

AKADEMIE MÚZICKÝCH UMĚNÍ V PRAZE

FILMOVÁ A TELEVIZNÍ FAKULTA

Filmové, televizní a fotografické umění a nová media

Produkce

DIPLOMOVÁ PRÁCE

**VYUŽITÍ DAT PŘI VÝVOJI, VÝROBĚ A
DISTRIBUCI AUDIOVIZUÁLNÍHO OBSAHU NA
VYŽÁDÁNÍ**

BcA. Jakub Rálek

Vedoucí práce: Mgr. Jakub Šiler

Oponent práce:

Datum obhajoby:

Přidělovaný akademický titul: MgA.

Praha, 2020

ACADEMY OF PERFORMING ARTS IN PRAGUE

FILM AND TV SCHOOL

Film, Television, Photographic Arts and New Media

Producing

MASTER THESIS

**THE USE OF DATA IN DEVELOPMENT,
PRODUCTION AND DISTRIBUTION OF
AUDIOVISUAL CONTENT ON DEMAND**

BcA. Jakub Rálek

Supervisor: Mgr. Jakub Šiler

Reviewer:

Date of defense:

Academic degree: Master degree

Praha, 2020

Prohlášení

Prohlašuji, že jsem bakalářskou/magisterskou/disertační práci na téma

**VYUŽITÍ DAT PŘI VÝVOJI, VÝROBĚ A DISTRIBUCI
AUDIOVIZUÁLNÍHO OBSAHU NA VYŽÁDÁNÍ**

vypracoval(a) samostatně pod odborným vedením vedoucího práce a s použitím uvedené literatury a pramenů.

Praha, dne

.....

podpis diplomanta

Upozornění

Využití a společenské uplatnění výsledků diplomové práce, nebo jakékoliv nakládání s nimi je možné pouze na základě licenční smlouvy tj. souhlasu autora a AMU v Praze.

Abstrakt

Uživatelská data se využívají v současnosti napříč všemi možnými odvětvími. V době, kdy je tržní západní svět orientovaný na potřeby spotřebitele a personifikovanou nabídku produktů a služeb, je práce s nimi nedílnou součástí marketingových nástrojů malých i velkých podniků a předmětem podnikání mnoha IT korporací. Cílem práce je popsat proces jejich získávání, analyzování a vyhodnocování v kontextu audiovizuálního průmyslu – zejména audiovizuálních služeb na vyžádání. V práci jsou popsány obecné principy práce s big data i konkrétní metody jejich aplikace. Na příkladu několika VoD platform je ilustrováno jejich praktické využití a vliv na jejich fungování i na celý online trh s audiovizuálními díly.

Práce odpovídá na otázky využití dat při vývoji, výrobě a distribuci audiovizuálních děl, a to jak z hlediska marketingového, tak zejména z hlediska producentského a manažerského. Výzkum vedle analýzy využití dat v AV tvorbě není primárně zaměřen pouze na technické řešení jejich sběru, ale spíše na právní, etické a politické aspekty sběru a jejich následného využití pro zmíněné účely na lokálním i světovém trhu.

Klíčová slova

Uživatelská data, Big data, Datová analýza, Video Genom, SVOD, VOD platformy, Algoritmus, Netflix, HBO, MALL.TV, Amazon, iVysílání, GDPR, Ochrana spotřebitele

Abstract

Nowdays, user data is currently gathered and used across all possible industries. At the time when the Western market oriented to consumer needs and a personalized products and services, working with user data is an integral part of the marketing strategy of both small and large companies. The aim of this thesis is to describe the process of their acquisition, analysis and evaluation in the context of the audiovisual industry especially in audiovisual services on demand. The paper describes the general principles of working with big data as well as specific methods of their application. The example of various VoD platforms illustrates their practical use and an impact on the entire online market of audiovisual products.

The thesis answers questions regarding the use of data for development, production and distribution of audiovisual products in terms of marketing and especially in terms of producing and management. The research is not only focused on the technical solution used for the data collection, but rather on the legal, ethical and political aspects of their collection and their subsequent use for mentioned purposes on the local and global market.

Keywords

User data, Big data, Data analysis, Video Genome, SVOD, VOD platforms, Algorithm, Netflix, HBO, MALL.TV, Amazon, iVysilani, GDPR

Poděkování

Můj dík patří především všem konzultantům, za jejich neutuchající vytrvalost odpovídat na mé dotazy. Jmenovitě bych rád poděkoval: Josefu Zukalovi a Dorotě Vašíčkové, Petru Szczepanikovi, Vojtěchu Matouškovi v neposlední řadě Jakubu Šilerovi.

Za nasměrování, rady a podporu děkuji vedoucí KP FAMU Karle Stojákové.

Osobní poděkování patří také Lindě Štencové a za velkou podporu děkuji Ivetě S. Jakubčíkové.

Obsah

	Úvod	1
1	Audiovizuální tvorba v současnosti	3
1.1	Proměna globálního trhu audiovizuální tvorby.....	3
1.2	Audiovizuální tvorba a VOD.....	4
1.2.1	Ekonomické faktory audiovizuální tvorby a VOD.....	11
2	Data jako nástroj pro práci s VOD	16
2.1	Big data	16
2.2	Jak VOD portál získává data.....	18
2.3	Využití big data v audiovizuální tvorbě.....	25
2.3.1	Využití big data v marketingu.....	29
2.3.2	Štítky a metadata.....	30
2.3.3	Video genom jako identita obsahu	32
2.3.4	Doporučování obsahu.....	38
2.3.5	A/B testování.....	45
3	Případové studie využití dat v audiovizuální tvorbě	48
3.1	Amazon	48
3.1.1	Amazon Prime a seriál Alpha House	49
3.2	Netflix a seriál House of Cards.....	50
3.2.1	Netflix Prize	54
3.2.2	Typy dat, které o uživatelích Netflix shromažďuje a jejich využití na příkladu House of Cards	56
3.3	Mall.TV	64
3.4	Specifika videoplatformy ČT jako VOD platformy veřejné služby.....	73
3.5	Další příklady	82
3.5.1	HBO	82
3.5.2	Stream	89
4	Politické a etické faktory využití dat v audiovizuální tvorbě na vyžádání	94
4.1	Regulace EU a možný dopad na provozovatele VOD	94
4.2	GDPR.....	95
4.3	Etický rozměr sběru a využití dat.....	100
4.3.1	Eticky pozitivní a negativní příklad práce s big data	101
4.3.1.1	Cambridge Analytica a ovlivňování politických výsledků	101
4.3.1.2	Google a mapování výskytu chřipky	102
5	Komentář k autorskoprávním řešením v kontextu big data	103
6	Možnosti potenciálního využití big data v audiovizuální tvorbě v budoucnosti	107
6.1	Vyhodnocení výzkumných otázek.....	107

6.2	Vyhodnocení hypotetických předpokladů.....	110
	Závěr.....	113
	Seznam použitých zdrojů	115

Seznam příloh

Graf 1: Věková struktura jednotlivců v ČR, kteří sledují pořady na internetu, podle typu sledovaných stránek (%), zdroj: Český statistický úřad, 2018.....	7
Graf 2: Výzkum sledovanosti OMG bus online reprezentativní populace ČR 15–64 let, sběr dat 1.–6. 2. 2019, zdroj: Vojtěchovská, 2019	9
Graf 3 Podíl kulturního průmyslu na HDP ČR, zdroj: Deloitte, 2020	11
Graf 4 Podíl na celkové přidané hodnotě kulturního průmyslu v ČR, zdroj: Deloitte, 2020	12
Graf 5 Obraty audiovizuální produkce v mil. Kč, zdroj: Deloitte, 2020.....	14
Graf 6 Vykázané výnosy Netflixu za roky 2006–2019, zdroj: Netflix, 2020	15
Graf 7: Proces, který probíhá, když uživatel přijde na web, kde se měří hity pomocí Google Analytics a načte se stránka, Zdroj: Digishuffle.com, 2018.....	21
Graf 8: Použití konkrétních nástrojů v různých fázích procesu práce s daty uživatelů Netflixu, zdroj: Perez, 2017.....	23
Graf 9: Proces zpracování big data v podniku, zdroj: Siddiq et al., 2016	25
Obrázek 1: Ukázka záběru ze softwaru pro rozpoznávání AV pomocí AI, zdroj: Vallosa.com.....	34
Obrázek 2: Ukázka využití umělé inteligence v jednom z pořadů na Mall.TV, zdroj: Mall.TV, 2020	37
Graf 10: Časová osa použití různých metod doporučení na základě množství dostupných uživatelských dat, zdroj: Kapustina, 2015.....	40
Tabulka 1: Příklad hodnocení pořadů uživateli k ilustraci systému kolaborativního filtrování, zdroj: Pelánek, 2011	42
Obrázek 3: Různé varianty cover grafiky pro pořad Stranger Things podle země distribuce, Zdroj: Uphoff, 2018	46
Graf 11: Životní cyklus pořadu v návaznosti na dostupnost získávaných dat a význam rozhodování o procesu, zdroj: Dye, 2019	52
Obrázek 4: Mapa zájmové síly diváckých vazeb mezi pořady v jednotlivých kategoriích nebo žánrech z června 2019, zdroj: Šiler, 2020.....	68
Obrázek 5: Rozložení pořadů z nabídky Mall.TV do clusterů v červnu 2019, zdroj: Šiler, 2020	69
Obrázek 6: Uživatelské zobrazení kategorií na webu www.mall.tv, zdroj: Mall.TV, 2020	71
Obrázek 7: Systém doporučení u konkrétního pořadu v pravé sekci Následující videa na webu www.mall.tv, Zdroj: Mall.TV, 2020	72
Obrázek 8 Ukázka uživatelských zásahů do děje a použití interaktivních prvků v online dokumentární sérii Do Not Track, zdroj: Upian, 2020	81

Obrázek 9 Ukázka interaktivních prvků vyhodnocovaných v reálném čase na základě uživatelských dat v sérii Do Not Track, zdroj: Upian, 2020	81
Obrázek 10 Schéma segmentů zákazníků na základě jejich vazby na VOD platformu HBO, zdroj, Boulton-Wallace, 2019	84
Obrázek 11 Testování uživatelské odezvy HBO filmu Nesmrtelný život Henrietty Lacksové na časové ose pořadu, zdroj: Boulton-Wallace, 2019	85
Obrázek 12 Spektrum diváckých návyků při sledování HBO Go, zdroj: Boulton-Wallace, 2019.....	86
Obrázek 13 Graf počtu předplatitelů (v milionech) a množství hodin obsahu v knihovně vybraných VOD platforem, zdroj: Boulton-Wallace, 2019	87
Tabulka 2 Poptávka po obsahu v jednotlivých žánrech na HBO VOD, zdroj: Boulton-Wallace, 2019.....	88
Obrázek 14 Efektivita uvádění nových a archivních pořadů na HBO VOD z hlediska diváckého přínosu, zdroj: Boulton-Wallace, 2019	89
Obrázek 7: Dopad GDPR na Netflix, zdroj: Sobers, 2020	99

Seznam použitého označování a zkratek

AI	z aj.	<i>Artificial Intelligence</i> – umělá inteligence
AV		audiovizuální
Big data	z aj.	veledata
CEO	z aj.	<i>Chief Executive Officer</i> – výkonný ředitel obchodní společnosti
ČT		Česká televize
CNC		společnost Czech News Center
Download		stažení online obsahu
GDPR	z aj.	<i>General Data Protection Regulation</i> – Obecné nařízení o ochraně osobních údajů
HBO		společnost Home Box Office
HDP		Hrubý domácí produkt
NM		Nová média
On demand	z aj.	na vyžádání, zpravidla obsah přístupný na základě jeho vyhledání uživatelem na internetu
Stream		vysílání online obsahu
TV		televizní stanice / televize
VOD	z aj.	<i>Video on Demand</i> – audiovizuální obsah na vyžádání
- AVOD	z aj.	Advertising-based VOD – audiovizuální obsah na vyžádání, který je k zpravidla volně přístupný, ale jeho nedílnou součástí je přítomná reklama
- SVOD	z aj.	Subscribtional VOD – předplacený audiovizuální obsah na vyžádání
- TVOD	z aj.	Transactional VOD – audiovizuální obsah na vyžádání, který je dostupný po uhrazení mikrotranskace

Úvod

Tato práce se zabývá problematikou využití dat v audiovizuální tvorbě. Big data jsou jedním z klíčových pojmů dnešního digitálního světa, a téměř každý se v nějaké formě s tímto pojmem vědomě či nevědomě setkal. Existuje předpoklad, že se využití big data do budoucna dále rozšíří do mnohých, zatím nedotčených oblastí lidského konání. Tento předpoklad je založen na možnostech efektivního využití big data v různých oborech lidské činnosti, včetně kulturní oblasti, a to konkrétně v rámci audiovizuální tvorby na internetu. Tato forma tvorby prochází neustálým a dynamickým vývojem, a její postup kupředu je výrazně ovlivněn právě technologickým rozvojem.

Tvůrci a distributoři audiovizuálního obsahu na vyžádání, kteří dokáží využít big data ke zkvalitnění výstupů svojí práce, získávají významnou konkurenční výhodu. Zároveň však z využití big data těží i divák (spotřebitel), který získává kvalitnější službu a lepší zážitek z konzumace audiovizuálního díla. Z tohoto důvodu předkládaná práce řeší velmi aktuální a potřebnou problematiku.

Cílem práce je zhodnotit možnosti využití big data a dat všeobecně v audiovizuální tvorbě, a to zejména v oblasti VOD. Struktura práce je rozdělena do šesti kapitol, které řešenou problematiku rozebírají z různých úhlů pohledu a za použití konkrétních příkladů. Snahou je nahlížet na zvolený fenomén očima filmového producenta, manažera, investora nebo marketéra, který přichází do styku s VOD platformou ať už na straně tvůrce, tak na straně distributora.

V práci nechybí informace z primárních zdrojů a sekundárních zdrojů, vlastní komentáře a poznatky, ale také studie uplatnění big data v audiovizuální tvorbě v praxi. Důraz je kladen na využití relevantních aktuálních zdrojů, neboť se jedná o obor, který se rozvíjí velmi dynamicky a má relativně

krátkou historii. Zároveň jsou využité i statistické informace které se vztahují k fenoménu dat a audiovize.

1 Audiovizuální tvorba v současnosti

1.1 Proměna globálního trhu audiovizuální tvorby

Halada a Osvaldová (2017, s. 38) ve *Slovníku žurnalistiky* vysvětlují, že se pojem audiovizuální skládá z latinských slov audio (slyšet) a video (vidět). Pojem audiovizuální se pak vztahuje k „označení specifické vlastnosti všech sdělení, u kterých se prostřednictvím technických vyjadřovacích prostředků transformuje kinetická a opticko-akustická hodnota zaznamenávané skutečnosti do ikonofonických obrazů, které člověk vnímá sluchem a zrakem v jejich audiovizuální, syntetické a časoprostorové podobě (Halada, Osvaldová, 2017, s. 38).“

Český statistický úřad (2019, s. 41) (dále jen ČSÚ) popisuje, že: „rozhlasové a televizní vysílání v druhé dekádě nového milénia již není limitované vlastnictvím klasického přijímače (televize či rádia) a stále významnější roli hrají jiná technologická zařízení (počítač, tablet, chytrý telefon), přes která mohou diváci prostřednictvím internetu sledovat či poslouchat živé vysílání. Díky existenci hudebních a audiovizuálních katalogů není divák odkázán pouze na poslech či sledování obsahu v reálném čase, ale může si vybrat z televizní či hudební nabídky obsah, který chce ve zvolenou dobu přehrát.“

Halada a Osvaldová (2017, s. 38) dále tvrdí, že audiovizuální sdělení umožňuje velmi přirozené a komplexní zprostředkování informací v mezilidské komunikaci.

Digitální televizní obsah není prezentován klasickým lineárním způsobem, ale umožňuje zákazníkům (resp. divákům), ať si vytváří vlastní plány toho, co chtějí sledovat. Zákazník získává větší prostor pro volbu pořadů. Zprostředkovatel televizního vysílání zase získává informace o chování zákazníků. V konečném důsledku pak digitální obsah generuje výhody pro obě strany (Marr, 2016, s. 144).

1.2 Audiovizuální tvorba a VOD

Subías, Laverón a Molina (2018) chápou jako významnou změnu v rámci konzumace audiovizuálního obsahu rozvoj internetu, který vedl k tomu, že televizní společnosti začínají poskytovat audiovizuální obsah prostřednictvím internetu. Tímto došlo ke změně paradigmatu celého odvětví, ve kterém se novou výzvou stává právě vybudování silnějšího vztahu s publikem. Audiovizuální spotřeba se tímto přibližuje spotřebě jiných druhů zboží v tržním prostředí (Subías, Laverón, Molina, 2018).

Halada a Osvaldová (2017, s. 260) ve *Slovníku žurnalistiky* vymezují pojem VOD (video on demand) jako video na vyžádání, což je název služby, při které si divák vybírá audiovizuální pořad dle vlastní volby. Služba je poskytována na internetu a nabízí ji specializované internetové placené i bezplatné služby nebo televizní stanice, které tímto umožňují dodatečné zhlédnutí odvysílaných pořadů.

VOD platforma umožňuje více zákazníkům na stejném místě sledovat různé programy na různých zařízeních, a to ve stejném čase. Jedná se o velmi pohodlné řešení, které umožňuje konzumaci preferovaného audiovizuálního obsahu dle konkrétní preference zákazníka (Mordor Intelligence, 2020).

Subías, Laverón, Molina (2018) ve svojí výzkumné studii potvrzují, že se v posledních letech dostupnost audiovizuálního obsahu pro koncové uživatele exponenciálně znásobila. V takovém prostředí je pro poskytovatele a tvůrce audiovizuálního obsahu nutné potenciálního diváka zaujmout a přesvědčit ke konzumaci obsahu. Vysoká rivalita konkurence vede k tomu, že poskytovatelé musí bojovat prakticky o každého diváka, k čemuž se stále častěji využívají i big data. Big data a jejich využití slouží ke zlepšení uživatelské zkušenosti, což má vést k posílení loajality těchto uživatelů (Subías, Laverón, Molina, 2018).

Ze statistických informací o televizním trhu ve Spojených státech vyplývá, že v roce 2018 počet uživatelů VOD platform přibližně rostl, a to na úkor počtu uživatelů kabelového televizního vysílání. Paradigma amerického audiovizuálního trhu se tímto výrazně proměňuje. VOD platformám se daří nejen získávat zákazníky klasických televizí, či kabelového vysílání, ale také oslovit zákazníky, kteří takové formy sledování audiovizuálního obsahu vůbec nevyužívají. V roce 2016 tak například využívalo VOD platformu 32 milionů zákazníků s nulovou zákaznickou historií s kabelovou či televizní službou. Příčinou tohoto jevu je široká dostupnost VOD platform a nižší ceny v komparaci s konkurencí (Park, Kwon, 2019).

Velká Británie je dalším příkladem země, kde provozovatelé platform VOD mění strukturu celého trhu televizního vysílání. Počet uživatelů těchto služeb zde roste, a to právě na úkor počtu zákazníků klasických televizních stanic – klesá mimo jiné i počet lidí, kteří vlastní televizor. To vše je v souladu s předpokladem, že platformy VOD mají velké výhody oproti klasickým provozovatelům televizních stanic (Daniel, 2015).

Macek (2017) ve svojí výzkumné studii uvádí, že po celá desetiletí je sledovanost populárního audiovizuálního obsahu jednou z hlavních aktivit českých domácností, každodenní rutinou a rituálem, který je hluboce zakořeněn v české kulturní identitě. V České republice patří sledování filmů a televizních seriálů k nejrozšířenějším domácím aktivitám, kdy více než 95 % občanů sleduje filmy a více než 76 % občanů sleduje televizní seriály. V současnosti však dochází k výrazné proměně konzumovaného audiovizuálního obsahu, protože dochází k rozvoji digitálních a síťových technologií, ke změnám obsahu, k jeho digitalizaci, takže už není vázán na konkrétní médium či distribuční prostředek (Macek, 2017).

ČSÚ (2019, s. 13–14) uvádí, že stále vyšší počet občanů České republiky využívá internet pro realizaci kulturních aktivit či k nákupu kulturního zboží,

za což se považuje také sledování videí na internetu nebo sledování televize na internetu. V roce 2018 například sledovalo videa na internetu 50 % populace ve věku nad 16 let, internetovou televizi potom 24 % osob. Jak dále uvádí ČSÚ (2019, s. 13–32), využití internetu k přehrávání filmů, videa a hudby se liší i dle věkových kategorií. Nejvíce tuto možnost využívá skupina uživatelů ve věku 16 až 24 let (tj. přibližně 87 % populace v této věkové kategorii). S rostoucím věkem pak četnost klesá (Český statistický úřad, 2019, s. 13–32).

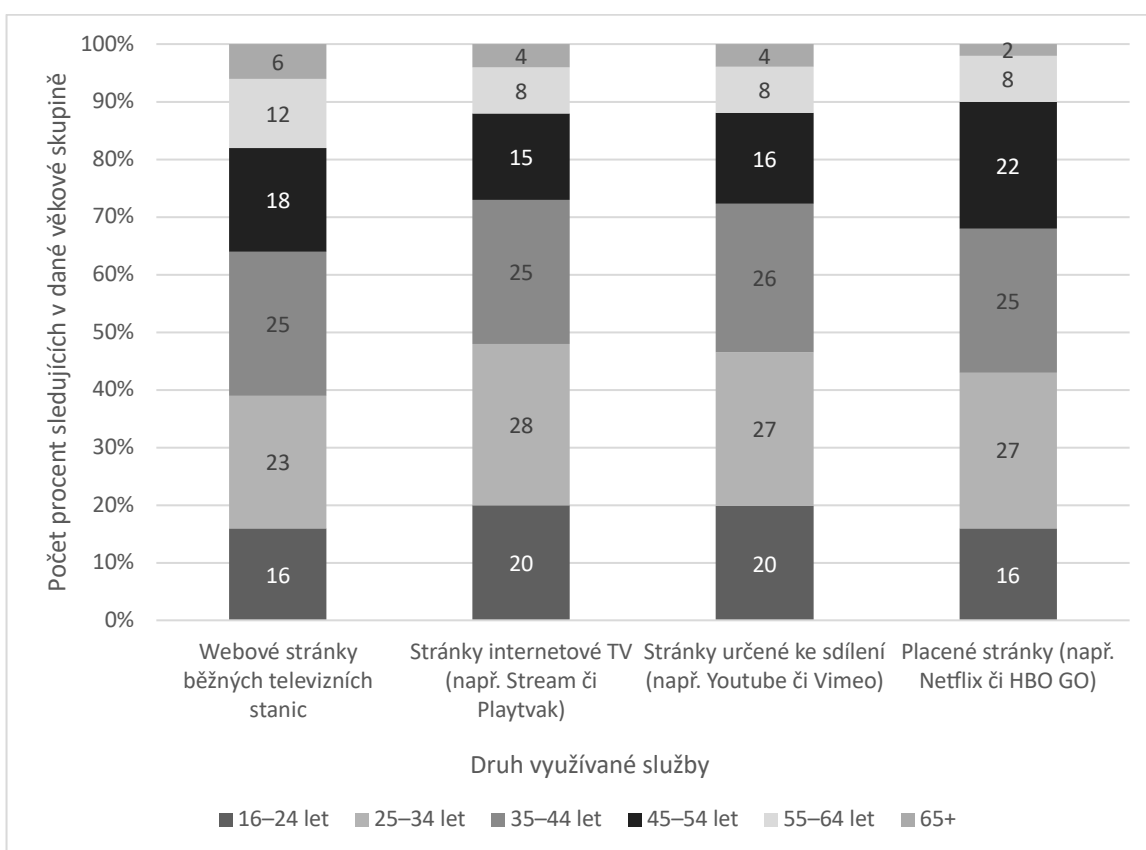
ČSÚ (2019, s. 28–29) rovněž zmiňuje, že i v zahraničí se objevuje trend stále častějšího využívání zpoplatněného stahování filmů a online sledování filmů z předplacených katalogů. Například ve Spojených státech došlo v letech 2017 až 2018 k růstu tržeb z digitálních služeb (stream a download) o 24 % (Český statistický úřad, 2019, s. 28–29).

„Sledování placených filmů, seriálů a jiných pořadů není v České republice příliš rozšířené. Této službě v roce 2018 využívají 4 % obyvatel, mezi nimi dominují příslušníci mladších věkových skupin, vysokoškolsky vzdělaní a studenti. Ve srovnání s ostatními zeměmi Evropské unie představuje 4% podíl jednotlivců sledujících placené pořady online vůbec nejnižší hodnotu. Průměr EU činí 17 %. Nejaktivnější jsou v této činnosti Dánové, mezi kterými za tuto službu zaplatilo 48 % osob (údaje za rok 2016) (Český statistický úřad, 2018, s. 72)“

ČSÚ (2019, s. 33–34) upozorňuje, že o aktivitách a tržbách společností poskytující digitální stažení filmů či možnost zhlédnutí filmů online nejsou v České republice dostupná data. V současnosti provozuje filmové online katalogy bezplatně několik televizních stanic (např. Prima nebo Česká televize). Za poplatek tuto službu nabízí televizní stanice Nova nebo další soukromí poskytovatelé. V tuzemsku pak nejrozsáhlejší databázi

zahraničních filmů a seriálů poskytuje od roku 2016 společnost Netflix (Český statistický úřad, 2019, s. 33–34).

ČSÚ (2019, s. 52–53) uvádí informaci, že v roce 2018 přibližně 24 % celkové populace starší 16 let sledovalo televizi prostřednictvím internetu. Videá na stránkách určených ke sdílení pak sledovalo 44 % populace a videá z neplacených katalogů pořadů sledovalo 26 % populace. Nejméně byla využívána možnost sledování videa z placených katalogů pořadů (tj. například Netflix), kterou využívala 4 % populace (Český statistický úřad, 2019, s. 52–53).



Graf 1: Věková struktura jednotlivců v ČR, kteří sledují pořady na internetu, podle typu sledovaných stránek (%), zdroj: Český statistický úřad, 2018

I přes to, že trend sledovanosti a ochoty platit ze prémiový obsah není v České republice zatím převládající, VoD služby typu Netflix či Amazon začínají podle názoru generálního ředitele Českých Radiokomunikací Martina Gebauera vytlačovat postupně placené televizní kanály satelitní či kabelové (Vojtěchovská, Gebauer: VOD služby začínají nahrazovat placené TV kanály, 2018).

Z globálního hlediska je role VoD platforem úplně jiná. Například na americkém trhu za Netflix platí podle Deloitte až 55 % všech domácností. Až 90 % předplatitelů HBO Go a 80 % předplatitelů Hulu na americkém trhu jsou zároveň předplatiteli služby Netflix. Celkový nárůst předplatitelů mezi lety 2009 a 2017 narostl o 450 % (UserTesting, 2018).

Netflix uvádí, že má v současné době celosvětově 167 milionů předplatitelů a toto číslo stále roste (Netflix, Inc., 2020). Reálných uživatelů je ještě daleko více, protože mnoho z nich má sdílený účet. Podle studie MoffetNathanson využívá 5 % uživatelů službu ve zkušební době zdarma, 27 % procent uživatelů používá předplatné jiného člena domácnosti, 14 % používá sdílený účet s někým mimo domácnost (přátelé a známí) a pouze zbylých 55 % jsou reální předplatitelé (Kafka & Molla, 2019).

V roce 2018 platilo, že Netflix ovládá z hlediska množství konzumace obsahu přibližně 40 % amerického VOD trhu, dále YouTube 18 %, Hulu 14 % a Amazon Prime 7 % (Park, Kwon, 2019)

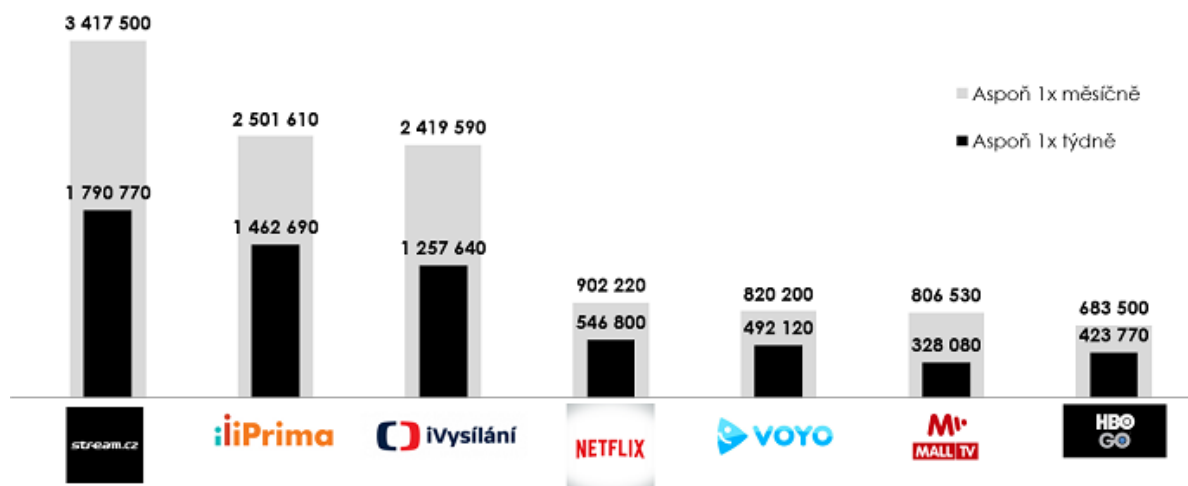
Podle výzkumu OMG Research¹ společnosti Omnicon Media Group patří mezi největší VoD služby na českém trhu z hlediska počtu diváků iVysílání, iPrima.cz a Stream.cz. *„Alespoň jednou týdně sleduje Stream.cz 1,79 mil. lidí, jednou za měsíc si pak Stream.cz pustí 3,4 mil. lidí. Následují*

¹ Jedná se o pravidelný online výzkum a data jsou zjišťována na reprezentativním panelu internetové populace ČR na vzorku 500 lidí ve věku 15–64 let.

platformy iPrima.cz s týdenní sledovaností 1,46 milionu a iVysílání od ČT s týdenní sledovaností 1,26 mil. lidí. Z VOD platforem měřených v průzkumu je globální služba Netflix čtvrtou nejsledovanější (Vojtěchovská, 2019)."

