a to primárne zo sociálnych sietí a sekundárne v tomto užívateľovi vzbudiť záujem o danú problematiku a zamyslenie sa nad ňou.

I. TEORETICKÁ ČASŤ

1 DÁTA

Úvodné strany kapitol teoretickej práce budú venované vymedzeniu problematiky, vysvetleniu základných pojmov a oblasti, ktorej sa budem venovať. V prvej kapitole sú definované dáta ako pojem v dnešnej dobe. V tej druhej sa zaoberám už priamo vizualizáciou dát a skúmam ich po výtvarnej stránke.

Som si vedomý, že pre odbornú verejnosť, ktorá bude túto bakalársku prácu čítať prípadne posudzovať, môžu byť tieto informácie známe, no rozhodol som ich začleniť, nakoľko si myslím, že pre mňa a moje skúmanie boli tieto zistenia potrebné a zásadné. Na druhej strane si myslím, že môžu byť pre spomínanú odbornú verejnosť niektoré informácie obsiahnuté v tejto práci nové, a zároveň obohacujúce.

1.1 Big data

Dáta alebo aj informácie – údaje – poznatky, takto nám definujú pojem synonymické slovníky. V poslednom období sa však môžeme stretnúť tiež s názvom „Big Data“. S definíciou som si pomohol na stránke spoločnosti Oracle, ktorá patrí medzi lídrov v tejto oblasti (SAS Institute, 2018). Big data, alebo aj “veľké data“, ako ich môžeme označovať ako veľké a zložité súbory dát (data sety), ktoré štandardné programy nie sú schopné spracovávať a spravovať. Toto obrovské množstvo dát môže napomáhať spoločnostiam zlepšovať ich biznis znižovaním nákladov, predikovaním poruchovosti strojov, predpovedaním chovania zákazníkov, analyzovaním produktov na základe spätnej väzby od zákazníkov a tak ďalej. Na stránkach Oracle píš, že dáta sa stávajú kapitálom veľkých technologických spoločností, s čím je nemožné nesúhlasiť.

Big data charakterizujeme podľa troch kategórií. V angličtine ich označujeme na základe ich anglických názvov ako 3 V.

Volume (objem) – okrem štruktúrovaného ukladáme aj neštrukturalizovaný obsah. V prípade Big Data hovoríme o objemoch dát o veľkostiach petabyte.

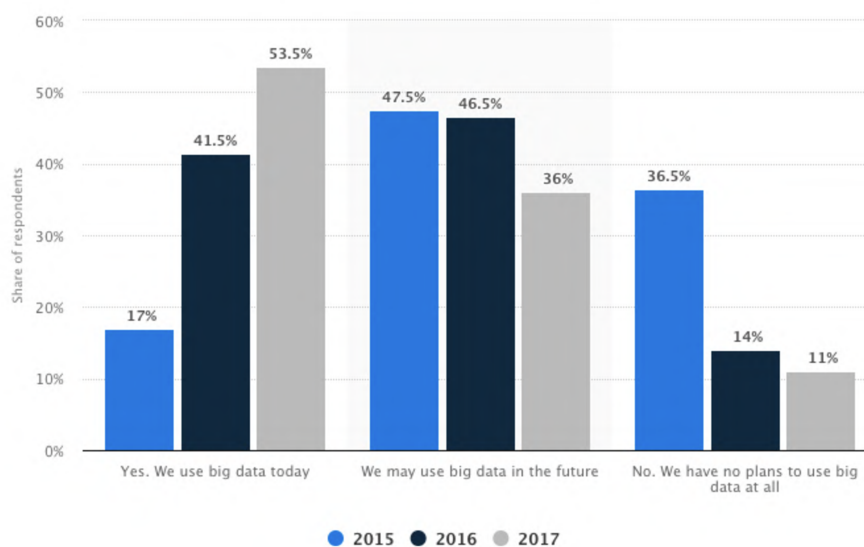
Velocity (rýchlosť) – niektoré aplikácie vyžadujú analýzu dát v reálnom čase, preto je rýchlosť veľmi dôležitá. Jedná sa napríklad o rozpoznávanie tváří z kamier.

Variety (rôznorodosť) – v dnešnej dobe nie sú dáta len štruktúrované, ale ukladáme aj veľké množstvo dát, ktoré sú neštrukturalizované, ako napríklad video, rozsiahle texty, obrázky, zvuky, ktoré sa následne musia spracovávať a vytvárať metadáta.

Čoraz častejšie sa používa aj štvrté V a to **Veracity** (vierohodnosť) – vzniká síce množstvo dát, ale je potrebné zistiť ich hodnotu a pravdivosť, a teda, či je možné ich ďalej využiť v náš prospech. (Oracle, 2014)

Aký ale majú veľký podiel „veľké dáta“ v súčasnosti? Nikdy totiž neboli tak veľkou témou ako v posledných rokoch, a ako môžeme vidieť v grafe nižšie (Statista, 2018), ešte v roku 2015 spoločnosti využívali adaptáciu dát v rámci firmy len z približne 17-tich percent a až 36 percent nemalo plán, respektíve nemali v pláne využívať tie technológie na spracovávanie „veľkých dát“, no od roku 2017 je vidieť obrovský obrat v týchto ukazovateľoch, čiže môžeme predikovať, že tieto technológie sú na obrovskom vzostupe.

A môžeme očakávať, že tieto veľké dáta, ktoré produkujeme aj my sami sa stanú súčasťou našich životov a ovplyvnia náš životný štýl a spôsob, ako interagujeme so svetom. (Viktor Mayer-Schönberger, 2014 str. 14)



Graf 1 – Rebríček najhodnotnejších spoločností v roku 2008 podľa Financial Times

1.2 Dáta ako nová komodita v 21. storočí

„Dáta sa zmenili na surovinu podnikania, cenný ekonomický vstup, ktorý umožňuje vytvárať nové ekonomické hodnoty.“ (Viktor Mayer-Schönberger, 2014 s. 13) Opäť som si pomohol knihou Big Data, kde to autori trefne zaznamenali.

Keď sa pozrieme na rebríček najhodnotnejších firiem z rokov 2008 a 2018 môžeme si všimnúť, že sa zásadne obmenil a vytratil sa veľké spoločnosti – hlavne teda ropné spoločnosti ako napríklad PetroChina, Exxon, Gazprom, Royal Dutch Shell, ktorých produktom je pochopiteľne ropa, a zamenili ich veľké technologické spoločnosti, ktorých hlavným produktom je hardware a software (Apple, Microsoft), online predaj (Amazon, Alibaba), online reklama (Alphabet, Facebook).

V obidvoch tabuľkách je zobrazená trhovú hodnotu firiem k dňu 31. marca. Prvú z roku 2008 zostavil magazín Financial Times vo svojom rebríčku FT Global 500 (Financial Times, 2008). Druhý rebríček zostavovala spoločnosť PwC. (Global Finance, 2018)

Ako je možné porovnať v tabuľke z roku 2018, až 7 z 10-tich spoločností je z oblasti technológií oproti roku 2008.

Global rank 2008	Global rank 2007	Company	Country	Market value \$m	Sector
1	1	Exxon Mobil	US	452,505.1	Oil & gas producers
2		PetroChina	China	423,996.2	Oil & gas producers
3	2	General Electric	US	369,569.4	General industrials
4	6	Gazprom	Russia	299,764.4	Oil & gas producers
5	16	China Mobile	Hong Kong	298,093.2	Mobile telecommunications
6	9	Indl & Coml Bank of China	China	277,235.5	Banks
7	3	Microsoft	US	264,131.9	Software & computer services
8	5	AT&T	US	231,168.1	Fixed line telecommunications
9	10	Royal Dutch Shell	UK	220,110.2	Oil & gas producers
10	13	Procter & Gamble	US	215,640.1	Household goods & home construction

Tabuľka 1 – Rebríček najhodnotnejších spoločností v roku 2008 podľa Financial Times

PwC Global Top 100

Rank	Company	Country	Industry	Market Cap (\$ bn)
1	Apple	USA	Technology	851
2	Alphabet	USA	Technology	719
3	Microsoft	USA	Technology	703
4	Amazon.com	USA	Consumer Services	701
5	Tencent	China	Technology	496
6	Berkshire Hathaway	USA	Financials	492
7	Alibaba	China	Consumer Services	470
8	Facebook	USA	Technology	464
9	JPMorgan Chase	USA	Financials	375
10	Johnson & Johnson	USA	Health Care	344

Tabuľka 2 – Rebríček najhodnotnejších spoločností v roku 2018 podľa PwC Global Top 100

Z týchto poznatkov môžeme vyčítať, že dáta sa stali novým druhom tovaru/komoditou, novým zlatom prípadne ropou. Akurát zdrojom ťažby už nie je primárne zemská kôra, ako tomu je v prípade ťažby ropy, ale hlavným zdrojom ťažby a produktom sa stávajú dáta. Otázkou je, prečo práve dáta získavajú na takej hodnote? V knihe Big Data, hovoria o tom, že v dátach je obrovský potenciál, a to hlavne kvôli tomu, že sú „recyklovateľné“, takže pokiaľ ich vlastnime a majú samozrejme potrebnú kvalitu, je možné s nimi pracovať opakovane. Čiže keď som vyššie spomínal prirovnanie k nerastným bohatstvám, tak dáta majú oproti nim tú výhodu, že sú znovu použiteľné, a teda môžeme ich predáť viac krát. (Viktor Mayer-Schönberger, 2014 str. 115)

1.3 Sociálne médiá a ochrana osobných údajov

Keď som sa rozhodoval, na čo sa chcem primárne sústrediť v mojej bakalárskej práci, tak to nebolo vôbec jednoduché. Už na začiatku som vedel, že téma dát je veľmi zaujímavá, no venovať sa minimálne teoreticky celej tejto oblasti je v rozsahu určite viac ako rozsah bakalárskej práce. Vedel som však, že jedným z najväčších držiteľov dát sú práve známe technologické spoločnosti, ktoré prevádzkujú rovnako známe sociálne médiá. Ak sa zamyslíme nad tým, čo vyplýva z tabuliek vyššie, že najhodnotnejšími spoločnosťami sú prevádzkovatelia týchto sociálnych sietí ako napríklad YouTube (vlastní Google), Facebook, Instagram (vlastní Facebook), ktorých hlavné produkty sú zadarmo, tak ako je možné, že tieto spoločnosti majú tak obrovskú hodnotu. Odpoveď sú zrejme dáta. Naše dáta,

ktoré my poskytujeme zadarmo. Ďalším aspektom, ktorý ma zaujímal a bol témou roku 2018 (a zrejme aj bude počas najbližších rokov) je ochrana osobných údajov. Práve tieto fakty ma nabádajú venovať sa tejto oblasti dát podrobnejšie.

1.3.1 Dvaja veľikáni

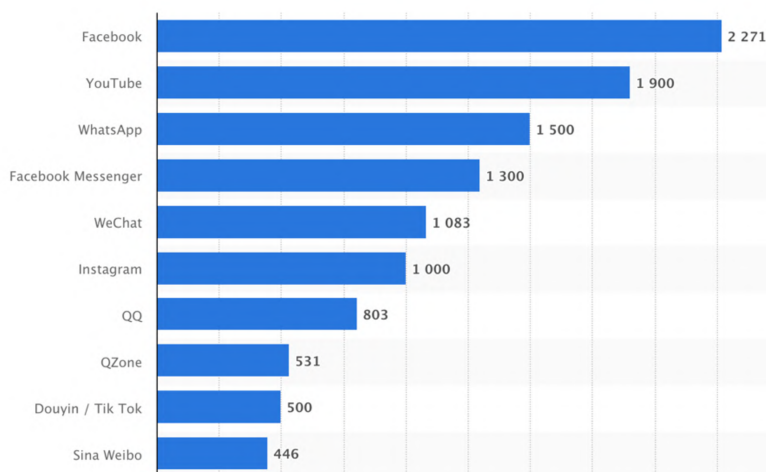
Medzi najväčšie globálne spoločnosti v oblasti sociálnych sietí patria jednoznačne Google a Facebook, preto som svoj fókus zamerlal práve na týchto dvoch veľikánov.

V portfóliu Facebooku je samozrejme samotný Facebook, ďalej Instagram, Messenger, WhatsApp (dve posledné sú komunikačné platformy), a ďalšie menšie produkty, medzi ktoré patria nástroje pre ostatné spoločnosti, ktoré ich vyžívajú na správu svojich reklám. (Facebook, 2019)

Portfólio spoločnosti Alphabet, do ktorej spadá Google, je podstatne pestrejšie a komplexnejšie. Ak spomenieme len tie najdôležitejšie, tak sú to produkty ako Google Search, Chrome, Google Translate, Google Maps, YouTube, Gmail, Google Messages, Photos, Google Calendar, netreba zabudnúť na OS Android a nehovoriac o Google Pixel telefóne, Google Home alebo Google Chromecast – každá z týchto platforiem drží obrovské množstvo našich údajov, takže môžu zaznamenávať prakticky celý náš deň. Ak sa však zameriame len na sociálne siete, tak vlajkovou loďou v Google je YouTube. Okrem týchto produktov majú v portfóliu taktiež platformy pre tretie strany, ktoré môžu spravovať svoje reklamy a využívať dáta v rámci siete Alphabet. (Google, 2019)

Pri čítaní jedného z článkov a infografiky na webe Visual Capitalist autor poukazuje z akej formy získavajú príjmy najväčšie technologické spoločnosti ako Apple, Alphabet, Microsoft alebo Facebook. A podľa výkazov príjmov v prípade Alphabet tvorí príjem z reklamy 85% a v prípade Facebooku je to na úrovni až 98,5%. (Desjardins, 2019)

V tabuľke nižšie môžeme vidieť, že v prvej šestke najpoužívanejších sociálnych sietí vládnu práve produkty týchto dvoch spoločností. Zvyšné 4 platformy z tabuľky pochádzajú z Čínskej ľudovej republiky a sú používané najmä na tomto trhu.



Graf 2 – Svetové najpoužívanejšie sociálne siete, január 2019, zobrazuje počet aktívnych používateľov v miliónoch

1.4 Aké dáta o nás platforma Facebook zbiera

Facebook na svojej stránke o zásade používania údajov uvádza druhy informácií, ktoré o nás ukladá. Informácie ukladajú do segmentov ako „Veci, ktoré vy a iní robíte a poskytujete“ a „Informácie o zariadeniach“ kde majú podrobne rozpísané, ktoré informácie o nás zbierajú. (Facebook)

Ideálnym spôsobom je však stiahnuť si svoje dáta priamo z Facebooku a vidieť reálne, čo zhromažďujú.