Sledování vybraných online VoD platforem

Počty diváků, projekce na populaci



Graf 2: Výzkum sledovanosti OMG bus online reprezentativní populace ČR 15–64 let, sběr dat 1.–6. 2. 2019, zdroj: Vojtěchovská, 2019

Vojtěchovská dále uvádí, že ze studie sledovanosti OMG vyplývá, že SVOD HBO Go a Netflix sledují zejména mladší diváci do 34 let. V případě Netflixu jsou to spíše muži z větších měst a diváci, kteří se běžně nedívají na klasické televizní vysílání. Oproti tomu Voyo, které má podobné výsledky sledovanosti jako dvě předchozí zmíněné služby, naopak sledují zejména diváci ve věku 55 – 64 let ze středních měst, kteří jsou pravidelnými diváky klasického TV vysílání (Vojtěchovská, 2019).

Rozložení podílu audiovizuálního obsahu na trhu ovlivňují také práva k jednotlivým převzatým filmům nebo seriálům a jsou značně ovlivňována především nadnárodními společnostmi jak uvádí autoři studie *Mapa audiovizuálního pole v České republice z hlediska digitalizace a strategie pro jednotný digitální trh*: „Některá studia nabízejí svou novější produkci v rámci

online distribuce v ČR výhradně prostřednictvím nadnárodních portálů (Sony Pictures na GooglePlay a iTunes; Paramount navíc i na HBO Go). Nabídka lokálních portálů je ovlivňována také změnami ve smluvních vztazích mezi nadnárodními portály a studii. Například online nabídku studia Disney v ČR do minulého roku zprostředkoval jen globální portál Netflix, popř. ještě iTunes, GooglePlay nebo HBO Go. Nicméně poté co Disney v srpnu 2017 odstoupil ze spolupráce s portálem Netflix, nabízí novější Disneyho filmy v ČR již jen O2 Videotéka (Szczepanik & Zahrádka (eds.), 2018)."

Mezi další VoD služby na českém trhu kromě výše zmíněných působí také Aerovod, Dafilms, Alza Media a mnohé další. Trh se neustále proměňuje a služby zanikají (Obbod, Kamil Ouška a Banaxi GO, Banaxi Limited)², transformují se (Stream.cz, Seznam a Playtvak, iDNES)³ nebo vznikají nové (Edisonline, Film Europe)⁴.

Obecně je můžeme rozdělit do tří hlavních skupin (Szczepanik & Zahrádka (eds.), 2018, str. 25):

- Okrajové VoD služby, které se zaměřují na tematický nebo artový obsah;
- mainstreamové VoD služby, u kterých převažuje obsah převzaté zahraniční tvorby nebo vlastních pořadů, případně kombinace obojího;
- veřejnoprávní VoD služby, kde je v současné době jediný významný hráč Česká televize se svým iVysíláním⁵.

² <https://www.mediar.cz/obbod-konci-rozlouci-se-filmem-vysehrad/>

³ <https://www.lupa.cz/aktuality/samostatny-web-stream-cz-skoncil-seznam-slucuje-videa-na-jedno-misto/>

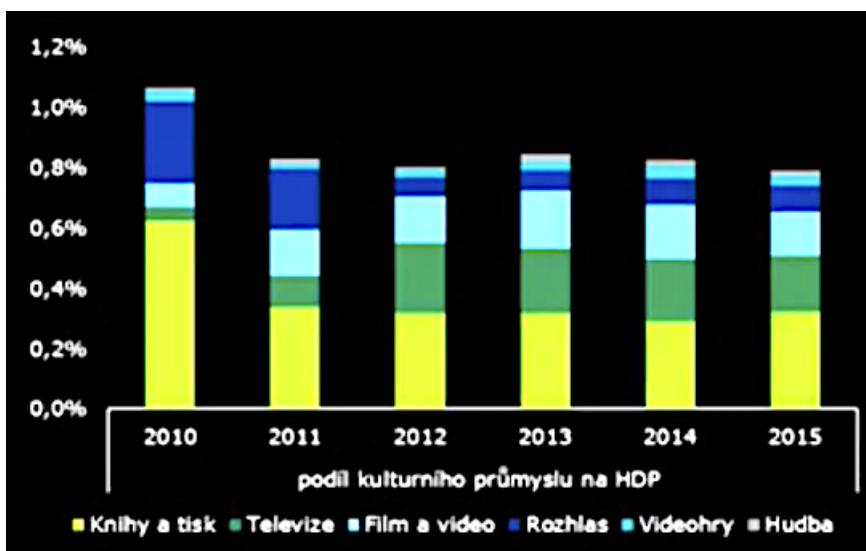
⁴ <https://www.totalfilm.cz/2020/04/edisonline-za-cenu-vstupenky-kina-nabizi-filmy-i-televizni-kanaly/>

⁵ Specifika právě poslední zmíněné skupiny jsou dále popsána v kapitole: 3.4

1.2.1 Ekonomické faktory audiovizuální tvorby a VOD

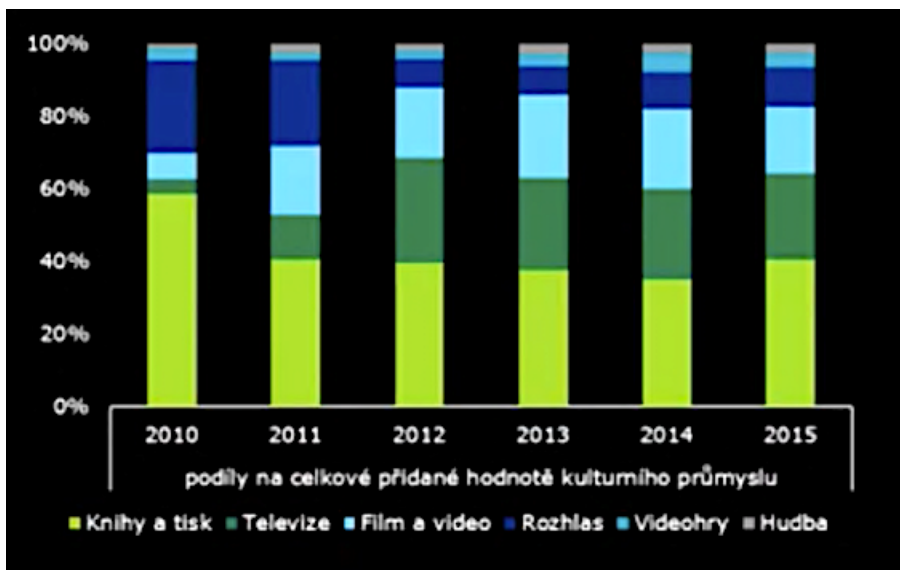
Deloitte (2020) v závěru svojí výzkumné studie českého a globálního audiovizuálního trhu potvrzuje, že dochází ke stále znatelnější provázanosti světů audiovizuální zábavy, ať už se jedná o interaktivní videohry, film, nebo hudbu. Filmová studia se snaží využít co nejvíce dostupných kanálů k vytvoření co nejvyššího zisku. Na úspěšné filmy navazuje vývoj počítačových her, vydávají se filmové soundtracky, prodávají se různé upomínkové předměty atd. Klíčovou složkou těchto děl je potom jejich vizuální stránka (Deloitte, 2020).

Následující obrázek popisuje podíl jednotlivých typů kulturních průmyslů na hrubém domácím produktu České republiky v letech 2010 až 2015. Informace vychází z výzkumné studie Deloitte (2020). Z vývoje vyplývá, že došlo ke zvýšení podílu televize na HDP, zatímco u ostatních složek lze sledovat spíše stagnaci vývoje (kromě knih a tisku). Dlouhodobě se tak význam jednotlivých kulturních průmyslů nemění.



Graf 3 Podíl kulturního průmyslu na HDP ČR, zdroj: Deloitte, 2020

Na dalším grafu se nachází podíly jednotlivých kulturních průmyslů na celkové přidané hodnotě kulturního průmyslu. Je čerpáno z výzkumné studie Deloitte (2020). Od roku 2012 do roku 2015 je situace téměř identická a nedochází zde k výrazným změnám.



Graf 4 Podíl na celkové přidané hodnotě kulturního průmyslu v ČR, zdroj: Deloitte, 2020

Z výzkumné studie organizace Deloitte (2020) vyplývá, že v roce 2018 dosáhly tržby filmového průmyslu v České republice 54 miliard korun a činnost tvůrců potom přispěla k produkci za dalších 50 miliard korun v dalších odvětvích. Celkový vliv filmového průmyslu na českou ekonomiku byl na úrovni 104 miliard korun. Přidaná hodnota filmového průmyslu a souvisejících odvětví dosáhla 1,1 % českého hrubého domácího produktu. Na daních bylo filmovým průmyslem odvedeno přes 19 miliard korun. Filmový průmysl zaměstnával 33 tisíc lidí, a to přímo 12 tisíc (Deloitte, 2020).

Z výzkumné studie organizace Deloitte (2020) plyne, že v globálním prostředí klesá podíl 3D filmů na celkových tržbách. Tento trend je tedy na ústupu. Například ve Spojených státech a v Kanadě klesl podíl 3D filmů na tržbách z 21 % v roce 2010 na 12 % v roce 2017. Naproti tomu k dalším trendům

v audiovizuální tvorbě patří využívání technologických inovací, nových médií a tvorba pro OTT⁶, mezi které spadají i VoD služby (Deloitte, 2020).

Na rozdíl od distribuce AV obsahu terestriálně, kde lze mezi jednotlivými lineárními kanály přepínat, některé služby OTT, jako je iTunes, vyžadují, aby se video nejprve stáhlo a poté se přehrало, zatímco jiné přehrávače OTT, jako jsou například Netflix, HBO Go, Hulu, Disney+ a Amazon Prime Video, nabízejí přehrávání filmů on-line, které jsou tím pádem k přehrání okamžitě za pomoci streamování (Cansado, 2017).

*Ve výzkumu *Mapa audiovizuálního pole v České republice z hlediska digitalizace a strategie pro jednotný digitální trh* je zmíněn důležitý jev, který má pozitivní ekonomické dopady na lokální růst odvětví VOD služeb: „Široká základna předplatitelů audiovizuálních služeb na vyžádání umožňuje nadnárodním poskytovatelům těchto služeb také v rostoucí míře investovat do výroby původních lokálních obsahů, které plní roli klíčového nástroje konkurenčního boje na jednotlivých trzích (Szczepanik & Zahradka (eds.), 2018).“*

Další graf poskytuje informace o obratu audiovizuální produkce v České republice v letech 2002 až 2016 a vychází z výzkumné studie společnosti Deloitte (2020). Z vývoje vyplývá, že se průběžně zvyšuje podíl české filmové produkce, ale také se v České republice vytváří zahraniční audiovizuální produkce, která je významným faktorem.

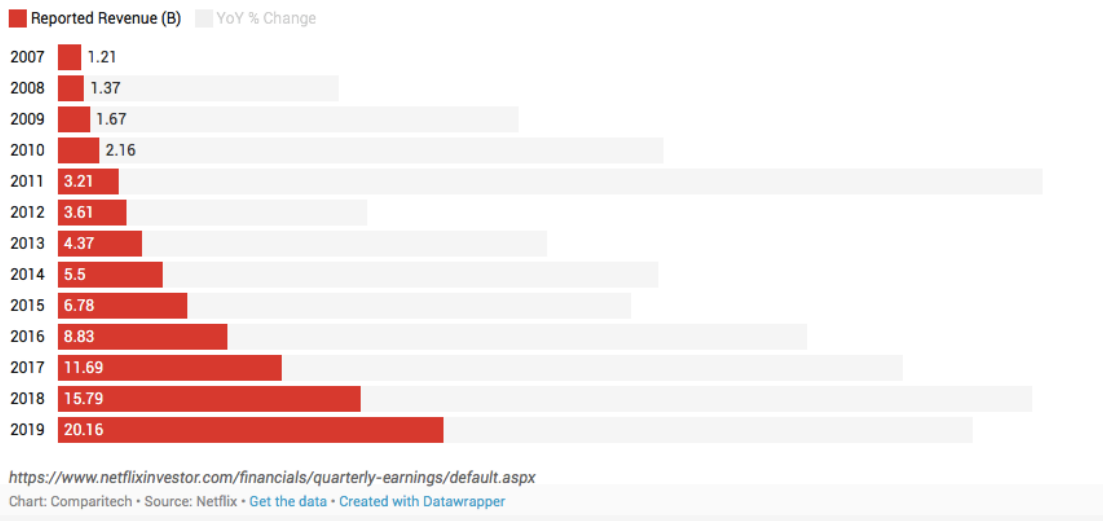
⁶ Jako over-the-top (OTT) je označována distribuce audio nebo video obsahu a dalších médií přes internet. Zdroj: en.wikipedia.org/wiki/Over-the-top_media_service



Graf 5 Obraty audiovizuální produkce v mil. Kč, zdroj: Deloitte, 2020

Jak bylo zmíněno již výše, objem vyráběného obsahu i počet předplatitelů se například u Netflixu zvyšuje rok od roku. Vykázané výnosy za rok 2019 dosáhly na základě finanční zprávy Netflixu z prvního čtvrtletí tohoto roku 2020 přes 20 miliard amerických dolarů což je oproti roku 2007 téměř dvacetinásobný nárůst. Přesto je nutné podotknout, že společnost je kvůli všem investicím do rozvoje a tvorby vlastního obsahu stále ve ztrátě (Netflix, Inc., 2020).

Netflix Reported Revenue 2006 - 2019



Graf 6 Vykázané výnosy Netflixu za roky 2006–2019, zdroj: Netflix, 2020

Výzkumná studie organizace Mordor Intelligence hovoří o tom, že byl celkový objem globálního trhu VOD služeb v roce 2019 ve výši 56,55 miliard amerických dolarů a do roku 2025 by měl dosáhnout velikosti 120,91 miliard amerických dolarů. Na trhu panuje vysoká rivalita konkurence, která nutí poskytovatele VOD k neustálému rozšiřování nabídky a k inovacím. Zejména vzestup poskytovatelů VOD jako Apple TV, Amazon, Netflix či Hulu je výzvou pro tradiční poskytovatele televizního vysílání, protože právě těmto může v budoucnosti přebrat výraznou část publika (Mordor Intelligence, 2020).

„Všechny streamovací služby potřebují obrovský objem obsahu, aby přitáhly předplatitele....Ke konkurenčnímu boji nadnárodních platforem (Netflix, Amazon, Hulu, Apple TV či třeba chystané Disney+) se přidává zvýšená produkce televizních vysílatelů (BBC či Canal+). Obojí způsobilo nevídaný boom seriálové tvorby a dramatické globální navýšení investic do originální seriálové produkce (Aust, 2019).“

2 Data jako nástroj pro práci s VOD

Jakub Šiler, který působil jako obsahový analytik pro společnost Seznam, především v redakci Stream.cz, následně jako analytik digitálního obsahu pro Mall.TV a v současné době v CNC, uvádí, že společnost Seznam pracuje denně až s miliardami datových bodů (Šiler, 2020). V takovém případě můžeme hovořit o termínu big data, které znamenají specifický přístup v práci a přináší komplexnější údaje o uživatelských službách a jejich chování.

2.1 Big data

Burian (2014, s. 90) popisuje, že big data využívají velké sady dat, tj. zvláště takové, které nemají dostatečnou strukturu, aby je bylo možné uložit v tradičním datovém skladu, databázi. Prostřednictvím robotů se získávají informace procházející webovým prostředím, kanály sociálních sítí a protokolů serverů či údaje z dodavatelských řetězců průmyslového prostředí, nebo prostřednictvím dohledu snímačů apod. V těchto případech jsou big data výstupem složitějších podnikových a korporátních dat, než tomu bylo kdykoliv předtím (Burian, 2014, s. 90).

Holubová a kol. (2015, s. 20) vysvětlují, že big data jsou revolucí v informačních technologiích, a tedy souvisí s rozvojem nových technologií, služeb a jejich kombinací. Příkladem jsou například senzorové sítě nebo vědecké přístroje, které zkoumají přírodní jevy, sociální sítě nebo mobilní technologie a související aplikace. Takovéto typy technologií a aplikací vygenerují s pomocí uživatelů každou vteřinu obrovské množství dat, která je nutné efektivně uložit a účelně zpracovat (Holubová a kol., 2015, s. 20).

Big data mají svoje specifické vlastnosti, kterými se liší od „tradičních“ podnikových dat. Důvodem je skutečnost, že datové sklady a nástroje pro správu dat obvykle nejsou připravené na zpracování a analýzu velkých objemů dat ve velmi krátkém čase či efektivním způsobem z hlediska nákladů (Burian, 2014, s. 90).

Specifikem big data je také jejich různorodost z hlediska formátů. Nejedná se pouze o strukturovaná (utříděná a jednoduše čitelná) data, ale také o různé kombinace dat, která jsou v odlišných formátech, mají jiné charakteristiky apod. Příkladem jsou například databázové informace s multimediálními daty, obsahy textových dokumentů či vzorců chování uživatelů apod. (Berger, 2016).

Floridi (2019, s. 31–32) ve své publikaci popisuje big data jako obrovskou a komplexní sestavu datových souborů, které je složité zpracovat s využitím dostupných uživatelských nástrojů či tradičních aplikací na zpracování dat. Data jsou stále prospěšnou věcí a zdrojem, který je vhodné využít. Zúžitkování big data je však dáno jejich specifiky (Floridi, 2019, s. 31–32).

Gemignani, Gemignani a Galentino (2015, s. 25) ve své publikaci shrnují, že *„třicet let po uvedení osobního počítače jsme vstoupili do éry veledat (big data) a rozhodování na základě dat, kterou přirovnáváme k ropnému boomu. Nadšení je rozhodně velké a doufáme, že data povedou k lepšímu porozumění a chytřejšímu podnikání.“*

V dnešním tržním prostředí musí organizace umět prodávat kvalitní produkty, které odpovídají přáním a potřebám zákazníků. Při stále zvyšující se síle konkurence jde o velmi složitý úkol. Jsou to však právě big data, která usnadňují splnění této ambice, a to zejména v prostředí velkých organizací s dostatečnými zdroji pro efektivní využití datových sad o zákaznících. Údaje shromážděné o užívatelích umožňují generovat množství výhod, které pak organizace využívají v rámci marketingu (Marr, 2016, s. 69–71).

DeSouza a Smith (2014) uvádějí výstup z výzkumu organizace IBM, dle kterého se každý den vytváří až 2,5 bilionu dat. Tato data slouží k různorodému účelu a velmi často rozhodují i o obchodním úspěchu různých organizací v různých odvětvích, protože se následně využívají k analýze a rozhodování.

Analýza big data v reálném čase nebo i s krátkým zpožděním představuje nejrychlejší a nejúčinnější způsob, jak získat užitečné znalosti o tom, co se momentálně děje. Na základě tohoto lze rychle a flexibilně reagovat na vznikající problémy nebo příležitosti v podobě nově vznikajících trendů. Správná reakce je prostředkem pro zlepšení výkonnosti (Bifet, 2013).

V kontextu tvorby audiovizuálního obsahu neznamení big data velikost audiovizuálních souborů samotných. Technické zajištění distribuce souborů audiovizuálních děl samotných

prostřednictvím sítě internetu není proto předmětem této práce.

Big data se od ostatních typů dat liší také tím, že se u nich zkoumá pět základních parametrů (Jain, 2016):

- objem – množství údajů či měřených položek;
- typ – charakterová povaha dat;
- rychlost – v tomto kontextu rychlost, s jakou jsou data generována a zpracovávána;
- variabilita – konzistentnost datové sady, která napomáhá procesům při správě dat;
- kvalita – výpovědní hodnota.

(V angličtině 5V – Volume, Variety, Velocity, Variability, Veracity.)

2.2 Jak VOD portál získává data

S příchodem uživatele na stránku VoD platformy se ihned spouští celá řada měřících procesů, které uživatel ve většině případů vůbec nezaznamená, neboť běží tzv. na pozadí webu.

Spolu s příchodem na stránku odešle webový server datovou informaci prohlížeči v zařízení uživatele, která se v tomto zařízení uloží. Při každé další návštěvě

serveru pak prohlížeč opět odesílá data zpět. Tato datová informace se označuje anglickým výrazem *cookie*⁷.

Sklenák (2003) definuje cookies tak, že se jedná o informaci, která: *„...vzniká při komunikaci mezi webovým serverem a klientem (webovým prohlížečem) a která se uchovává na vyhrazeném místě na straně klienta jako textový soubor. Soubor obsahuje údaje o klientovi (konfigurace, přihlašovací jméno, uživatelská nastavení prostředí aj.), jež mohou být využívány při opakovaném připojování k danému serveru. Jejich účelem je identifikace uživatele při prohlížení dalších stránek na serveru.“*

Autoři Barvíř, Hampl, Melišová (2011) uvádějí: *„Cookies neznamenaají žádné nebezpečí pro počítač jako takový. Přesto cookies mohou být nebezpečné pro ochranu soukromí. Navštívený web si totiž může ukládat do cookies jakékoliv informace, které o návštěvníkovi zjistí a může tak postupně zjišťovat zájmy konkrétního návštěvníka – které stránky navštěvuje, jaké informace vyhledává, jak často daný web navštěvuje apod. (Barvíř, Hampl, & Melišová, 2011)“*

Ke konkrétním VoD platformám Mall.TV uvádí datový analytik Vojtěch Matoušek, který dříve pracoval pro Seznam a nyní působí v CNC, následující: *„Od doby, co uživatel přijde na náš web, tak začínáme měřit jeho aktivitu a pokud si cíleně cookies nemaže, tak o něm jsme schopní zjistit, jestli už na našem webu někdy byl, odkud na něj přišel a podobně. Měřit se dá téměř cokoli, otázkou je, co chce firma přesně zjistit, protože s množstvím*

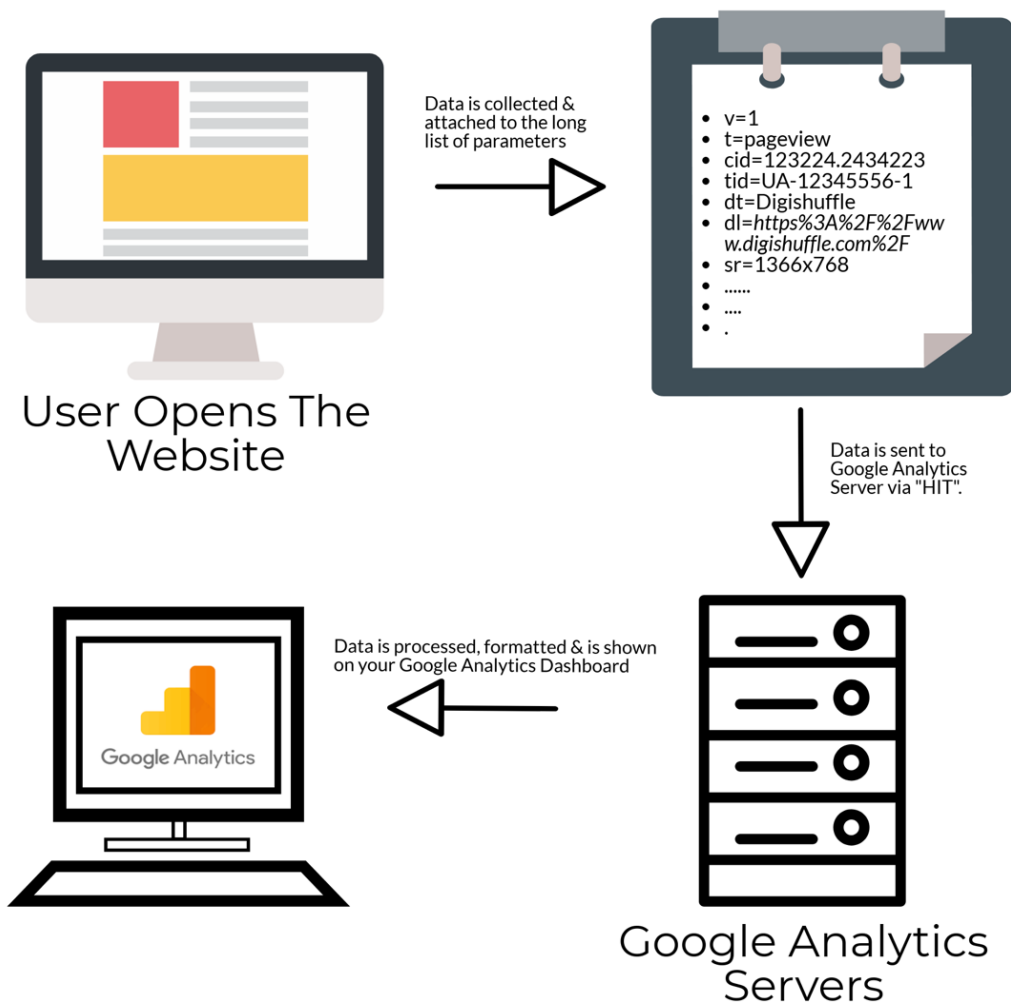
⁷ Název cookie – *sušenka* asociuje zvyklost ze Spojených států nebo Velké Británie nabídnout účastníkům určitého zájmového spolku nebo skupiny jejich oblíbenou sušenku pro vytvoření příjemnější atmosféry. (Zdroj: https://cs.wikipedia.org/wiki/HTTP_cookie)

uživatelů a množstvím měřených parametrů roste obrovským způsobem zátěž na zpracování získávaných dat, a to bývá problém."

Miller (2019) uvádí, že každá jednotlivá informace, která se zaznamenává za účelem měření výkonnosti nebo chování, se označuje jako *hit*. Společnost Google například ve svém nástroji Google Analytics rozlišuje druhy hitů podle kategorií (Miller, 2019):

- Pageviews – zaznamenání každé návštěvy webové stránky;
- Devices – informace o zařízení, které uživatel využívá k zobrazení webu;
- Events – zaznamenává to jakým způsobem uživatel interaguje s webovou stránkou;
- Social Interactions – zaznamenává interakci na stránce skrz sociální sítě, pokud je takový obsah k dispozici;
- E-commerce – zaznamenává informace o procesu nákupu v případě e-shopu;
- Exception Hit – zaznamenává různé druhy neobvyklého chování webu nebo chyby, ke kterým může dojít;
- User Timing – zaznamenává dobu načítání obsahu na zařízení jednotlivého uživatele.

Schéma procesu sběru a uložení dat zobrazuje graf níže.



Graf 7: Proces, který probíhá, když uživatel přijde na web, kde se měří hity pomocí Google Analytics a načte se stránka, Zdroj: Digishuffle.com, 2018

Matoušek dále uvádí, že existují společnosti, které zajímají i takové údaje, jako je pohyb kurzoru myši uživatele po obrazovce, což je příklad obrovské datové zátěže, který se v Seznamu nebo v CNC neměří. „Mezi standardní zájmy, které se u nás běžně hitují, patří *Page Views* (počet zobrazení/návštěv na stránce za stanovené časové období), *Real Users* (počet unikátních uživatelů za stanovené časové období) nebo například *kliknutí na nějaké prvek na stránce* nebo *scrollování* (pohyb na stránce nahoru a dolů). Podle

scrollování jsme schopni odvodit zájem o konkrétní obsah a v případě článků i třeba jejich dočtenost." (Matoušek, 2020)

Metoda neustálého měření by rovněž znamenala obrovskou datovou zátěž, a proto většina firem (včetně Seznamu i CNC) neměří v době, kdy uživatel není aktivní, data, ale z hlediska časového rozlišení pouze časové rozestupy mezi jednotlivými měřenými hity. Z toho následně dovozují, zda je uživatel stále aktivní, nebo zda třeba neodešel od otevřeného počítače (Matoušek, 2020).

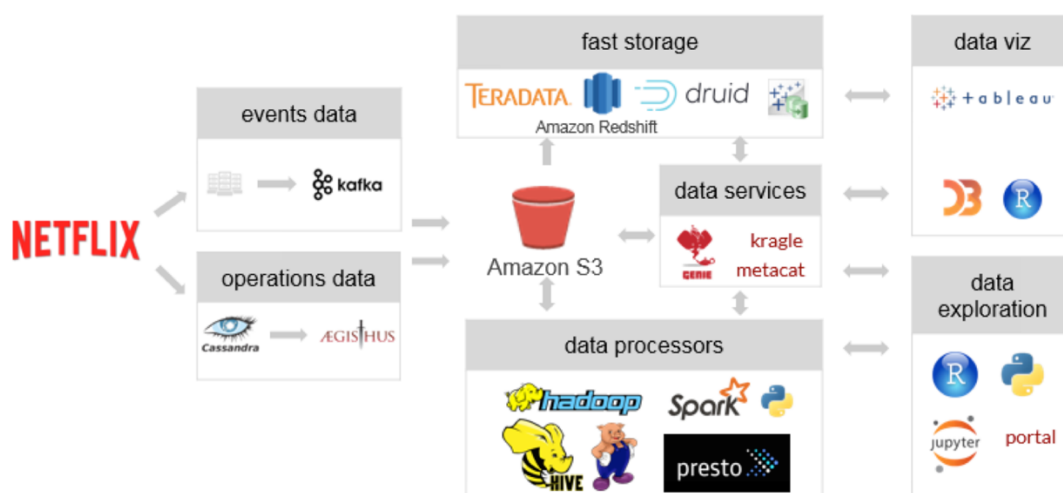
Google Analytics nebo Reportér jsou nástroje, kde se data shromažďují a měří nebo vizualizují. Jejich sběr a ukládání, které standardně probíhá přidáním do datového skladu, však neslouží pouze ke zpětnému vyhodnocování datovými analytiky, ale také k zajištění zobrazení položek z nabídky VoD platformy na základě tzv. systému doporučení (viz kapitola 2.3.4). Právě k doporučení se nejvíce používá strojové učení a algoritmizace (Plummer, 2017).

Společnost Netflix je jedním z leaderů na poli práce s daty u typu služby, jakou je VoD platforma. Pro zpracování big data používá Netflix software business intelligence⁸ typu Hadoop či Teradata a také má vlastní řešení s otevřeným zdrojovým kódem. Software využívá pro shromažďování, ukládání a zpracování velkého množství informací primárně nasbíraných z uživatelských aktivit na svém webu. Netflix nevyužívá tradiční datový sklad založený na datových centrech. Aby mohl ukládat a zpracovávat rychle rostoucí množinu dat, využívá Amazon S3 (službu konkurenčního výrobce),

⁸ Business intelligence software je typ aplikačního softwaru, navrženého k analýze, třídění, reportování a transformaci dat, která slouží ke zlepšení business intelligence. Business intelligence jsou dovednosti, znalosti, technologie, aplikace, kvalita, rizika, bezpečnostní otázky a postupy používané v podnikání pro získání lepšího pochopení chování na trhu a obchodních souvislostech., Zdroj: BI & Data Science slovník – Stránky o Business Intelligence & Data science. *janpansky.cz* [online]. [cit. 2020-05-03]. <http://janpansky.cz/business-intelligence/bi-data-science-slovník>

takže rozděluje více clusterů Hadoop pro různá pracovní zatížení s přístupem ke stejným datům. V ekosystému Hadoop využívá Hive pro ad hoc dotazy a analytiku a nástroj Pig pro práci s algoritmy. Dále Netflix vytvořil svůj vlastní program Genie, který se orientuje na škálování stále rostoucích objemů dat (Sadeh, 2019).

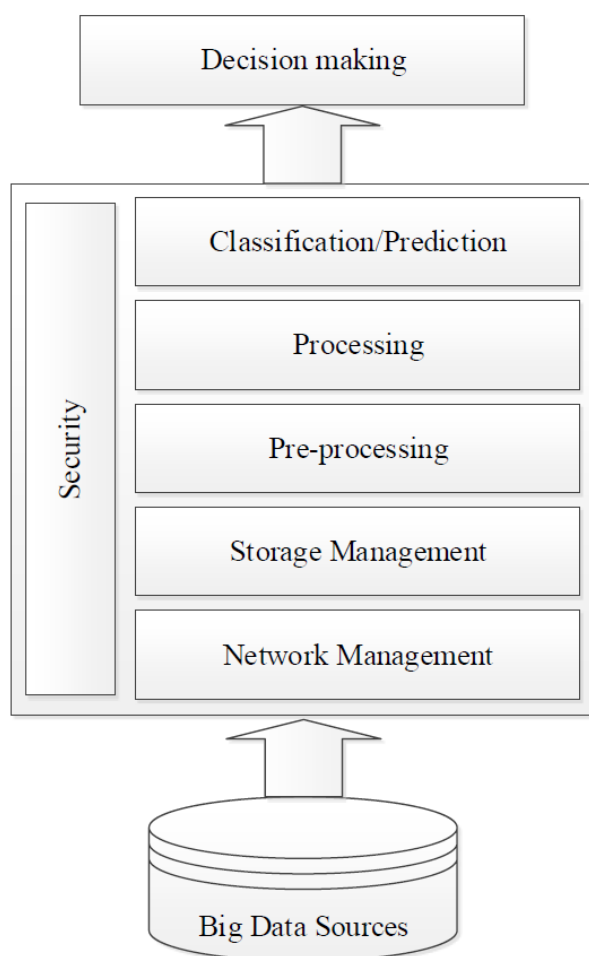
Proces, jakým Netflix s daty pracuje, je daleko komplikovanější oproti předchozím zmíněným příkladům, a to především z toho důvodu, že měří v úhrnu větší množství dat, u více uživatelů a globálně. Používá k tomu v různých fázích procesu desítky hardwarových i softwarových produktů různých dodavatelů i programy, které si na míru vyvíjí. Schéma je vyobrazeno na grafu níže.



Graf 8: Použití konkrétních nástrojů v různých fázích procesu práce s daty uživatelů Netflixu, zdroj: Perez, 2017

Ambicí této práce není věnovat se příliš detailnímu popisu technologických řešení, ale spíše nastínit možnosti, které jsou technologicky dostupné, protože data takto získaná tvoří pomyslnou stavební hmotu při práci s datovou analytikou u VoD platforem. Podrobnější specifiky typologie zkoumaných dat právě u VoD platforem jsou popsány v práci dále (zejména v podkapitole 0), ochrana uživatelských dat pak v kapitolách 4.2 a 0.

Obecně lze říct, že big data a jejich efektivní využití se stává základním předpokladem úspěchu na dnešních globálních trzích, a to v celé řadě různých oborů. Jde o významnou výzvu pro všechny organizace, které pracují s internetem, a větší skupinou uživatelů. Zároveň se však jedná o poměrně nové odvětví, které ještě nemá ustálené procesy a přístupy k využití big data ve svůj prospěch. Následující obrázek popisuje procesní tok správy big data z obecného hlediska. Celý proces lze obecně popsat tak, že začíná získáním dat, tj. zdrojem dat (big data sources), následně dochází ke správě sítě (network management), ke skladování dat (storage management), předzpracování dat (pre-processing), zpracování dat (processing), ke klasifikaci a predikci (classification/prediction), a také jsou po celou dobu data chráněna (security), následně dochází ke konečnému rozhodnutí na základě výstupů z procesu řízení big data (decision making) (Siddiqa et al., 2016).



Graf 9: Proces zpracování big data v podniku, zdroj: Siddiq et al., 2016

2.3 Využití big data v audiovizuální tvorbě

„Prvními oblastmi, kde se využívala Big Data, byly výzkumné instituce a marketing. Jejich průkopníky se staly společnosti jako Google, eBay nebo Amazon s cílem zmapovat, jak co nejlépe provést uživatele svými stránkami tak, aby se zvýšila výtěžnost z klikání (Berger, 2016).“

Internet a vzestup online streamingových platforem vytvořil podmínky pro vznik zcela nového věku audiovizuální tvorby. Platformy typu Amazon Prime nebo Netflix vykazují nadprůměrný růst a jejich vysoká úspěšnost je podpořena právě prostřednictvím správné a efektivní práce s big data. To

znamená, že právě big data jsou zdrojem růstu organizací, které se zabývají tvorbou a sdílením audiovizuálního obsahu ve formě seriálů a filmů (Goodworklabs, 2020).

Big data poskytují užitečné informace o chování publika, preferencích publika a vytvářejí prostředek pro strategický přístup k tvorbě audiovizuálního díla, ať už se jedná o jeho kreativní ztvárnění, marketingovou strategii, či další aspekty (Goodworklabs, 2020).

Pro diváky je více výhodné, když jsou služby a produkty přizpůsobené jejich individuálním preferencím, a proto se dlouhodobě ukazuje, že většina uživatelů je ochotných svoje data poskytnout. Poskytovatelé VOD služeb se s tímto zcela ztotožňují a nabízejí personalizovaný obsah, který dosahuje obecně lepších výsledků (Davies, 2019).

Fernández-Manzano, Neira a Clares-Gavilán (2016) si všímají, že aplikace nových technologií v audiovizuální oblasti vedla k řadě různých systémových změn v obchodních modelech. Zatímco některá odvětví se rychle přizpůsobila trendu big data, tak další nikoliv. Odvětví audiovizuální tvorby je právě jedním z těch, ve kterém mohou tvůrci audiovizuálního obsahu získávat významnou konkurenční výhodu, když se jim podaří správně big data využít (Fernández-Manzano, Neira, Clares-Gavilán, 2016).

U společností jako Amazon či Netflix je práce s big data jádrem strategie. Obě společnosti investují do využití těchto dat stovky milionů dolarů za účelem pochopení přání a potřeb zákazníků. Prostřednictvím těchto investic a této strategie pak vytvářejí originální a prémiový audiovizuální obsah. Interní big data o zákaznících umožňuje získat exkluzivní pohled na chování diváků (zákazníků) a potom dle tohoto vytvářet prémiový obsah. Těmito investicemi a systémy se daří organizacím Amazon a Netflix vytvářet klíčovou konkurenční výhodu. Získávají přehled o tom, co diváci audiovizuálního díla

požadují, co se jim líbí, dále kdy pravděpodobně přehrávání zastaví nebo co přeskočí, jaké žánry a obsah vyžadují (Musso, 2020).

Big data vytvářejí výzvy a příležitosti pro posílení vztahů mezi televizními společnostmi, publikem a inzerenty v rámci digitálního televizního vysílání, a tedy i audiovizuální tvorby. Televizní vysílání zažívá další ze svých revolucí, která je vyvolána změnami širokopásmového televizního připojení, rozvojem internetu, možnostmi digitálních technologií, rozšiřováním dostupnosti digitálních technologií, zvyšující se popularitou sociálních sítí apod. Mění se zejména povaha distribuce televizního vysílání, a to směrem k širšímu využití internetu v rámci spotřeby audiovizuálních děl. Audiovizuální průmysl musí na tyto změny aktivně reagovat a realizovat nové strategie pro oslovení a získání cílového publika, což často souvisí právě i s nutností využití nových nástrojů a prostředků typu big data (Murschetz, Schlütz, 2018).

Strategie využití big data v rámci analýzy cílového publika audiovizuálního díla umožňuje generovat zejména následující výhody (Murschetz, Schlütz, 2018):

- Umožňuje porozumět přáním a potřebám cílového publika, přizpůsobit se požadavkům cílového publika.
- Umožňuje zlepšit práci s databázemi.
- Zprostředkovává možnost pro využití velkého množství dat za účelem lepšího cílení reklamních kampaní.
- Umožňuje automatizaci ve výrobě obsahu.

Organizace, které chtějí efektivně využívat big data v rámci audiovizuální tvorby, musí splňovat zejména následující podmínky (Murschetz, Schlütz, 2018):

- Implementace využití big data do strategie podniku, aby se vytvořilo povědomí o možnostech jejich použití.

- Změna organizační struktury, aby se mobilizoval potenciál využití big data.
- Pomoc zaměstnancům při práci s big data, při podpoře pochopení big data.
- Použití relevantních měřících technik pro vyhodnocení efektivity využití big data.

Řada odborníků na big data potvrzuje, že tvůrci audiovizuálního obsahu mohou využívat výhod datové analýzy ke zlepšení obsahu audiovizuálního díla a pro lepší oslovení cílového publika. Objevují se však i negativní názory, které upozorňují na nevýhody, jež s sebou nese využití big data v audiovizuální tvorbě. Nejčastěji se zmiňují následující nevýhody využití big data v audiovizuální tvorbě (Davies, 2019):

- Potřeba talentovaného a kvalifikovaného personálu – v případě, kdy chce tvůrce audiovizuálního obsahu využít big data, musí disponovat dostatečně kvalitním personálem, tj. odborníky, kteří umí s big daty pracovat a provádět jejich vyhodnocení. V řadě případů hledají tvůrci pomoc u externích agentur a velkých datově analytických organizací, protože jde o levnější variantu, než zaměstnání vlastních expertů.
- Problémy s kvalitou dat – tvůrci audiovizuálního obsahu se musí ujistit, že využívají spolehlivá data, která jsou ve správném formátu a umožňují jejich plnohodnotnou analýzu a interpretaci. Využití špatných dat může být neefektivní a příliš nákladné.
- Ochrana osobních údajů – tvůrci audiovizuálního obsahu musí zajistit ochranu osobních údajů zákazníků a dodržování legislativy související s touto problematikou. Ukládání osobních a citlivých informací o zákaznících není možné bez zajištění dostatečné ochrany a zabezpečení.

Využití big data se stále rozvíjí a vyvíjí, neustále se objevují nové výzvy a nelze konstatovat, že se jedná o vědu, která by už byla plně rozvinutá.

Některé otázky vyplývající z big data budou ještě dlouhou nerozluštěné. Nadále platí, že u výroby audiovizuálního díla bude muset být přítomen člověk v pozici odborníka, který nejen vyhodnotí daná data s ohledem na konkrétní okolnosti (Macnab, 2018).

2.3.1 Využití big data v marketingu

V současnosti již nelze pochybovat o tom, že mají big data obrovský dopad na zábavu a reklamu. Analýza těchto dat umožňuje tvůrcům audiovizuálního obsahu publikovat relevantní obsah pro jejich cílovou skupinu, což povede k dosažení značného finančního zisku. Big data umožňují pochopit cílovou skupinu, a to velmi hluboce. Marketingové a reklamní příležitosti plynoucí z využití big data jsou potom výrazným oživením pro celý audiovizuální průmysl, který byl ještě před deseti lety na pokraji propasti a hrozil mu výrazný ústup ze slávy. Technologie big data tomuto vývoji zabránila a vytvořila prostor pro další rozvoj (Goodworklabs, 2020).

Gemignani, Gemignani a Galentino (2015, s. 29) potvrzují, že se big data v dnešní době aktivně využívají v celé řadě oborů a činností, a to tedy nejen v rámci obchodu, ale také například ve veřejné správě, ve školství, zdravotnictví, ve sportu atd. Big data se stávají nástrojem pro předvídání, snižování nákladů, vyhledávání obchodních příležitostí, zlepšování výsledků výzkumů či výkonů, zefektivnění prodejní a výrobní činnosti.

Analýza big data může být využita pro různé účely. V rámci audiovizuální tvorby či realizace marketingové strategie se může jednat o následující (Goodworklabs, 2020):

- Cílení reklamy – big data poskytují informace o tom, kdy si publikum prohlíží konkrétní obsah, jak reaguje na tento obsah, jaké reklamy přeskakuje apod. Na základě těchto informací lze potom cílit reklamní sdělení tam, kde bude nejúspěšnější, případně identifikovat body, které v reklamě zaujmou a motivují k prohlédnutí celého obsahu.

Tvůrci reklamy se potom řídí konkrétními informacemi, a nikoliv dohady.

- Predikce chování diváka – big data poskytují informace o zvycích a chování diváků či poskytují informace demografického charakteru. Jejich kombinací pak lze vytvářet zajímavější obsah či doporučení ke konzumaci obsahu.
- Hodnocení audiovizuálního obsahu – streamingové služby typu Netflix či Amazon Prime využívají big data k hodnocení obsahu audiovizuálního díla. Toto následně ovlivňuje rozhodnutí o tom, jestli ve výrobě pokračovat, či nikoliv, případně, jaký rozpočet přidělit tvůrcům daného díla. Stejný princip je realizován při rozhodování o zahájení nové show.
- Plánování programu – big data se v tomto případě využívají zejména při živém streamování, ve kterém se platforma spojuje s živým publikem v rámci vysílání. Big data určují, kdy je nejvhodnější takové vysílání zahájit.

Netflix je běžně využíván jako příklad využití big data a algoritmů pro získání diváků a zajištění jejich loajality. Jedná se však o soukromou společnost, která svoji činnost realizuje za účelem dosažení zisku a big data chápe jako investici. Na druhé straně stojí některé veřejné televizní společnosti (tj. například klasičtí poskytovatelé lokálního veřejnoprávního televizního vysílání), které si mnohdy využití big data nemohou v takové míře dovolit, protože je to příliš nákladné, nemají k tomu personální zdroje a celkově se musí soustředit na jiné priority (Sørensen, Hutchinson, 2017).

2.3.2 Štítky a metadata

S rostoucím množstvím obsahu, který přibývá v knihovnách VoD platform, roste i objem práce spojený s jeho *tagováním* (štítkováním). Štítkování, nebo z hlediska informačních technologií lépe řečeno přiřazování parametrů, slouží

k tomu, aby byly stroj nebo umělá inteligence zodpovědná za vyhledávání, třídění, nebo měření kvality obsahu, schopny poskytovat relevantní výstupy. Štítků existuje celá řada. Od zcela základních, jako jsou:

- jména tvůrců;
- stopáž pořadu;
- klíčová slova;
- žánry a případně subžánry;

až po daleko sofistikovanější, se kterými nástroje na třídění a řazení obsahu pracují na pozadí.

Vždy je třeba rozlišovat, čemu, respektive komu mají *tagy* sloužit. Uživatele nejvíce zajímá pohodlnost konzumace obsahu. Potřebuje požadovanou položku co nejnáze najít, ať už v připravené kategorii v knihovně VoD portálu, nebo třeba pomocí jeho vyhledávače. Datové analytiky naproti tomu zajímají tagy k vyhodnocení některých nasbíraných dat a jejich filtrování.

Podle Jakuba Šilera jsou základní štítky, kterých existuje například v Mall.TV okolo třiceti, přiřazovány jednotlivým položkám vzhledem k jejich relativně k nízkému počtu, a tedy i objemu práce při nahrávání do administračního systému ručně. Přiřazování štítků se podobně tomu, jako je to u videí ve videotéce, realizuje i u mnoha jiného obsahu, který je dostupný online. Klasickým příkladem mohou být položky v e-shopu nebo novinové články na zpravodajské stránce. V tomto případě je podle Šilera už do jisté míry využívána i umělá inteligence, respektive strojové učení. Nástroj pro přiřazování štítků je schopný analyzovat text redaktorova článku a podle vývojářem nastavených kritérií sám nabízet vhodná klíčová slova například podle četnosti jejich výskytu v článku. (Šiler, 2020)

Podle společnosti Dell, která se věnuje řadu let vývoji a výrobě IT řešení, je vysoká poptávka ze strany provozovatelů VoD platforem po automatizaci

procesů štítkování a všeobecně práci s metadaty. Společnost Dell uvádí, že průměrně je trénovaný pracovník schopný za jednu osmihodinovou směnu zpracovat důkladně a podle potřeb provozovatele VoD platformy pouze dvě hodiny audiovizuálního materiálu. Efektivita je tedy 25 %. Na druhou stranu je strojová umělá inteligence schopná vyhodnotit stejný materiál s efektivitou 67 % (Burns, 2019).

2.3.3 Video genom jako identita obsahu

S pojmem *The Video Genom* se poprvé setkáváme v roce 2010, kdy autoři Kimmel, Bronstein a Bronstein publikovali studii, ve které zkoumají, jestli a jakým způsobem je možné přečíst „genetickou“ identitu video obsahu. Ve své práci přirovnávají DNA a jeho konkrétní složení z nukleotidů k video obsahu, který je složen z jednotlivých prvků audiovizuálního díla, které může divák vnímat a konzumovat. Každé dílo má tak jedinečný klíč složený z jednotlivých metadat a lze tak přečíst jeho verzi nebo celý obsah (Kimmel, Bronstein, & Bronstein, 2010).

Jak později publikoval magazín *Wired*, jejich snahou bylo primárně vytvořit nástroj, který by lépe napomáhal automatickému skenování a rozšifrování stovek tisíců hodin video obsahu, které se nacházejí na internetu, konkrétně například na stránce YouTube, a pomohl tak snadno detekovat pirátský obsah. Obsah je podle autora článku možné detekovat až s 99% přesností. (Welsh, 2010)

Autoři studie si v roce 2010 možná ještě neuvědomili, že taková práce s metadaty bude později sloužit k daleko širšímu použití všem velkými provozovatelům VOD platform.

Podle Bretta Danahera (2019) je použití skutečného genomu v praxi daleko víc o nastavení správné optiky. Konkrétně uvádí, že rozdělení na žánry nestačí, ale zabývat se daty na úrovni přítomnosti konkrétního objektu na

scéně nebo stopáže jednotlivé skladby soundtracku není efektivní. Dále hovoří o důležitosti v rozklíčování následujících faktorů:

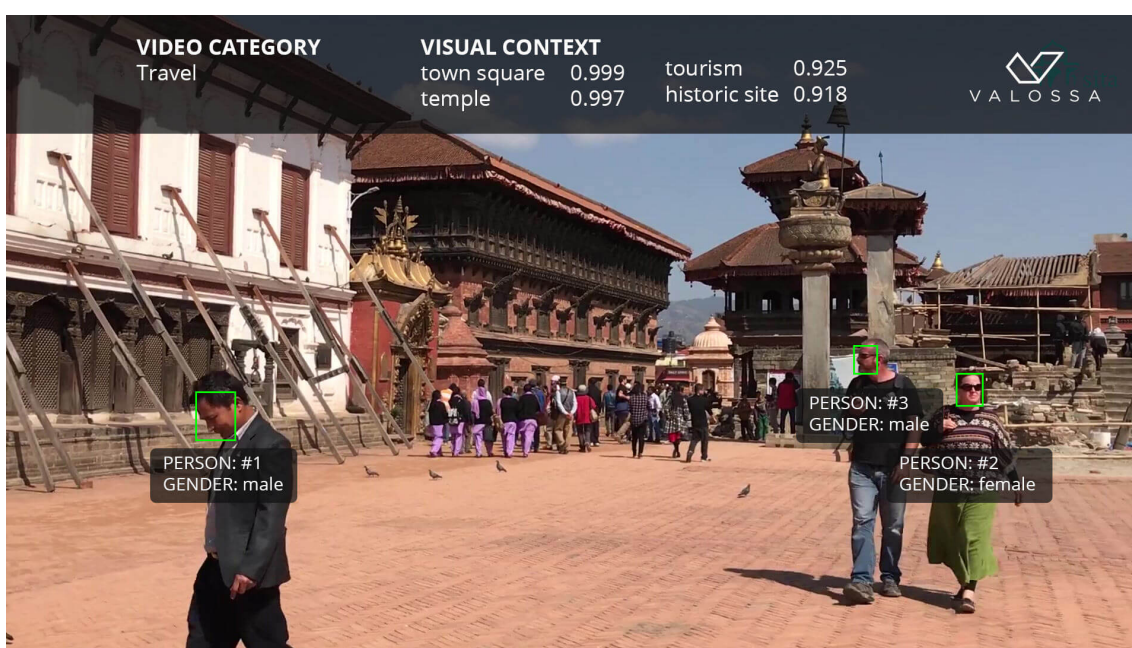
- téma;
- celková nálada nebo vyznění;
- příběhové oblouky;
- hodnocení;
- pohlaví, národnost, etnicita nebo příslušnost ke kultuře hlavních protagonistů;
- prostředí.

S přibývajícím snahou cílit produkt co nejpřesněji na koncového zákazníka rostou i nároky na analýzu produktu u všech druhů VoD. Aby mohla VoD platforma nabídnout efektivně koncovému uživateli personalizovaný produkt, potřebuje k tomu dvě věci. Tou první jsou co nejpřesnější údaje o zákazníkovi, o jeho potřebách a návycích nebo lépe řečeno o jeho touze po konkrétním obsahu a o jeho ochotě za něj zaplatit (ať už peníze, nebo čas strávený sledováním reklamy). Druhá, a v této části podstatnější, je pak právě znalost vlastního produktu. Například prosté rozřazení do žánrů se podle Thomase Burnse, CTO Dell pro média a zábavu nejeví jako úplně dostačující. Burns uvádí, že například nástroj Valossa⁹ je schopný rozpoznávat celou škálu parametrů, a to rychleji než v reálném čase. Konkrétně hovoří o schopnosti:

- rozpoznat jednotlivé herce na scéně;
- rozpoznat prostředí, rekvizity, dekorace;
- rozpoznat jazyk jakým protagonisté hovoří;
- určit barevné spektrum a nejvíce zastoupené barvy v obraze;

⁹ <https://www.valossa.com>

- určit v procentech nebo celkové stopáži díla výskyt produktu nebo představitele;
- určit v procentech nebo celkové stopáži díla množství hudebních doprovodů, ruchů nebo dialogů;
- vyhledat konkrétní obrazové nebo slovní zmínky určitého produktu, značky nebo termínu;
- určit množství a druh obsahu nevhodného dětem (nevhodný jazyk, nahota, násilí apod.) a mnoho dalšího. (Burns, 2019)



Obrázek 1: Ukázka záběru ze softwaru pro rozpoznávání AV pomocí AI, zdroj: Vallosa.com

Získaná data o již systémem zpracovaných filmech pak navíc služba používá k provozování bezplatného prohlížeče WhatIsMyMovie.com, kde je možné vyhledat AV položku podle jakýchkoliv klíčových slov na základě výše zmíněných parametrů nebo štítků.

Na základě takto podrobných metadat uspořádaných na časové ose je pak provozovatel VOD nejen schopen s velkou přesností analyzovat produkt, ale

zároveň i s každým dalším zhlédnutím zpřesňovat i velmi podrobně profil každého diváka.

Burns dále uvádí, že plánované množství investic streamingových platforem (např. Netflix, Amazon Prime, Hulu a další) do vytváření vlastního obsahu v roce 2022 bude znamenat, že tyto platformy vyrobí v úhrnu až jednu třetinu veškerých nabízených seriálů a filmů na trhu. To v konečném důsledku přinese další investice do datových analýz a lepších nástrojů. (Burns, 2019)

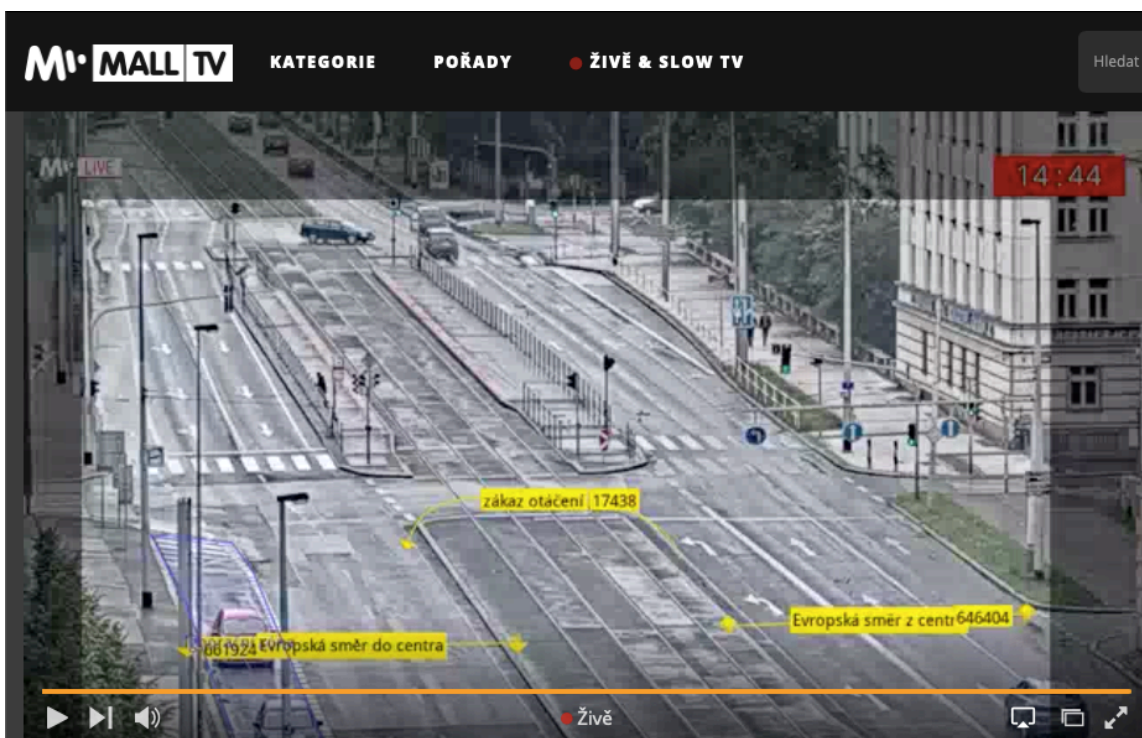
Brett Danaher (2019) během přednášky *The Frontier of Entertainment Analytics* uvedl, že není pochyb o tom, že VoD platformy s daty aktivně pracují při nastavování cenové politiky za službu v oblasti distribuce a v oblasti marketingu. Větší výzvou současnosti jsou podle něj fáze, které tomu předcházejí, a sice licencování, nebo lépe řečeno akvizování obsahu a rozhodování o vývoji nových vlastních pořadů.

Webová služba Hulu byla spuštěna v roce 2007, aby poskytovala zákazníkům streamovaná videa z různých televizních pořadů a filmů prostřednictvím internetu. V roce 2009 se většinovým vlastníkem VoD služby stává společnost Walt Disney. V současné době zažívá Hulu růst, ať už z hlediska nabídky audiovizuálního obsahu, či počtu zákazníků (Hosch, 2020).

Hulu v roce 2016 realizovalo akvizici start-upu Video Genome Project, který identifikuje, zpřesňuje a analyzuje data a metadata audiovizuálních děl, typu filmů, televizních pořadů, online pořadů. Technologie projektu využívá big data za účelem identifikace jasného a komplexního obrazu každé části audiovizuálního díla. Hulu touto akvizicí chce dosáhnout lepšího algoritmu pro personalizovaná doporučení audiovizuálního obsahu svým zákazníkům. Tato strategická akvizice by také měla Hulu pomoci k lepšímu a efektivnějšímu využití big data, což má zase zajistit další rozvoj poskytovaných služeb a růst organizace. Video Genom Project dokáže automaticky agregovat big data související s video obsahem a přiřadit „geny“ audiovizuálnímu obsahu na

základě stovek kontextových atributů identifikovaných v metadatech, které se potom integrují za účelem identifikace vztahů a souvislostí mezi jednotlivými audiovizuálními tituly. Výsledkem jsou relevantnější doporučení, která umožňují mnohem lepší propojení mezi zákazníkem a audiovizuálním obsahem. Hulu potvrzuje, že 75 % audiovizuálního obsahu je spotřebováno právě na základě doporučení, takže je využití big data pro zlepšení algoritmu velmi žádoucí (Spangler, 2016).

S mapováním obsahu pracuje i Mall.TV, i když zatím jen v omezené podobě. Jakub Šiler (2020) zmiňuje dva příklady, kde dochází k měření a zobrazování obsahových dat přímo v přímém přenosu v rámci vysílání. V živém slow TV pořadu *Umělá inteligence počítá přestupky* dochází k měření dopravních přestupků na jedné z pražských ulic a divákům se zobrazují přímo v obraze počty aut, která dopravní přestupek spáchala. Druhým příkladem je měření počtu osob, které na sobě mají nebo nemají roušku v pořadu *Umělá inteligence počítá roušky na Václaváku*. Oba případy jsou spíše zajímavým příkladem použití technologie, než že by sloužily k reálnému rozboru dat obsahu, jak je uvedeno v některých příkladech výše.



Obrázek 2: Ukázka využití umělé inteligence v jednom z pořadů na Mall.TV, zdroj: Mall.TV, 2020¹⁰

Příspěvek společnosti Dell Technologies na konferenci *Filmmaking by the Numbers* uzavírá Burns prognózou budoucnosti audiovizuální tvorby s využitím big data a moderních technologií následujícím způsobem (Burns, 2019):

- Shromažďování a zpracování dat bude probíhat prostřednictvím cloudu.
- Využití big dat bude stále více integrováno s umělou inteligencí a strojovým učením.

¹⁰ <https://www.mall.tv/stavby-a-technologie/umela-inteligence-pocita-prestupky>

- Shromažďování a zpracování dat bude probíhat prostřednictvím cloudu.
- Využití big dat bude stále více integrováno s umělou inteligencí a strojovým učením.
- Dnes je nutné přesouvat obrovské soubory digitálních médií a dat. V budoucnosti budou big data zpracovávána bez nutnosti přesunu, takže jejich zpracování bude podstatně rychlejší.