To urobil autor článku Todd Haselton na webe CNBC, ktorý sa na to pozrel podrobnejšie. V článku spomína najzaujímavejšie zistenia. Facebook o nás ukladá všetky správy, ktoré sme kedy poslali, komentáre, zmeny statusov a všetky interakcie. Okrem toho aj údaje na rozpoznávanie tváre, ďalej zoznam priateľov, kontakt list, navštívené udalosti, logy z prihlásenia, zaznamenávajú našu polohu, vedia na akú reklamu sme klikli a čo nás zaujíma. (Haselton, 2018)

1.5 Aká je hodnota ľudí v digitálnom priestore

Svoje skúmanie som smeroval aj k otázke, aká je hodnota dát, a aká je naša hodnota v digitálnom priestore? Ak si uvedomíme, že využívame služby najhodnotnejších technologických firiem ako napríklad Google alebo Facebook zadarmo, čo je ich najväčším produktom? Zaujímavý titulok je možné nájsť v študentskej práci „Study case: Facebook“,

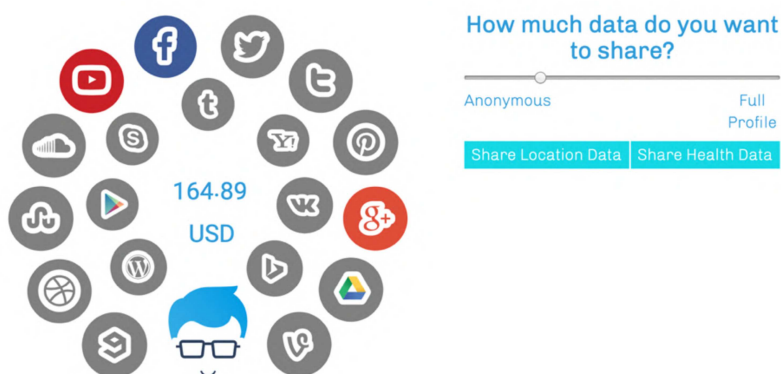
ktorý znie „If you are not paying you are the product“, vo voľnom preklade „ak neplatíš, si produktom“. V kapitole ďalej porovnáva Facebook a YouTube, kde síce obidve platformy zarábajú na našom obsahu, no YouTube sa snaží do istej miery odmeňovať svojich tvorcov. (Zewdu Melkamu, 2016 s. 21)

Svoj pohľad do témy vnášajú aj autori knihy Big Data, ktorý poukazujú na vstup Facebooku na burzu. Večer pred vstupom spoločnosti na burzu, banky a experti stanovili cenu 38\$ za akciu, čo zodpovedalo cene 104 miliárd dolárov. No skutočná hodnota bola iná. Vo finančných správach priznali aktíva v hodnote 6,3 miliardy dolárov. Hodnota dát, ktoré uchovávajú však nemala prakticky žiadnu hodnotu, pritom firmu tvoria prakticky len dáta. Môžeme si všimnúť veľký nepomer medzi tržnou kapitalizáciou cca. 104 miliárd dolárov a účtovných zostatkov v hodnote 6,3 miliardy dolárov. Avšak spoločnosť Gartner zhromaždila 2,1 biliónov speňažiteľného obsahu, ako sú diskrétné údaje, ale aj reakcie ako „páči sa mi to“, komentáre atď., pričom každý dátový bod mal hodnotu približne 5 centov. V konečnom dôsledku mal každý užívateľ hodnotu v priemere 100 dolárov. (Viktor Mayer-Schönberger, 2014 str. 131)

1.5.1 Priemerný príjem na užívateľa

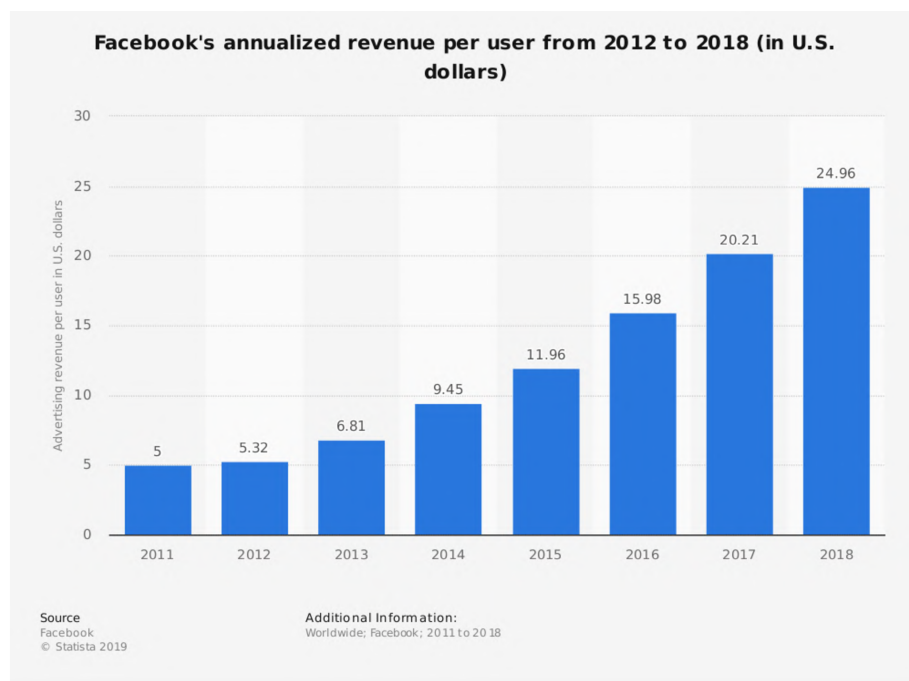
Jednou z metodík oceňovania je metrika „Average revenue per user“ (ďalej len ARPU). Využíva sa hlavne pri sociálnych sieťach a je to celkový príjem spoločnosti vydelená počtom užívateľov. (Wikipedia, 2019)

S touto metodikou pracovali na projekte Data Calculator, kde na základe používaných sociálnych sietí a frekvencie ich používania vypočítajú cenu Vašich dát v digitálnom priestore.



Obrázok 1 – Data Calculator

Podľa autora Felixa Richtera z portálu Statista, odkazujúceho sa na finančnú správu Facebooku, bolo priemerné ARPU spoločnosti Facebook v roku 2018 24,96 \$ (suma je vrátane ďalších produktov Facebooku – Instagram, Messenger, WhatsApp). Ak sa pozriem na tabuľku nižšie, tak ARPU sa z roka na roka zvyšovalo a podľa zdroja tento rast príjmu ani užívateľov nezastavili ani medializované problémy spoločnosti v súvislosti s kauzou Cambridge Analytica. (Richter, 2019)



Graf 3 – Priemerná ARPU spoločnosti Facebook

ARPU za rok 2018 spoločnosti Google je podstatne vyšší, v článku žurnalistu Frederica Fillouxa, ktorý sa taktiež odvoláva na finančné reporty a v článku porovnáva aj ďalšie veľké digitálne médiá je hodnota pre Google a jej ďalšie produkty stanovená na 137,- \$.

1.5.2 Ktoré dáta ovplyvňujú našu hodnotu

Problematikou oceňovania dát sa zaoberali aj v magazíne Financial Times, kde vytvorili online kalkulačku, v ktorej zohľadňujú rôzne informácie od demografických údajov, veku, pohlavia, povolania, rodinného stavu ale aj zdravotných informácií, vlastníctve nehnuteľností, voľno-časových aktivít a zákazníckeho správania. Dobré to poukazuje na fakt, ktoré vstupné dáta ovplyvňujú našu hodnotu.

What is your data worth?

DEMOGRAPHICS FAMILY & HEALTH PROPERTY ACTIVITIES CONSUMER

Data brokers scour public documents, such as birth records and motor vehicle reports, to compile basic data about individuals. It is likely they already know your:

- Age
- Gender
- ZIP code
- Ethnicity
- Education level

Are you a millionaire?

No
 Yes

What is your job?

President

Are you engaged to be married?

Yes
 No

Are you?

- Recently married
- Recently divorced
- Empty nester

\$0.219
Current value of my data

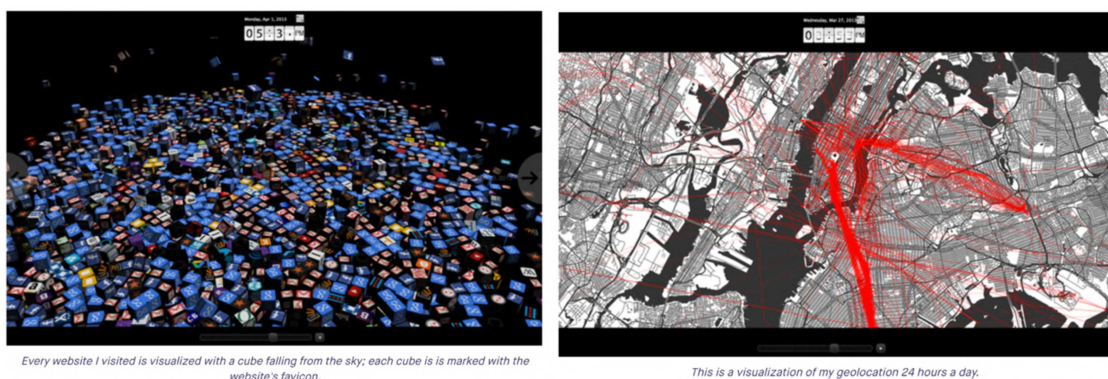
NEXT ►

Obrázok 2 – Kalkulačka hodnoty dát podľa Financial Times

1.5.3 Abstraktné poňatie

V roku 2013, prišiel so zaujímavým projektom, a mohli by sme ho kategorizovať ako konceptuálne umenie, Federico Zannier, ktorý „vydoloval“ svoje dáta zo svojej online aktivity. Dáta zahŕňali od navštívených stránok, cez geo-lokácie, logy z počítača, ktoré aplikácie používal atď.

V projekte chcel poukázať na fakt, že v reklame sa točí obrovské množstvo peňazí, a práve tieto spoločnosti zarábajú na našich reklamných profiloch. Rozhodol sa predávať tieto svoje „vydolované“ dáta na crowdfunding platforme Kickstarter. (Zannier, 2013)



Obrázok 3 – Projekt Federica Zanniera

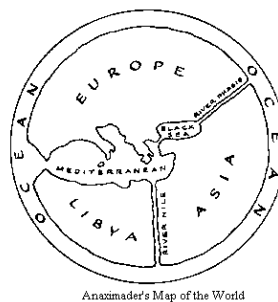
2 VIZUALIZÁCIA DÁT

Na začiatku svojho skúmania som si položil otázku “Čo je design informácií?”. Mohlo by sa zdať, že pre budúceho designera to bude jednoduchá odpoveď, avšak prehĺbovaním znalostí v tejto téme som dospel k záveru, že na to nie je jednoznačná odpoveď. V snahe zadefinovať to jednou vetou by sme mohli povedať, že sa jedná o vizuálne spracovanie informácií alebo dát, avšak tento pojem môže znamenať aj omnoho viac.

Design informácií môžeme chápať ako informačnú grafiku, tiež pod známejším pojmom infografika, a môže to byť vizualizácia dát – v angličtine data visualisation a taktiež pojem informačná architektúra.

2.1.1 Informačný design a historická nadväznosť

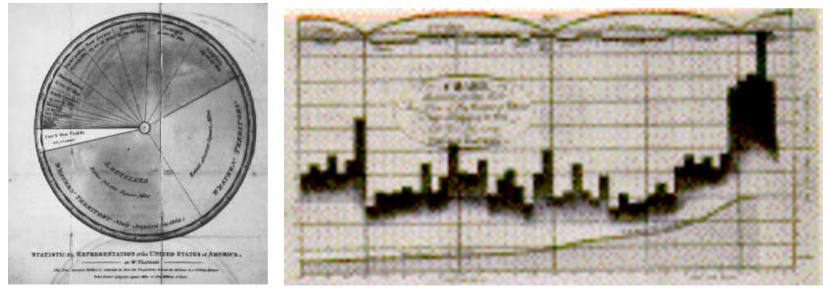
Prvé náznaky vizuálneho vyobrazenia informácií môžeme pozorovať už pred tisíckami rokov. Boli nimi prvé jaskynné maľby. Neskôr to boli mapy, vek prvej svetovej mapy odborníci datujú na približne 550 rokov pred naším letopočtom. (FRIENDLY, 2009 s. 4)



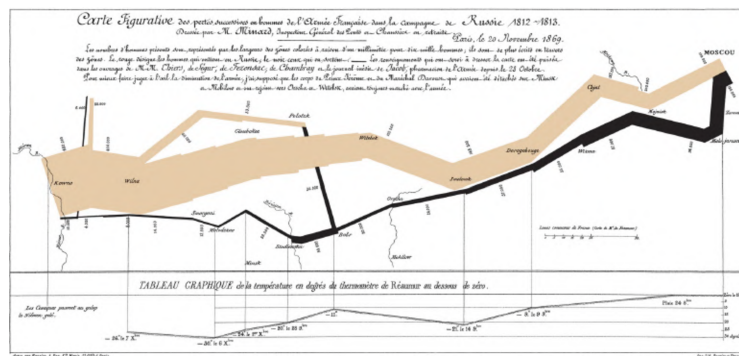
Anaximander's Map of the World

Obrázok 4 – Prvá svetová mapa z roku 550 p.n.l.

Ďalej sú to rôzne rodokmene či diagramy. Od 18. st., v dobe osvietenstva, prichádza k rozvoju vedy, astronómie a matematiky. O dôležitý míľnik sa na prelome 18. a 19. sa postaral britský vedec Wiliam Playfair, ktorý vynašiel základne formy zobrazovania štatistík. Bol to líniový graf, stĺpcový graf a koláčový graf. (Spence, 2004) Ďalšia ukážka je z roku 1869, v ktorej Charles Joseph Minard zobrazil v kartografických mapách diagramy a grafy. Na jeho práci zobrazoval napoleonské ťaženie do Ruska aj so stratami na životoch. (Mapping Complex Information, 2012)



Obrázok 5 – Grafy Williama Playdaira



Obrázok 6 – Mapa Charlesa Minarda

Naplno sa začal informačný design prejavovať v 20. storočí. Z množstvá ukážok spomeniem niekoľko dôležitých mien a míľnikov. Za priekopníka informačného designu bol považovaný český rodák Ladislav Šutnár, ktorý bol názoru, že funkcionlita by mala byť nad estetiku a informácie by mali byť efektívne od komunikované. Jedným z jeho diel bol americký telefónny zoznam, ktorý významne uľahčil jeho používanie. (Mapping Complex Information, 2012)

Za skvelú inšpiráciu z 30. rokov 20. storočia môžeme považovať prvú mapu londýnskeho metra od Harryho Becka. Pri tvorbe mal v zásade voľnú ruku, musel dodržiavať len grafické štandardy no nie kartografické. Beck mal skúsenosť s technickým kreslením elektrických obvodov, v ktorých sa používali horizontálne línie, 90° a 45° uhly. Tieto konvencie preniesol aj do mapy metra. Mapa tak bola oslobodená od akýchkoľvek geografických pravidiel a naberala abstraktnú formu. Avšak v podzemí zobrazovala reálnu podobu trás čo v tomto prípade malo logiku, keďže cestujúceho v metre, čiže pod zemou, nezaujíma ako to vyzerá na povrchu. (Iliinsky, 2010 s. 4)



Obrázok 7 – Mapa londýnskeho metra, Harry Beck

Termín informačný design neskôr spopularizoval Edward R. Tufte, ktorý patrí medzi súčasných odborníkov v oblasti informačného designu a ktorý v roku 1982 vydal knihu „Vizuálne zobrazenie kvantitatívnych informácií“ (voľný preklad). Primárne sa teda jedná o zobrazovanie informácií a diagramov z kvantitatívnych dát. (Wikipedia, 2018)

2.1.2 Infografika

Pojem infografika popisuje Mark Smiciklas vo svojej knihe „The Power of Infographics“. Smiciklas poukazuje, že infografika môže byť zafinovaná ako akási vizualizácia dát či nápadov, no musí v sebe niesť informáciu, ktorá sa stáva hlavnou myšlienkou vizualizácie a zároveň by mala byť vizualizáciou predaná ľahko a zrozumiteľne danému divákovi. (Smiciklas, 2012 s. 3)

Čiže spojením ilustrácií, koláží a samozrejme informácií a dát vytvárame informačnú grafiku a vizuálny príbeh, ktorého cieľom je pomôcť sa čo najrýchlejšie a najjednoduchšie zorientovať v danej problematike.