2.3.4 Doporučování obsahu

„Více než 80 % televizních pořadů, které lidé sledují na Netflixu, uživatelé objevují prostřednictvím systému doporučovací platformy. To znamená, že většina toho, co se rozhodnete na Netflixu sledovat, je výsledkem rozhodnutí učiněných na základě výpočtu algoritmu (Plummer, 2017).“

Výzkumná studie společnosti UserTesting, Inc. z roku 2018 naopak zjistila, že v USA pouze 29 % zákazníků SVOD služeb sleduje obsah na základě doporučení, což znamená, že z obecného hlediska se musí VOD platformy v tomto směru hledat prostor pro zlepšení algoritmů doporučení (Pressnell, 2019). Studie, která zkoumala pět SVOD platforem (Netflix, HBO Now, Hulu, Amazon Prime a prémiovou Youtube TV) však pracovala v dotazníkovém šetření pouze s malým vzorkem pěti set respondentů a výsledky proto mohou být zavádějící. V rozhovoru pro *Bussines Insider* uvádí Carlos Gomez-Uribe – někdejší viceprezident produktového vývoje a doporučovacích algoritmů – rovněž číslo 80 % (McAlone, 2016).

Společnost UserTesting ve studii uvádí: *„Přestože většina zákazníků dostala nějakou formu doporučení (s výjimkou HBO Now), většina služeb nenaplnila zcela očekávání, pokud jde o poskytování příslušných doporučení na základě zákazníkovi historie sledování a jeho preferencí. Netflix poskytl nejlepší*

zážitek z pěti zkoumaných služeb, ale zákazníkům poskytl pouze jedno, někdy dvě doporučení, která zákazníci chtěli sledovat. (UserTesting, 2018)"

Pressnell (2019) komentuje výsledky studie tak, že lídři tohoto trhu si tuto skutečnost uvědomují a aktivně investují do zlepšení poskytovaných služeb. Toto zlepšení je pak často podmíněno právě schopností pracovat co nejvíce s relevantními daty.

Nehledě na to, že oba zmíněné údaje ukazují jiné procentuální využití doporučení uživatelem, trend posílení role doporučení je zcela zjevný.

„Nepříliš dávno bylo při práci s informací hlavním problémem dostat se k tomu, co člověk chtěl. Dnes hlavní problém spočívá v tom, aby si ujasnil, co vlastně chce, a vybral si z obrovské nabídky. Jistě to znáte, vybíráte si knížku nebo film a stojíte před regálem (nebo internetovou stránkou), který obsahuje tolik možností, že vás to paralyzuje a nejste schopni se rozhodnout. Oblast, která nám může pomoci, se nazývá doporučující systémy. (Pelánek, 2011)"

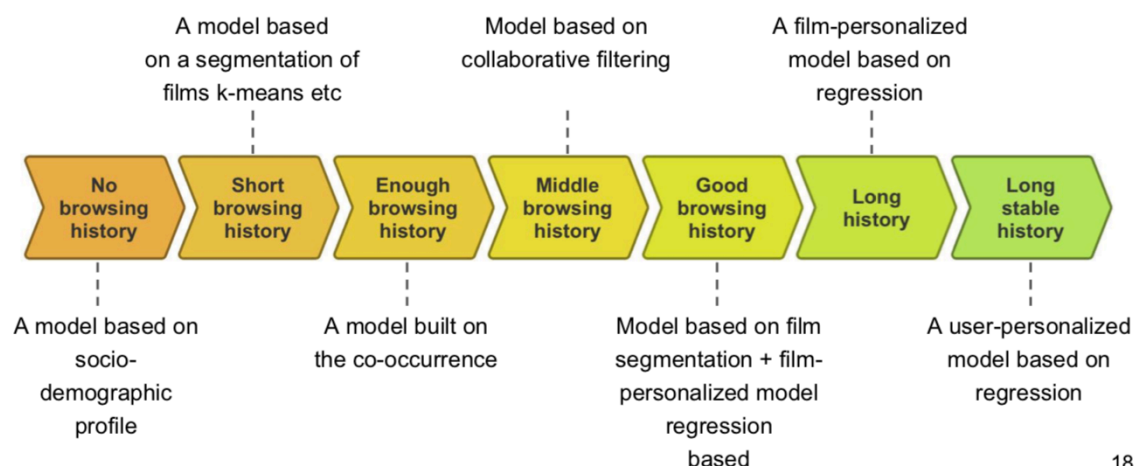
Doporučovací systémy jsou dnes součástí všech pokročilých e-shopů a využívá je i většina VOD platforem. Tento logický důsledek vychází i ze snahy zlepšit *User Experience* (uživatelský zážitek), neboť s globálně rostoucím množstvím nabídky produktů VOD platforem na internetu může vznikat u uživatelů rozhodovací paralýza. Třídění a doporučování je tedy systém, který pozorujeme napříč odvětvími.

Obecně rozlišujeme dva druhy doporučovacího systému. Prvním z nich je tzv. *Content Based system* (postavené na analýze obsahu) doporučování. Tento systém pracuje s produktem (videem v knihovně pořadů VoD platformy) a s jeho parametry (štítky). Na základě požadavku uživatele tak systém doporučí co nejbližší výsledek definovaný štítky. Druhým druhem doporučovacího systému je *Collaborative Filtering system* (kolaborativní

filtrování), který pracuje se skupinovým chováním a zájmy podobných uživatelů (Pelánek, 2011).

Štítky zpřesňují typologii obsahu, ale pro sofistikovanější řešení doporučení samotného nejsou tak stěžejní. Obecně platí, že čím více informací z hlediska obsahu (objem) a po čím delší dobu jsou měřeny (kontinuita), tím může dostat uživatel lépe doporučený obsah.

Záleží ovšem na doporučovací metodě, která je v daný moment zvolena. Graf 10 níže zobrazuje použití různých metod systému doporučení v různých fázích sběru dat, který vede k ucelenější a přesnější představě o uživatelské zájmu.



18

Graf 10: Časová osa použití různých metod doporučení na základě množství dostupných uživatelských dat, zdroj: Kapustina, 2015

Z laického úhlu pohledu by bylo ideální pro provozovatele VOD platformy vědět při doporučení jakou má uživatel v době sledování náladu, co ho přesně zajímá a proč na web přišel, případně kolik přesně času na něm může strávit, aby mu pak na konci zobrazil jednu jedinou položku: „dokonalý“ produkt servírovaný v „dokonalý“ čas. Praxe je však jiná.

Pokud uživatel vstoupí na web VoD platformy poprvé nebo v režimu anonymního prohlížení, dostane obsah v téměř výchozím zobrazení – ovlivněném například pouze trendy aktuální sledovanosti. Zdálo by se, že uživatel přihází jako nepopsaný, nebo lépe řečeno jako nepřečtený list papíru, avšak díky možnosti získávat informace z webů třetích stran je již VoD platforma při použití správné metodiky schopná na základě jeho předchozí aktivity na internetu odhadnout, jaké jsou jeho zájmy a také socio-demografický profil.

„Prvním krokem je získání dat, na základě kterých se budou doporučení dělat. Data většinou udávají hodnocení jednotlivých položek uživateli, přičemž hodnocení může být buď explicitní, tj. uživatel přímo produkt ohodnotí (hodnocení 1 až 5 hvězdičkami podle toho, jak se mu líbil film), nebo implicitní, tj. sledováním toho, co uživatel dělá (sledováním toho na jaké položky na webových stránkách klikl). (Pelánek, 2011)“

Pro doporučování obsahu není ale zcela tak důležité, jaké je chování jednotlivce, jako spíše jaké je chování uživatelů skupiny uživatelů před ním. Od první interakce s obsahem se uživatel pomyslně přiřadí ke skupině, která měla zájem o stejný obsah, a na základě metody *Co-Occurance* je mu doporučen další obsah, o který projevila zájem největší skupina dále (Kapustina, 2015).

Pelánek (2011) ilustruje systém kolaborativního doporučování na příkladu šesti filmů a hodnocení, které může být, jak bylo zmíněno výše, i implicitní a záleží pouze na kritériích pro jeho měření.

	Pelíšky	Pupendo	Kolja	Titanic	Terminátor	Schindlerův seznam
já	5	4	5	3	1	?
Petr	2	3	1	4	5	3
Jana	5	4	4	4	2	5
Matěj	3	2	3	5	4	2

Tabulka 1: Příklad hodnocení pořadů uživateli k ilustraci systému kolaborativního filtrování, zdroj: Pelánek, 2011

Pelánek (2011) se zamýšlí nad tím, jaký film by se mu mohl dále líbit ve vztahu k hodnocení ostatních filmů jeho kolegy. Zatímco na základě *Content Based* analýzy mu vychází, že další film by byl například *Musíme si pomáhat*, protože v něm rovněž účinkuje v hlavní roli Bolek Polívka, na základě *Collaborative Filtering* analýzy mu ale může vyjít film úplně jiný. Z tabulky výše by se dalo očekávat, že vzhledem k tomu, že uživatel *Jana* hodnotil nejpodobněji uživateli *já* u předchozích filmů, pak je statisticky pravděpodobné, že hodnocení *Schindlerova seznamu* v tabulce bude u uživatele *já* také vysoké.

Analýza obsahu málokdy doporučí něco překvapivého – oblíbeného herce nebo režiséra dokážeme určit sami a na filmy, které nám tento systém doporučí, bychom nejspíš časem přišli. Kolaborativní filtrování naproti tomu dokáže vygenerovat překvapivé návrhy, protože k doporučení využívá informace, které nemáme k dispozici (podobný vkus s neznámými osobami) (Pelánek, 2011).

Je však důležité podotknout, že i metoda kolaborativního filtrování může uzavírat uživatele ve vkusové a zájmové bublině. Strojové učení a quasi-umělá inteligence mohou v doporučování napomáhat, ale redakční práce je nenahraditelná pro udržení objektivitu a pestrosti. Sofistikovanější systémy proto pracují v rámci doporučování i s přidáváním nových položek, které se

zavedenou vazbou tolik nesouvisí. Strojové učení se tedy zpřesňuje na základě

Matoušek uvádí, že pro přesnější měření a zlepšování doporučování nebo pro testování nových pořadů se někdy uměle ohne doporučovací algoritmus, aby se tím zpřesnil výsledek celku. (Matoušek, 2020)

„Uvedená tabulka (Tabulka 1) je samozřejmě jen ilustrativní, z takto malého vzorku dat nelze dělat použitelné předpovědi. Reálné doporučující systémy, jako je třeba Netflix, mají k dispozici statisíce hodnocení, která navíc dále zpracovávají a filtrují... Netflix považuje doporučující systém za natolik důležitý¹¹, že vypsala v roce 2006 soutěž, jejímž cílem bylo vylepšit přesnost jejich předpovědácího algoritmu o 10%. Hlavní odměna v soutěži byla 1 milion dolarů. (Pelánek, 2011)“

Jakub Šiler (2020) vysvětluje na příkladu uživatelů, kteří přišli na web Mall.TV a dá se o nich zjistit, že měli v minulosti zájem o tematiku auto-moto, že projeví s největší statistickou pravděpodobností zájem o pořad, který do této tematické kategorie v rámci nabídky katalogu spadá (např. *Mistři volantu*). Čím větší skupina diváků pak klikne na další konkrétní pořad (např. *Darwinovy ceny*), tím více se v principu každému novému dalšímu uživateli se zájmem o pořad *Mistři volantu* doporučí také pořad *Darwinovy ceny* (Šiler, 2020).

V praxi musí být doporučovací algoritmus daleko sofistikovanější, aby se například všem nezobrazoval stejný obsah, o který projeví zájem největší skupina uživatelů. Vliv na doporučování má také to, pokud uživatelé na webu dříve konzumovali stejnou položku obsahu, zda odebírají nějakou sérii pořadu nebo zda jsou zaregistrovaní a přihlášení a web o nich sbírá zmíněné zájmy

¹¹ Další podrobnosti k systému doporučování Netflixu jsou popsány v případové studii v podkapitole 3.2.1

a z toho určuje i dlouhodobě vyplývající vazby. Systému doporučování a jeho inovaci se věnuje jak v Seznamu, tak v Mallu specializované oddělení, protože nespadá zcela do kompetence analytiků datového obsahu (Matoušek, 2020).

Kapustina (2015) uvádí další příklad metody doporučování: *Binary Logistic Regression* (systém Binární logické regrese). Tento pokročilý systém zpřesňuje doporučování na základě dlouhé historie sběru uživatelských dat a na krátké časové úseky doporučování není vhodně aplikovatelný. Doporučovací systémy se neustále zlepšují invencí matematiků a vývojářů a ukazuje se, že je vhodné používat jejich kombinace. Zároveň je však potřeba nahlížet na doporučovací systémy vždy jen jako na doplňkovou část větší úvahy o distribuci produktů (Kapustina, 2015).

Protože jsou webové katalogy VoD platforem živým organismem, který se neustále mění s přibývajícím obsahem, přistupuje se zvláště k novým položkám, u kterých prokliky, a tedy i vzorec doporučování na základě kolaborativního filtrování ještě neexistuje. Většinou se podle Matouška přidává položce nějaký zvláštní parametr (např. kategorie „nové“) a taková položka má zpravidla období hájení, než se přepne do režimu běžného doporučování. K přepočítávání a zpřesňování vazeb v doporučovacím systému dochází (na Mall.TV) v pravidelném časovém intervalu – každých 8 minut. (Matoušek, 2020).

Snaha o doručení obsahu na míru, a zejména pak toho reklamního, se však netýká pouze VoD platforem. Podle The Wall Street Journal například kabelový vysílatel Comcast v roce 2015 odkoupil společnost Visible World a přidruženou společnost AudienceXpress, které od roku 2000 jako první na světě vyvíjely systém, jenž inzerentům, reklamním agenturám a mediálním společnostem umožňuje dodávat programatickou, adresnou a měřitelnou reklamu. Tato služba měla zastoupení ve třech milionech domácností a ruku

v ruce s tím šlo i o výzkum a následně lepší schopnost predikce uživatelského chování (Vranica, 2015).

2.3.5 A/B testování

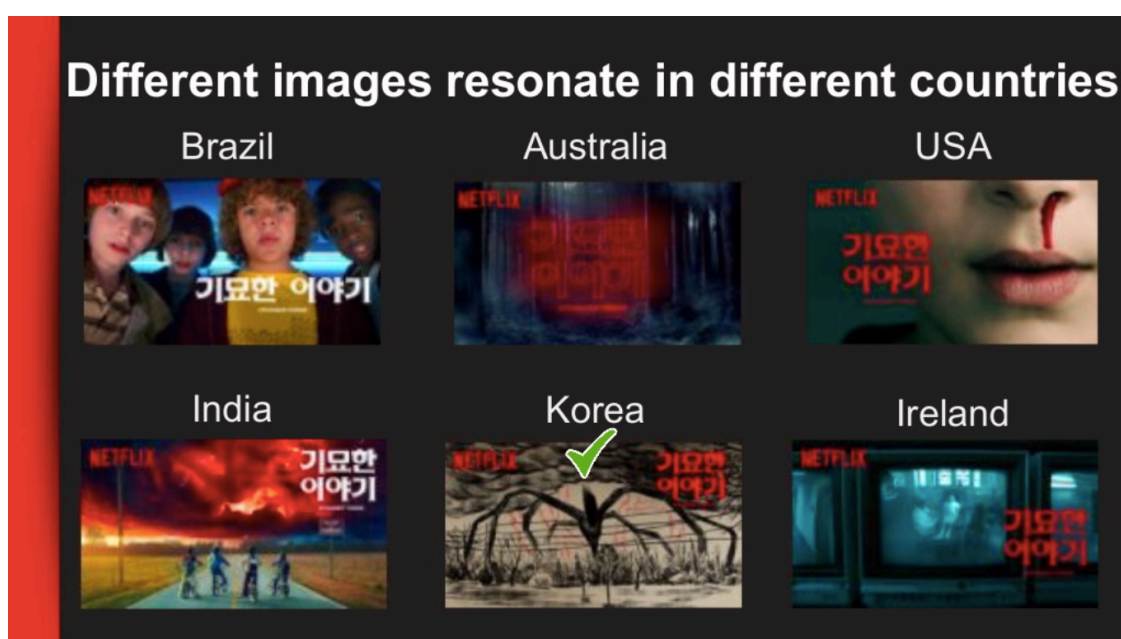
A/B testování je marketingová metoda, která se využívá ke zvýšení konverzí uživatelů prostřednictvím změny obsahu nebo zobrazovaných prvků na stránce. Její průběh spočívá v tom, že části návštěvníků webové stránky se zobrazuje varianta A (označována jako šampion nebo vzor), a části uživatelů se zobrazuje varianta B (označovaná jako vyzyvatel). Následná analýza dat, získaných na základě měření aktivity uživatelů obou variant za stanovený čas, vede k rozhodnutí, která varianta bude nadále preferovaná. Vliv na výsledky A/B testování má velikost zkoumaného vzorku (návštěvnost na webu), počet zkoumaných prvků a rozdílnost mezi nimi a významnost prvku v dopadu na konverzi (vliv měřený cíl) (Snížek, 2011).

Podle Nicka Nelsona, který se ve společnosti Netflix věnuje strojovému učení, je A/B test jedním z hlavních nástrojů, který ovlivňuje konečnou podobu stránek Netflixu (Usabilla, 2017).

„Klíčové je, aby obsah, který dostávají uživatelé, byl co nejvíce přehledný a příjemný na konzumaci (Goyal, 2019)“ Obsah, který na Netflixu uživatelé dostávají, je seřazený na základě různých parametrů, jako jsou například žánry a subžánry, tvůrci, země původu, rok vydání, věkové omezení, jazyk a hodnocení. Položky k přehrání jsou rozděleny automaticky na základě metadat do sbírek nejen podle žánru, ale i podle příležitosti, při jaké by se na ně uživatel mohl dívat (např. večer s přáteli) nebo nálady kterou navozují. Vše vyhodnocuje algoritmus, který zajišťuje požadované pořadí a umístění na stránce (Goyal, 2019).

Pro uživatele je důležité, aby se mohl rychle a snadno rozhodnout, jakou položku si z knihovny vybere. K tomu slouží čtyři hlavní nástroje: obrázek, ukázka nebo trailer, anotace a shoda vyjádřená procenty na základě dat o

uživateli o tom jaký obsah v minulosti konzumoval. Konkrétní výsledky testů se liší celosvětově napříč populací. Ovlivňují je kulturní a sociodemografické faktory. „V případě testování vizuálních výstupů dochází k propojování umění a technologie za účelem dosažení nejlepších výsledků“ uvádí na příkladu testování propagačních materiálů ředitelka marketingu a obsahové analytiky Kelly Uphoffová (2018).



Obrázek 3: Různé varianty cover grafiky pro pořad Stranger Things podle země distribuce, Zdroj: Uphoff, 2018

„Pro každou položku v našem katalogu testujeme stovky náhledových obrázků – zvláště v každém regionu a na každém typu zařízení. Není možné, aby pro takové množství obsahu vybíral obrázky člověk. Proto jsme sestavili algoritmus, který vybírá správné obrázky na základě toho, zda obsahují akci, zda jsou ostré, zda mají správný poměr kontrastu a světelnosti, a jsou celkově kvalitní (Goyal, 2019).“

Goyal (2019) rovněž poukazuje na to, že významná část uživatelů si přehraje před spuštěním samotného filmu nebo seriálu také video ukázkou, na základě které se následně rozhoduje, zda pořad shlédne či nikoliv. Tyto ukázky, stejně

jako celou řadu externích reklamních výstupů pro marketing a pro promování jednotlivých pořadů, Netflix rovněž sestavuje pomocí automatizovaného procesu a A/B testování.

Služba zároveň mapuje časovou osu každé položky a uživatel má tak možnost přeskočit jednoduše pomocí kliknutí na jedno tlačítko úvodní a závěrečné titulky.

Metriky, které zajímá Netflix a další SVOD platformy v současné době pro posouzení úspěchu nejvíce jsou:

1. Počet uživatelů
2. Doba, po kterou zůstanou předplatiteli
3. Počet hodin obsahu, které konzumují

3 Případové studie využití dat v audiovizuální tvorbě

3.1 Amazon

Amazon patří mezi největší elektronické obchody na světě, které se zaměřují na maloobchodní prodej různého zboží koncovým zákazníkům. V současnosti provozuje více než 175 logistických center po celém světě, zaměstnává přes 250 000 pracovníků a denně expeduje přes milión položek zboží. Tímto se zároveň vytváří obrovské množství dat, s kterými organizace aktivně pracuje, aby zlepšila svoji výkonnost, hospodaření a konkurenceschopnost. Výzvy, s kterými se Amazon potýká v oblasti big data, jsou podobné jako u ostatních velkých společností. V případě vzniku VOD služby Amazon Prime došlo k využití technologické vyspělosti společnosti pro účely rozšíření portfolia produktů. (Vogels, 2019).

Pro práci s big daty využívá Amazon tzv. datové jezero (data lake). Jde o centralizované a zabezpečené datové úložiště, které umožňuje ukládat, spravovat, objevovat a sdílet všechna strukturovaná a nestrukturovaná data v libovolném měřítku. Tato datová jezera umožňují zpracovávat nestrukturovaná data, takže datový analytik musí přesně vědět, jaká data chce prozkoumat a získat (Vogels, 2019).

Výhodou datového jezera je zprostředkování přístupu k datům těm týmům, které aktuálně daná data potřebují. V rámci rychlého růstu jsou data klíčovým prostředkem pro dosažení úspěchu, ale musí být k dispozici relevantním osobám, pracovníkům. Datové jezero umožňuje ukládat data na různá místa na světě, a to různými způsoby v reálném čase (Vogels, 2019).

„Amazon dokáže z dat o nákupním chování uživatelů vytáhnout veškeré souvislosti, takže lze jen s velmi malou nadsázkou říci, že ví o přání svých zákazníků dříve, než jim přijdou na mysl. Sklady Amazonu se tak plní v předstihu zbožím, o které bude dle jeho předpovědi v dané lokalitě největší

zájem během následujících týdnů, čímž se ušetří miliony dolarů na logistice (Janča, 2015)."

Amazon využívá big data i v rámci cenové politiky, kdy vyhodnocuje poptávku po produktech a v reálném čase upravuje cenu dle očekávaného prodeje. Prostřednictvím big data umí Amazon odhadnout ochotu zákazníka ke koupi, a tedy na základě tohoto nastavit cenu. K tomuto ještě přidává hodnocení cen konkurence, skladovou dostupnost apod. To vše vede k dynamické cenotvorbě formou big data (Danzinger, 2019).

Další příklad využití big data lze vysledovat u Amazonu v oblasti zajištění bezpečnosti uživatelů a eliminace podvodných transakcí. Systém společnosti umožňuje efektivní datovou analytikou každoročně zabránit podvodným transakcím ve výši přesahující milion dolarů (Danzinger, 2019).

Pro Amazon jsou dále big data důležitá v rámci doporučení produktů uživatelům. Personalizace přístupu k zákazníkům umožňuje nabídnout zákazníkům přesně to, co momentálně hledají, co potřebují. Tímto roste jejich spokojenost a loajalita (Danzinger, 2019).

Tyto a mnohé další případy technologické vyspělosti ve využití dat a práce s nimi stojí za tím, že Amazon rozšířil portfolio svých služeb do dalších oblastí, kde je může potenciálně využít.

3.1.1 Amazon Prime a seriál Alpha House

Big data v rámci audiovizuálního obsahu využívá Amazon při poskytování svojí služby Amazon Prime (Goodworklabs, 2020).

Big data využívá nejen v rámci audiovizuální tvorby, ale také k integraci této tvorby se svými dalšími produkty a službami. Příkladem může být snaha o efektivní propojení obsahu audiovizuální tvorby a maloobchodní nabídky, kdy dochází k analýze toho, jak diváci reagují na přítomnost některých produktů

v audiovizuálním vysílání (product placement), a do jaké míry pak produkty u Amazonu poptávají a kupují (Johnson, 2018).

Amazon díky těmto efektům ve stále větší míře navrhuje produkty tak, aby se jejich prostřednictvím dalo získávat co nejvíce informací o uživateli, a tedy, aby vytěžil co nejvíce big data. Příkladem je právě služba Amazon Prime, která je založena na nutné registraci zákazníka (členství), a která poskytuje přístup nejen k audiovizuálnímu obsahu, ale také umožňuje nákupy na jiných platformách Amazonu, takže organizace získává přehled o tom, co zákazník kde nakupuje. Amazon Prime přímo motivuje k co největšímu počtu aktivit a nákupů dalších služeb a produktů, což vytváří prostor pro růst tržeb a podrobnější představu o jednotlivých zákaznících. Big data pak využívá Amazon i při návrhu dalších produktů a služeb, mimo jiné pořadů pro Amazon Prime (West, 2019).

Konkrétním příkladem využití dat v Amazon Prime, s přímým dopadem na vývoj pořadů, bylo vytvoření osmi konceptů televizních seriálů a natočení jejich pilotních epizod. Amazon Prime následně umožnil jejich přehrání zdarma všem uživatelům internetu, což znamenalo miliony amerických diváků (v podstatě základ cílové skupiny). Amazon pak mohl sledovat jejich chování, shromažďovat datové body (co při sledování dělají, kdy obsah zastavují apod.) Na základě těchto dat v kombinaci s uživatelským hodnocením, Amazon určil, který koncept televizní show bude pravděpodobně nejúčinnější, a zahájil jeho produkci. Tímto vznikl seriál Alpha House, který byl ovšem po dvou sezónách stažen z vysílání, a rozhodnutí se tedy posléze ukázalo jako chybné (Macnab, 2018).

3.2 Netflix a seriál House of Cards

Subías, Laverón a Molina (2018) se ve svém článku věnují historickému vývoji společnosti Netflix, která byla původně založena jako online služba, kde si mohli zákazníci od roku 1999 vypůjčit DVD prostřednictvím pošty. Od roku

2007 začala služba se streamováním audiovizuálního obsahu, který získávala pomocí akvizic, a následně v roce 2012 přidala nabídku vlastních pořadů (Subías, Laverón, Molina, 2018).

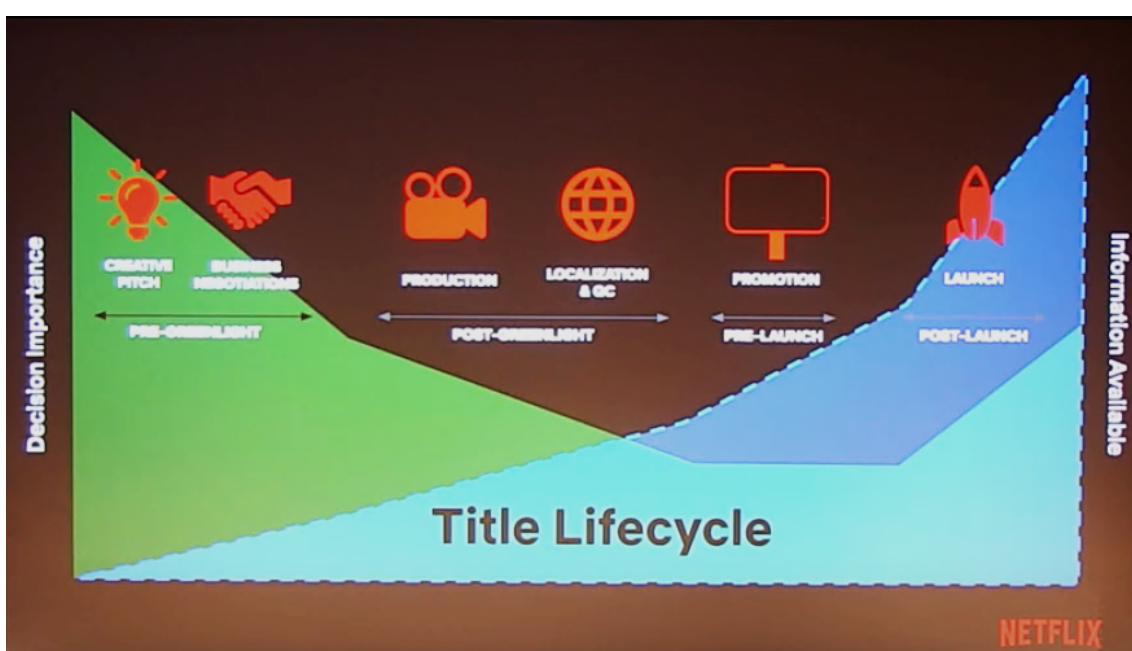
Dle autorek Subías, Laverón a Molina (2018) klade Netflix důraz na využití big dat, protože tato platforma nabízí vysoce rozmanitý obsah, který musí neustále zdokonalovat a nabízet čím dál širším skupinám potenciálních zákazníků (Subías, Laverón, Molina, 2018).

Marr (2016, s. 17-18) popisuje, že big data využívá a shromažďuje Netflix pro pochopení zákaznického chování. Rozvoj využití big data nastal se změnou primární podnikatelské činnosti. V momentě, kdy se Netflix začal orientovat na streamování televizních pořadů, vznikla potřeba pochopení zákazníků a jejich chování ve větší míře, než tomu bylo, když docházelo k distribuci DVD. Streaming zpřístupnil nová data o zákaznících a umožnil vytvářet modely, které predikují chování zákazníků, jejich tendence k využití různých služeb atd. (Marr, 2016, s. 17-18).

Dle Marra (2016, s. 18-19) se v posledních letech Netflix profiluje i jako výrobce audiovizuálních děl, a v tomto směru využívá big data rovněž velmi aktivně. Strategie vývoje audiovizuálních děl je ovlivněna daty, s kterými Netflix pracuje. Prvním, a často skloňovaným případem, byl seriál *House of Cards*, který byl ovlivněn prediktivním datovým modelem, kdy každý aspekt výroby televizního seriálu vycházel z výstupů analýz big data (Marr, 2016, s. 18-19).

Jak potvrzuje Melody Dye v rámci TED přednášky *Stanford at Netflix: From Script to Screen*, s daty se pracuje během celého životního cyklu pořadu. Na ose jednotlivých fází, od kreativního zadání přes výrobu a propagaci až po zveřejnění a vysílání, která je vyobrazená níže, jsou vyobrazené dvě křivky. První s označením Decision Importance značí důležitost rozhodování na základě dopadu na výslednou podobu pořadu a počtu zásahu diváků

respektive konzumovaných hodin pořadu. Druhá nese označení Information Available a znázorňuje růstovou tendenci, která značí, že informací, které jsou tvůrci schopní o daném pořadu získat, rostou v čase. Logicky největší (masivní) sběr dat probíhá až po zveřejnění položky v seznamu pořadů, avšak k průběžné práci s daty a jejich vyhodnocení už probíhá podle Dye daleko dříve během celého procesu (Dye, 2019).



Graf 11: Životní cyklus pořadu v návaznosti na dostupnost získávaných dat a význam rozhodování o procesu, zdroj: Dye, 2019

Melody Dye ve své přednášce dále uvádí konkrétní příklady¹², u nichž Netflix s daty pracuje (Dye, 2019):

1. Doporučování obsahu – skladba programu pro každého uživatele na míru v daný časů;

¹² Některé z nich jsou popsány v jiných částech této práce.

2. Strojové učení – náhledové obrázky pro jednotlivé filmy jsou různé v každém regionu a pro každý typ uživatele;
3. Monitoring sítě – automatické předstahování dat do lokálních serverů podle kvality připojení a zájmu o konkrétní položky z knihovny v daném regionu;
4. Preprodukce – mapování a vyhodnocování míst vhodných pro natáčení z hlediska dostupných a oblíbených lokací, obsazení herců na základě diváckých preferencí (jazyka, etnicity, povědomí atd.), mapování možných úspor v produkčních výdajích;
5. Postprodukce – AI dává poznámky ke střihu, pokud narazí na potenciálně problematické místo.¹³

Od chvíle, kdy se začíná uvažovat o tvorbě nového audiovizuálního díla, se objevují možnosti pro využití big data. V rámci pre-produkce dochází k rozhodnutím, která mají v konečném důsledku významný dopad na celkové náklady, harmonogram či kreativní výsledky. Tradičně se rozhoduje dle zkušeností a intuice, ale stále častěji se využívá i big data a jejich analýzy pro rozhodnutí. Netflix využívá big data pro vyhodnocení různých scénářů vývoje (například celkové náklady v závislosti na lokaci, kde se bude seriál natáčet apod.) Tímto se big data promítají i do oblasti produkce audiovizuální tvorby. V rámci postprodukce se big data využívají například pro identifikaci

¹³ Pozn. aut.: Zejména body 4. a 5. v uvedeném výčtu jsou z hlediska filmové produkce a producentství nejzajímavější, avšak ukazuje se, že není možné o tomto tématu v souvislosti s Netflixem dohledat více konkrétních informací. Buď není doposud v plné míře využíván, nebo zůstává zatím předmětem obchodního tajemství.

nedostatků a doporučení k vhodným změnám. V rámci lokalizace dochází k přizpůsobení obsahu jednotlivým trhům, tj. dabování audiovizuálního díla apod. Big data v tomto případě nabízí vhodný postup založený na specificích každé lokality, resp. každého trhu (Kumar et al., 2018).

Zmíněný bod 3. Monitoring sítě ve výčtu výše ukazuje, že podobně jak zmiňuje Janča v předchozí kapitole v případě Amazonu v souvislosti s prediktivním doplňováním skladů, i Netflix pracuje s předpovídáním chování svých uživatelů, a to dokonce přímo u AV produktů. Předpovídání je důležité zejména při zajištění kvalitního streamingového přenosu bez zpoždění. Při současném rozšíření služeb Netflixu na trhu a při objemu dat, který představuje desítky tisíc hodin audiovizuálního obsahu v plné kvalitě, není možné, aby byla všechna data uložena fyzicky na jednom místě. Po celém světě proto existuje několik velkých serverů, které distribuují obsah na základě předpovědi uživatelských preferencí do menších regionálních celků. Jinými slovy, Netflix dopředu na základě dat ví, jaký pořad bude v danou chvíli a na daném území nejvíc sledovaný, a připraví na zátěž svou distribuční síť.

Takové opatření je nezbytné, neboť v roce 2015 Netflix ve Spojených státech nárazově využíval až 37 % celkové kapacity průtoku dat v internetové síti, což je obrovské množství (MacDonald, 2015). I po sérii inovativních opatření v oblasti kompresí a světovém vývoji přenosu dat se v analýze z roku 2019 se ukazuje, že až 12,6 % veškeré světové konzumace dat z hlediska objemu způsobuje právě Netflix (Sprangler, 2019).

3.2.1 Netflix Prize

V roce 2006 vypsal Netflix soutěž o nejlepší algoritmus na doporučování obsahu. V rámci soutěže byla rozdána částka ve výši 1 milionu dolarů programátorům, kteří přišli s nejlepším algoritmem pro předvídaní toho, jaký obsah budou chtít zákazníci sledovat. V dnešní době pomáhají algoritmy

společnosti Netflix ušetřit až 1 miliardu dolarů ročně, kterou by jinak musel Netflix investovat do udržení stávajících zákazníků (Insidebigdata, 2018).

Tato soutěž reagovala na zjištění, že typický zákazník společnosti Netflix ztratí zájem o sledování určitého pořadu po 60 až 90 vteřinách od doby, kdy jej začne sledovat. Ve stejné době se rozhoduje mezi 10 až 20 pořady. Z tohoto důvodu bylo nutné zlepšit algoritmus pro doporučení audiovizuálního obsahu (Marr, 2018).

Výsledek soutěže byl zveřejněný v roce 2009, kdy cenu získala skupina programátorů označovaná jako BellKor's Pragmatic Chaos tým. (Netflix, 2009)

Algoritmus pro doporučení audiovizuálního obsahu nebyl na příliš dobré úrovni. Řada zákazníků si například stěžovala, že jim je doporučován film, který už zákazník viděl apod. Bylo tedy nutné přistoupit ke změně algoritmu a vyhlášení soutěže se jeví jako dobrá příležitost (Fritz, 2018).

Výsledek soutěže byl zveřejněný v roce 2009, kdy cenu získala skupina programátorů označovaná jako BellKor's Pragmatic Chaos tým. (Netflix, 2009)

Podle CEO společnosti Netflix Reeda Hastings se vedly tehdy výsledky dosažené pomocí implementace vítězného projektu Netflix Prize ke zlepšení doporučování o 10 % oproti stávajícímu internímu systému Cinematch. Jak ale sám záhy dodává, vnímal celý projekt jako dlouhodobou investici, která odstartovala technologický růst firmy a dlouhodobě se vyplácí. (TED, 2018)

Jiný zdroj uvádí, že za vítěze soutěže byl sice prohlášen takový algoritmus, který dokázal zlepšit přesnost doporučení minimálně o 10 %, avšak v praxi Netflix algoritmus nikdy neimplementoval. Pouze se inspiroval jeho výkonností a nastavením pro tvorbu vlastního algoritmu a pro tvorbu strategie (Bulygo, 2019).

Společnost Netflix měla v plánu v roce 2009 soutěž Netflix Prize zopakovat a vyhlásit druhé kolo, nicméně to už se nikdy nestalo. Skupina uživatelé Netflixu podala žalobu kvůli úniku dat a neoprávněnému nakládání s osobními údaji třetí stranou, neboť s nimi v rámci vývoje pracovali v roce 2007 také dva výzkumníci z University of Texas v Austenu, kteří byli schopni anonymním datům z Netflixu přiřadit konkrétní identity v databázi serveru IMDb. (Demerjian, 2007) Soutěž byla tedy z obav o další soudní spory předčasně ukončena a společnost Netflix se s žalobci domluvila na mimosoudním vyrovnání. (Singel, 2019)

3.2.2 Typy dat, které o uživatelích Netflix shromažďuje a jejich využití na příkladu House of Cards

Netflix o uživatelích nejčastěji shromažďuje právě tato big data (Insidebigdata, 2018):

- Hodnocení AV obsahu uživatelem
 - Vyhledávání AV obsahu uživatelem
 - Datum sledování AV obsahu uživatelem
 - Zařízení, na kterém je audiovizuální obsah sledován
 - Čas zastavení sledování AV obsahu
 - Změna povahy sledovaného AV obsahu v závislosti na zařízení
 - Opakované sledování AV obsahu
 - Přeskakování titulků

Tato nashromážděná big data Netflix využívá následujícím způsobem (Insidebigdata, 2018):

- Výroba audiovizuálních děl
- Personalizované doporučení AV obsahu ke sledování
- Top ranking AV obsahu, který vychází z personalizovaného doporučení

- Doporučení AV obsahu na základě trendů
- Pokračování ve sledování (po jeho přerušení)
- Vyhledání podobného AV obsahu, jaký uživatel sleduje

Při využití big data v rámci tvorby seriálu House of Cards se Netflix inspiroval příkladem Amazonu. Analýzu big data však Netflix více propracoval a snažil se identifikovat různé vzorce chování, které vykazují diváci při sledování audiovizuálního díla. Tento postup přinesl větší úspěch z hlediska výnosů, počtu shlédnutí i získaných ocenění. Ukázalo se, že big data mohou výrazně pomoci audiovizuálnímu průmyslu, ale je nutné provádět promyšlenou a systematickou analýzu i těch nejmenších detailů (Macnab, 2018).

Informace společnosti Netflix potvrzují (jak uvádí Marr, 2016, s. 18-19), že se realizace strategie založené na využití big data vyplácí. V prvním čtvrtletí roku 2015 služba získala 4,9 milionu nových odběratelů, a to právě díky neustále se zlepšujícímu obsahu služby. Akvizice nových zákazníků probíhala především na základě vlastní audiovizuální tvorby. Schopnost využití big data k předpovědi toho, co se bude divákům líbit, je tedy možné považovat za stěžejní podmínku úspěchu (Marr, 2016, s. 18-19).

Dle autora Marr (2016, s. 22-23) je další výzvou pro Netflix a ostatní streamingové služby využití big data k předpovědi toho, co budou chtít diváci sledovat. Výstupem by měla být situace, kdy bude mít každý zákazník na základě individuálních preferencí přístup k individualizované nabídce pořadů (Marr, 2016, s. 22-23).

DeSouza a Smith (2014) popisují, že Netflix každý měsíc shromažďuje uživatelská data z milionů hodin konzumovaného obsahu. Dochází k analýze titulů a žánrů, které zákazníci sledují, a to za účelem identifikace jejich preference. Netflix následně dle těchto preferencí průběžně aktualizuje svoji nabídku (DeSouza, Smith, 2014).

Jak uvádí DeSouza a Smith (2014), právě seriál House of Card využil kombinaci údajů o chování zákazníků a analytické metody k vytvoření příběhu v jednotlivých epizodách. Netflix do vývoje této technologie investoval 100 milionů dolarů za účelem pochopení preferencí svých 33 milionů zákazníků tehdejších předplatitelů. (DeSouza, Smith, 2014).

Havens (2014) potvrzuje, že využití big data vedlo ke zlepšení zážitku zákazníků a k lepšímu porozumění divákům prostřednictvím zacílení na jejich vkus. Obrovské množství údajů o divácích, které shromažďuje Netflix, včetně více než 50 druhů datových bodů o spotřebě každého filmu a televizního pořadu, poskytuje prostor pro efektivní předvídání toho, co diváci chtějí vidět (Havens, 2014).

Amazon při tvorbě Alpha House hodnotil big data o chování diváků v průběhu pilotního dílu, tj. kolik lidí díl sledovalo, jak díl hodnotili, jak dlouho jej sledovali, jaké části přeskočili apod. Netflix využil jiný přístup a prostřednictvím širší škály dat dokázal identifikovat, co se bude divákům skutečně líbit (Lazzaro, 2016).

Důvod, proč zejména výsledky na základě hodnocení uživatelů nejsou tak přesné, jako měření chování uživatelů, rozebírá CEO Netflix Reed Hastings. Uvádí zejména, že při hodnocení obsahu hraje roli představa o „ideálním já“, které se dívá pouze na společensky přijatelné a uznávané pořady. *„Toužíme být lepšími osobnostmi. Schindlerův seznam má na spoustě filmových databází hodnocení od uživatelů pět hvězdiček z pěti a hloupou komedii s Adamem Sandlerem pak daleko častěji označí veřejně dvěma nebo třemi hvězdičkami. Průzkum sledovanosti ale ukazuje, že diváci si daleko častěji a daleko radši pustí hůře hodnocený, ale přístupnější pořad. Uživatelské hodnocení preferencí jednotlivých filmů nebo seriálů tedy není dobrý předpoklad k tomu, jak sestavit nebo vytvořit program pro co nejvíce diváků ... Na rozdíl od Facebooku nebo Youtube má však Netflix stále ambici dělat*

obsah kurátorovaný a algoritmus pro doporučení používá pouze jako nástroj. Kvalita obsahu je podle Hastingsse důležitá z toho důvodu, že divák si spojuje obsah se značkou Netflixu, a kdyby byl obsah pouze zcela konzumního charakteru, divák bylo nebyl ochotný za službu platit předplatné. (TED, 2018)"

Scénáře House of Cards a celý vývoj se nezakládaly pouze dle přání producentů, režisérů či herců, ale byl primárně ovlivněn sběrem a vyhodnocením big data. Na počátku tvorby seriálu stál požadavek na využití dat o chování 33 milionů zákazníků Netflixu. Výstupy z analýzy se pak reflektují do rozhodnutí o obsahu seriálu, ale také do rozhodnutí o hercích a režisérech. Například bylo zjištěno, že velká část zákazníků je fanouškem filmů režiséra Davida Finchera a herce Kevina Spaceyho, takže bylo rozhodnuto o tom, že seriál bude režírovat tento režisér, a že hlavní roli obsadí Kevin Spacey. Big data také odhalila, že většina zákazníků sledovala původní seriál BBC, kterým se obsah House of Cards inspiruje, a tedy existuje potenciální poptávka po jeho výrobě. Big data také způsobila, že nebylo nutné investovat do pilotní epizody, na jejímž základě by se s odstupem času rozhodlo o natočení dalších dílů AV díla. Analýza big data prokázala, že existuje publikum, které bude seriál sledovat, a na které může Netflix zacílit, a proto byla ihned zahájena výroba 26 epizod (Smith, Telang, 2017).

Později však Michael Smith, který se na pozici profesora na Carnegie Mellon University věnuje analýze chování firem a spotřebitelů na trzích s digitálními informacemi a produkty digitálních médií, jsou informace o tom, že Kevina Spaceyho a Davida Finche můžeme považovat za populární tvůrce, nebo že byl seriál House of Cards na BBC úspěšným počinem, jednoduše dosledovatelnými fakty na základě preferencí diváků a jejich ohlasů. Smith tak věří, že využití big dat Netflixu neposkytlo zásadní konkurenční výhodu. (Smith, 2018).

Ať tak, či onak, odborníci se shodují, že v případě House of Cards, ale i v mnoha dalších případech, Netflix pracoval a pracuje s daty od vývoje až po distribuci, i když není možné úspěch či neúspěch těchto seriálů spojovat pouze s využitím big data, ale je nutné konstatovat, že zde rozhodují i další faktory jako například kreativní ztvárnění, marketing, herecké výkony, demografické rozdíly mezi diváky Netflixu a Amazon Prime, různorodý algoritmus doporučení programů atd. (Lamont, 2016).

Big data hrají podstatnou roli v marketingu audiovizuálních děl, kdy poskytují informace, které lze využít v rámci marketingové strategie, taktiky, či operativy. Například lze vyhodnotit, jaké zákaznické chování uživatel vykazuje, a dle toho mu nabídnout marketingové sdělení, resp. reklamu atd. (Johnson, 2016).

Autoři Smith a Telang (2017) ve své knize uvádějí, že Netflix vytvořil stovky propagačních materiálů jako jsou obrázky, grafiky, trailery apod. které pak přesně cílil na své uživatele podle jejich preferencí. Big data umožnila Netflixu poznat své zákazníky jako jednotlivce, což byla výrazná konkurenční výhoda. Jiné platformy mohou vycházet pouze ze souhrnných výzkumů či statistických informací. (Smith, Telang, 2017).

Analýza big data promlouvala i do toho, jaké scény jsou v seriálu House of Cards použité a jaký je děj seriálu. V úvodní scéně seriálu hlavní představitel uškrtní psa přejetého automobilem. Tato scéna je sama o sobě dost kontroverzní, a v tvůrčím týmu nedošlo ke shodě o tom, jestli ji použít. Netflix však využil analýzu reakcí testovacího vzorku diváků na tuto scénu a rozhodl o jejím zachování v rámci seriálu. Dále bylo díky analýze big dat zjištěno, že řada uživatelů preferuje audiovizuální obsah se silným ženským charakterem, a proto Netflix využil této informace při tvorbě výrazné ženské postavy (manželka hlavního hrdiny Franka Underwooda Claire) i v seriálu House of Cards (Smith, Telang, 2017).

Podle Smitha je hlavní důvod, proč bylo zařazení některých kontroverzních scén v případě House of Cards možné je to, že Netflix není na rozdíl od běžného TV vysílání financovaný z reklamy. Nemusel se tedy obávat toho, že by se dotkl svých inzerentů tím, že zařadí takto kontroverzní obsah do svého vysílání. Co je však podstatnější, Netflix se nemusí tolik obávat toho, že jeho divák jednoduše přepne na jinou televizní stanici, protože nabízí takové množství obsahu, že si divák jednoduše vybere jiný obsah z jeho databáze. Tím se pouze zpřesní to, co mu Netflix příště doporučí nebo nedoporučí, a ještě více tak zpřesní doporučovací systém (Smith, 2018).

Scénárista House of Cards Beau Willimon na panelové diskuzi při festivalu Aspen Ideas v roce 2014 k tématu kontroverzní scény v úvodu seriálu uvádí, že kreativita má u SVOD daleko větší prostor než v běžném televizním vysílání, kde si v praxi nikdo tolik nedovolí riskovat po obsahové stránce jakoukoliv invencí a ztratit tím zejména v prvních minutách zájem diváků (Couric, 2017).

I přes všechny zmíněné příklady, analýza a využití big data v audiovizuální tvorbě ještě nezaručuje, že se podaří vytvořit skutečně kvalitní audiovizuální dílo, které přiláká miliony diváků. Velmi dobře je to vidět právě na příkladu seriálů Alpha House a House of Cards. V obou případech došlo k masivnímu využití big data, ať už pro vyhodnocení poptávky po seriálu a velikosti cílového publika, či za účelem ovlivnění děje seriálu, aby co nejvíc odpovídal preferencím cílové skupiny. Alpha House společnosti Amazon Prime však nedosáhl takového úspěchu jako konkurenční House of Cards od společnosti Netflix. Za tímto rozdílem stojí skutečnost, že při tvorbě House of Cards došlo nejen k využití většího množství dat, ale také jiným způsobem. Amazon analyzoval menší množství dat a v podstatě se zaměřil pouze na hodnocení toho, jak zákazníci reagují na pilotní epizodu každého z osmi testovaných seriálů, a na komparaci reakce s ostatními pilotními epizodami jiných seriálů. Omezená velikost vzorku dat pak vedla k tomu, že Amazon nedokázal zcela

pochopit publikum a jeho preference, takže nevytvořil skutečně zajímavý děj pro větší cílovou skupinu (Dai, 2017).

Reed Hastings popisuje rovněž zavedení do této doby téměř nepoužívaného systému zveřejnění všech dílů ve stejný čas ke shlédnutí okamžitě (tzv. binge watching). Tato myšlenka podle něj vzešla právě ze zkušenosti z distribuce DVD, kdy se Hastings setkal se skutečností, že mnoho klientů rádo sleduje na DVD více filmů nebo dílů z jedné série po sobě. Na rozdíl od jiných převzatých seriálů si uživatelé Netflixu spojili House of Cards právě se službou Netflix, protože na ní byl tento seriál přímo navázaný, což přivedlo mnoho předplatitelů. Podle stejného klíče pak Netflix přišel s myšlenkou realizovat další pořady stejným způsobem. Jednalo se například o pořady Narcos, Orange is the new Black, Jessica Jones, The Crown, Black Mirror, Stranger Things. Z původní investice 100 milionů dolarů tedy v roce 2017 Netflix investoval ročně do všech svých vlastních produkcí okolo 8 miliard dolarů (TED, 2018).

Netflix od doby vydání House of Cards vyrobil přes 1600 pořadů z celkového počtu 15400 titulů. Zdaleka ne všechny jsou však dostupné ve všech zemích (What's on Netflix, 2020).

Na velký úspěch House of Cards Netflix navázal dalšími seriály, které dosahují podobných výsledků, a kromě výše zmíněných to byly seriály Bojack Horseman, The Unbreakable Kimmy Schmidt, Master of None, či Stranger Things, čímž se plnohodnotně zařadil mezi špičkové tvůrce audiovizuálního obsahu jako jsou společnosti HBO, Showtime, FX (Fritz, 2018).

Z růstu celé platformy z hlediska množství konzumovaného obsahu, ročních výnosů společnosti, počtu uživatelů a množství originálního obsahu, lze předpokládat, že metody, které využil Netflix na úspěšném projektu House of Cards využívá hojně i v jiných případech.

Eddy Wu (2019), ředitel oddělení vývoje Netflixu, uvádí asi jeden z nejznámějších případů rozvoje technologií VOD platformou současné doby, a sice interaktivní pořad Bendersnatch.

Tento krátký film, který tematicky navazuje na úspěšný seriál Black Mirror, odstartoval rodící se divizi Netflix Interactive. V rámci jeho děje je uživatel daleko víc vtažen do děje, protože má v klíčových momentech seriálu možnost ovlivnit, jaká s příběhových linek se bude dále rozvíjet, a to rozhodnutím ze dvou nabízených variant. Podle Eddyho Wu (2019) bude v následujících letech čím dál významnější snaha rozvíjet nové druhy zábavy, a interaktivní prvky u VOD platformou jsou jedním z nich (Wu, 2019).

Netflix na svých stránkách uvádí seznam, dalších originálních i převzatých interaktivních pořadů s informací, že další jsou připravované (Netflix, Inc., 2020):

- Puss in Book: Trapped in an Epic Tale
- Buddy Thunderstruck: The Maybe Pile
- Stretch Armstrong: The Breakout
- Minecraft: Story Mode
- Tvůj souboj s divočinou s Bearem Gryllsem
- Velká volba Kapitána Bombardáka
- Carmen Sandiego: Krást či nekrást
- Nezdolná Kimmy Schmidt: Kimmy versus reverend

Eddy Wu (2019) vidí inovaci Netflixu především ve zlepšování uživatelského zážitku. Podle něj je snaha zlepšovat s rozvojem technologií, a především kvality internetového připojení i streamingovou kvalitu obsahu, zlepšovat personalizační algoritmus a učit zákazníky, jak mají produkt využívat naplno. Zároveň uvádí rizika a slepé cesty, kterými Netflix za dobu své existence prošel. Ukázalo se například, že uživatelé mají sice zájem vědět, co jaké pořady sledují na Netflixu jejich známí a přátelé, ale sami už nejsou ochotní

o sobě totéž veřejně sdílet. Wu popisuje, že v mnoha uživatelích převládá pocit viny za druh obsahu a množství času věnovanému zábavě, a proto tyto údaje Netflix neumožňuje v rámci svojí služby sdílet, tak jako je tomu například u služby s hudebním obsahem Spotify.

Wu (2019), také uvádí snahu zavést lineární vysílání a napodobit tak klasickou TV stanici (tak, jako je to třeba u české Televize Seznam). *„Funguje systém toho, že služba nabídne rovnou další položku ze seznamu, ale když byl program pevně stanoven dopředu, tak mnoho lidí odešlo, protože jim sestavený obsah nevyhovoval nebo pro ně nebylo pohodlné sledovat ho ve stanoveném čase.“*

3.3 Mall.TV

Mall Group, která od roku 2018 provozuje VoD platformu Mall.TV, vznikla v roce 2015 a v současné době působí kromě České republiky také na Slovensku, v Polsku, Maďarsku, Slovinsku, Chorvatsku a Rumunsku.

Mall Group vlastní ze 40 % skupina PPF (Petr Kellner), ze 40 % EC Investments (Daniel Křetínský a Patrik Tkáč) a Rockaway (Jakub Haverlant), který s 20% podílem vykonává manažerskou kontrolu. (PPF, 2020)

U zrodu Mall.TV stál tým lidí, který v čele s šéfproducentem Lukášem Záhořem, kreativním producentem Milanem Kuchynkou a šéf dramaturgem Martinem Krušinou přešel z platformy Stream.cz (Seznam). Po nich následovali jak někteří tvůrci, tak i jádro původního týmu, které se stará o produkční, technické i administrativní zajištění platformy. (Šiler, 2020)

„Vyložil jsem si to tak, že Seznamu, respektive novému obsahovému řediteli, začalo jít už jenom o jednoduchou matematiku vytrženou z kontextu. To znamená o peníze nebo o ekonomickou rozvahu nad každým jednotlivým pořadem. A to byl asi ten základní problém, protože pro mě byl Stream jako organismus, v rámci kterého jsou části, které vydělávají peníze, ale i části,

„které ty peníze spíš spotřebovávají a vydělávají v nějakém dlouhodobém horizontu, protože budují značku,“ uvedl Lukáš Záhoř v rozhovoru pro server Lupa.cz v březnu 2018 (Slížek, 2018).

Před vybudováním vlastního webu pro distribuci zvažovalo vedení nově vzniklé Mall.TV jako distribuční kanál rovněž YouTube. Od tohoto nápadu záhy sešlo, neboť YouTube není dostatečně flexibilní a neumožňuje požadované funkce a nároky z hlediska kategorizace, vizuální podoby, ani možností umístit a měřit vlastní reklamní formáty. Webový portál Mall.TV tedy nakonec realizovala kosovská společnost Gjirafa, která se stala předmětem investice společností spoluvlastnicích Mall Group.

„Další interní vývoj této platformy (Gjirafa) by měl vést k možnosti nabídnout ostatním mediálním společnostem plně konkurenceschopné a komplexní state-of-the-art technologické řešení. Z něho budou profitovat jak čtenáři, kteří získají atraktivnější, částečně personalizovaný obsah, tak i inzerenti, kteří poptávají reklamní zásah konkrétních cílových skupin. Růst čtenosti a lepší cílení reklamy by mělo vést k vyšším inzertním příjmům vydavatelských domů.“ uvedlo tiskové oddělení společnosti Czech Media Invest (Daniel Křetínský), která do Gjirafy investovala v rozhovoru pro server Lupa.cz (Sedlák, 2020).

Uživatelé obsah v současné době nekonzumují však pouze na webových stránkách Mall.TV, ale podobně jako u většiny jiných velkých provozovatelů mají celou škálu možností. Mezi další distribuční kanály patří mobilní aplikace pro iPhone nebo Android, aplikace ve SmartTV a aplikace na konzoli Xbox. Ze všech těchto kanálů může Mall čerpat uživatelská data (Šiler, 2020).

Snahou Mall.TV bylo stavět programovou skladbu na třech pilířích, kterými jsou seriálová tvorba, publicistika a infotainment¹⁴, a produktová videa (např. nákupní rádci). „Pořady a seriály slouží zábavě, produktová videa jsou servisní složka, která je rozprostřena v rámci aktivit Mall Group,“ uvedl Lukáš Záhoř v roce 2018 pro MediaGuru (Vojtěchovská, 2018).

Na rozdíl od etablovaného Streamu a potažmo TelevizeSeznam.cz, které měly dlouhodobě výsadní pozici v podobě zobrazování videí na hlavní stránce Seznam.cz, se Mall.TV v počátcích potýkala s přirozeným nedostatkem diváků a stále dosahuje menších čísel. Sledovanost Mall.TV se s postupem času a s počtem rostoucích pořadů v knihovně však postupně zvyšuje (Šiler, 2020).

„V případě nejmladší služby na trhu Mall.TV je sledovanost rovnoměrně rozložená mezi slabší i silné televizní diváky. Je zde velký poměr diváků z nejmladší cílové skupiny do 24 let, ale i starší skupiny uživatelů. Sledování je nezávislé na velikosti místa bydliště.“ (Vojtěchovská, 2019)

Podle Jakuba Šilera, analytika digitálního obsahu, je práce s daty při budování VoD platformy (jakou je Mall.TV) zásadní. Data pomáhají řadit tvorbu do jednotlivých clusterů podle diváckého profilu, pomáhají v marketingu pořadů (cílení a doporučování) a nově po integraci, respektive posílení spolupráce s vydavatelstvím Czech News Center, budou sloužit k lepšímu využití potenciálu v mediálním zásahu prostřednictvím cross-promování obsahu (Šiler, 2020).

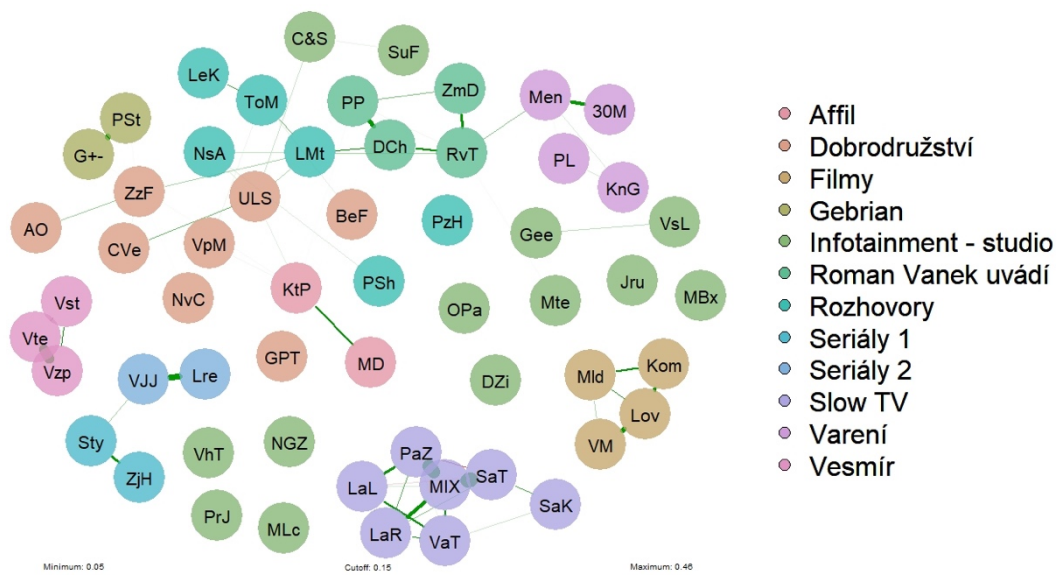
¹⁴ Infotainment je druh zpravodajství, které podřizuje výběr témat a jejich zpracování účelu vyvolat emoce a pobavit. (Zdroj: <https://cs.wikipedia.org/wiki/Infotainment>)

Mall.TV tímto krokem poskytuje CNC možnosti pokročilého cílení a práci s velkými daty pod vedením výkonného ředitele Mall.TV Juraje Felixe (Sedlák, 2020).