2.1.3 Vizualizácia dát

Vizualizáciu dát oproti infografike môžeme chápať ako surové zobrazenie dát, zväčša bez linearity a interpretované hlavne v digitálnom prostredí.

Svoj postoj opieram o vyjadrenia Matthewa Pritcharda a jeho porovnanie medzi infografikou a vizualizáciou dát, kde hovorí o ich rozdieloch. Podľa Pritcharda pri vizualizácii dát

nemusíme zobrazovať akýsi „vizuálny príbeh“, a na rozdiel od infografiky zobrazujeme „len“ surové dáta, čo nám dovoľuje objektívnejšie zhodnotenie témy. (Pritchard, 2016)

To znamená, že informácie a dáta, ktoré zobrazujeme nie sú podporované žiadnymi vizuálnymi prostriedkami a nevnímame ich ako celok či príbeh. Grafické prvky nám však len pomáhajú dáta pochopiť a pohybovať sa v nich podľa nášho uváženia.

Okrem toho, oproti infografike je väčšinou proces výroby generovaný na základne naprogramovaných algoritmov. Síce sú vizualizácie z počiatku navrhované ľuďmi, finálny produkt je kreslený algoritmicky a s grafickými softvérmi. Vizualizácia tak vždy môže byť ľahko aktualizovaná novým prísunom dát. Rozdiel nastáva pri zobrazovaní veľkého množstva dát v informačnej vizualizácii, kde infografika vychádza esteticky bohatšia. (Iliinsky, a iní, 2011 s. 7)

Vizualizácia dát sa taktiež stala atraktívnou disciplínou pre digitálnych umelcov. Za obľúbený nástroj pre kreatívne programovanie môžeme považovať napríklad programovací jazyk Processing, o ktorom sa bližšie zmiňujem v ďalších kapitolách tejto práce.

2.1.4 Informačná architektúra

Informačná architektúra (ďalej IA) nie je vyslovene vizuálnou záležitosťou, a preto viac spadá do oblasti User Experience¹ (ďalej UX). No jej prínos je (nielen) v digitálnom prostredí obrovský. IA chápeme ako rozvrhnutie informácií v rámci daného média – napríklad web, software, mobilné aplikácie a, samozrejme, to nemusí byť len digitálne médium. Tak sa jedná o akúsi informačnú a obsahovú hierarchiu, s ktorou sa stretáva a pracuje koncový užívateľ.

Ako sa uvádza na webe Inštitútu informačnej architektúry, dobrou IA pomáhame používateľom nájsť to, čo hľadajú. Táto zručnosť je tiež dôležitou súčasťou UX, ale tiež ďalších disciplín ako tvorba obsahovej stratégie, technického písania a taktiež v interakčnom designe. (Institute)

¹ User Experience – spôsob akým ľudia interagujú s našim produktom, a aký pocit z jeho používania majú.

Aj keď sa niekomu môžu zdať jednotlivé vyššie zmieňované pojmy rozdielne, prienik designovania dát a informácií je u každého zrejmy a ich vysvetlenie je pre ďalšie skúmanie problematiky dôležité.

2.2 Uplatňované zásady pri tvorbe

Noah Iliinsky vo svojej knihe „Beautiful Visualization“ hovorí o kráse vizuálneho spracovania (nie len dát). Čo ale musí spĺňať, aby sme mohli o výsledku hovoriť, že je pekný? Je to hypotéza, ktorú môžeme len ťažko definovať, keďže pre každého môže slovo pekné znamenať niečo iné. No pri používaní niekoľkých konvencií, o ktorých píše Noah, môžeme získať očakávaný „krásny“ výsledok. Tými zásadami nie je len estetická časť projektu, ktorá je samozrejme dôležitá, no dôležitá sa zdá byť tiež informatívnosť, efektívnosť – zobrazenie správnych a inovatívnych. Postupne si rozoberieme tieto pravidlá. (Iliinsky, a iní, 2010 s. 1)

2.2.1 Informatívnosť

Kľúčom k úspechu každého vizuálu, či už pre oko krásneho alebo nie, je umožniť prístup k informáciám, takže používateľ z neho získa určitú vedomosť. Vizuál, ktorý toto nespĺňa, zlyhal. Keďže je to najdôležitejší faktor pri určovaní celkového úspechu, schopnosť poskytnutia informácie, tak musí byť primárnou hnacou silou designu vizuálu. (Iliinsky, a iní, 2010 s. 2)

2.2.2 Efektívnosť

Tento bod nadväzuje na ten predošlý. Ako je zjavné, každý vizuál by mal odovzdávať nejakú správu. Prístup k tejto správe by mal byť jasný, nemal by byť komplikovaný, to znamená, že by mal byť zbavený rušivých vplyvov, prípadne by nemal zobrazovať irelevantné dáta, ktoré by tomu mohli zamedzovať. Ak splníme tieto podmienky, môžeme hovoriť o efektívnej interpretácii vizuálu. (Iliinsky, a iní, 2010 s. 2)

Dobrym príkladom by mohla byť informačná tabuľa na vlakovej stanici. Je nedeľa večer, utekám na vlak a ponáhľam sa. Takýto scenár väčšinou nezažívam sám. Pri tom zhone sa však viem zorientovať na čiernej tabuli, ktorá viac-menej vždy vyzerá rovnako. A práve tu môžeme vidieť efektívne spracovanie dát. Mentálny model, ktorý máme v sebe zakódovaný

a jednoduchá interpretácia v jednom riadku nám umožňuje rýchle zorientovanie sa v aktuálnej dopravnej situácii. Nejedná sa síce o esteticky krásnu ukážku, a mohli by sme polemizovať, či sa to dá spraviť lepšie, no myslím si, že tú základnú myšlienku, a tou je funkčnosť a efektívne vyjadrenie dát, to spĺňa.

ODCHOD DEPARTURE	SMER DIRECTION	DOPRAVCA OPERATOR	CIEĽOVÁ STANICA DESTINATION	SMER JAZDY DIRECTION	MIESTO PLATFORM	MEŠKA DELAY
z dôvodu nehody na trati je doprava pozastavená na neurčito						
9:25 R		ZSSK	Košice	Štrba	2	130
11:25 R		ZSSK	Košice	Hargecany		30
11:32 Os		ZSSK	Žilina	Púchov	1	20
12:35 R		ZSSK	Bratislava hl. st.	Trnava		
12:43 Z		ZSSK	Košice	Spišská Nová Ves		

Obrázok 8 – Informačná tabuľa – vlaková stanica

2.2.3 Estetika

Základné grafické prvky, ako sú tvary, farby, čiary, textúry a typografia sú samozrejme nevyhnutnou súčasťou vizuálu. Ich účelom je napomáhať čitateľovi v orientácii a správne ho chápať. Treba však myslieť na to, že v prípade nekorektného používania môžu skôr tieto prvky spôsobovať potencionálnu prekážku a následný problém v interpretácii. Treba za každých okolností pozerieť na primárny cieľ, a to prezentovanie informácií. Takže aj v tomto prípade platí, že menej býva viac. (Iliinsky, 2010 s. 2)

Často sa môžeme stretávať s rôznymi trendami, novými vizuálnymi prístupmi, kde môžeme vidieť, že cieľom autorov nie je dodať zrozumiteľné informácie čitateľovi/užívateľovi, ale primárnym cieľom je byť iný a vytvoriť niečo vizuálne atraktívne. Samozrejme nemôžeme hovoriť len o tom, že vizuálnu interpretáciu treba potláčať, nie je predsa nič krajšie ako vidieť veci, ktoré fungujú a sú pekné.

2.2.4 Nadčasovosť

Platí to vo všetkom. Pokiaľ chceme byť odlišní, musíme byť inovátori a prísť s niečím novým. V prípade vizualizácie dát to môže byť napríklad výsledný formát.

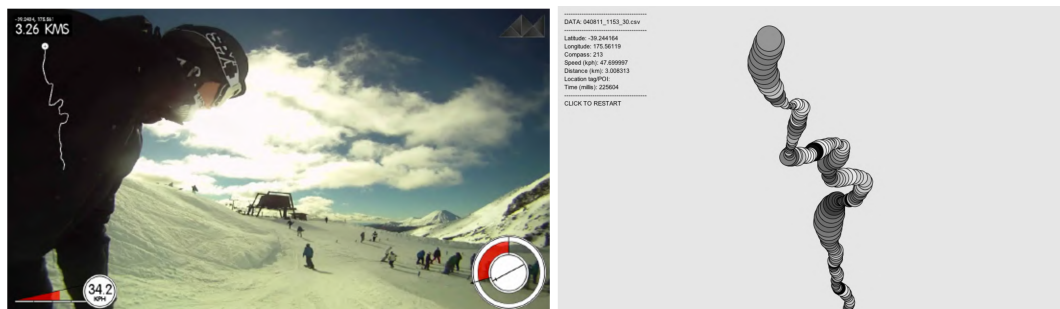
Niekedy bežný papier ako výstup postačoval, ale postupne začali byť tvorcovia odvážnejší, príkladom môžu byť rôzne pop-up knihy, v ktorých už môžeme vidieť náznaky interakcie medzi čitateľom. V posledných rokoch sa forma mení na digitálnu a dáta vizualizujeme pomocou naprogramovaných algoritmov.



Obrázok 9 – Ukážka z knihy „Budapest POP-UP Book“, Livia Hasenstaub

S príchodom nových technológií vstupujú do výsledných prác rôzne interakčné prvky, ktoré reagujú na pohyb, hlas alebo hmat a práve tieto veci môžu výsledok obohatovať a vnášať doňho prvky nadčasovosti.

Zaujímavým príkladom je projekt AntiMap, ktorý už je síce von istú dobu, no môžeme v ňom badať rôzne inovácie. Mobilná aplikácia nahráva vašu aktivitu a na základe vstavaných funkcií ako GPS alebo „compass sensor“ zbiera dáta, ktoré následne vizualizuje, a tu sa dostávame k spomínanej interakcii s používateľom. Tiež je zaujímavé real-time zobrazovanie dát. (Visnjic, 2012)



Obrázok 10 – Ukážka z aplikácie AntiMap

2.2.5 Animácie

Danyel Fisher v knihe „Beautiful Visualization“ hovorí o ďalšom princípe, a tým je animácia. Nie je nutná, ale veľmi napomáha celkovému dojmu.

Dnešné technológie ako JavaScript, Processing alebo video-editačné nástroje ako After Effect, Motion, atď. napomáhajú tvorcom k tvorbe a distribúcii animovanej vizualizácii dát. Pomocou animácie môžeme ešte jednoduchšie porozumieť kontextu a zo statického vizuálu vieme naopak dynamikou vytážiť ešte viac.

Danyel tvrdí, že pokiaľ je dvojrozmerný statický obraz dobrý, potom pohyblivý obraz bude ešte lepší. Pohyb všetci dobre poznáme a je nám známy, vidíme ako sa všetko mení v reálnom čase. Animácia zobrazuje, ako sa údaje menia v priebehu času, a aj to môže používateľa viac vtiahnuť do obrazu. (Iliinsky, a iní, 2010 s. 329)

2.2.6 Zhrnutie

Musíme si uvedomiť, že pokiaľ chceme dosiahnuť pozitívneho výsledku je dôležité brať ohľad aj na isté zásady, a tými hlavnými sú informatívnosť a efektívnosť. Pokiaľ sa nám to podarí a pridáme k tomu kúsok inovatívnosti a kvalitné vizuálne spracovanie, môžeme očakávať aj skvelý výsledok.

2.3 Ako postupovať pri efektívnej vizualizácii

Dodržiavanie správneho procesu pri tvorbe je dôležité v každej oblasti. Tak isto je tomu aj pri tvorbe vizualizácie dát. Nie je to len zásadami, ktoré sú uvedené vyššie, ale aj správne kroky v procese tvorby môžu dopomôcť nášmu projektu.

2.3.1 Proces tvorby pri vizualizácii dát

V knihe „Beautiful Visualization“ autor Ben Fry v procese tvorby vizualizácie dát taktiež identifikuje 7 fáz. Tieto fázy sa pokúsím v nasledujúcej podkapitole v jednoduchosti vysvetliť. Ben Fry tiež uvádza, že samozrejme sa môže stať, že niekde tieto fázy môžeme aplikovať všetky a inde len časť s z nich. Záleží od projektu.

- Acquire – môžeme chápať ako získavanie dát pre náš konkrétny účel
- Parse – analyzovanie dát, ich štrukturalizácia a kategorizácia
- Filter – filtrovanie dát a vymazávanie nepotrebných údajov
- Mine – aplikovanie metód zo štatistiky a data miningu
- Represent – výber štýlu vizuálnej prezentácie, a teda či použijeme jednoduché grafy napríklad stĺpcové, koláčové, stromovú štruktúru alebo si vyberieme zložitejšie vizualizačné nástroje a výsledok interpretujeme, napríklad pomocou Processingu
- Refine – zdokonaľovanie vytvorenej vizualizácie po funkčnej a estetickej stránke
- Interact – pridanie interakcií a animácií pre zlepšenie užívateľského zážitku

Tieto jednotlivé fázy si vyžadujú vždy špecifický talent respektíve tie správne schopnosti a mnoho krát je to v rámci jednej osoby zložitý proces. (Fry, 2008 s. 5)

Dajme vizualizácii príbeh

Na tvrdenie Ben Fry reaguje a nadväzuje Matthias Shapiro v knihe „Beautiful Visualization“, kde hovorí, že okrem skĺbenia schopnosti, ktoré vedia zastrešiť dané fázy, by mal autor ovládať ešte jednu vec a to je umenie „storytellingu“, teda rozprávania príbehu.

Podľa Matthiasa majú príbehy obdivuhodnú silu sústrediť našu pozornosť a pomáhajú nám rozpoznať, ktoré dáta a prečo sú pre nás dôležité a relevantné. Dáta sú totiž zmysluplné, len ak rozumieme kontextu, a práve preto si myslí, že využívanie dát v príbehu je vynikajúcim spôsobom, ako používateľa dostať do deja a umožniť mu porozumieť im.

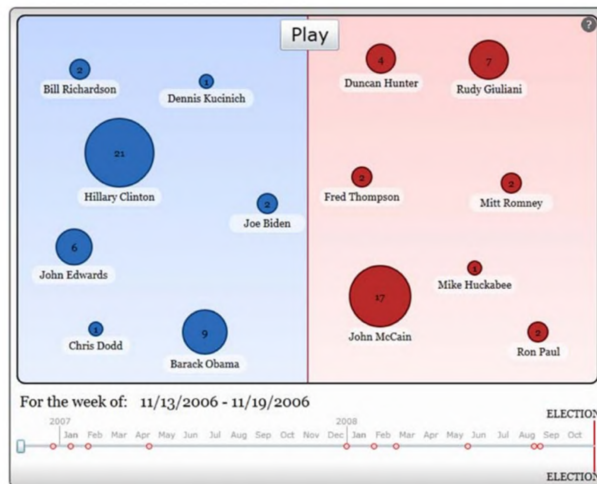
No upozorňuje, že príbeh nie je vždy nevyhnutný, ak sa jedná o jednoduchú esteticky pekne spracovanú vizualizáciu, prípadne nemá byť umelecké charakteru, kde skôr môžeme vidieť abstraktné formy vizualizácie dát. No väčšinou sa vizualizácia dát robí za nejakým hlbším účelom, a vtedy je príbeh potrebný. (Iliinsky, a iní, 2010 s. 15)

Najskôr si položíme otázku, na ktorú chceme hľadať odpoveď

Aby sme mohli používateľa uviesť do danej oblasti a nášho kontextu, je potrebné nájsť otázky, ktoré mu predložíme. Matthias hovorí, že otázky by mali obsahovať nejaký úvod do príbehu a viesť používateľa postupne, až kým dáta neprevezmú príbeh do vlastných rúk.