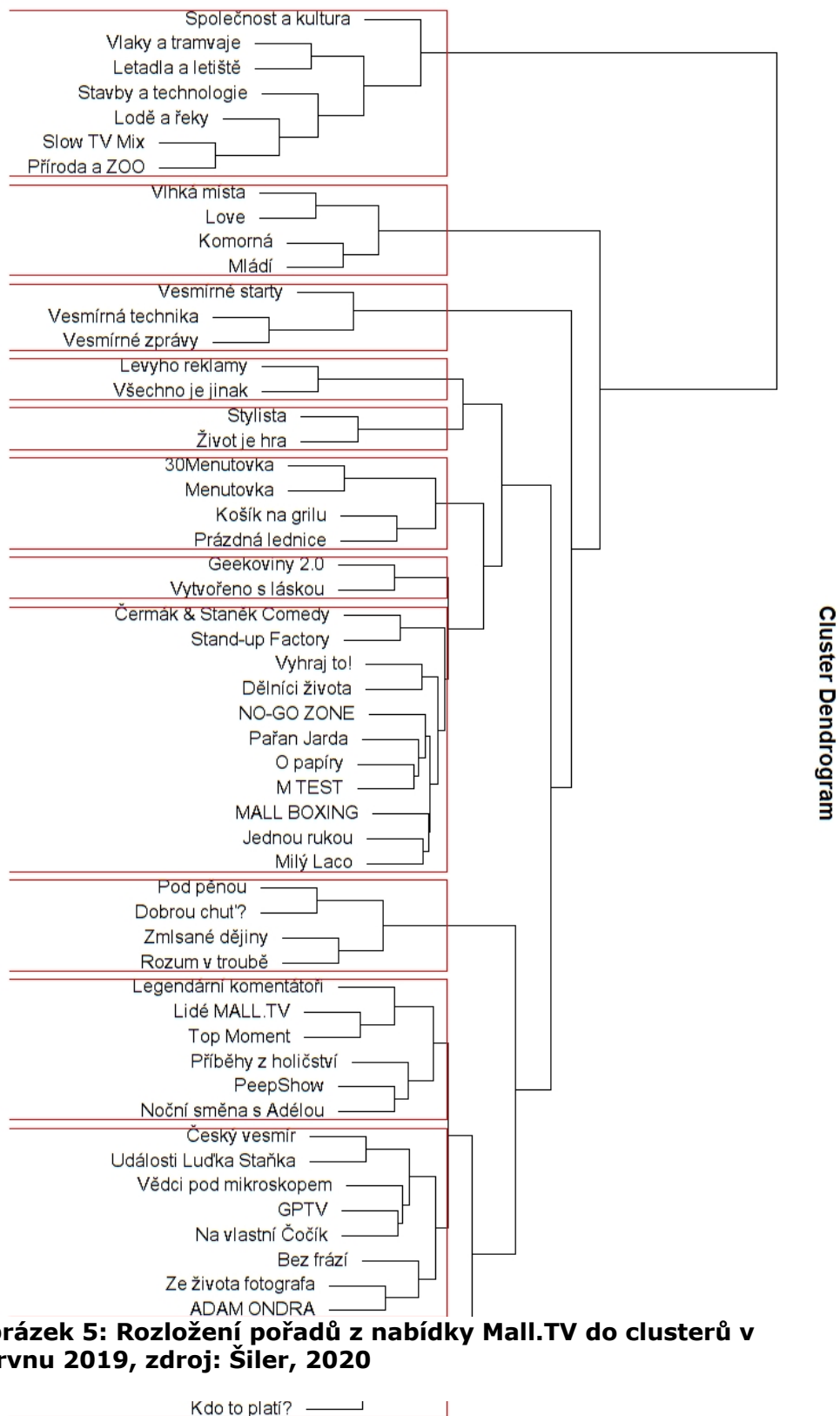
Velmi důležitou analytickou metodou, jak již bylo zmíněno, je podle Šilera právě práce s clustery. Ty řadí diváky do různých skupin podle jejich zájmu. Vazby mezi pořady vznikají tak, že přibližně 50 % diváků během sledování obsahu na Mall.TV projeví zájem o jiný pořad v nabídce. Zájem je převeden do datové podoby a tu je pak datový analytik schopen vyhodnotit za použití správného nástroje (např. software RStudio) a vazbu vizualizovat. (Šiler, 2020)

Pokud mají pořady správně přidělené štítky, které jim datový analytik přiřadí, utváří se z nich podle zadaných kritérií skupiny. V případě Mall.TV jsou pořady rozděleny do skupin podle:

- osobnosti tvůrce (např. Gebrian, kam spadají pořady: Gebrian PLUS/MINUS nebo Překvapivé stavby);
- obsahového zaměření (např. Vesmír, kam spadají pořady: Vesmírné starty, Vesmírná technika, Vesmírné zprávy a Český vesmír);
- formátu (např. Filmy nebo Slow TV);
- atp.



Obrázek 4: Mapa zájmové síly diváckých vazeb mezi pořady v jednotlivých kategoriích nebo žánrech z června 2019, zdroj: Šiler, 2020



Obrázek 5: Rozložení pořadů z nabídky Mall.TV do clusterů v červnu 2019, zdroj: Šiler, 2020

Provozovatelům platformy pomáhají tato data zejména k rozvaze o budoucím vlastním nebo převzatém obsahu, který by doplnil knihovnu platformy a k pochopení vazby uživatelů na jednotlivé pořady.

Podle Šilera (2020) se ukazuje, že se z hlediska dat o sledovanosti a věrnosti diváků spíše nevyplácí vyrábět nebo akvírovat obsah k distribuci, který by byl úplně mimo zavedené struktury Mapy zájmových sil (viz Obrázek 4). Takové rozhodnutí musí být podle Šilera opřené buď o:

- snahu přilákat novou diváckou skupinu (respektive rozšířit tu stávající);
- nějakou konkrétní situaci způsobenou vnějším vlivem, která vyvolá organický zájem;
- vyšší záměr a přinést zajímavé a důležité kurátorované téma nebo producentský vzhled.

Příkladem snahy přilákat diváckou skupinu mimo zavedené struktury byly například pořady s cestovatelskou tematikou, kdy se povedlo spojit pořady do jedné kategorie označené jako „Cestování & Dobrodružství“, a dát jim tak z diváckého úhlu pohledu jasnější tvář v rámci kampaně na jejich sledování, i v rámci organického dosahu.

U druhého případu, jakým je reakce na koronavirovou krizi v podobě záznamů divadelních a hudebních vystoupení s názvem #kulturažije, však Šiler uvádí, že bez kontinuálního budování vazeb na další pořady z knihovny Mall.TV hrozí, že diváci platformu opět v krátkém čase opustí. Jedná se tak v lepším případě zejména o posílení povědomí o značce.

Posledním příkladem snahy o obsah v tomto případě s edukativní obsahem byla minisérie #martyisdead. Tato minisérie s původními clustery příliš nesouvisela, ale otevírala téma kyberšikany u mládeže a pomohla přilákat

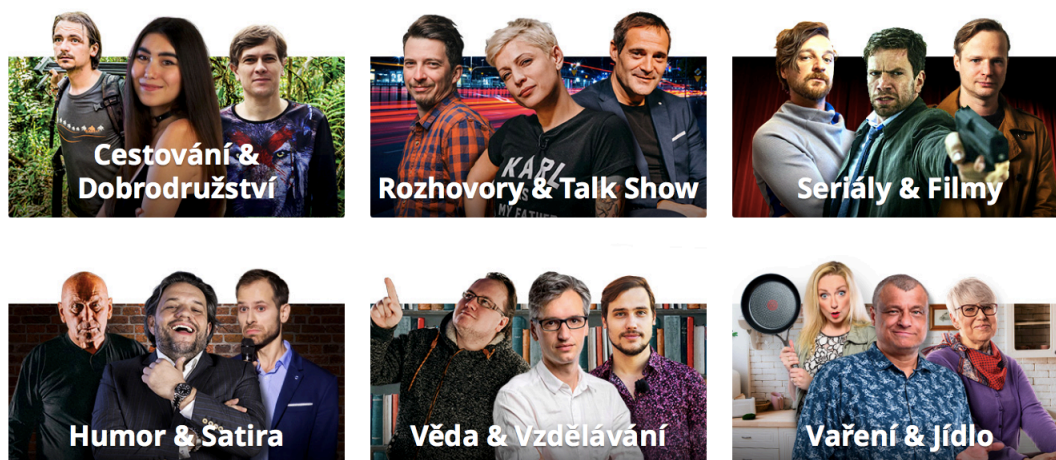
novou cílovou skupinu i na další bonusový obsah: Deník predátora, Jak na internet a další vzdělávací videa. (Šiler, 2020)

Ukazuje se, že práce s daty má přímý dopad na sledovanost pořadů. Snaha redakce o ruční zařazení konkrétního pořadu do systému automatizovaného doporučování navzdory algoritmu za účelem jednorázového navýšení počtu shlédnutí mívá často dlouhodobě opačný efekt v podobě odlivu diváků. (Šiler, 2020)

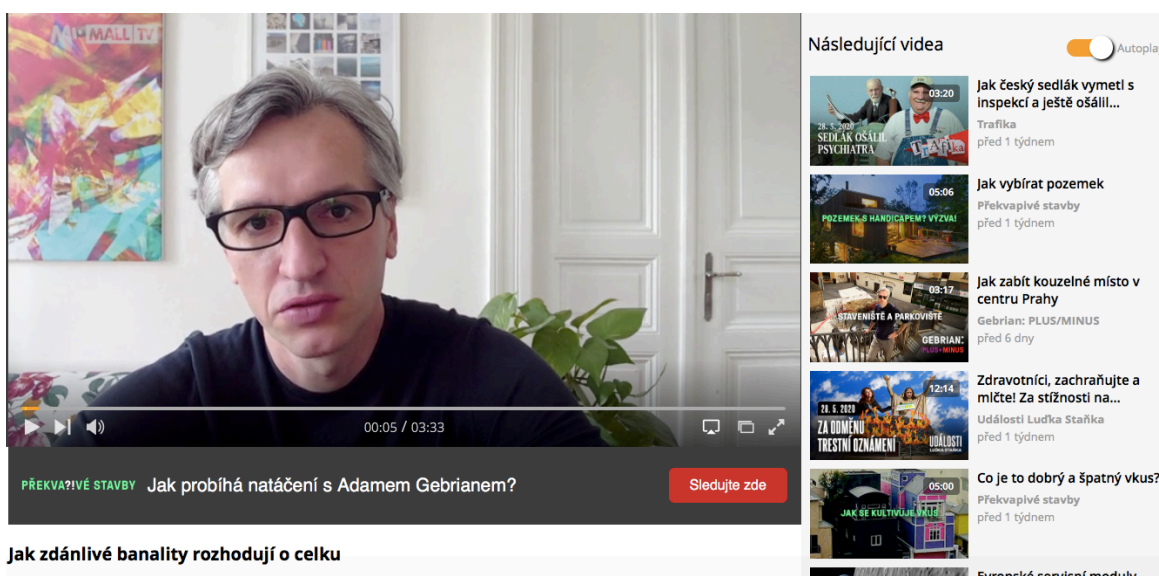
Doporučovací algoritmy mají v případě Mall.TV na starosti zejména:

- doporučování obsahu uživatelům v pravém postranním sloupci (viz. Obrázek 6);
- řazení obsahu do přehledných skupin podle zájmů a preferencí uživatelů (viz. Obrázek 7);

Hlavní kategorie



Obrázek 6: Uživatelské zobrazení kategorií na webu www.mall.tv, zdroj: Mall.TV, 2020



Obrázek 7: Systém doporučení u konkrétního pořadu v pravé sekci Následující videa na webu www.mall.tv, Zdroj: Mall.TV, 2020

U nových pořadů pak hrají roli analýza nasycenosti trhu, ze které v případě Mall.TV vzešlo, že je prostor i pro velmi tematicky okrajové pořady v kategorii Vesmír. V tomto ohledu Jakub Šiler nepopírá, že data slouží jako pomocný nástroj, který doplňuje intuici producentů. (Šiler, 2020)

Skladba příjmů Mall.TV, kterou tvoří zejména reklama, ovlivňuje požadavky na práci s daty. Kromě product placementu na produkty z e-shopů Mall Group, mezi které patří například CZC.cz, Košík.cz, Mall.cz, Vivantis.cz a další, poskytuje Mall.TV nabídku inzerce i externím klientům. Mezi hlavní benefity, které je Mall.TV schopné nabídnout, jsou data o historii chování zákazníků, historii nákupů zákazníků a chování diváků Mall.TV. Všechna tato data shromažďuje v nástroji Data Management Platform. (MallTV, 2019)

Jakub Šiler dále uvádí, že existuje snaha monitorovat a vyhodnocovat téměř veškerou uživatelskou aktivitu na stránce Mall.TV. Analytické nástroje měří na pozadí zejména:

- množství přehraného obsahu;

- časy spuštění a časy ukončení přehrávání;
- kliknutí na další obsah v rámci webu;
- kliknutí na reklamu nebo akční prvek, který vybízí k přejítí do e-shopu;
- pohyb po jednotlivých podstránkách;
- uživatelské hodnocení (like/dislike);

Všechny tyto údaje jsou průběžně vyhodnocovány a slouží ke zlepšování, respektive zvyšování výkonu služby. (Šiler, 2020)

V roce 2020 kdy se ukazuje, že mnohé české VoD platformy zanikají nebo brzdí tvorbu svého obsahu je Mall.TV jednou z mála, která dále produkuje vlastní obsah. *„Mall.tv postavená na provázanosti se skupinou e-shopů Mall Group jako víceméně jediná stále pokračuje v tvorbě vlastních seriálů, čímž navazuje na dlouholeté působení Lukáše Záhoře a jeho týmu na Streamu.“* (Jetmar, 2020)

3.4 Specifika videoplatformy ČT jako VOD platformy veřejné služby

Česká televize provozuje platformu pro odložené přehrávání iVysílání v plné míře od roku 2011. Dostupná byla v testovací podobě už koncem roku 2008, nicméně plného provozu a současné podoby se jí dostalo až o tři roky později. V letech 2010 a 2011 byl vybudován nový útvar, který měl za cíl dostat obsah ČT na nové platformy, zejména na internet a do mobilních telefonů, pověřen Milan Fridrich, který uvedl: *„Nové weby by měly ukázat ČT jako veřejnou službu, která umí dělat atraktivní pořady i pro internetového diváka. Nechceme jen hezčí weby, smyslem je modernizovat a prohloubit online komunikaci televize a diváků.“* (Bém, 2011)

Martina Vojtěchovská (2019) v analýze audiovizuálních služeb na vyžádání na serveru Mediaguru o téměř desetiletí později uvádí na základě studie OMG Research: *„Přestože je služba iVysílání provozována Českou televizí, mezi jejími uživateli jsou spíše slabší televizní diváci. Tedy ti, kteří sledují televizi*

jednou týdně anebo i méně. Průzkum eviduje u iVysílání velký poměr diváků z cílové skupiny 35-54 let a nejvyšší sledovanost ze středně velkých obcí." (Vojtěchovská, 2019) Alespoň z hlediska záměru cílit na ne zcela běžného televizního diváka, ale rozšířit diváckou skupinu o diváka internetového byla prognóza Milana Fridicha tedy úspěšná.

iVysílání ČT podle výzkumu společnosti Omnicon Media Group dosahuje výsledků sledovanosti týdně průměrně 1,26 mil. lidí. (Vojtěchovská, 2019)

Podle společnosti SPIR NetMonitor jsou týdenní výsledky návštěvnosti webu iVysílání ale daleko menší. V květnu 2020 byl například týdenní výsledek návštěvnosti 166 tisíc unikátních uživatelů a počet návštěv stránek pak 448 tisíc. (NetMonitor, 2020)

Samotná Česká televize uvádí v jednom ze svých reportů, že *„oproti roku 2018 se zvýšil počet průměrných měsíčních unikátních návštěvníků jakékoli z našich služeb o 225 000 uživatelů."* a dále zmiňuje, že průměrný počet unikátních uživatelů webových stránek za měsíc iVysílání byl v uplynulém roce 624 000. (Kollárová, 2019)

Česká televize v současné době negeneruje ve větší míře zvláštní obsah, který by byl určen přímo pro VOD portál iVysílání, přesto archiv této služby čítá již 94 000 hodin obsahu. Toto číslo roste s každou vteřinou o průměrně 2-4 vteřiny nového obsahu, mimo jiné i proto, že se zde ukládá kromě archivních pořadů také běžné lineární televizní vysílání v reálném čase. (Česká televize, 2020)

Česká televize – jako médium veřejné služby – si dala za cíl po vzoru jiných veřejnoprávních televizních médií přinášet divákům rovněž zvláštní obsah, který je určen pouze pro VOD platformu a nevejde se do vysílacího času televizních kanálů. Pro tento účel a s vidinou zlepšení stávající služby ustanovila Česká televize tým, který pracuje na vzniku nové videoplatformy,

jež by více odpovídala světovým standardům a uživatelským požadavkům. Ve zprávě ČT z roku 2019 o činnosti úseku Nová média, pod který spadá i videoplatforma, je uvedeno, že: „*Nejdůležitějším momentem celého roku s ohledem na zachování a udržitelný rozvoj nelineární nabídky ČT v budoucnosti bylo přijetí Iniciační studie projektu Nové videoplatformy ČT a následné zahájení realizace projektu. Vstupní analýza a definice celého projektu vznikly v úzké spolupráci s celou řadou oddělení napříč ČT.*“ a dále pak „*Cílem oddělení je realizovat a provozovat Videoplatformu ČT - nový produkt, který v roce 2021 zcela nahradí současný online videoarchiv "ivysilani.cz" a stránky pořadů na webu...*“

...První internetové vysílání spustila ČT v roce 2000. Pět let předtím než vznikl YouTube, jsme již z Kavčích hor streamovali online několik, zejména zpravodajských, pořadů.“ (Kollárová, 2019)

Doplňkový obsah už v současné době vzniká, ale zejména v podobě textových příspěvků, fotografií z natáčení, online přenosů, rozhovorů s tvůrci a podobně. „*Kromě zajištění technologií a provozu webových služeb zajišťují Nová média celou řadu obsahových aktivit na podporu vysílání nebo odložené sledovanosti pořadů ČT. Oddělení Podpora systémů a obsahu NM také zajišťuje stovky online přenosů, které rozšiřují obsahovou nabídku nelineárních služeb ČT; spravuje také mobilní aplikace.*“ (Kollárová, 2019)

V závěru zprávy uvádí Kollárová (2019), že: „*Za úspěchem stojí zejména angažovanost kmenových pracovníků ČT, kteří v rámci vlastních rolí a svého profesního kreditu dokázali přesvědčit a nadchnout pro myšlenky a cíle projektu své kolegy z oboru a také intenzivní podpora ze strany oddělení Lidských zdrojů České televize. Souběžně s nábořem jsme zefektivnili řadu procesů spojených s příchodem nových pracovníků, a to i díky spolupráci s útvarem Informatiky.*“

Právě neveřejná *Iniciační studie projektu Nové videoplatformy ČT*, která vznikla v roce 2019 dle zadání managementu televize, je výchozím materiálem, ve kterém tým videoplatformy zkoumal způsoby, jak celý projekt realizovat. Na studii se podílel dílčím způsobem rovněž výzkumný tým studie *Strategie pro udržitelnost a rozvoj televize veřejné služby v éře internetu: příklady dobré praxe na základě mezinárodního srovnání* v čele s Petrem Szczepanikiem. Tento tým při Univerzitě Karlově, čítající mimo jiné sociologa, etnografa a datového analytika obsahu, byl podpořen v roce 2020 v Programu ÉTA Technologické agentury ČR.

Podle Josefa Zukala a Doroty Vašíčkové (2020) z týmu Nové videoplatformy, je snaha, aby spolupráce s výzkumným týmem nadále fungovala, byť na sobě platforma ani zmíněná externí studie nejsou přímo závislé. (Vašíčková & Zukal, 2020)

Zukal a Vašíčková dále uvádějí, že cílem Nové videoplatformy je nahradit současné iVysílání, které neodpovídá standardům a požadavkům na současnou VOD platformu. Obsah je sice možné dále přidávat do současného systému, ale snahou Nové videoplatformy má být budování principu personalizace, která umožní nabízet videoobsah na základě preferencí jednotlivých uživatelů. Zcela zásadní má však být role aktivní dramaturgie, tedy propojování pořadů ve větší celky a jejich strukturování podle témat, žánrů a dalších vlastností, a to jak u stávajícího archivního obsahu, tak i u nových tzv. online only¹⁵ pořadů. *„V rámci vývoje nových pořadů budou hrát data roli, ale nechceme se spoléhat pouze na algoritmus. Pro zachování smyslu ČT jako veřejnoprávního média hraje klíčovou roli redakční systém. Jeho záměr však není diktovat divákům na jaký obsah se mají dívat, pouze hlídat, zda je obsah vyvážený a pro všechny divácké cílové skupiny, ať už ty*

¹⁵ Obsah, který vzniká pouze pro distribuci na videoplatformě a není určen nebo nepochází původně z běžného televizního lineárního vysílání.

stávající nebo ty potenciálně nové. Zároveň má zajistit pestrost skladby obsahu." (Vašíčková & Zukal, 2020)

Množství online only obsahu, který bude na Nové videoplatformě k dispozici není podle Vašíčkové a Zukala (2020) v současné době známo.

V současné době je na webu ČT pro tvůrce vyvěšena výzva k podávání námětů¹⁶, ve které je uvedena možnost zasílat náměty i pro webovou platformu, nejedná se ovšem o výzvu přímo spojenou s Novou videoplatformou. Náměty by měla v budoucnu schvalovat programová rada. V příloženém dokumentu *Základní principy vývoje originálního online obsahu v ČT*, jenž doposud vypracoval právě tým videoplatformy ČT, je zmíněno pět základních bodů, které má originální obsah splňovat. Kromě zaměření na nová menšinová publika a snahu o autenticitu obsahu, je jedním z principů také otevřenost novým formám (Česká televize, 2020).

Zukal (2020) uvádí, že v případě menšinového publika nejde o okrajovost, ale o zmnožení perspektiv při zachování ambice oslovit v součtu co nejširší spektrum publika.

Podle Vašíčkové (2020) by měla platforma sloužit k distribuci obsahu, který se nevejde do lineárního vysílání žádného ze stávajících kanálů ČT, a to nejen po obsahové stránce, ale také z hlediska technického provedení formátu, či stopáže. *„Obsah, který bude vznikat pouze pro účely videoplatformy bude ještě dlouhou dobu z hlediska počtu pořadů nebo celkového objemu stopáže minoritní, neboť je v archivu uloženo obrovské množství obsahu převzatého z on-air vysílání a další stále přibývá. O to více chceme, aby se online obsah odlišil formou, sdělením, zaměřením a autenticitou. (Vašíčková & Zukal, 2020)“*

¹⁶ <https://www.ceskatelevize.cz/vse-o-ct/podavani-nametu-a-projektu/pro-autory/>

Celou kostru i konečné technické provedení Nové videoplatformy si zajišťuje Česká televize interně. Za vznikem platformy stojí tým mnoha interních zaměstnanců včetně vývojářů. Vašíčková a Zukal poukazují na to, že webová platforma je ze své podstaty velmi nestálý produkt, který potřebuje neustálé úpravy na základě vývoje trhu a technických inovací (Vašíčková & Zukal, 2020).

„Technologicky je videoplatforma rozložena do malých, vzájemně nezávislých služeb, jež jsou snáze popsatelné, mají jednodušší vývoj, lze je samostatně testovat a kdykoliv vyměnit nebo vylepšit část, která je již zastaralá nebo nevyhovující, a to bez dopadů na celý ekosystém. Nová videoplatforma bude dostupná také jako aplikace pro mobily a tablety.“ (Rožánek, 2020)

Členové týmu videoplatformy Zukal a Vašíčková (2020) připouští, že se s uživatelskými daty pracovat bude. Ne všechny formy využití jsou však v současné době známé. Josef Zukal dodává, že jejich primární využití by nemělo sloužit k navyšování sledovanosti na konkrétních pořadech, protože tím by se jen podporovala divákova uzavřenost do konkrétního typu pořadu nebo zábavy. *„Dosáhnout vysokého zásah počtu diváků je přirozenou ambicí. My však věříme, že to není jediné kritérium a důležitější je samotný nabízený obsah. Každý divák si zároveň bude moci vypnout personalizaci a doporučení, respektive se do služby nepřihlašovat a zabránit sběru dat“* (Vašíčková & Zukal, 2020).

Hlavní obavou je podle Zukala to, zda se povede Nové videoplatformě zajistit skutečně vyvážený zásah napříč všemi socio-demografickými skupinami včetně minorit. I přes to, že existuje dlouhodobý trend, kdy čím dál více diváků běžného vysílání konzumuje obsah *on demand* (viz. 1.1), s rostoucím množstvím cílových skupin musí být webové rozhraní a obsah platformy univerzálnější, aby byl přístupný všem. Exponenciálně v důsledku toho tak

rostou požadavky na kvalitu poskytované služby a množství distribuovaného obsahu. (Vašíčková & Zukal, 2020)

Zukal nevidí přímou konkurenci v podobě ostatních konkrétních pořadů VOD platforem působících na českém trhu. „Z našeho úhlu pohledu je zde omezený počet diváků a ti mají omezený čas, který věnují konzumaci (video) obsahu na internetu. Cílem ČT by neměla být primární snaha bojovat o tento prostor a navyšovat neustále čas, který divák stráví konzumací. Primární snaha by měla být o to, aby měl divák jednoduše k dispozici obsah, který bude svou kvalitou, zpracováním a volbou témat představovat atraktivní alternativu komerční nabídce.“ (Vašíčková & Zukal, 2020)

Podle datového analytika obsahu Jakuba Šilera je velký potenciál Nové videoplatformy právě ve využití uživatelských dat, se kterými by se mohlo po technické přestavbě pracovat v daleko větší míře. Na rozdíl od Mall.TV, která má podle něj v porovnání s ČT daleko omezenější prostředky na výrobu obsahu, může být ČT velmi výrazným hráčem na trhu VOD služeb v České republice. Rozhodující je zda správně využije svůj potenciál v podobě objemu prostředků a v současné době na trhu dostupných technologických možností. (Šiler, 2020)

Příkladem stávající fungující veřejnoprávní VOD služby je platforma francouzsko-německé TV ARTE dostupná na webu Arte.TV. Ta produkuje nebo koprodukuje vlastní obsah se zaměřením na kulturu, technologie, záznamy a přenosy živého umění a zpravodajství.

Podle Régine Hatchondo, ředitelky ARTE France, dosahuje Arte až 20 % všech zásahů skrz VOD služby. Síla digitálních médií podle ní neustále roste a to především u cílové skupiny do 35 let. Kromě vlastních webových stránek a aplikací umísťuje Arte obsah i na sociální sítě. Hatchondo dále zmiňuje vznik nové společné francouzské VOD platformy Salto, která má být zprovozněna na podzim roku 2020 a konkurovat velkým komerčním hráčům jako jsou

Netflix, Amazon Prime Video, Apple TV+ a Disney+ a jejich tlaku na trh. (Dredge, 2019)

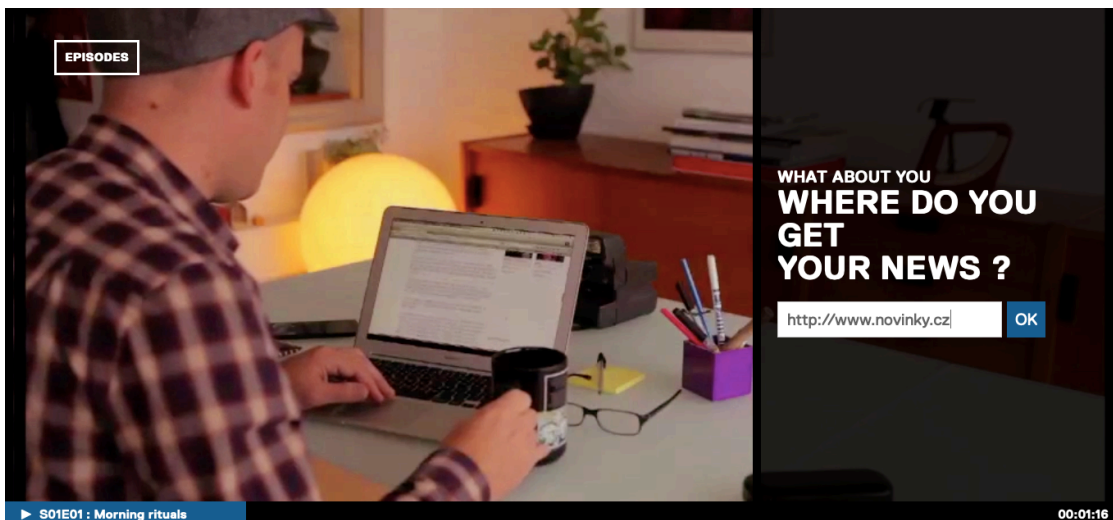
Kay Meseberg (2019) z oddělení nových médií TV ARTE, který se zaměřuje na vývoj, výrobu a distribuci nových pořadů v kategorii Virtuální realita a 360° video obsah zmiňuje, že data o uživateli sbírá Arte.TV v omezené míře a především za účelem lepšího doporučení obsahu během aktivity na stránce. Snahou je spíše budovat technologická řešení pro zajímavější a interaktivnější obsah a pro lepší a pohodlnější konzumaci obsahu pro diváky (Mesenberg, 2019).

Arte analyzuje práci s využitím dat také v online dokumentární minisérii „Do Not Track“¹⁷ z roku 2015, která vznikla v koprodukcí ARTE France, francouzské společnosti Upian, kanadské National Film Board's (NFB) a německé Bayerischer Rundfunk (BR). V každé epizodě divák sleduje nejen natočený materiál, ale zároveň svojí aktivitou spouští interaktivní prvky, které se přímo tematicky vztahují k problematice sběru uživatelských dat na internetu. Divák má tak mimo jiné možnost vyplnit, na jaké webové stránce nejčastěji tráví čas, nebo kde se v poslední době zaregistroval, a seriál mu v reálném čase zobrazí, komu všemu tím divák poskytnul o sobě data. Zároveň je zde kromě principu sběru dat vysvětleno i to, že data ve většině případů neslouží k identifikaci konkrétních osob, ale k predikci spotřebitelského chování, čímž vzniká riziko cyklení poptávky a nabídky u jakéhokoliv produktu nebo služby. (Gaylor, 2015)

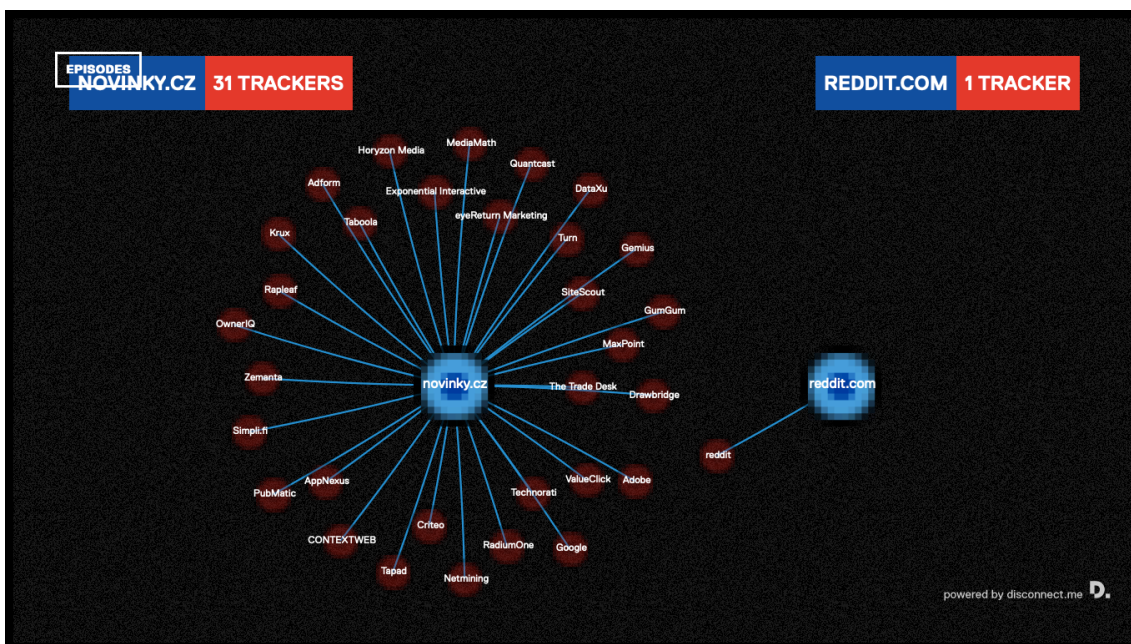
Samotní tvůrci seriálu transparentně uvádějí, jaké služby a nástroje k monitorování používají a vyzývají diváka k tomu, aby se zamyslel, jak

¹⁷ <https://donottrack-doc.com/en/intro/>

mohou být data o něm zneužita velkými společnostmi na trhu za účelem cílení obsahu, prodeje reklamy atp. (Upian, 2020).



Obrázek 8: Ukázka uživatelských zásahů do děje a použití interaktivních prvků v online dokumentární sérii Do Not Track, zdroj: Upian, 2020



Obrázek 9: Ukázka interaktivních prvků vyhodnocovaných v reálném čase na základě uživatelských dat v sérii Do Not Track, zdroj: Upian, 2020

Mimo TV ARTE má svojí VOD platformu celá řada nejen evropských veřejnoprávních médií. Mezi další pozitivní příklady, které zmiňuje i studie *Strategie pro udržitelnost a rozvoj televize veřejné služby v éře internetu: příklady dobré praxe na základě mezinárodního srovnání* patří i VOD portál BBC, belgická RTBF.be, norská NRK TV, nebo švédská SVT Play (Vašíčková & Zukal, 2020).

3.5 Další příklady

S daty pracuje v nějaké míře celá řada VOD poskytovatelů služeb. Organizace pracující efektivně s big daty jsou zpravidla výkonnější než jejich konkurence. Nejvíce pak k výkonnosti přispívá využití big data pro zkvalitnění vztahů se zákazníky (Bughin, 2016).

3.5.1 HBO

Společnost Home Box Office (HBO) byla založena v roce 1972 v USA. Ve svém začátku provozovala nejprve běžné satelitní, a následně prémiové kabelové televizní vysílání. V roce 1980 představila HBO svůj druhý kanál Cinemax, který měl soupeřit s konkurenčním Showtime (Viacom Inc.). Od roku 1982 se pak společnosti CBS Inc. a Columbia Pictures aktivně podílely na založení Tri-Star Pictures ve snaze vytvářet vlastní prémiový obsah. V roce 2010 HBO spustila také VOD platformu s názvem HBO Go (Gregersen, 2020).

HBO Go byla původně zamýšlená jako služba pro stávající předplatitele kabelového vysílání ke sledování pořadů z archivu. Pro uživatele, kteří nebyli předplatiteli kabelového vysílání HBO, zpřístupnilo HBO službu Now. Později se ukázalo, že se však jedná o téměř identickou knihovnu pořadů. V některých zemích pak byla dostupná pouze služba HBO Go, která však fungovala stejně jako HBO Now (Blumenthal, Katzmaier, & Pendlebury, 2020).

HBO, ve které měla od jejího vzniku podíl společnost Time Inc. (později Time Warner a v současné době jeden z největších světových mediálních domů WarnerMedia, LLC) spustila ve snaze konkurovat stávajícímu tlaků velkých provozovatelů VOD platforem na trhu v květnu roku 2020 novou platformu HBO Max, která má postupně nahradit stávající HBO Now a HBO Go, a doplnit svůj obsah o nabídky televizních stanic a studií DC, CNN, TNT, TBS, truTV, Cartoon Network, Adult Swim, Rooster Teeth, Looney Tunes a Warner Bros (Gilbert & Webb, 2020)

Prezentace společnosti HBO (2019) z konference Filmmaking by the Numbers představuje, jakým způsobem tento – jeden z největších hráčů na trhu s VOD obsahem – pracuje s daty.

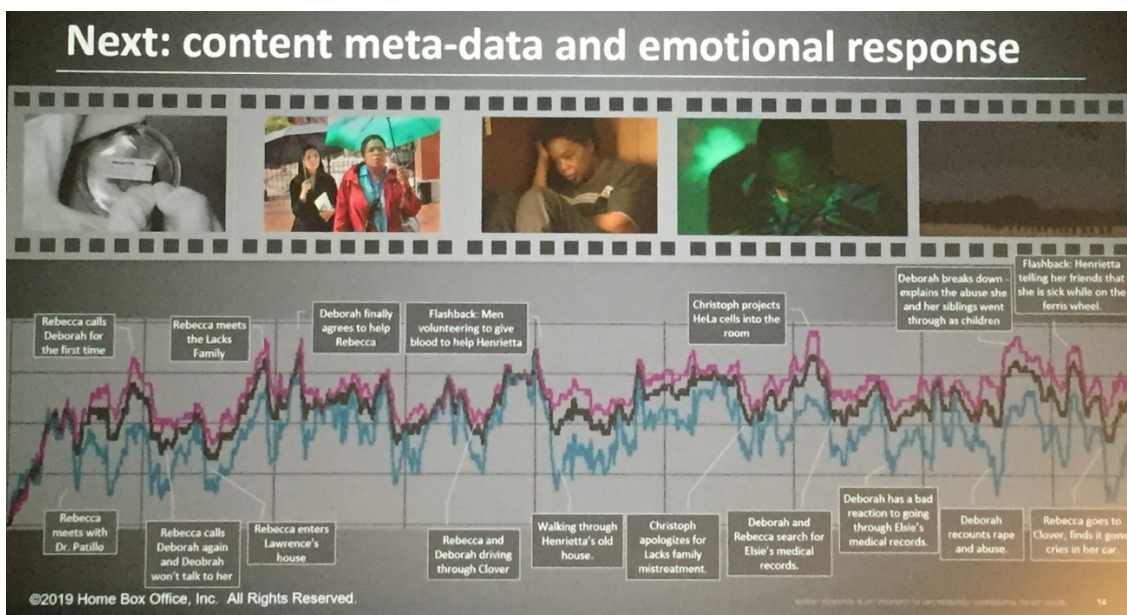
Souvztažnost čtyř klíčových konceptů je vodítkem k využití big data v rámci tvorby audiovizuálního díla. Obsah audiovizuálního díla je ovlivněn následujícími prvky (HBO, 2019):

- Preference cílového segmentu – jaká je cílová skupina? Jak je cílová skupina velká? Jak je možné cílový segment zasáhnout? Jak lze cílový segment obsloužit a kým?
- Nákupní chování (okolnosti zhlédnutí audiovizuálního díla) – při jaké příležitosti budou zákazníci audiovizuální dílo sledovat? Jak tyto příležitosti ovlivňují celkovou sledovanost?
- Efektivita obsahu – jaký obsah zvyšuje zájem zákazníků o audiovizuální dílo? Jaký obsah zvyšuje zájem zákazníků o využití dalších služeb?
- Získávání nových zákazníků – jak HBO získává zákazníky? Jaké jsou vztahy mezi spotřebou audiovizuálního díla, zaujetím a retencí? Jakou roli hraje HBO v domácnostech? Jaké mají zákazníci od HBO očekávání?



Obrázek 10: Schéma segmentů zákazníků na základě jejich vazby na VOD platformu HBO, zdroj, Boulton-Wallace, 2019

Na základě těchto informací pak HBO vytvoří segmenty zákazníků, ke kterým přiřadí vhodná audiovizuální díla ze své produkce. Například v segmentu *Can't live without* se nachází fanoušci seriálů, kteří konzumují audiovizuální díla denně, a VOD v jejich životech hraje důležitou roli. Tomuto segmentu nabízí HBO mimo jiné seriál *Hra o trůny* apod. Taková rozdělení a zacílení provádí HBO právě prostřednictvím analýzy big data. Dále se u této VOD platformy využívá analýzy big data pro sledování reakcí diváků na děj seriálu či filmu. Sledují a vyhodnocují se různé metriky a reakce na konkrétní děj, aby mohlo docházet ke zlepšování podoby audiovizuální díla v budoucnosti (Boulton-Wallace, 2019).



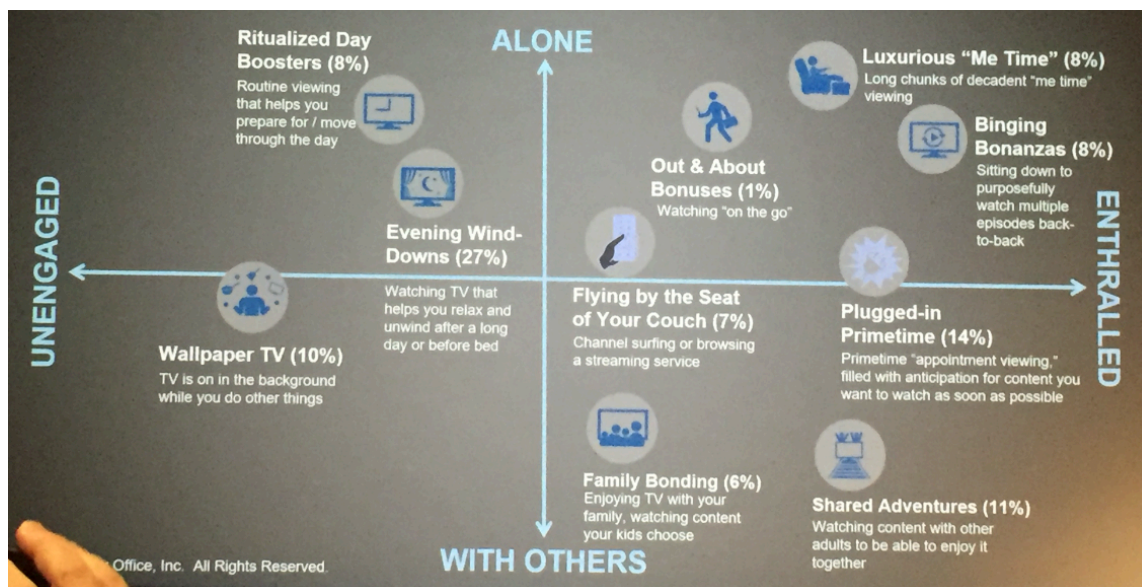
Obrázek 11: Testování uživatelské odezvy HBO filmu Nesmrtelný život Henrietty Lacksové na časové ose pořadu, zdroj: Boulton-Wallace, 2019

HBO zároveň pracuje s údaji o chování svých uživatelů v době sledování HBO Go. Kombinací dotazníkových šetření a sběru dat o používání jednotlivých účtů na různých zařízeních například společnost vyhodnotila, že 10 % uživatelů používá službu v momentě, kdy potřebuje jakousi kulisu do pozadí k jiné aktivitě, nebo že v 6 % případů službu využívá celá rodina ke společnému sledování, a HBO Go tak slouží k posílení pocitu společně stráveného času (Boulton-Wallace, 2019).

Data o chování uživatelů, respektive jejich zájmech, napomáhají provozovatelům služby, aby rozklíčovali preference diváků a zpřesnili co nejvíce skladbu svojí knihovny pořadů. Jak uvádí Boulton-Wallace, při tvorbě nebo akvírování nových pořadů je vždy důležité přemýšlet nad tím, pro koho je pořad určen, a kde a jak ho bude daný divák konzumovat.

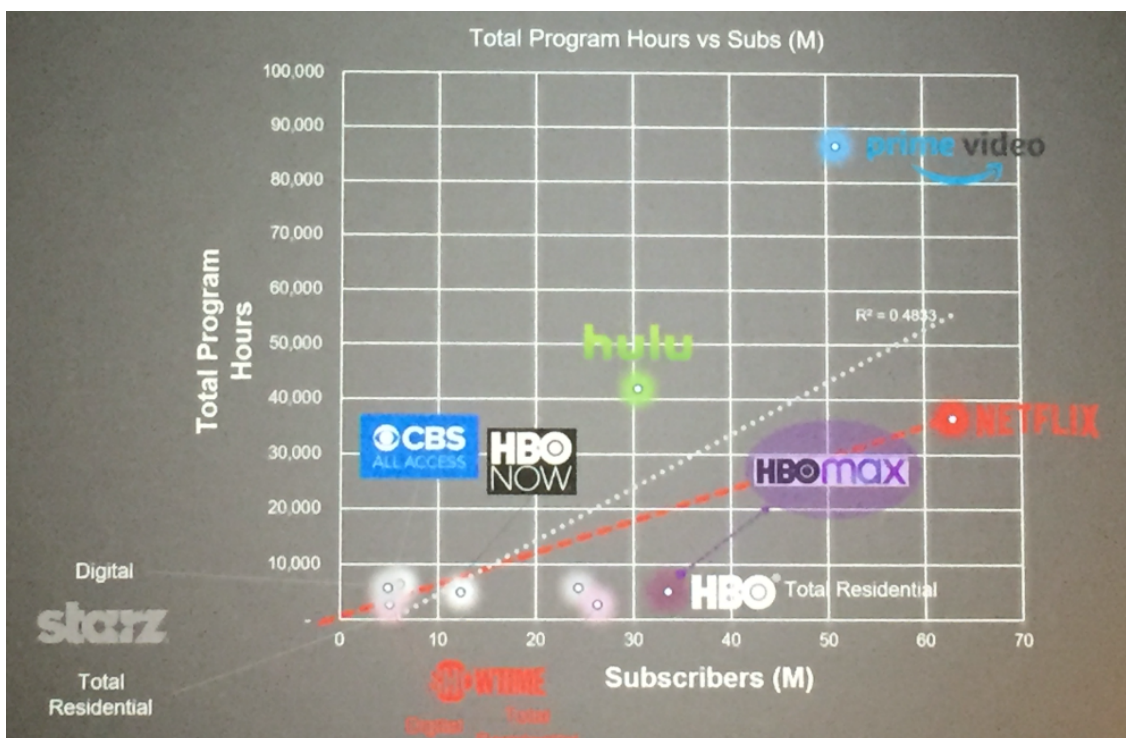
Podle Boultona-Wallace dává tedy smysl zaměřovat se na všechny čtyři kvadranty z následujícího grafu, a nabízet možnosti jak divákům koncentrovaným, tak těm příležitostným apod. Zásadním rozdílem podle

Boultona-Wallace je, že chování uživatelů je možné na rozdíl od běžné televize velmi efektivně monitorovat (Boulton-Wallace, 2019).



Obrázek 12 Spektrum diváckých návyků při sledování HBO Go, zdroj: Boulton-Wallace, 2019

Mezi počtem hodin ke shlédnutí a náklady není sice přímá úměra, ale HBO se snažila obě položky vyčíslit a průměrovat a podle toho určit nastavení svého obchodního modelu (viz graf níže).



Obrázek 13 Graf počtu předplatitelů (v milionech) a množství hodin obsahu v knihovnách vybraných VOD platform, zdroj: Boulton-Wallace, 2019

Jak uvádí v přednášce na konferenci TED (2018) CEO Netflixu Reed Hastings, snahou SVOD platformy není (na rozdíl od TVOD nebo AVOD) nasbírat co nejvyšší počet shlédnutí na jedné položce v knihovně, ale najít optimální poměr mezi počtem položek, respektive počtem hodin obsahu ke shlédnutí za dosažení co nejvíce počtu předplatitelů.

HBO si tedy pokládá otázku, jaké množství jakého obsahu je potřeba pravidelně vyrobit nebo nakoupit, aby byla naplněná uživatelská poptávka. V této souvislosti pracuje také s pojmem *fresh*, což můžeme přeložit jako čerstvý nebo aktuální obsah. Nejlepším příkladem vysokého požadavku na *fresh* obsah jsou pak položky v žánrových kategoriích: zpravodajství, sport, dokument a reality.

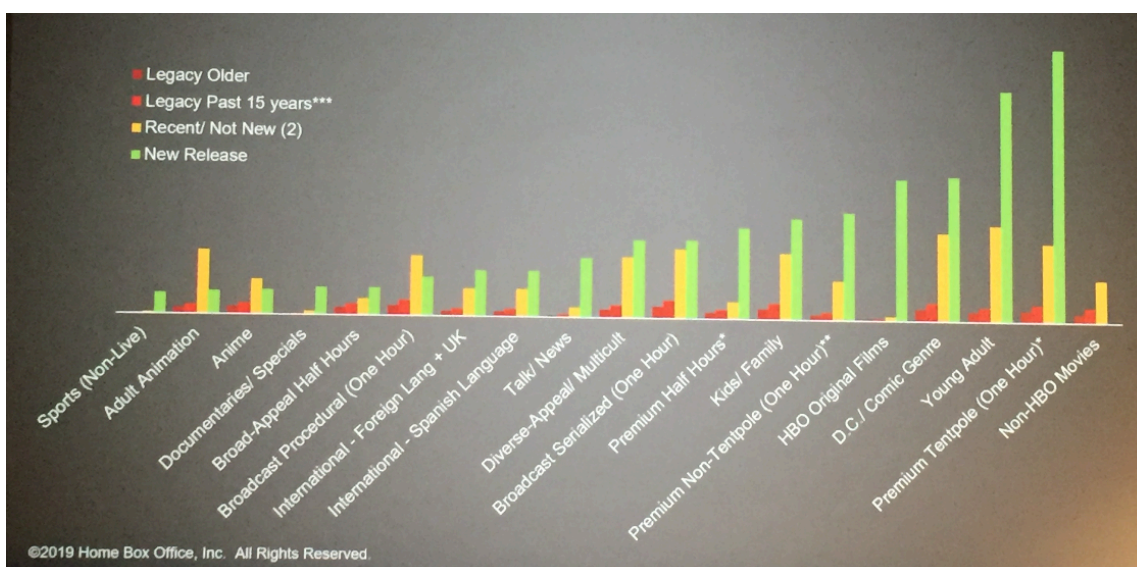
Žánr	Poptávka po jednotlivých žánrech¹⁸	% podíl na poptávce po „fresh“ obsahu v jednotlivých žánrech	% podíl na průměrném množství nově vznikajícího obsahu
Drama mini-série	40%	31%	33%
Komedie / Půl hodinové pořady	10%	24%	11%
Rozhovory / Zpravodajství	3%	95%	10%
Děti / Rodina	7%	35%	10%
Animace	3%	24%	5%
Akční / Dobrodružné / Superhrdinské	3%	31%	6%
Reality show / Bez scénáře	5%	75%	5%
Dokumenty	3%	75%	3%
Sport	1%	74%	4%
Film (všechny žánry)	25%	66%	25%

Tabulka 2 Poptávka po obsahu v jednotlivých žánrech na HBO VOD, zdroj: Boulton-Wallace, 2019

Tento model tedy na základě uživatelských statistik a preferencí predikuje, kolik obsahu jakého žánru by měla služba nově pravidelně vyrábět, nebo přebírat a nabízet, pokud chce uspokojit z největší míry poptávku uživatelů.

¹⁸ Vypočteno z podílu na průměrném čase, který tráví uživatel služby sledováním. Tento čas je spočítaný na 5 minut a 39 vteřin.

Po započtení socio-demografických proměnných a nákladů na nákup nebo výrobu pořadů do modelu se ukazuje, že největší divácký zájem je o nově zveřejněné originální seriály v hodinové stopáži, originální filmy a filmy a pořady podle komiksů. Naopak nejmenší zájem je o archivní sportovní pořady, archivní dokumenty a trochu překvapivě i o archivní filmy mimo produkci HBO (Boulton-Wallace, 2019).



Obrázek 14 Efektivita uvádění nových a archivních pořadů na HBO VOD z hlediska diváckého přínosu, zdroj: Boulton-Wallace, 2019

3.5.2 Stream

Televize Stream prošla těsně před dokončením této práce transformací, proto se jí autor věnuje pouze stručně s přihlédnutím k její dřívější dominantní pozici na trhu a průkopnictví v oblasti využití dat v oblasti AV na českém trhu.

Původní tým, který stál u nejsilnějšího období Streamu vedl od roku 2011 Lukáš Záhoř, který: „se svým týmem vybudoval svébytnou podobu jak českých online seriálů, tak publicistiky a infotainmentu. Tvůrčí vedení webu upustilo od bulvárního schématu, které na konci první dekády symbolizovaly pořady Pavla Novotného nebo Agáty Prachařové, tehdy

Hanychové. Místo nich přišly cykly reagující na společenské problémy, které se staly poznávací značkou Streamu (Jetmar, 2020)."

Po odchodu Lukáše Záhoře a velké části původního týmu do Mall.TV, přichází do Streamu Dan Strejc, který dříve působil v Obvodu Kamila Oušky. Stejně jako Lukáš Záhoř se svým týmem přetáhl část původních tvůrců do nově vznikající Mall.TV, tak i Strejc přináší seriál Lajna na Stream.

S odchodem Lukášem Záhoře v roce 2018 se postupně mění i programová skladba, která nakonec vede na základě požadavků managementu k úplnému pozastavení produkce vlastní tvorby. Poslední dva originální pořady, které na Streamu vznikli jsou Skvrna a druhá série Lajny.

Na odchod publicistických tváří a celkové úspory ve vlastní tvorbě Stream v roce 2018 reaguje návratem k videu od třetích stran. Nevrací se ale přitom k uživatelskému videu, nýbrž nabízí prostor kanálům a tvůrcům, kteří dosud publikovali jinde. *„Představitelé Seznamu to popisovali jako analogii s „nekonečným“ feedem. Další zajímavé články z českého internetu, který Seznam umístil na svou domovskou stránku a odkazuje z něj na stovky českých webů, s nimiž si pak dělí zisky z reklamy. V tomto případě nasmlouvání tvůrci nahrávají na Seznam svá videa, která nadále publikují i jinde. Dostane se jim za to možnost návštěvnosti z domovské stránky Seznamu a provize z inzerce. „Zapojení třetích stran je něco, co je v českém online světě revoluční. Víme, že videoobsahu tady není zas až takové množství, ale stejně jako se do textového feedu postupně zapojuje spousta webů s kvalitními texty, je možné, že podobně tomu bude v případě tvorby videa“ vysvětloval v rozhovoru pro Médiář Kalaš. (Jetmar, 2020)*

„Seznam.cz proměnu přístupu k původnímu videu dovršuje letos na jaře, když představuje nové schéma pozemní Televize Seznam. Ta se od svého spuštění v roce 2018 potýká s nízkou sledovaností, loni zaujala 0,23 % diváků starších 15 let, na přelomu letošního dubna a května činil share 0,41 %.

Televize deklaruje, že se zaměřuje na diváky od 25 až 55 let, nicméně loni plnou polovinu jeho publika činili lidi starší 55 let." (Jetmar, 2020)

Vojtěch Matoušek, který pro Seznam pracoval mezi lety 2018 a 2020 jako datový analytik obsahu, uvádí, že v souvislosti s video obsahem a daty byl požadavek vedení na řadu dílčích úkolů. Jako příklad zmiňuje analýzu „dokoukanosti“ pořadů, která vedla k rozhodnutí zkrátit u konkrétních pořadů titulky na konci videa, aby se zvýšila počet prokliků na další pořady v nabídce. Dalším příkladem byla studie, která zkoumala data zveřejňování jednotlivých epizod pořadů k dosažení co největšího zásahu, respektive počtu shlédnutí.

Matoušek dále uvádí, že Seznam používá ve svých redakcích vlastní analytické nástroje a systém doručování reklamy uživatelům Sklik. V českém prostředí je podle Matouška reklama, daleko častěji zobrazována na základě burzovního systému, než že by byla výrazněji ovlivněna uživatelskými daty.

Zajímavé jsou také výsledky statistiky chování uživatelů při používání mobilních zařízení. *„Konzument na mobilu má oproti konzumentovi na počítači daleko kratší dobu návštěvy a tím, že se jedná o menší displej, tak se ukazuje, že je i menší šance, že bude konzumovat video obsah. U sledování obsahu na počítači je také větší tolerance vůči obsahu. To znamená, že i když uživatel nemá pocit, že je obsah přímo určen pro něj, vydrží se na něj dívat déle nebo se o něm dozvědět více v porovnání s mobilním zařízením.“*

Dlouhodobě se ukazuje, že v případě online videí dochází k největší ztrátě zájmu (10 – 20% diváků) během prvních 10 vteřin sledování videa. Pak už je opadnutí zájmu a vypnutí videa nebo přepnutí na další video statisticky menší. *„Vždy však velmi záleží na cílové skupině a obsahu konkrétního videa. Někdy je snaha firmy tlačit na výsledky reklamního zásahu a dosáhnout co*

největší počtu prokliků na video prostřednictvím clickbaitů¹⁹, ale to může být riskantní, protože uživatel už se příště ke službě vrátit nemusí. Proto nám vychází, že daleko efektivnější z je z dlouhodobého hlediska důležité budování uživatelské vazby na produkt a přímý vztah mezi divákem a určitou skupinou videí podle zájmu. (Matoušek, 2020)"

Mezi hlavní kvalitativní metriky patří zejména:

- čas strávený na stránce;
- dosledovanost videa;
- zapojení v diskuzích pod videem;
- přehrání dalšího videa automaticky versus jeho manuální výběr;
- Na rozdíl od klasického TV vysílání slouží data také k vyhodnocování času a počtu zobrazované reklamy. Snahou tvůrců video obsahu je nasazovat reklamu do předělů mezi tématy tak, aby došlo k co nejmenšímu zásahu do příběhu a aby nebyla narušena pozornost v rámci jednoho pořadu, avšak na základě datových analýz dochází k vyhodnocení, že nejlepší výsledky dosledovanosti dosahují právě videa, kde je reklama nasazená naopak do prostřední části tématu, neboť divák udrží pozornost i v průběhu reklamy a nemá pocit, že už se dozvěděl celou informaci a může tak z videa odejít (Matoušek, 2020).
-
- Jak bylo zmíněno výše, seznam vyvinul vlastní analytický nástroj, který se jmenuje Reporter. Ten je od roku 2016 veřejně dostupný v beta

¹⁹ Clickbait (z ang. *click*, kliknutí a *bait*, návnada, tedy návnada na kliknutí) je většinou poutavý titulek, který nutí čtenáře na něj kliknout. Může se jednat o titulek který má pouze lidi nalákat a po prokliknutí se sice návštěvník většinou nic, co bylo slíbené v titulku, nedozví, nicméně zvýšil návštěvnost webu, a to i v případě, že web ihned opustil. (Zdroj: <https://cs.wikipedia.org/wiki/Clickbait>)

verzi, která obsahuje pouze vybrané statistické údaje a vizualizovaná data (Čížek, 2016).

Podle Matouška (2020) nechtěl Seznam na svých webech používat externí metriky, ale mít možnost uzpůsobit si systém pro vlastní potřebu. Proto dlouhodobě vyvíjí vlastní software, který k datovým analýzám používá. Matoušek také zmiňuje, že systém Reporter počítá na rozdíl od základní verze konkurenčních Google Analytics každého uživatele zvlášť. Google Analytics naopak vytváří statistiku na základě reprezentativního chování vzorku (tzv. smplování), což může být zejména u měření objemově malých statistik a specifických cílů problematické, a tento postup tak může vést k zavádějícím výsledkům. Hlavní benefitem vlastního softwaru je přehlednost sledovaných trendů, díky které je systém relativně jednoduchý na používání. Nevýhodou použití vlastního nástroje pro práci s daty v případě Reporteru může být mimo jiné specifická, která znemožňuje univerzální použití. Pronové zaměstnance může být problematické na specifický systém přejít. Existuje také neustálá potřeba inovace a investice na základě posunu technologických trendů, podle Matouška však k výrazným inovacím již více než 10 let nedochází. *„Reporter není špatný nástroj, to určitě ne, ale není aktivně vyvíjen tak rychle jako třeba Google Analytics a trochu stagnuje. (Matoušek, 2020)”*

4 Politické a etické faktory využití dat v audiovizuální tvorbě na vyžádání

Využití big data je mimo jiné spojeno i s řadou obav a etických otázek. Jak uvádí ve svém článku Havens (2014), objevují se zejména obavy spojené s bezpečností a zneužitím dat. Spotřebitelé poskytují data o svém chování a často není jasné, jakým způsobem budou tato data využita. Je to právě nedostatečná informovanost, která vede k výše zmíněným obavám. (Havens, 2014).