Kontext nám poskytne informácie, ktoré napríklad odpovedajú na otázky typu, „Na aké dáta sa pozeráme?“, „V akom sú časovom rámci?“, „Ktoré a ako udalosti ovplyvňujú dáta?“

V príklade ukazuje na vizualizáciu, ktorú vytvorili v New York Times k príležitosti prezidentských volieb. Jej cieľom bolo ukázať koľko krát bol daný prezidentský kandidát spomenutý v rôznych súvislostiach v danom týždni.



Obrázok 11 – Ukážka vizualizácie

V prípade, ak aj užívateľ nevedel vôbec nič o danej problematike, tak minimálne chápal to, že je tam časová os, ktorá označovala a postupne zobrazovala udalosti. Ak poznal niečo z kontextu, v našom príklade to môžu byť mená kandidátov, vedel, že to súvisí s prezidentskými voľbami a až po kliknutí na otáznik sa užívateľ dozvedel celý kontext tejto vizualizácie.

Čiže vizualizácii musíme ako užívateľa chápať a rozumieť, aj keď na začiatku možno nepoznáme celý kontext, postupným zoznamovaním a interakciou s ňou by nás projekt mal vtiahnuť do celého kontextu a vysvetliť nám ho. Z pohľadu tvorcu nám práve vytvorené otázky majú dopomôcť v priebehu procesu tvorby dosiahnuť takýto výsledok.

Získavanie dát

Dôležitou a neľahkou úlohou je práve získavanie toho, čo chceme vizualizovať, a to sú dáta. Slúžia k tomu rôzne platformy, kde sú dáta voľne dostupné. Zaujímavé dáta môžeme nájsť na platformách spoločnosti Google², prípadne zaujímavé môže byť využiť API známeho periodika New York Times.³

Samozrejme, ale záleží, čo chceme vizualizovať, pokiaľ sa jedná o verejná dáta, tak vyššie spomenuté zdroje by nám mohli pomôcť. Ak sa jedná napríklad o interné dáta firmy, je to zložitejšie a dáta je potrebné získavať inou cestou. Môže to byť napríklad pomocou systémov a platforiem, ktoré sa v rámci firmy používajú. V takom prípade je vhodné, ak je k týmto systémom možné pristupovať napríklad pomocou API, prípadne dáta exportovať.

Ešte pred tým, ako chceme získať informácie, je zložité prísť na to, ktoré informácie chcem zobrazovať. Keď vieme toto, môžeme prejsť k tomu, ako tieto dáta získame, a k tomu by už nám mohli postačiť vyššie zmienené postupy. (Iliinsky, a iní, 2010 s. 19)

2.3.2 Vizuálna interpretácia dát

Opäť sa budem opierať o poznatky z knihy „Beautiful Visualization“. Pokiaľ vieme, aké dáta chceme zobrazovať a vieme ich aj získať je potrebné sa zaoberať otázkou, ako tieto dáta budeme nasledovne interpretovať tak, aby ich užívateľ jednoducho pochopil.

Rozlišujeme ich podľa viacerých faktorov. Môžeme použiť bežné vizuálne prvky, ktoré nám napomáhajú utvárať hierarchiu a dôležitosť. Rozlišovať môžeme pomocou veľkosti, farieb, času, často môžeme vidieť vizualizácie pomocou máp alebo rôznych plánov. Ďalej, ako uvádza autor, to môže byť prezentácia pomocou siete.

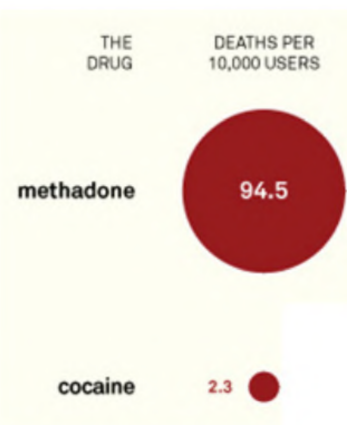
² (<https://trends.google.com/trends/>, <https://www.google.com/publicdata>)

³ (<http://developer.nytimes.com/>).

Veľkosť

Veľkosť je zrejme najbežnejším prvkom vo vizuálnej interpretácii. Pomocou veľkosti vieme rozlišovať množstvo a poukázať na rozdiel medzi danými dvoma a viacerými objektami.

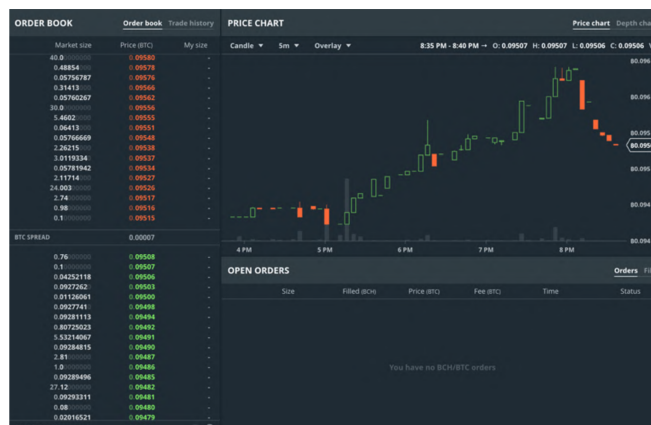
V knihe uvádza príklad vizualizácie, ktorá zobrazuje počet úmrtí v dôsledku drog. Pomocou veľkosti kruhov si vieme jednoducho a rýchlo vyjadriť tento rozdiel.



Obrázok 12 – Príklad znázornenia informácií na základe veľkosti

Farba

Farba je ďalším bežne používaným grafickým princípom, ktorý môžeme použiť. Pomocou farby môžeme jednoducho interpretovať pozitívny rast alebo naopak úbytok a prepád. Tento typ je obľúbený napríklad pri zobrazovaní prepádu respektíve rastu cien akcií, prípadne pri znázorňovaní počasia.



Obrázok 13 – Ukážka z burzy GDAX, kde sú farebne rozlíšene prepady prípadne rast cien

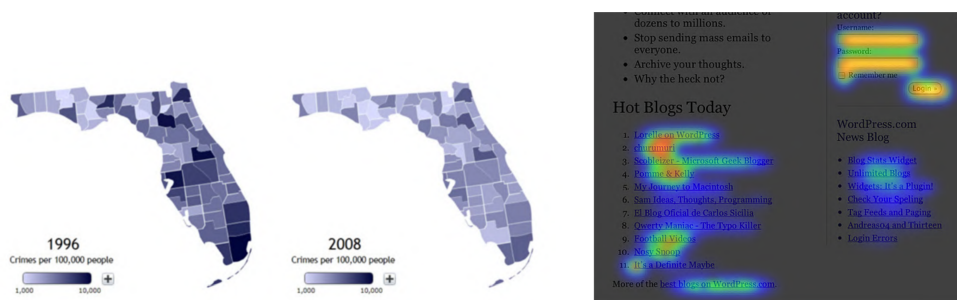
Bitcoinu

Autor tiež uvádza, že je potrebné dbať na prístupnosť pre slabozrakých, ktorým farba môže spôsobovať značné problémy. A taktiež je lepšie zobrazovať ukazovatele, kde sú rozdiely väčšie a tým pádom bude aj odtieň farby kontrastnejší, menšie odtiene farieb si môžeme horšie všímať.

Lokalizácia

Často sa môžeme stretnúť s používaním máp. To je tiež jedna z možností, ako môžeme efektívne zobrazovať dáta v rámci danej lokality. Nemusia to byť samozrejme len mapy, ale napríklad aj plán sedenia v lietadle. V príklade autor na mape vizualizuje kriminalitu v rozmedzí desiatich rokov na Floride.

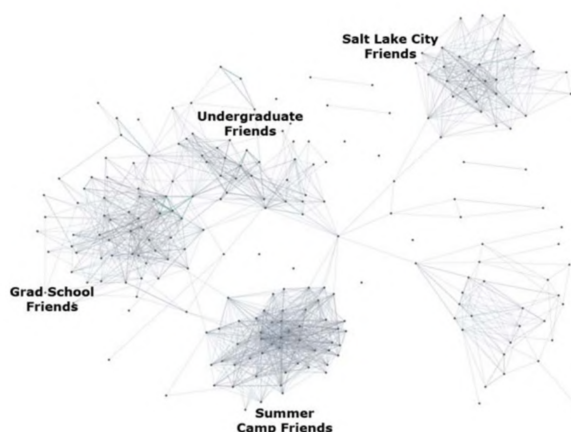
Ďalším príkladom môžu byť aj obľúbené klikacie mapy, ktoré zobrazujú kde užívatelia na stránke klikajú.



Obrázok 14 – Ukážka z vizualizácie, ktorá zachytáva kriminalitu na Floride. Na obrázku vedľa zas môžeme vidieť klikaciu mapu na stránke

Sieť

Pomocou siete môžeme zobrazovať spojenia medzi určitými dátovými bodmi. Je to užitočné pri zobrazovaní vzájomných vzťahov medzi dvomi a viacerými bodmi. Môžeme takto zobrazovať napríklad lokáciu našich priateľov na sociálnych sieťach. O niečo podobné sa pokúsil autor a zobrazil svojich priateľov zo sociálnej siete Facebook a ich vzájomné prepojenia.



Obrázok 15 – Ukážka z vizualizácie, ktorá zachytáva kriminalitu na Floride

Čas

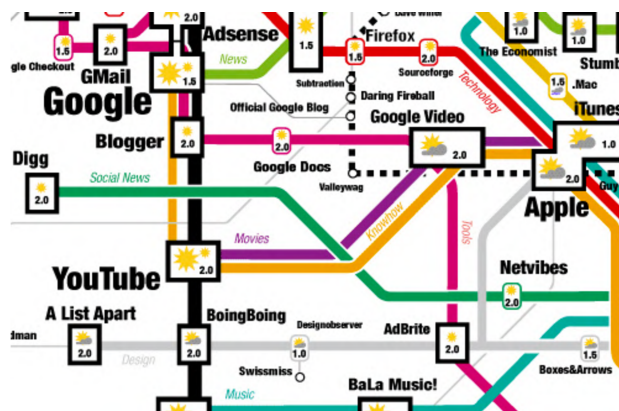
V dnešnej dobe nám technológie dovoľujú zobrazovať dáta, ktoré sa menia v priebehu času a sú zaujímavé z dôvodu, že môžeme sledovať ich pohyb a každú zmenu. Toto animované stvárnenie je jednoznačne zaujímavejšie ako statický obraz. Môžeme to často vidieť pri zobrazovaní pohybov na burze, zmenách v anketách. Jedným zo zaujímavejších príkladov, kde môžeme vidieť aj predošlý princíp máp, je vizualizácia od New York Times, ktorá zobrazovala aktuálne tweety na mape počas zápasu Super Bowl, ktoré sa menili v priebehu času. (Iliinsky, a iní, 2010 s. 24)

Vitaly Friedman v článku na portály Smashing Magazine uvádza ďalšie možnosti interpretácie. Spomína bežne dostupné koláčové, stĺpcové grafy, histogramy, rôzne tabuľky, no v článku sa skôr zameriava na nevšedné no modernejšie prístupy, ktoré dokážu taktiež efektívne interpretovať dáta. Uvádza napríklad využitie princípu myšlienkových máp, zaujímavé prístupy vizualizácie noviniek a správ a mnoho ďalšej inšpirácie.

Príklady z tých najbežnejších typov tu nie je potrebné rozpisovať, no uvediem niekoľko zaujímavých ukážok z článku. Viac sa podobným príkladom budem zaoberať v druhej časti „Oblasti využitia a prezentácie dát“.

Myšlienkové mapy

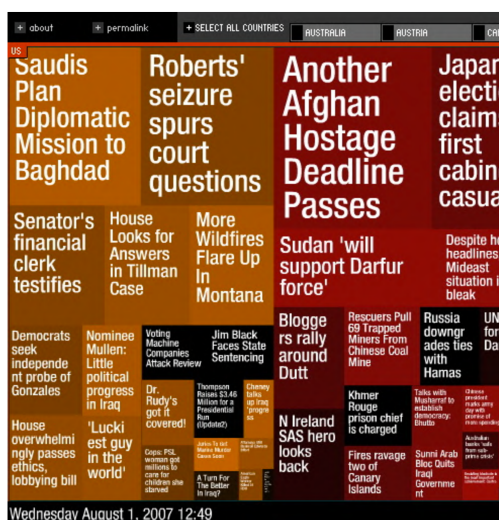
Myšlienkovú mapu premietli do mapy japonského metra, na ktorej zobrazili približne 200 odporúčaných webových stránok, ktoré kategorizovali podľa úspešnosti, popularity a veľkosti.



Obrázok 16 – Mapa japonského metra prerobená na myšlienkovú mapu odporúčaných webových stránok

Zobrazovanie správ a noviniek

Aplikácia zobrazovala správy z Google News agregátora, ktoré rozlišovala pomocou veľkosti písma, blokov a farby. Princíp rozlišovania fungoval na základe popularity článkov.



Obrázok 17 – Vizualizácia zobrazujúca správy. Popularita článku ovplyvňuje veľkosť bloku.

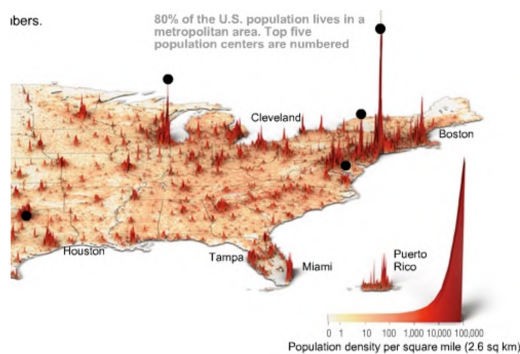
Ďalšie príklady zobrazovania informácií

Zaujímavým príkladom je aplikácia, ktorá zobrazovala nahrané obrázky v službe Flickr v reálnom čase. Tie následne vyskladala do aktuálneho času.



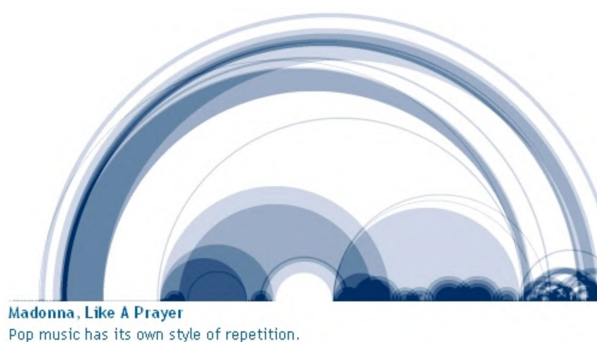
Obrázok 18 – Nahrané obrázky v službe Flickr sa vykresľovali do aktuálneho času.

Magazín Time použil vo svojej vizualizácii kopce na mape aby zdôraznil hustotu americkej populácie.



Obrázok 19 – Vizualizácia hustoty obyvateľstva.

„Shape Of Song“ zisťovali, aký môže mať tvar hudba. Pokúsili sa to vizualizovať pomocou na mieru vytvoreného programu. (Friedman, 2007)



Obrázok 20 – Vizualizácia hudby.

2.3.3 Vizualizácia dát pomocou grafov

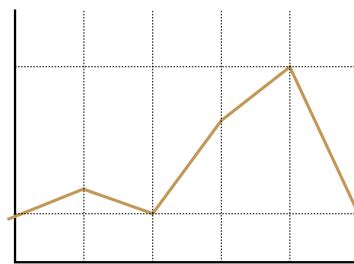
Táto časť bude venovaná základným typom grafov, s ktorými sa bežne stretávame a umožňujú nám ich vytvárať známe programy ako Excel alebo Word. Avšak aj tak pri ich

čítaní a porozumeniu nie každý dokáže pochopiť a je nutné mať vyššiu informačnú gramotnosť.

Na druhej strane máme programovacie jazyky ako napríklad Processing alebo Impure, ktoré síce dokážu vytvárať zložitejšie a vizuálne atraktívnejšie diagramy no ich pochopenie je ešte podstatne zložitejšie ako v prípade tých základných. (Jaroš, a iní, 2017 s. 72)

Spojnicový graf

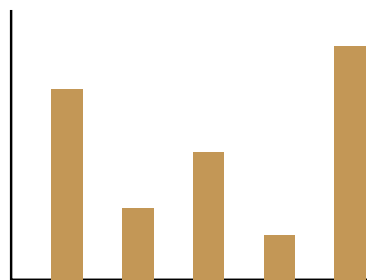
Tento diagram vynesie dátové body k mriežke a zobrazuje ich spojite pomocou liniek. Využíva sa zobrazenie množstva v nadväznosti na nejaký časový priebeh. Čítame ho tak, že sa najskôr sústredíme na osy aby sme pochopili, čo diagram reprezentuje. Potom skúmame dátové body a ich vzťah k spojniciam. Môžeme z neho vysledovať trend a prípadnú predikciu trendu do budúcnosti. (Jaroš, a iní, 2017 s. 76)



Obrázok 21 – Spojnicový graf

Stĺpcový a skupinový stĺpcový graf

Slúži na komparáciu hodnôt rôznych kategórií v určitom časovom bode. Graf čítame tak, že porovnávame výšky stĺpcov, ktoré zodpovedajú množstvu danej premennej alebo teda kategórie. (Jaroš, a iní, 2017 s. 77)

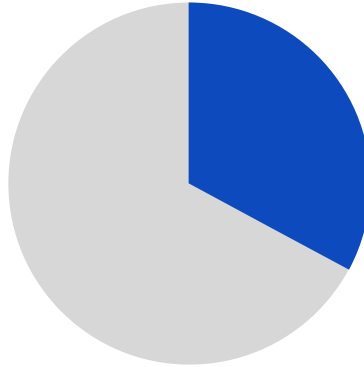


Obrázok 22 – Stĺpcový graf

Kruhový graf

Alebo aj koláčový graf slúži na zobrazenie pomerného množstva viacerých kategórií, ktoré porovnávame a da sa tento pomer vyjadriť percentuálne. Následne sú pomerné časti danej

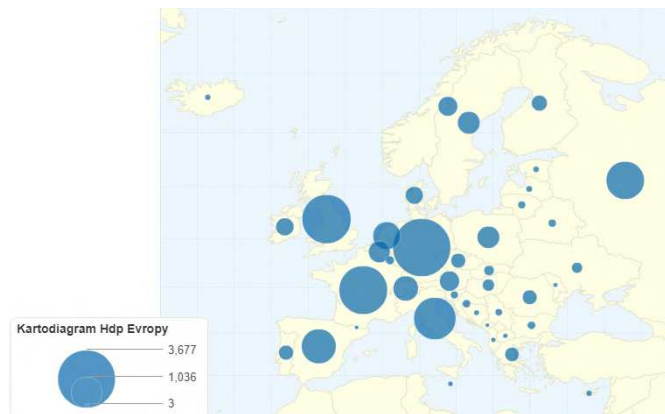
kategorie vyjadrene do kruhového objektu a tým vzniká teda spomínaný koláčový graf. Pri čítaní postupujeme tak, že porovnávame pomerové časti v rámci grafu. (Jaroš, a iní, 2017 s. 79)



Obrázok 23 – Kruhový graf

Kartodiagram

Znázorňuje hodnoty pomocou veľkosti zástupného objektu, napríklad množstvo zobrazíme väčší alebo menší kruh, ktorý relatívne zobrazí hodnotu. Základom býva geografická mapa. Tento diagram zobrazuje kvantitatívne dáta, ktoré napríklad týkajú územných celkov. (Jaroš, a iní, 2017 s. 85)



Obrázok 24 – Kartodiagram

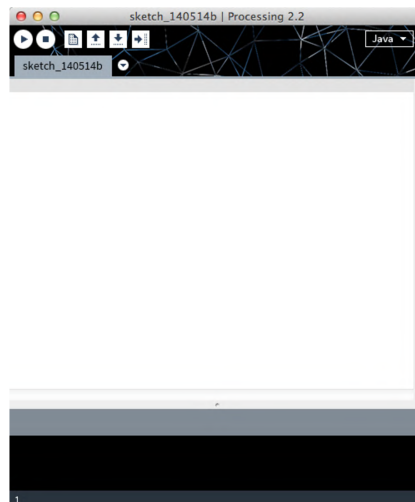
2.3.4 Technológie a software na spracovanie a vizualizáciu dát

Častokrát sa na webových stránkach môžeme stretnúť s rôznymi interaktívnymi infografikami a inými prezentáciami dát. Tie nám pomáhajú generovať rôzne nové technológie a vznikajúce knižnice na spracovanie dát. Pozrime sa na niekoľko z nich.

Processing

Asi najznámejšou technológiou je open-source⁴ programovací jazyk Processing. Jeho základ je postavený na programovacom jazyku Java. Processing je veľmi obľúbený medzi umelcami a vlastne za tým účelom je aj vytvorený, ako nástroj pre kreatívne programovanie. Nie je veľmi zložitý, disponuje vlastným užívateľským rozhraním a kompilátorom a taktiež obsahuje množstvo knižníc. Výsledky je možné sťahovať či už do rastrových alebo aj vektorových formátov. Jeho tvorcovia sú Ben Fry a Casey Reas.

Od Processingu sa odvodilo niekoľko ďalších variant a nastavieb pre ďalšie programovacie jazyky. Je to napríklad o Processing.js, p5.js, ktoré sú písane v JavaScripte a výsledok sa zobrazuje do webového canvasu⁵, takže výsledok je možné zobrazovať na webe. Ďalej to je Processing.py, čo je mód pre programovací jazyk Python. (Processing Foundation)



Obrázok 25 – Vývojové prostredie v Processingu

D3.js

Z JavaScriptových knižníc je tiež veľmi obľúbená knižnica D3.js, ktorá je zameraná práve na vizualizáciu dát a môžeme v nej vytvárať od základných grafov po zložité vizualizácie. (D3.js)

openFrameworks

⁴ Otvorený a slobodne šíriteľný program

⁵ Element, ktorý slúži na vykresľovanie vo webovom prehliadači

Ďalším zaujímavým nástrojom je open-source toolkit openFrameworks, ktorý bol podobne ako Processing tiež vyvinutý na účely kreatívneho programovania. openFrameworks napísaný v programovacom jazyku C++ a postavený na OpenGL technológii.

Jeho výhodami je taktiež jednoduchosť, takže je vhodný aj pre začiatočníkov, no na druhej strane je veľmi výkonný. Má dobrú užívateľskú základňu čiže komunitu, ktorá sa jednak stará o ďalší vývoj a tvorbu knižníc ale aj o podporu pre tvorcov. Ďalšou výhodou je, že je cross-platformový, čiže je možné ho používať na rôznych operačných systémoch. (openFrameworks)

Cinder

Cinder je podobne ako openFrameworks open-source založený na programovacom jazyku C++ a tiež bol vytvorený za účelom kreatívneho programovania. Taktiež sa jedná o cross-platform. (Cinder)

V prospech openFrameworks hrá fakt, že disponuje väčšou popularitou. Okrem toho openFrameworks má výhodu v tom, že veľké množstvo závislostí na open-source knižniciach, takže komunita má väčšiu kontrolu nad celým nástrojom. Cinder je oproti tomu viac závislý na knižniciach, ktoré sú súčasťou operačného systému.

Arduino a Raspberry Pi

Určite netreba zabudnúť aj na hardware-ové možnosti, ktorými sú napríklad Arduino a Raspberry Pi. Jedná sa o malé jednodesktopové počítače vo veľkosti približne kreditnej karty. Obidve platformy sú obľúbené a majú za sebou silnú komunitu.

Arduino však pôsobí jednoduchším dojmom. Má vlastné vývojové prostredie, ktoré je podobné Processingu a práve spojenie napríklad Processingu a takéhoto počítača môže byť zaujímavým. Arduino ani Raspberry Pi priamo neslúžia na vizualizáciu, no ich sila je v doplnkoch, ktoré je k nim možné pripojiť. Môžu byť rozšírené o rôzne pohybové čidlá, senzory a tieto údaje môžeme vizualizovať pomocou napríklad vyššie spomenutého Processingu.

II. PROJEKTOVÁ ČÁST

3 CIEĽ A KONCEPT VÝSTUPNEJ PRÁCE

Keď som si vyberal tému BP, tak pôvodným cieľom nebolo zamerať sa na vizualizáciu osobných dát a sociálne siete, no vedel som len, že oblasť dát ma veľmi zaujíma. Dlho som premýšľal, ako uchopiť tak rozsiahlu tému, a zároveň v nej nájsť zanieťenie. Síce to trvalo relatívne dlhú dobu, ale som rád, že som nakoniec šiel cestou vizualizácie osobných dát.

Teoretické zistenia a rešerš, ktoré som spravil na tému dát bol však veľmi obohacujúci, a zámer sústrediť sa na moje osobné dáta bol dobrý. Zistenia ako fungujú veľké spoločnosti prevádzkujúce najväčšie sociálne siete rovnako aj to, ako pracujú s našimi dátami boli pre mňa veľmi dôležité a mohol som naviazovať na skúmanie, ako funguje samotná vizualizácia dát.

3.1 Cieľ práce

Pri stanovení cieľa som sa zameril na dva aspekty, a to osobný a verejný. Osobne som sa chcel zdokonaľiť v tejto oblasti ako po vizuálnej stránke, tak aj k pochopeniu, ako tvoriť vizualizácie dát, a aký proces je nutné dodržiavať. Verejný cieľ bol poukázať ľuďom, ktorí budú konfrontovaní s mojou prácou, že dôležitosť súkromia môže byť dnes dôležitejšia ako kedykoľvek v minulosti – nepochybujem o tom, že ľudia to vnímajú, no častokrát sa podľa toho aspoň podľa môjho názoru nesprávajú. Rovnako častokrát nevieme, a sám som nevedel, aké informácie o nás zbierajú.

Pôvodný zámer bol zozbierať dáta zo všetkých platforiem, ktoré zbierajú moje informácie, avšak rozsah všetkých dát by bol nad rozsah bakalárskej práce. Preto som sa rozhodol zamerať len na jednu platformu (Facebook) a venovať sa jej viac do hĺbky. Zber dát prebiehal od mojej registrácie z 2. marca 2009 do 29. marca 2019, kedy som dáta sťahoval. Aj preto som prácu nazval „10 years on Facebook“.

3.2 Koncept

Musím povedať, že koncepcná časť bola pre mňa jednoznačne najťažšia časť projektu. Bolo potrebné nájsť vhodné riešenie, aby boli vyššie spomenuté ciele splnené a idea bola vhodne odprezentovaná.

V koncepte samozrejme tvoria hlavnú časť grafy, ktoré sú segmentované do viacerých celkov (kategórií). Tieto informácie sú rozdelené na segmenty ako „Base data“ – základné informácie o autorovi (o mne), „Friends data“ – pohľad na pomer priateľov a ďalších kontaktov, ktoré o autorovi Facebook drží. „Interaction data“ – zobrazuje súkromné a verejné interakcie autora. Interakciami sa myslí každá zmienka, ktorá bola autorom zaznamenaná na Facebooku. V prípade súkromných dát, tak sú to správy alebo vyhľadávania. V prípade verejných sú to statusy, komentáre alebo reakcie „páči sa mi to“. Ďalšou kategóriou sú „Location data“, ktoré zaznamenávajú najčastejšie navštívené miesta. Treba podotknúť, že nakoľko nebola aplikácia sociálnej siete používaná väčšinu zo sledovaného obdobia, tak tieto údaje nemôžeme hodnotiť ako úplne presné. Jednou z alternatív na lokalizáciu je využiť dáta, ktoré zaznamenávajú z prihlasovania pomocou IP adresy. No problém je, že tieto dáta nie sú úplne presné, nakoľko IP adresa v prípade napríklad prihlasovania cez mobilný internet nemusí korešpondovať s miestom, kde som sa aktuálne nachádzal. Ďalšou kategóriou sú „Advert data“, kde je zobrazený autorov referenčný reklamný profil. V závere sa nachádza už len krátke zhodnotenie.

Druhým výrazným prvkom je autorská časť, ktorej úlohou je zobrazit' moju reakciu ako autora a vlastníka dát na niektoré výsledky, ktoré sa dajú z grafu vyčítať, sú pre mňa kontroverzejšie a mali by zvýrazniť tento problém u publika. Táto časť je doplnená o „chatbota“, ktorý reaguje v mene Facebooku a snaží sa odpovedať na daný problém.

4 PROCES TVORBY

4.1 Zber dát

Dôležitou súčasťou bolo zozbieranie všetkých potrebných dát. To prebiehalo postupne z každej platformy samostatne. Konkrétne sa jednalo o platformy Facebook (4,69 GB), Google (6,7 GB + približne 9 GB dát z Gmailu), Instagram (106,5 MB) a ostatné platformy (Spotify, Garmin, Apple), ktorých súčet bol minimálny (do 30 MB). Konkrétne posledne menovaný Apple ma zaujal, keďže moje očakávania boli výrazne vyššie. Veď predstava, že zariadenia tejto spoločnosti so mnou strávia počas dňa najviac času sa výrazne líšila od reality. Dáta na stiahnutie boli vo veľkosti 22,5 MB. Pátral som, prečo tomu tak je, a podľa zistení z ich stránky sa práve Apple snaží dbať na bezpečnosť svojich používateľov a tieto informácie ďalej nepoužíva (Apple.com). Otázka do akej miery sa tomu dá veriť, no myslím, že pri tak veľkej spoločnosti, ako je Apple, si nemôžu dovoliť zavádzať svojich používateľov. Na druhej strane Google veľkú časť dát ukladá, a to napríklad aj údaje z OS Android, pokiaľ zariadenie s týmto OS vlastníte.

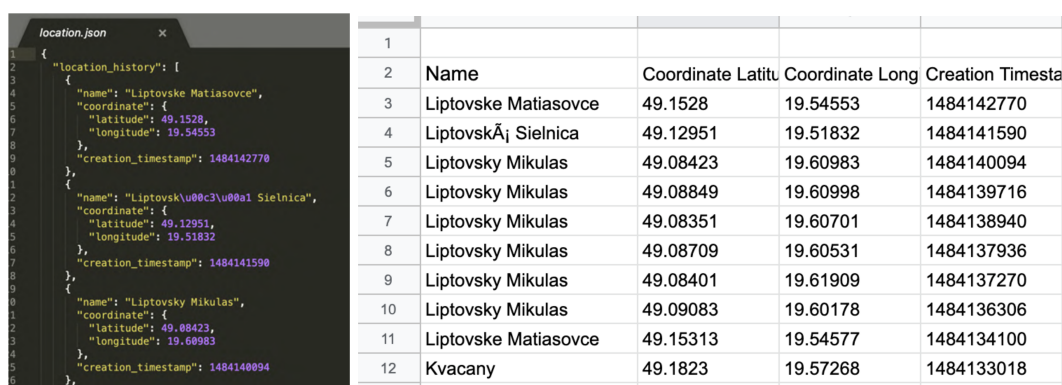
4.2 Rozdelenie dát

Potom, ako som mal k dispozícii všetky potrebné informácie, bolo potrebné rozanalyzovať dáta a vybrať tie vhodné. Facebook umožňuje stiahnutie dát v HTML alebo JSON formáte. HTML formát je síce pohodlnejší a dá sa lepšie čítať, no pre moje potreby som potreboval surové dáta, a preto som zvolil JSON formát, čo je v jednoduchosti objekt, ktorý umožňuje ukladať, prijímať a posilať texty alebo polia, a nad nimi môžeme pomocou napríklad JavaScriptu vytvárať potrebné operácie. Napríklad môžeme prijať tieto súbory zo serveru, vytiahnuť si potrebné informácie a zobrazit' ich na webe užívateľovi. (w3schools.com)

Keďže tieto JSON súbory obsahovali aj veci, ktoré pre mňa neboli relevantné a vedel som, že ich nebudem zobrazovať kvôli napríklad zložitosti alebo kvôli súkromiu ľudí (napríklad obsah správ), bolo potrebné tieto súbory „rozpársovať“ na to potrebné. Využívajú sa nato rôzne programovacie jazyky, no nakoľko ovládam na nejakej pokročilejšej úrovni len JavaScript, použil som ten.

V ďalšom procese zobrazovania dát som zistil, že skvelým nástrojom je Excel, a moje nadobudnuté zručnosti zo strednej školy a hodín informatiky mi teraz padli naozaj vhod. Používal som najmä Excel na tvorbu kontingenčných tabuliek a tiež jeho verziu pre Google Docs – Google Spreadsheet, ktorý umožňuje vytváranie JavaScript funkcií nad jednotlivými poľami. To mi veľmi zjednodušilo samotné spracovanie dát, a nemusel som všetko „rozparšovať“ pomocou čistého JavaScriptu. Google Spreadsheet umožňoval pomocou tohto scriptovania jednak import JSON súborov, čo základná verzia Excelu neumožňuje, a tiež napríklad výpočet dátumu, ktorý Facebook ukladá v milisekundách, a bolo potrebné ho prekonvertovať.

Tieto dáta boli postupne vyselektované o tie nepotrebné, a tie potrebné boli rozdelené do kategórií, ktoré som spomínal v kapitole „4.2 Koncept“.



1				
2	Name	Coordinate Latitu	Coordinate Long	Creation Timesta
3	Liptovske Matiasovce	49.1528	19.54553	1484142770
4	Liptovská Sielnica	49.12951	19.51832	1484141590
5	Liptovský Mikulas	49.08423	19.60983	1484140094
6	Liptovský Mikulas	49.08849	19.60998	1484139716
7	Liptovský Mikulas	49.08351	19.60701	1484138940
8	Liptovský Mikulas	49.08709	19.60531	1484137936
9	Liptovský Mikulas	49.08401	19.61909	1484137270
10	Liptovský Mikulas	49.09083	19.60178	1484136306
11	Liptovske Matiasovce	49.15313	19.54577	1484134100
12	Kvacany	49.1823	19.57268	1484133018

Obrázok 25 – Prevod JSON formátu do CSV

4.3 Definovanie cieľového publika a vytvorenie príbehu

Toto boli dva nutné kroky, ktoré som si potreboval ujasniť pre ďalší priebeh. Publikum, ktoré bude konfrontované s mojou prácou, nie je to len vedúci práce a oponent, ale aj ďalší, ktorí sa zúčastnia obhajob a následnej výstavy ako komisia a návštevníci. Vedel som, že koniec koncov to budú prevažne študenti a takmer všetci budú z kreatívneho prostredia.

Je to dôležité preto, aby som mohol nájsť vhodný spôsob, akým budem interpretovať zistenia a dáta. Bolo nutné obaliť koncept do príbehu, ktorý bude pre danú skupinu publika vhodný, a nielen pre publikum, ale príbeh je potrebný aj pre samotné pochopenie grafov a podanie finálnej myšlienky, a tiež som chcel dodržať istú linearitu. Namieste preto boli otázky – má to byť viac vizuálne? Alebo naopak pragmatické? Bude to skôr data-estetika? Alebo

zobrazenie pomocou napríklad typografie? Má to byť interaktívne alebo vytlačené na papiery? Myslím, že vo finále sa jedná o nejaký kompromis všetkého.

4.4 Výber výstupného média

Ako výstupné médium som zvolil webovú prezentáciu, nakoľko mojím cieľom bolo priniesť do výsledku dynamiku a interaktivitu. S tým súvisel ďalší postup prác ako vytvorenie drôtených modelov, tvorba prototypu, vymyslenie animácií, a tak ďalej.

4.5 Štruktúra webu

Nakoľko som sa rozhodol, že výstup bude interaktívnym médiom, a teda web, bolo potrebné si ujasniť, akým štýlom bude web robený. Práve preto som vytváral wireframes, čo je základný model bez akýchkoľvek vizuálnych prvkov a slúži na ukážku rozloženia prvkov na webe. Web sa skladal z obsahovej časti a časti, kde sa mala objavovať moja autorská reakcia.

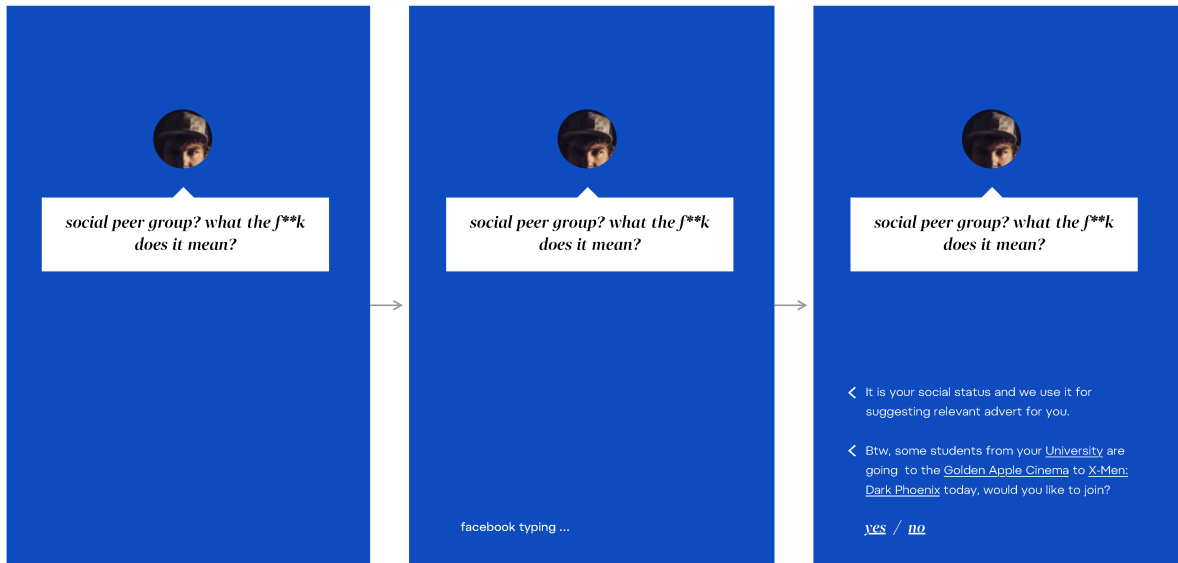
Obsahová časť

V obsahovej časti sú zobrazené grafy a všetky ostatné vizuálne a navigačné prvky.

Autorská reakcia + Chatbot

Táto časť ma byť vizuálne odlišená a výraznejšia. Slúži na zobrazenie autorskej reakcie, ktorú dopĺňa chatbot, ktorý reaguje na autorovu otázku/reakciu.

Pri navrhovaní chatbotu bolo nutné sa zamyslieť možnými stavmi a hlavne, ako to bude v princípe fungovať. Pri zobrazení danej sekcie sa v pravej časti zobrazí moja reakcia, a následne sa spustí chatbot, ktorý sa snaží odpovedať alebo nejak reagovať na moju autorskú reakciu. Snažil som sa vierohodnosť zachytiť v detailoch, teda najskôr sa zobrazí stav, keď chatbot začne reagovať a ešte len píše, následne prichádza stav, keď sa správa zobrazí. Toto sa opakuje, kým sa nezobrazia všetky predpripravené správy. Možné je, že sa chatbot niečo opýta. V takom prípade sa zobrazia možné odpovede, na ktoré môže človek obsluhujúci prototyp odpovedať.

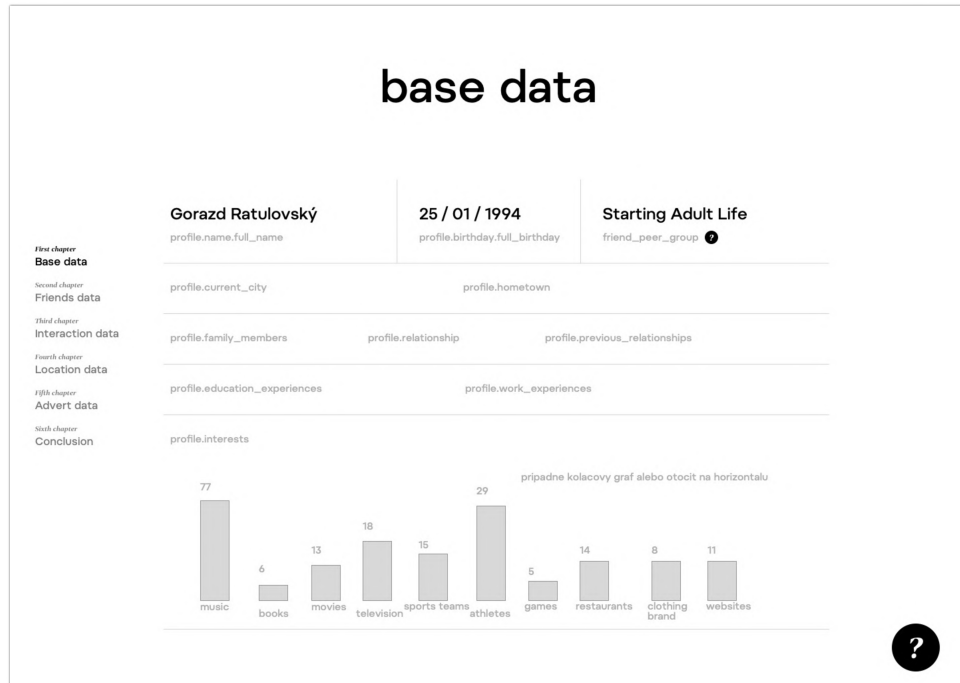


Obrázok 26 – Ukážka stavov (reakcia, Facebook píše, zobrazenie odpovede)

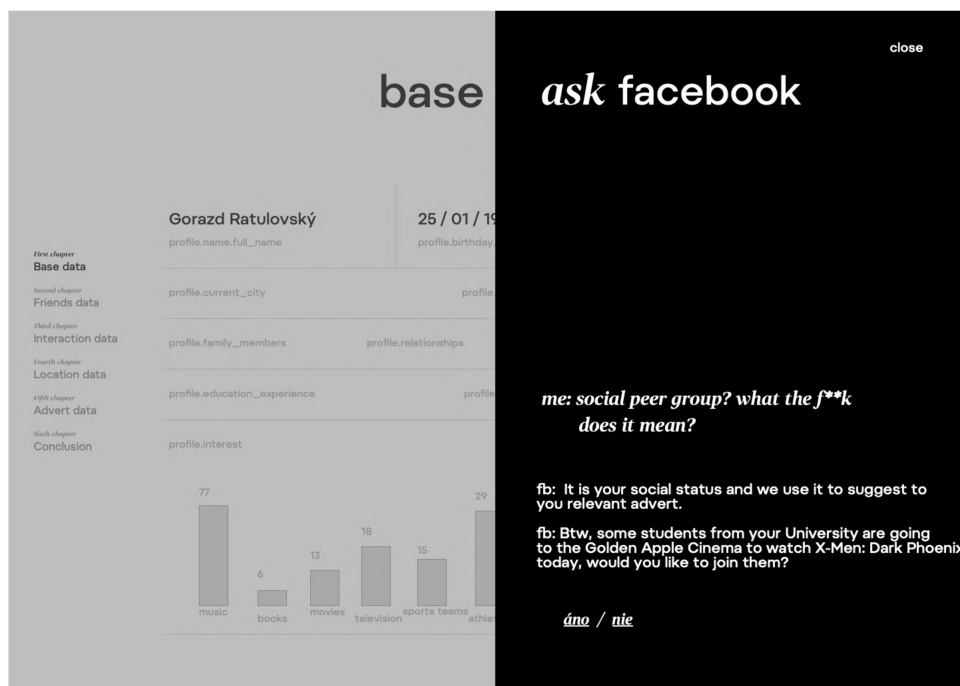
4.6 Textová časť webu

Keď už som mal vytvorené wireframes bolo potrebné sa zamyslieť nad presnými textami, ktoré použijem. Je pravda, že v obsahovej časti, kde sa nachádzali grafy, tých textov veľa nie je, no dôležitou súčasťou bola časť s autorskou reakciou a chatbota, kde bolo potrebné napísať reakcie, no taktiež nasimulovať odpovede chatbota. Dáta, ktoré chatbot zobrazuje, boli inšpirované oficiálnymi zdrojmi Facebooku prípadne inými médiami, no tie bolo potrebné nejak zaobaliť.

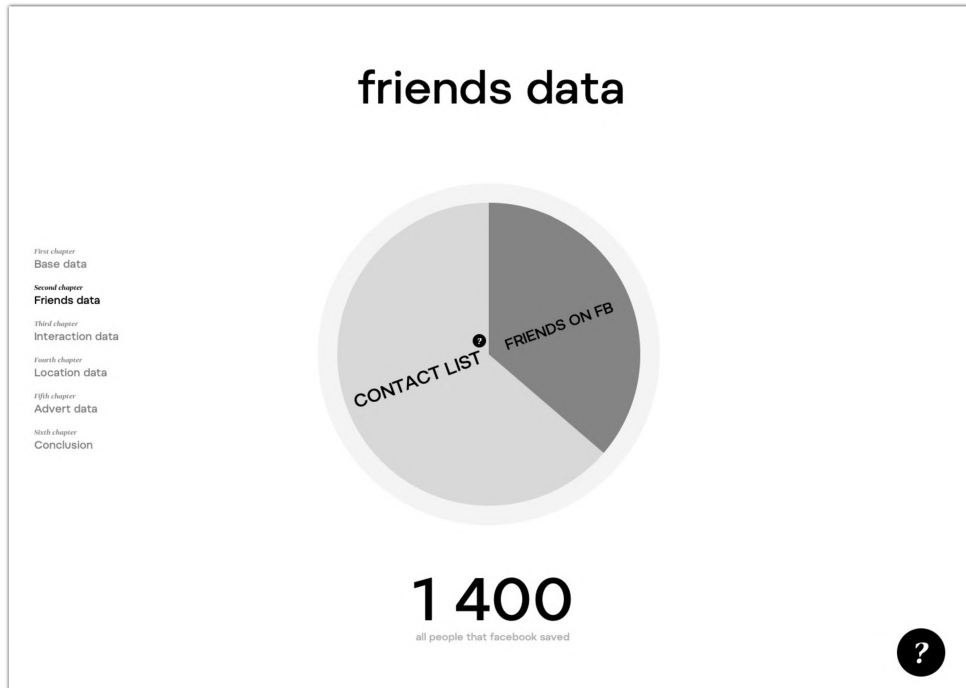
Snažil som sa jednak do reakcii ale aj do odpovedí chatbotu vniesť neformálny až ironizujúci štýl, ktorý by mal mať miestami vtipnejší podtón. Odpovede chatbotu sú na jednej strane písané reálne, no tiež bolo zámerom, aby bolo vidieť, že je to len pokus o realnosť a že je to len nedokonalý chatbot, ktorý nie je človek.



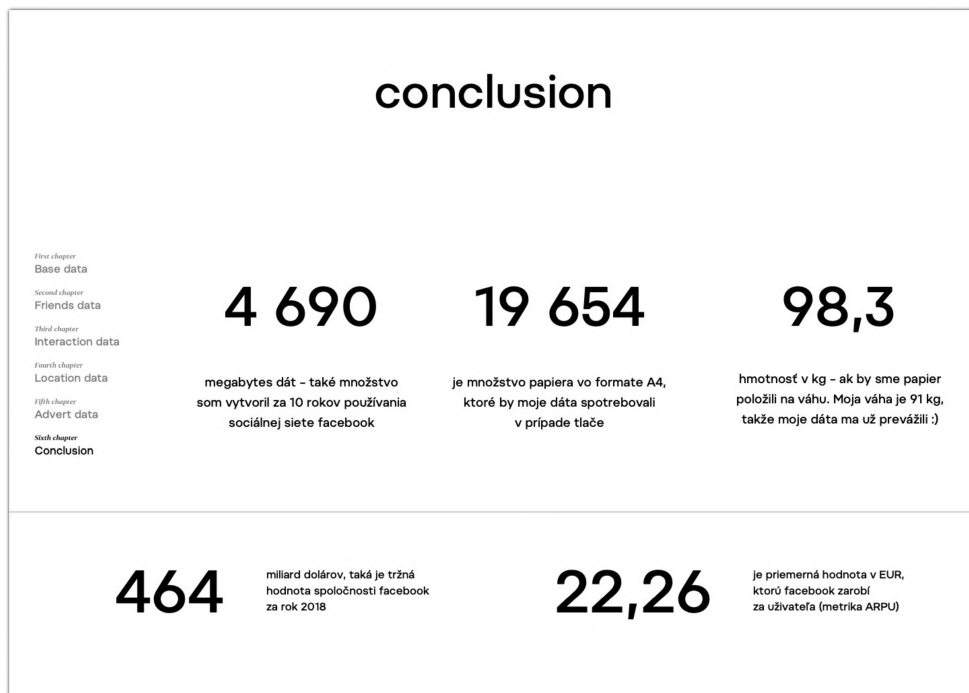
Obrázok 27 – Ukážka wireframe (základné dáta o autorovi)



Obrázok 28 – Ukážka wireframe (ukážka ako bude vyzerat' reakcia autora a chatbot)



Obrázok 29 – Ukážka wireframe (zobrazenie dát priateľov v pomere s kontakt listom)

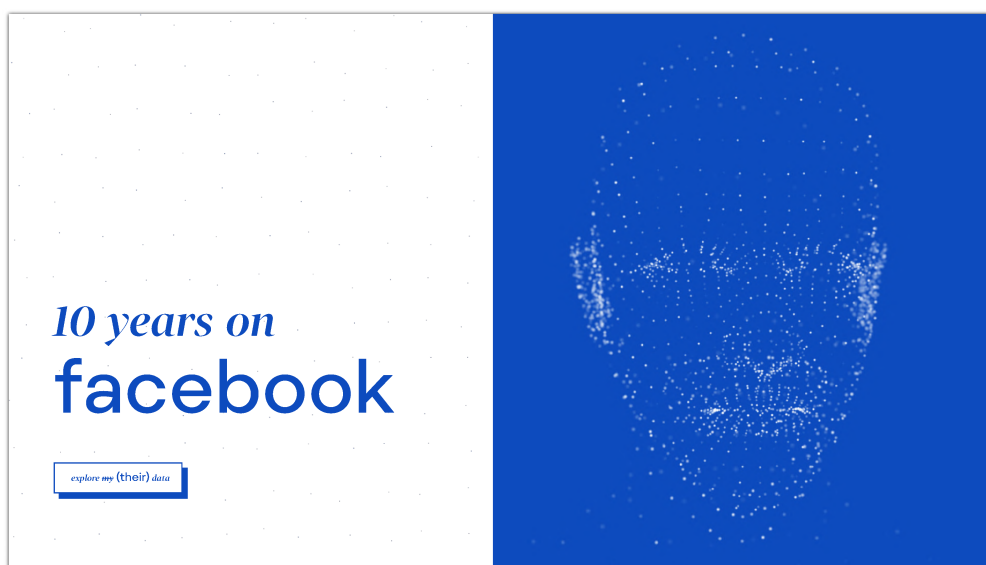


Obrázok 30 – Ukážka wireframe (ukážka ako by malo vyzerať záverečné zhodnotenie)

4.7 Vizuálne spracovanie

Výsledkom predchádzajúcich krokov je finálne spracovanie už aj s grafickými prvkami. Jednak bolo potrebné vytvoriť vizuálny štýl celému konceptu, ktorý sa týkal celkovej prezentácie, ale dosť zásadné bolo vytvoriť grafické princípy použité na grafoch. Veľmi nápomocným mi bol Illustrator, ktorý mi dokázal vytvoriť niektoré grafy, a ja som ich už len graficky upravil podľa mojich predstáv a potrieb.

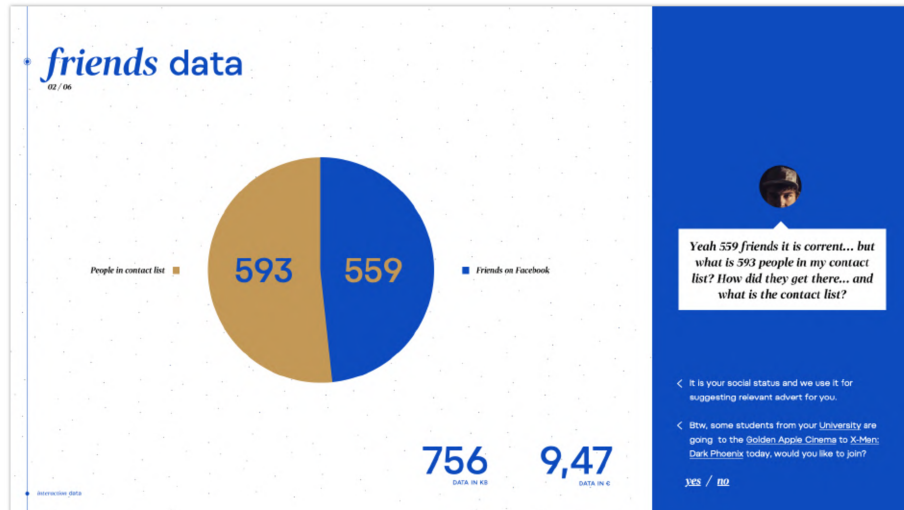
V kapitole „Vizuálne spracovanie konceptu a vizuálne princípy“ sa jednotlivým grafickým prístupom venujem podrobnejšie.



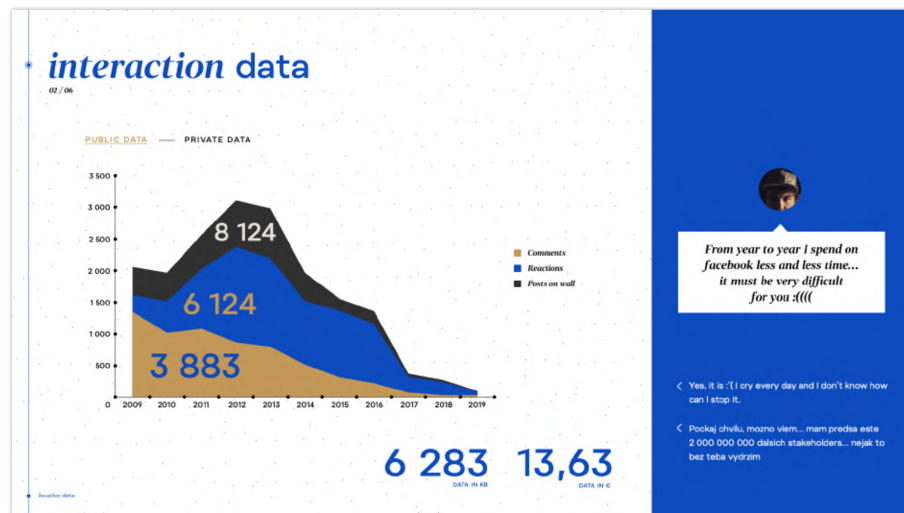
Obrázok 31 – Vizuálne riešenie úvodnej obrazovky



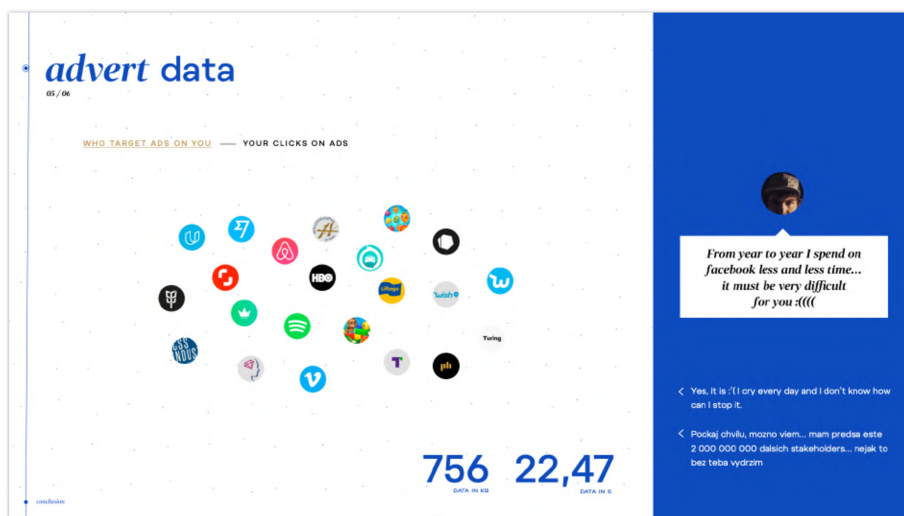
Obrázok 32 – Finálne riešenie obrazovky základných informácií



Obrázok 33 – Finálne riešenie obrazovky pomeru priateľov vs. kontakt list



Obrázok 34 – Finálne riešenie zobrazenie trendu interakcií dát od založenia po súčasnosť



Obrázok 35 – zobrazenie aké spoločnosti na mňa zámerne cielili svoje reklamy

5 GRAFY A VIZUALIZÁCIA DÁT

Dôležitou súčasťou boli samozrejme grafy a ich použitie. Ukazovalo sa viacero ciest, ako tie dáta zobraziť – ísť viac do data-estetiky a niečoho experimentálnejšieho alebo pracovať so základnými typmi grafov. Keďže cieľom bolo podať výsledok čo najjasnejšou formou, zvíťazila druhá možnosť.

Rozmanitosť dát ma prinútila zamyslieť sa nad rôznymi možnosťami, ako zistené informácie interpretovať. Pre nie všetky typy údajov bolo vhodné použiť graf, ako ho poznáme, niečo bolo potrebné vizualizovať pomocou typografie a skôr sa zamerať na to, aby to pôsobilo ucelene.

5.1 Interpretácia pomeru a množstva

V mojom prípade som musel riešiť interpretáciu množstva pri vyjadrení množstva záujmov v jednotlivých kategóriách záujmov. V tomto prípade som použil horizontálny stĺpcový graf.

Ďalším príkladom zobrazenia pomeru a zároveň množstva bolo zobrazenie mojich priateľov na Facebooku vs. kontaktov, ktoré si Facebook automaticky sťahuje a ukladá. V tomto prípade som použil kruhový diagram.

5.2 Interpretácia lokality

Keďže Facebook zaznamenáva aj to, kde sa nachádzame v daný čas, chcel som poukázať aj na tento fakt. Keďže v dátach je možné nájsť od, samozrejme, miesta a GPS súradníc až po deň, kedy sme sa v tej lokalite nachádzali. Na toto zobrazenie som použil mapu, v ktorej som interpretoval zistené informácie.

5.3 Interpretácia času (množstva v čase)

V prípade sekcie „interaction data“ bolo mojim cieľom zobraziť akúkoľvek moju interakciu vytvorenú na Facebooku od registrácie. Za interakciu je považované každé vyhľadávanie, každá odoslaná správa, každá reakcia na príspevok, zmena statusu na „Facebook wall“ alebo komentár.

Zaujímavý mi prišiel pohľad na tieto dáta v čase a pohľad, ako sa menilo moje správanie v rámci rokov a dospievania. V mojom prípade môžem povedať, že mladícka nerozvážnosť u mňa zohrala rolu, no čím som bol starší, tak som si uvedomoval hodnotu toho, čo na sociálnych sieťach vyslovím, a kde svoju stopu zanechám. Ďalším faktom určite je, že Facebook už nie je tak trendy sociálnou sieťou ako pred pár rokmi. Zaujímavý pohľad by sa mohol naskytnúť pri porovnaní viacerých ľudí za rovnaké obdobie a v rovnakom veku.

Pre tento typ informácií bolo prioritou ukázať trend používania v čase a množstvo pre daný rok alebo aj celkovo bolo až druhoradým kritériom. V tomto prípade som použil líniový graf.

5.4 Interpretácia ostatných dát

Facebook o nás ukladá obrovské množstvo dát a prístupov, akými ich môžeme zobrazit', je veľa. Každopádne sú typy dát, ktorých zobrazovanie v grafoch, ako ich poznáme, nie je nutné alebo možné. Môže sa jednať o klasický zoznam položiek alebo základné údaje o používateľovi, kde efektívnejším spôsobom je zobrazit' dáta v celku a jasne.

Ja som napríklad pracoval s dátami, ktoré vedú o mne. Sú to bežné informácie ako moje meno, vek, rodinný stav, pracovné skúsenosti ale aj pre mňa celkom zaujímavé informácie, ktoré o nás vedú ako sociálne zaradenie, alebo aj zoznam mojich ex-priateliek.

Tieto informácie som sa snažil zoradiť do relevantných skupín a hodnoty zobrazit' len pomocou textu.

Ďalším typom údajov, ktoré zobrazujem, sú spoločnosti, ktoré zámerne cielili na moju osobu. Je to klasický zoznam firiem a žiadne ďalšie údaje ako napríklad, kedy na mňa cielili, alebo koľko krát nemám k dispozícii. Preto som sa zamýšľal, ako priniesť možno trochu inakosti a nezobrazit' len klasický zoznam. Logá spoločností som dal do kruhových másk a vytvoril som akési pole, v ktorom sa tieto spoločnosti nachádzajú a pridal trochu dynamiky a interakcie. Pri prvom zobrazení sa toto pole rozbije a tieto drobné kruhy do seba môžu narážať.

6 VIZUÁLNE SPRACOVANIE KONCEPTU A VIZUÁLNE PRINCÍPY

6.1 Farby

Počiatočný zámer bol skúsiť použiť rozmanité palety, no nakoniec som sa dostal ku kombinácii modrej a bielej. Môže sa to zdať ako klišé, keďže podobnú paletu Facebook používa aj v brande, no myslím, že podobnosť nie je zlá, a práve naopak si to môžeme spojiť hneď s tým, čo komunikujeme. A zároveň použitá farba je výrazne sýtejšia ako brandová.

Doplňková farba oranžovo-hnedá s nádychom zlatej farby, ktorá mi napomáhala najmä pri odlišovaní grafov.



Obrázok 36 – ukážka farebnej palety

6.2 Typografia

Typografia zohráva v mojom riešení zásadnú časť, a to nie len z estetického hľadiska ale aj funkčného. Mojm cieľom bolo vo finálnom riešení rozlíšiť rolu autora a rolu Facebooku. A vhodné mi prišlo použiť dve typy písma, teda presnejšie dve rozličné písmové rodiny a z rozličnej klasifikácie.

Kritérium pre písmo použité najmä pre reakcie autora bolo, aby podčiarklo ideu, že je to reakcia reálneho človeka a nie nejakej digitálnej formy. Malo by byť preto dynamickejšie, tradičnejšie, ľudskejšie no tiež vizuálne výrazné. Zvolil písmovú rodinu „Tabac G1“ od „Suitcase Type Foundry“, ktorá je serifového typu. Pracoval som nakoniec s rezom SemiBold Italic, ktorý spĺňa kritériá.

Druhý typ písma má evokovať reakcie Facebooku. Kritérium bolo, aby písmo pôsobilo viac digitálnejšie, geometricky a strojovejšie. Zvolil som písmovú rodinu „Pepi“ taktiež od „Suitcase Type Foundry“, konkrétne rez Medium.

Princípom je vhodná kombinácia týchto dvoch písiev nie len v rámci odstavcov, ale aj v rámci slovného spojenia alebo vety. Takáto kombinácia vyzdvihuje základnú ideu, a zároveň to vyzerá príjemne aj z estetického hľadiska.

Pepi Medium

Tabac G1 SemiBold Italic

I like my data but Facebook loves it more

Obrázok 37 – návrh typografie

ZÁVER

„Keby som mal k dispozícii hodinu na zvládnutie problému, na ktorom by závisel môj život, strávil by som 40 minút jeho štúdiom, 15 minút jeho analýzou a 5 minút jeho riešením.“

Albert Einstein

Tento citát vystihuje proces k prístupu spracovania záverečnej práce – akurát v tomto prípade nešlo o život, a ak by sa ešte nejaká minútka navyše pridala k riešeniu, tak by bol každý autor rozhodne spokojnejší.

Teoretická časť práce mi umožnila nazrieť celkom podrobne do problematiky dát a ich vizualizácie a predstavovala obrovský prínos pre mňa samotného, a to nie len po profesijnej stránke týkajúcej sa designu spracovania, ale aj po stránke problematiky dát a súkromia ľudí v digitálnom priestore. Taktiež verím, že táto práca všetkých, ktorí mali možnosť si ju prečítať, obohatila aspoň o zlomok nových informácií a vyvolala v nich prípadnú otázku spojenú s ochranou vlastného súkromia.

Dovolím si povedať, že nakoniec sa mi podarilo dostať k ucelenému konceptu prezentácie samotných dát, čo bola rozhodne najťažšia časť celého projektu. Vzhľadom na to, že nájsť spôsob ako zobrazit' všetky dostupné dáta tak, aby to dávalo zmysel, predstavovalo tú najväčšiu výzvu. Keď som však zvládol spracovávanie surových dát, vizuálne spracovanie bolo už len pre mňa skrz profesijné skúsenosti príjemnou činnosťou na záver. Tým však nie je myslené, že projektová časť bola jednoduchá, ale analýza a štúdium teoretických poznatkov bolo zložitejšie, a naopak v tej praktickej časti som sa cítil viac doma. Môžem teda rozhodne povedať, že táto práca bola pre mňa zásadná minimálne v tom, že mi ukázala, kde sú moje prednosti a aj nedostatky, a na čom mám v budúcnosti pracovať.

ZOZNAM POUŽITEJ LITERATÚRY

- [1] Apple.com. Privacy. *Apple.com*. [Online] <https://www.apple.com/privacy/>.
- [2] Cinder. about. *Cinder*. [Online] [Dátum: 10. apríl 2018.] <https://libcinder.org/about>.
- [3] D3.js. Overview. *D3.js - Data-Driven Documents*. [Online] [Dátum: 10. apríl 2018.] <https://d3js.org/>.
- [4] Denyer, Simon. 2016. China's plan to organize its society relies on 'big data' to rate everyone. *The Washington Post*. [Online] 2016. https://www.washingtonpost.com/world/asia_pacific/chinas-plan-to-organize-its-whole-society-around-big-data-a-rating-for-everyone/2016/10/20/1cd0dd9c-9516-11e6-ae9d-0030ac1899cd_story.html?noredirect=on&utm_term=.6071e800ed72.
- [5] Desjardins, Jeff. 2019. How the Tech Giants Make Their Billions. *Visual Capitalist*. [Online] 29. marec 2019. <https://www.visualcapitalist.com/how-tech-giants-make-billions/>.
- [6] Facebook's annualized revenue per user from 2012 to 2018 (in U.S. dollars). *Statista.com* [online]. 2019 [cit. 2019-04-25]. Dostupné z: <https://www.statista.com/statistics/234056/facebooks-average-advertising-revenue-per-user/>
- [7] Facebook. Čo sú to produkty Facebooku? *Facebook*. [Online] 2019. <https://www.facebook.com/help/1561485474074139>.
- [8] Facebook. Zásady používania údajov . *Facebook*. [Online] <https://www.facebook.com/about/privacy>.
- [9] Financial Times. 2008. FT Global 500 2008. *Financial Times*. [Online] 28. jún 2008. [Dátum: 4. apríl 2019.] <https://www.ft.com/content/e01aab98-40c9-11dd-bd48-0000779fd2ac>.
- [10] Friedman, Vitaly. 2007. Data Visualization: Modern Approaches. *Smashing Magazine*. [Online] 2. August 2007. [Dátum: 9. Apríl 2018.] <https://www.smashingmagazine.com/2007/08/data-visualization-modern-approaches/>.

- [11] FRIENDLY, Michael. 2009. Milestones in the history of thematic cartography, statistical graphics, and data visualization. *University of York*. [Online] 2009. <http://www.math.yorku.ca/SCS/Gallery/milestone/milestone.pdf>.
- [12] Fry, Ben. 2008. *Visualizing Data*. O'Reilly Media, Inc, 2008. s. 5. 978-0-596-51455-6.
- [13] Global Finance. 2018. World's Largest Companies 2018. *Global Finance*. [Online] 30. november 2018. [Dátum: 4. apríl 2019.] <https://www.gfmag.com/global-data/economic-data/largest-companies>.
- [14] Google. 2019. Products. *Google*. [Online] 2019. <https://about.google/products/>.
- [15] Haselton, Todd. 2018. I checked and it turns out Facebook knows a lot about me. *CNBC*. [Online] 3. marec 2018. <https://www.cnbc.com/2018/03/27/facebook-knows-a-lot-about-me.html>.
- [16] Iliinsky, Noah a Julie, Steele. 2011. *Designing Data Visualizations*. O'Reilly Media, 2011. ISBN 978-1-449-31228-2.
- [17] Institute, IA. What is Information Architecture? *IA Institute*. [Online] [Dátum: 9. apríl 2018.] <https://www.iainstitute.org/what-is-ia>.
- [18] Julie Steele, Noah Iliinsky. *Beautiful Visualization*. O'Reilly Media, 2010, ISBN 9781449379889.
- [19] Mapping Complex Information. 2012. 20 Information Design Milestones. *Mapping Complex Information. Theory and Practice*. [Online] 16. január 2012. [Dátum: 9. apríl 2018.] <https://sheilapontis.wordpress.com/2012/01/16/20-information-design-milestones/>.
- [20] openFrameworks. about. *openFrameworks*. [Online] [Dátum: 10. apríl 2018.] <http://openframeworks.cc/about/>.
- [21] Oracle. 2014. What Is Big Data? *Oracle*. [Online] 2014. <https://www.oracle.com/big-data/guide/what-is-big-data.html>.
- [22] Pritchard, Matthew. 2016. Data Visualization Vs. Infographics. *Killer Infographics*. [Online] december 1, 2016. [Cited: apríl 10, 2018.] <http://killerinfographics.com/blog/data-visualization-versus-infographics.html>.

- [23] Processing Foundation. Projects. *Processing Foundation*. [Online] [Dátum: 10. apríl 2018.] <https://processingfoundation.org/projects>.
- [24] Richter, Felix. 2019. Facebook's Revenue Per User Rises Steeply. *Statista*. [Online] 31. január 2019. <https://www.statista.com/chart/7892/facebook-average-revenue-per-user/>.
- [25] SAS Institute. 2018. Big data and analytics software revenue worldwide by leading vendor from 2014 to 2017 (in billion U.S. dollars)*. *Statista*. [Online] September 2018. [Dátum: 4.. Apríl 2019.] <https://www.statista.com/statistics/472979/business-analytics-software-revenue-worldwide-by-vendor/>.
- [26] Smiciklas, Mark. 2012. *The Power od Infographics*. s.l. : Pearson education, inc., 2012. 978-0-7897-4949-9.
- [27] Spence, Ian. 2004. 'Playfair, William (1759-1823)', Oxford Dictionary of National Biography. *Oxford University Press*. [Online] 2004. [Dátum: 9. apríl 2018.] <http://www.oxforddnb.com/view/article/22370>].
- [28] Statista. 2018. Big data technology adoption plans in organizations worldwide from 2015 to 2017. *Statista*. [Online] 2018. [Dátum: 4. apríl 2019.] <https://www.statista.com/statistics/919670/worldwide-big-data-adoption-expectations/>.
- [29] Viktor Mayer-Schönberger, Kenneth Cukier. 2014. *Big Data*. Computer Press, 2014. ISBN 978-80-251-4119-9
- [30] Visnjic, Filip. 2012. AntiMap [oF, Android, iPhone, Javascript, Processing]. *CreativeApplications.Net*. [Online] 12. december 2012. [Dátum: 9. apríl 2018.] <http://www.creativeapplications.net/android/antimap-of-android-iphone-javascript-processing/>.
- [31] w3schools.com. JSON - Introduction. *w3schools.com*. [Online] https://www.w3schools.com/js/js_json_intro.asp.
- [32] Wikipedia. Information design. *Wikipedia*. [Online] https://en.wikipedia.org/wiki/Information_design.
- [33] Wikipedia. *Average revenue per user*. [Online] 2019. https://en.wikipedia.org/wiki/Average_revenue_per_user.

- [34] Zannier, Federico. 2013. Kickstarter. *A bite of Me*. [Online] 2013.
<https://www.kickstarter.com/projects/1461902402/a-bit-e-of-me>.
- [35] Zewdu Melkamu, a spol. 2016. Study case: Facebook. *Academia.edu*. [Online] 2016.
https://www.academia.edu/22547552/Case_study_on_Facebook_is_really_good_or_bad.

ZOZNAM POUŽITÝCH SYMBOLOV A SKRATIEK

JSON JavaScript Object Notation

HTML Hypertext Markup Language

IA Informačná architektúra

UX User Experience

ARPU Average Revenue Per User

ZOZNAM OBRÁZKOV

Obrázok 1 – Data Calculator	16
Obrázok 2 – Kalkulačka hodnoty dát podľa Financial Times	18
Obrázok 3 – Projekt Federica Zanniera	18
Obrázok 4 – Prvá svetová mapa z roku 550 p.n.l.	19
Obrázok 5 – Grafy Williama Playdaira	20
Obrázok 6 – Mapa Charlesa Minarda	20
Obrázok 7 – Mapa londýnskeho metra, Harry Beck	21
Obrázok 8 – Informačná tabuľa – vlaková stanica.....	24
Obrázok 9 – Ukážka z knihy „Budapest POP-UP Book“, Livia Hasenstaub.....	25
Obrázok 10 – Ukážka z aplikácie AntiMap.....	26
Obrázok 11 – Ukážka vizualizácie	28
Obrázok 12 – Príklad znázornenia informácií na základe veľkosti.....	30
Obrázok 13 – Ukážka z burzy GDAX, kde sú farebne rozlíšene prepady prípadne rast cien Bitcoinu	30
Obrázok 14 – Ukážka z vizualizácie, ktorá zachytáva kriminalitu na Floride. Na obrázku vedľa zas môžeme vidieť klikáciu mapu na stránke	31
Obrázok 15 – Ukážka z vizualizácie, ktorá zachytáva kriminalitu na Floride	32
Obrázok 16 – Mapa japonského metra prerobená na myšlienkovú mapu odporúčaných webstránok	33
Obrázok 17 – Vizualizácia zobrazujúca správy. Popularita článku ovplyvňuje veľkosť bloku.....	33
Obrázok 18 – Nahrané obrázky v službe Flickr sa vykresľovali do aktuálneho času.	34
Obrázok 19 – Vizualizácia hustoty obyvateľstva.	34
Obrázok 20 – Vizualizácia hudby.....	34
Obrázok 21 – Spojnicový graf.....	35
Obrázok 22 – Stĺpcový graf.....	35
Obrázok 23 – Kruhový graf.....	36
Obrázok 24 – Kartodiagram	36
Obrázok 25 – Prevod JSON formátu do CSV	43
Obrázok 26 – Ukážka stavov (reakcia, Facebook píše, zobrazenie odpovede).....	45

Obrázok 27 – Ukážka wireframe (základné dáta o autorovi)	46
Obrázok 28 – Ukážka wireframe (ukážka ako bude vyzerat' reakcia autora a chatbot)	46
Obrázok 29 – Ukážka wireframe (zobrazenie dát priateľov v pomere s kontakt listom)	47
Obrázok 30 – Ukážka wireframe (ukážka ako by malo vyzerat' záverečné zhodnotenie)	47
Obrázok 31 – Vizuálne riešenie úvodnej obrazovky	48
Obrázok 32 – Finálne riešenie obrazovky základných informácií	48
Obrázok 33 – Finálne riešenie obrazovky pomeru priateľov vs. kontakt list.....	49
Obrázok 34 – Finálne riešenie zobrazenie trendu interakcií dát od založenia po súčastnosť	49
Obrázok 35 – zobrazenie aké spoločnosti na mňa zámerne cielili svoje reklamy.....	49
Obrázok 36 – ukážka farebnej palety	52
Obrázok 37 – návrh typografie	53

ZOZNAM TABULIEK

Tabuľka 2 – Rebríček najhodnotnejších spoločností v roku 2008 podľa Financial Times	12
Tabuľka 3 – Rebríček najhodnotnejších spoločností v roku 2018 podľa PwC Global Top 100	13

ZOZNAM GRAFOV

Graf 1 – Rebríček najhodnotnejších spoločností v roku 2008 podľa Financial Times	11
Graf 2 – Svetové najpoužívanéjšie sociálne svete, január 2019, zobrazuje počet aktívnych používateľov v miliónoch.....	15
Graf 3 – Priemerná ARPU spoločnosti Facebook	17