4.1 Regulace EU a možný dopad na provozovatele VOD

Trh audiovizuálních mediálních služeb je součástí jednotného evropského trhu, a tedy je předmětem legislativní regulace ze strany Evropské unie. Pravidla pro audiovizi obsahuje směrnice Evropského parlamentu a Rady 2010/13/EU a v tuzemsku je reflektována v Zákoně o rozhlasovém a televizním vysílání. V roce 2018 Evropská unie novelizovala zmíněnou směrnici a členské státy musí změny do své národní legislativy promítnout do září 2020. Novelou se výrazně mění pravidla pro poskytování audiovizuálních služeb, což mění i pravidla poskytování audiovizuálních mediálních služeb na vyžádání (tzv. video-on-demand). V novém režimu mají jak tradiční televize, tak i nové audiovizuální služby fungovat na mediálním trhu za víceméně rovných podmínek, což doposud nebylo pravidlem (Tajbr, 2019).

Tato směrnice počítá s tzv. koregulací, což je kombinace státního dozoru a samoregulačních mechanismů. Samoregulační mechanismy musí spoludefinovat dotčené subjekty, tedy i poskytovatelé VoD služeb (Brychta, 2018).

Nová regulace Evropské unie mimo jiné vytváří podmínku podporující evropský obsah i u služeb VoD, a tedy implementuje 30% kvótu na evropská díla u těchto služeb. Služby VoD mají dále finančně přispívat na rozvoj evropské filmové produkce, ať už přímými investicemi do produkce či

neadresnými příspěvky do vnitrostátních fondů pro rozvoj kinematografie. Regulace počítá s implementací pravidla zvýšené ochrany dětí před násilím, nenávisí, terorismem, škodlivou reklamou, a provozovatelé VoD musí vytvořit efektivní mechanismus, který uživatelům zprostředkuje možnost nahlášení škodlivého obsahu, aby mohlo v případě potřeby dojít k jeho následnému vymazání (Tajbr, 2019).

Z tohoto vyplývá, že se toto nařízení výrazně dotkne také společností jako Netflix nebo Amazon Prime (Brychta, 2018).

Změny v důsledku směrnice Evropské unie pro provozovatele VoD se dále dotýkají reklamy, kdy reklama může tvořit nanejvýš 20 % vysílacího času od 6.00 do 18.00 hodin a od 18.00 do 24.00 hodin (Tajbr, 2019).

Nová směrnice Evropské unie je přirozeně předmětem kritiky ze strany poskytovatelů VoD služeb, kteří považují definované kvóty pro evropskou produkci za příliš protekcionistické, zatěžující a zabraňující volným tržním mechanismům (Tajbr, 2019).

4.2 GDPR

Big data jsou masivní objemy dat, které se ukládají většinou v datových skladech a obsahují všechny druhy digitálních stop, které za sebou zanechávají uživatelé internetu (jde mimo jiné často o dokumenty, fotografie, telefonní záznamy, elektronickou poštu apod.) Pro organizace je prospěšná zejména analýza dat, jejímž prostřednictvím mohou získávat informace a znalosti (Ježek, 2018).

„Model pokročilého zpracování velkých dat může při správném nastavení dodávat společností na denní bázi afinitní zákazníky k oslovení, upozornit na mezery v poskytovaných službách, upravit styl nejvýhodnější reklamy a mnoho dalšího (Ježek, 2018).“

Ve většině případů je podnikatelský subjekt také správcem osobních údajů, a proto se na něj GDPR vztahuje. Z tohoto důvodu je nutné provést precizní GDPR compliance analýzu²⁰, která zodpoví základní otázky související s využitím osobních dat v souladu s nařízením, a to i u big data, kdy v minulosti často docházelo k tvrzení, že nestructurovaná big data v podstatě nelze kontrolovat. Tento přístup od účinnosti GDPR není možný, a ochrana a účinky se vztahují také na big data (Malatinský, 2018).

Rozhodnutí systému či stroje je dáno existencí vstupních dat, která jsou vyhodnocena, a na jejichž základě se rozhodne o dalším postupu. Pokud se jedná o data, která nesouvisí s lidmi (například data o klimatických podmínkách), není jejich využití z legislativního hlediska problematické. Řada nejrůznějších strojů, zařízení a aplikací však pracuje právě s osobními údaji a musí tak reflektovat různá legislativní rizika (Škorníčková, 2018).

Legislativní rámec GDPR byl spojen s obavou řady organizací, že již nebude možné plnohodnotně využívat big data pro zajištění rozvoje podniku a získání konkurenceschopnosti, protože s využitím big data budou spojena významná legislativní rizika a případné sankce (Škorníčková, 2018).

Big data a jejich rozvoj jsou typickou situací, kdy technologický pokrok předbíhá legislativu a technické možnosti regulátorů (dozorujících na trhy). Legislativní rámce využití big data ještě není zcela vytvořen, a tak chybí také prostor pro regulaci. Přitom se neustále objevují otázky související s ochranou získaných dat a zachováním soukromí uživatelů internetu, kteří big data používají. Právě takové otázky by měla legislativa řešit (Ježek, 2018).

Oblast ochrany osobních údajů a dat je ovšem legislativně regulována, a to poměrně přísným způsobem. Od roku 2018 plošně na celoevropské úrovni

²⁰ Analýza, na základě se vyhodnotí, zda je jednání společnosti, zaměstnanců a vedení v souladu s právními předpisy týkajícími se GDPR.

prostřednictvím tzv. GDPR, tj. nařízení o ochraně a zpracování osobních údajů. V typickém případě obsahují big data i osobní údaje, pokud neprocházejí algoritmem, který je anonymizuje. Ke zpracování osobních údajů v souvislosti se zpracováním a analýzou těchto velkých dat může dojít poměrně jednoduchým způsobem. V takovém případě na zpracovatele dat dopadá i nařízení GDPR a riziko veškerých sankcí plynoucích z nařízení (Ježek, 2018).

GDPR definuje využití big dat jako automatizované individuální rozhodování, které je charakteristické tím, že (Škorníčková, 2018):

- Určitá rozhodnutí ve vztahu k subjektům údajů jsou založená výhradně na automatizovaném zpracování osobních dat, dále pak se na tvorbě rozhodnutí ve stěžejních fázích, zejména v rámci vyhodnocování informací, na kterých se rozhodnutí zakládá, nepodílí člověk.
- Tato rozhodnutí mají vůči danému subjektu právní účinky či se ho jinak významně dotýkají.

V řadě případů big data obsahují osobní údaje a jejich anonymizace není možná, protože nelze rozdělit data na osobní a neosobní. Případné využití pouze neosobních dat navíc nepovede k získání smysluplných informací. Organizace, které big data využívají, proto musí mít souhlas dané osoby s tím, že budou její data využita, aby došlo k plnění GDPR nařízení (Ježek, 2018).

„Analýza a využívání Big Data za pomoci umělé inteligence patří v poslední době k hlavním trendům na vyspělých trzích a připravit se na tuto možnost kvůli obavám ze striktního výkladu GDPR by mohlo snížit konkurenceschopnosti tuzemských firem vůči mimoevropským podnikům (Škorníčková, 2018).“

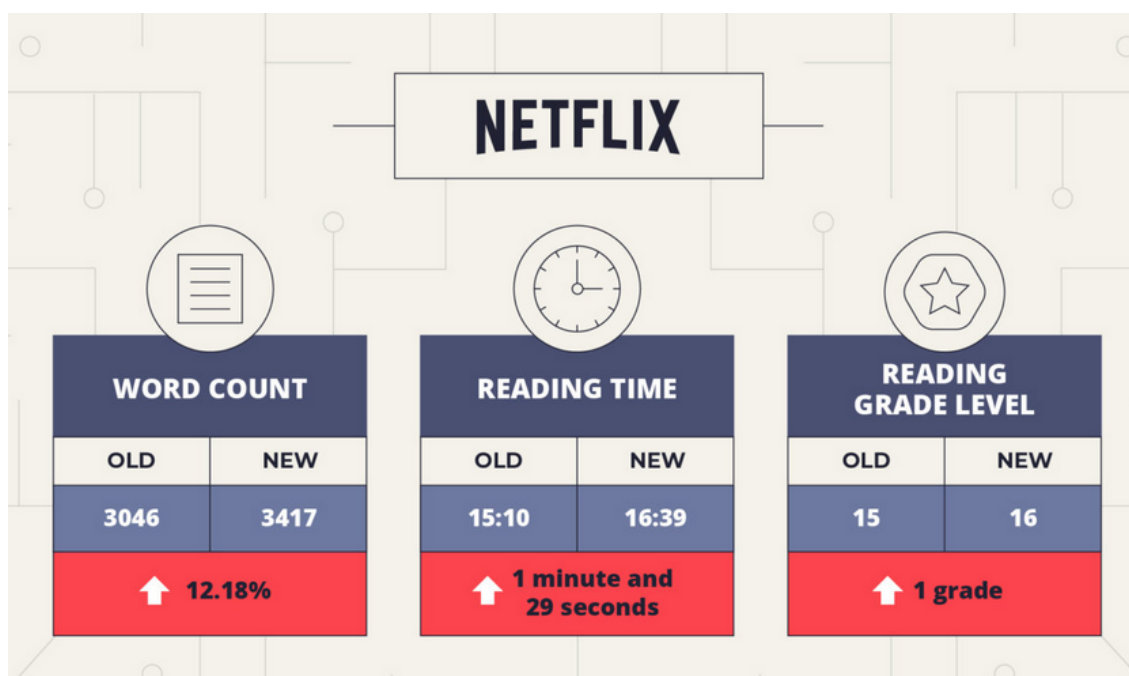
Netflix ukládá údaje o zákaznících, aby mohl vytvářet statistiky o jejich rozhodování v průběhu sledování či výběru audiovizuálního obsahu. Na

základě analýzy big data potom získává přehled o tom, jak diváci s příběhem interagují. Výsledky jsou následně využity pro zlepšení modelu vyprávění děje audiovizuálního díla. Za problematické lze považovat, že o tomto postupu společnost diváka explicitně neinformuje, což je problémem právě v oblasti GDPR a ochrany osobních údajů. Uživatel také nemá možnost odmítnout takové ukládání dat (Mikešová, 2019).

V souvislosti s GDPR čelí některé VoD platformy podezření z porušování tohoto předpisu a zneužívání osobních dat o uživatelích. Z testu rakouské neziskové organizace vyplývá, že organizace jako Netflix či Amazon neposkytují uživatelům žádné informace o tom, jestli nakupují či prodávají uživatelská data, nebo jak s daty pracují. Každá z platform byla v rámci testu požádána o informace o tom, jakým způsobem využívá zákaznická data, ale žádosti nebylo zcela vyhověno. Nedošlo k poskytnutí informace o tom, s kým jsou data sdílena apod. Každá z oslovených platform se vyhýbala svým povinnostem o poskytování údajů, či poskytla pouze nepřesné nebo nedostatečné odpovědi. Z tohoto důvodu nelze hovořit o dodržování GDPR jmenovanými společnostmi, za což jim hrozí sankce ve výši až 20 mil. EUR nebo 4 % z globálního obrátu. Netflix sice poskytnul data v podobě popisu toho, co daný uživatel sledoval, kdy a kde audiovizuální obsah sledoval, či na jakých zařízeních. Chybí však informace, jak s těmito daty Netflix dále nakládá. Společnost Netflix se však brání, že veškerá nařízení GDPR dodržuje (Murgia, 2019).

Dopady GDPR na poskytovatele VoD služeb se zabýval ve svém výzkumu také Sobers (2020), který potvrzuje, že GDPR bylo implementováno za účelem větší transparentnosti o tom, jak společnosti nakládají s osobními údaji uživatelů, a také bylo cílem implementovat přísnější požadavky na použití a sdílení osobních údajů s třetími osobami. Přestože se nařízení dotýká členských zemí Evropské unie, je platné i pro americké podniky, a to včetně poskytovatelů VoD služeb, řada z těchto společností tak musela aktualizovat

informace a podmínky pro nakládání s osobními údaji uživatelů. Z výzkumu vyplývá, že Netflix v důsledku GDPR rozšířil své obchodní podmínky o téměř 400 slov, čímž se zvýšil čas na jejich prostudování o minutu a dvacet devět sekund. Zároveň jsou pro uživatele obchodní podmínky Netflixu poměrně složité na pochopení. Pro analýzu složitosti tohoto textu bylo použito hodnocení Reading grade level. To posuzuje složitost obchodních podmínek na stupnici od 1 do 18, kdy 18 je text, který je srozumitelný pro osobu s právnickým vzděláním. Výstupy shrnuje následující obrázek (Sobers, 2020).



Obrázek 15: Dopad GDPR na Netflix, zdroj: Sobers, 2020

O problematice GDPR hovoří v případě českého prostředí a konkrétně Mall.TV také analytik datového obsahu CNC Vojtěch Matoušek (2020). Uvádí, že směrnice GDPR značně komplikuje datovým analytikům práci a znamenala velký zásah do způsobu získávání a zejména skladování dat. Většinu času nejsou podle něj data personalizovaná, respektive jsou naopak záměrně anonymizovaná, ale mohou nastat určité momenty, kdy je potřeba nakládat

s daty registrovaných uživatelů velmi citlivě a pravidelně je mazat, aby nedošlo k porušení zákona. Zmiňuje také, že v České republice došlo ke schválení Adaptačního zákona k nařízení EU k GDPR velmi narychlo kvůli tehdejší krizi vlády, a nebyly tedy využity zdaleka všechny možnosti odchýlit se na státní úrovni od úpravy GDPR a stanovit určité výjimky, tak jako to udělaly některé západní země EU, kde pravidla nejsou tak přísná. (Matoušek, 2020)

4.3 Etický rozměr sběru a využití dat

V rámci využití big data je klíčovou otázkou i problematika etického sběru a využití těchto dat, protože už nyní dochází ke zpracování velkého množství osobních dat a prostřednictvím propojení (často zdánlivě nesouvisejících a neškodných dat) lze získat velmi osobní a citlivé informace. To vzbuzuje otázky etického charakteru, zejména s předpokladem dalšího růstu míry využití big data. V budoucnosti bude růst automatické zpracování zdravotních dat, osobních finančních dat, dat o pohybu člověka, o jeho spotřebitelském chování. Toto zároveň povede k rozšiřování způsobů, kterými se budou data využívat. V řadě případů bude docházet k narušování soukromí, či nenápadnému posouvání zavedených etických hranic a norem (Akademie věd České republiky, 2019).

„Technologie je jedna stránka věci, na opačné straně je lidský rozměr a vyhodnocení užitečnosti pro jednotlivce. A etický rozměr. V digitálním světě jsme se s tím již postupně smířili, ale přesah do reálného světa část společnosti odmítá s odkazem na přílišný zásah do soukromí. I proto je nutné k technologiím big data přistupovat velmi citlivě a důsledně vážit přínosy a nevýhody. Jen tak mohou společnosti jako celek obohatit a posunout kupředu (Berger, 2016).“

4.3.1 Eticky pozitivní a negativní příklad práce s big data

Etika, která je filozofickým neexaktním oborem, je založená na morálním hodnocení případů a je vždy do jisté míry subjektivním pohledem. Následující dva příklady autor vybral jako společensky pozitivní a negativní příklad toho, jak mohou být data využívána, aby lépe přiblížil kladné i záporné dopady využití těchto dat.

4.3.1.1 Cambridge Analytica a ovlivňování politických výsledků

Cambridge Analytica byla organizace poskytující strategické poradenství pro politiky, vojenské a národně bezpečnostní složky různých států, či pro podnikatelské organizace. K tomuto aktivně využívala analýzu dat a zejména big data. Cílem Cambridge Analytica bylo vytvoření nástroje, který propojí charakteristiky uživatelů s jejich psychologickými profily, aby klient (nejčastěji politik) věděl, na koho se zaměřit v rámci kampaně, ale také, jak nejlépe daného voliče přesvědčit. Šlo o spojení psychologických profilů s osobními daty (Weiss, 2018).

K tomuto účelu využívala organizace psychometrii, umělou inteligenci, a právě big data jako digitální stopu, kterou jedinec nechává v online i offline prostoru (prostřednictvím návštěv internetových stránek, obchodních transakcí, aktivit na sociální síti). S využitím těchto dat pak dokázala organizace na 95 % určit barvu pleti, na 88 % určit příslušnost k politické straně apod. (Kříž, 2017).

Výsledky analýzy Cambridge Analytica ovlivnily volby téměř na všech kontinentech a společnost se tak stala symbolem morálně sporné práce s big data. Podezření na zneužití dat nakonec vedlo k zániku této společnosti. Cambridge Analytica zaplatila přes 1 milion dolarů, aby se dostala k datům o 50 milionech uživatelů sociální sítě Facebook, a to bez získání souhlasu ať už ze strany společnosti Facebook, tak jejích uživatelů. Data pak organizace využila pro svoji obchodní činnost. Big data byla využita k tvorbě algoritmu,

který následně cílil správná sdělení správné cílové skupině za účelem ovlivnění chování uživatelů (Michl, 2018).

„Cambridge Analytica, která se dostala k údajům o 50 milionech lidí na Facebooku, ovlivnila až 200 voleb v zemích po celém světě. Podle výpovědi jednoho z informátorů působila i v Česku. (Hlaváčová, 2018)“

4.3.1.2 Google a mapování výskytu chřipky

Prvním výraznějším příkladem využití big data byl postup společnosti Google v roce 2008, kdy došlo ke zveřejnění předpovědi šíření epidemie chřipky v závislosti na frekvenci hledání výrazů souvisejících s jejími příznaky a poradenstvím ohledně jejího léčení. Rychlost a přesnost organizace Google výrazně předčila standardní statistické metody založené na zpětném sběru diagnóz lékařů. Takovýto výsledek nebyl dán systematickou vědeckou prací, ale pouze statistickým vyhodnocením milionu vyhledávacích výrazů (Janča, 2015).

Tento projekt byl realizován v rámci činnosti nadace Google a dodnes slouží jako pozitivní příklad toho, jak se dají využít big data pro dobročinné účely. Google využil desítky tisíc současně pracujících počítačů, aby vytvořil správný algoritmus pro rozpoznání propuknutí epidemie do 24 hodin. Oproti 14 dnům, tedy době, jak dlouho trvalo než byly zachyceny kritické hodnoty ze zpětného sběru a diagnóz lékařů, to byl diametrálně jiný výsledek. Big data umožnila identifikovat, na co se soustředí kolektivní pozornost (Goleman, 2014).

Tyto dva konkrétní příklady případy nehodnotí etiku korporací jako celků, ale spíše mají za cíl poukázat na možné rizika nebo benefity, které nám tato technologie může přinést.

5 Komentář k autorskoprávním řešením v kontextu big data

Není ambicí této práce přinášet konkrétní autorskoprávní výklady k jednotlivým případům užití práce s big data, databázemi nebo softwary využívanými při práci s daty, jako spíše nastínit, že existují legislativní regulace, pod které tato užití spadají.

Jak již bylo zmíněno výše, existuje několik legislativních opatření, které se věnují práci s daty. Mezi nejzásadnější patří Nařízení Evropského parlamentu a Rady (EU) 2016/679 o ochraně fyzických osob v souvislosti se zpracováním osobních údajů a o volném pohybu těchto údajů (GDPR), dále pak Směrnice Evropského parlamentu a Rady (EU) 2019/790 ze dne 17. dubna 2019 o autorském právu a právech s ním souvisejících na jednotném digitálním trhu. Tyto dvě směrnice regulují mimo jiné ochranu spotřebitele a zajišťují dostupnost produktů, které jsou distribuovány digitálně, a to především pro vědecké a vzdělávací účely.

Databáze nebo softwary, které vznikají při práci s big data nebo které jsou při práci s big data využívány, požívají vlastní právní ochranu a to jak na evropské tak na české úrovni. Konkrétně je reguluje na evropské úrovni zejména Směrnice Evropského parlamentu a Rady 2009/24/ES ze dne 23. dubna 2009 o právní ochraně počítačových programů a Směrnice Evropského parlamentu a Rady 96/9/ES ze dne 11. března 1996 o právní ochraně databází. Ochrana databází a softwarů je pak na úrovni ČR upravena v zákoně č. 121/2000 Sb., obecně známém jako autorský zákon v relativně obecném znění. § 2 odst. 1 stanovuje, že: „Předmětem práva autorského je dílo literární a jiné dílo umělecké a dílo vědecké, které je jedinečným výsledkem tvůrčí činnosti autora a je vyjádřeno v jakékoliv objektivně vnímatelné podobě včetně podoby elektronické, trvale nebo dočasně, bez ohledu na jeho rozsah, účel nebo význam.“

„Stěžejním aspektem v kontextu big data projektů je, aby data byla použitelná a přenositelná. Je zcela pochopitelné, že k tomu, aby mohl být jakýkoliv projekt, který je postavený na dispozici s daty, realizovatelný, musí být zajištěna možnost data shromažďovat, rozmnožovat, třídit, obohacovat, slučovat nebo oddělovat, sdílet atd. a primárně musí být umožněno používat data jako zdroj informací. V tomto ohledu je klíčové, aby bylo pevně stanoveno, kdo je oprávněn s daty nakládat, a rovněž musí být dostatečně zajištěna přístupová práva k datům. (Čapek, 2019)“

Čapek dále uvádí ve své rigorózní práci k tématu právní ochrany při práci s big data, že navzdory všem normám, které jsou aktuálně platné ať už pro ČR nebo pro celou EU, *„neexistuje jednoznačná odpověď ve smyslu, jaký typ dat spadá pod takovou (autorskoprávní) ochranu a jaký ne. Jinými slovy, bude nutné zkoumat způsobilost dat spadat pod autorskoprávní ochranu, a to případ od případu a ve světle aktuálních právních norem a rozhodovací praxe každého státu. (Čapek, 2019).“*

Značná regulace ze strany EU nebo jednotlivých států a nepřehlednost některých opatření má za následek přirozené zbrzdění procesu vývoje dílčích subjektů v odvětví. Výrazný mediální ohlas přinesly v roce 2018 zejména kampaně serverů, které pracují s autorským obsahem, jako je například Facebook, Youtube, ale i Wikipedia nebo český Seznam. Všechny zmíněné subjekty se ohrazovaly zpravidla proti článkům 11 a 13 výše zmíněné směrnice EU o jednotném digitálním trhu. Tyto opatření znemožňovaly bez řádně uzavřených smluv s autory publikovat jejich obsah tak jako doposud.

Zvláštní právní ochrana, která postihuje mimo jiné i VoD platformy, se týká ochrany uživatele proti sběru a ukládání dat cookies bez jeho předchozího souhlasu. JUDr. Martin Karner uvádí, že opatření, které se týká cookies vychází z tzv. E-Privacy směrnice EU (resp. její novelizace směrnicí č. 2009/136/ES), která byla v ČR transponována zákonem č. 468/2011 Sb.,

jenž s účinností od 1. ledna 2012 novelizoval zákon č. 127/2005 Sb. o elektronických komunikacích (Kartner, 2015).

„Hlavní problém tkví v tom, zda mají být cookies považovány za osobní údaje. Evropská unie v posledních letech usiluje o to, aby se na cookies nahlíželo jako na osobní údaje. V této souvislosti bude důležité sledovat, v jaké podobě bude přijat návrh nařízení EU o ochraně osobních údajů. (Kartner, 2015)“

Výše zmíněné opatření mimo jiné zakazuje provozovatelům webů práci s uživatelskými daty bez předchozího souhlasu uživatele. V českém právním výkladu a ve výkladu Úřadu pro ochranu osobních údajů však funguje stále systém tzn. *opt-out* což znamená, že uživatel má naopak povinnost požádat provozovatele webu o vyškrtnutí z databáze uživatelů, u kterých se s cookies pracuje, aby tím jejich používání vzhledem ke své osobě zabránil (Kartner, 2015).

Mnozí provozovatelé webů přesto do svých uživatelských rozhraní přidali lištu s informovaným souhlasem a možností potvrdit sběr dat cookies, a dále doplnili podmínky k užívání služby a registraci, které uživatel vědomě stvrzuje.

„Současné modely předpovědí založené na analýzách big data jsou neustále posouvány směrem kupředu. Data nejsou užívána ve vztahu k osobám většinou již pouze pro jejich kategorizaci na základě osobnostních rysů a návyků, ale stále častěji jsou analyzovány prvky chování, resp. jejich odraz v běžném jednání (tón hlasu, řeč těla, intenzita psaní, zvýšený tep a teplota atd.) a jsou užívány v reálném čase. Kromě nejruznějších benefitů však využití těchto praktik zároveň přináší rizika pro soukromí jednotlivců, která nelze ignorovat (Čapek, 2019).“

Regulace za účelem ochrany konzumentů AV obsahu je důležitá a v éře big data bude čím dál důležitější, stejně jako je důležitá ochrana autorů AV děl a ochrana provozovatelů, správců a autorů databází, kteří s daty pracují.

6 Možnosti potenciálního využití big data v audiovizuální tvorbě v budoucnosti

Brett Danaher (2019) uvádí, že data a datová analytika nám může sloužit k lepšímu rozklíčování obsahu a k porozumění diváckých preferencí, avšak zároveň upozorňuje, že úspěšné dosažení cíle a posouvání hranic je vždy o správném poměru role datové analytiky a lidské odborné intuice a inspirace. *„V současné době jsou data nezbytným pomocníkem a je téměř nepředstavitelné, kolik práce je na nich závislé. Dá se předpokládat, že role datové analytiky a umělé inteligence jako pracovního nástroje se bude stále posilovat, ale je důležité mít na paměti, že obojí vždy podléhat zadání lidí, ať už v požadavcích na výsledky nebo při jejich samotném programování.“* (Danaher, 2019)

Možnosti potenciálního využití big data v audiovizuální tvorbě v budoucnosti jsou definované prostřednictvím vyhodnocení výzkumných otázek a hypotetických předpokladů.

6.1 Vyhodnocení výzkumných otázek

Jak a v jakém rozsahu Netflix a Amazon pracují s big daty?

Poskytovatelé VoD služeb Netflix a Amazon Prime využívají big data ve velkém rozsahu. V obou případech promlouvají big data do různých fází v rámci výroby audiovizuálních děl a významně ovlivňují jeho využití. Dále se big data využívají pro nastavení marketingové kampaně, nebo vyhodnocení různých scénářů vývoje v rámci preprodukce, produkce či post produkce. Big data představují nástroj, kterým si oba provozovatelé VoD služeb upevňují svoji pozici na trhu a účinně komunikují se zákazníky. Z hlediska zpracování práce a řešené problematiky je podstatné zjištění, že big data ovlivňují i kreativní ztvárnění audiovizuálního díla, tedy nikoliv pouze jeho propagaci, podmínky tvorby (například rozpočet, lokaci natáčení atd.) Big data umožňují vytvořit pro zákazníky prémiový obsah. Jak se dále prokázalo,

pravděpodobnost úspěchu je vyšší v případě, kdy dojde k využití lidského faktoru, a tedy nelze spoléhat pouze na big data, ale je nutné provést jejich relevantní vyhodnocení specialistou v oboru (lidskou silou).

Jedná se pouze o sledování aktivity na stránce samotné, nebo sledujete i jiné formy digitálních stop, které za sebou uživatelé zanechávají?

Netflix i Amazon nesledují pouze aktivity zákazníků na dané stránce a platformě, ale svoji pozornost orientují také na další digitální stopy uživatelů. Sledována je tedy například aktivita těchto uživatelů na sociálních sítích. Využití big data vykazuje lepší výsledky úměrně velikosti datového souboru – čím větší je množství relevantních informací o uživatelích, tím přesnější je také výsledek. Informace o sociálních vazbách doplňují informace o chování zákazníků, aby došlo k ještě přesnějšímu zacílení na jejich preference.

Jakou formou jsou výsledky převáděny do zadání pro kreativní tvůrčí skupiny zabývající se vývojem pořadů?

Výsledek vždy závisí na konkrétní situaci. V některých situacích je vhodnější nechat tvůrcům audiovizuálního díla větší autonomii při kreativním ztvárnění, u jiných případů tento postup naopak vhodný není, neboť se ukazuje, že dobře postavené datové analýzy mohou přinést daleko přesnější výsledky než odhad postavený na intuici. Oba přístupy je žádoucí zkombinovat tak, aby byl konečný výsledek co nejpřesnější. V praxi se však jedná o velmi složitý úkol, a to i z toho důvodu, že o úspěchu audiovizuálního díla rozhoduje velké množství faktorů, a pouhé využití big data není zárukou úspěchu.

Do jaké míry jsou tyto tvůrčí skupiny autonomní vůči výsledkům získaným z analýz?

Tvůrčí skupiny vykazují různorodou autonomii vůči výsledkům získaným z analýz big data. Například v případě seriálu House of Cards byla autonomie tvůrčího týmu spíše na nižší úrovni a audiovizuální dílo dosáhlo výrazného

úspěchu. Naopak, u seriálu Alpha House byla situace odlišná a nepodařilo se dosáhnout takového úspěchu. V případě redakce platformem Stream a Mall.TV slouží datové analýzy jako významný poradní orgán.

V jaké formě dostávají tvůrci obsahu zadání nebo případně jaké dostávají podklady pro výrobu nových pořadů?

Tvůrci často dostávají od analytických odborníků okruhy sestavené na základě uživatelských preferencí stávajících cílových skupin. Sledované jsou jak některé sociodemografické údaje, tak i zájmové preference a vazba na stávající pořady v knihovně VoD platformy. Zpravidla se nejedná o předepsané zadání, které by přímo ovlivňovalo konkrétní dílčí položky na úrovni dialogů ve scénáři, ale spíše o obecnou typologii produktu z hlediska stopáže nebo žánru.

Jaké další analýzy kromě sběru dat slouží k marketingu a k tvorbě obsahu u poskytovatelů AV obsahu?

Kromě tzv. sběru dat na pozadí jsou u některých provozovatelů VoD stále častěji využívány také focus groups. K jejich využití dochází, jak u zahraničních velkých provozovatelů jako je HBO, tak i u lokální jako je Česká televize nebo Mall.TV. Jak uvádí Matoušek (2020), jejich běžné využití v ČR je však méně obvyklé, neboť jejich realizace je zpravidla velmi nákladná a ukazuje se, že výsledky mohou být zavádějící. K využití focus groups dochází zpravidla v situacích, kdy neexistuje možnost sbírat data jiným způsobem, tedy ve chvíli, kdy se jedná o dosud neexistující nebo nezměřitelný inovativní produkt, přístup nebo řešení.

Provádí se práce s big data a také objevování ve způsobu jejich získávání interně, nebo tuto službu zajišťuje někdo externí?

Většina poskytovatelů VoD služeb při využití big data spolupracuje s externími partnery, protože je využití big data závislé na nutnosti dispozice kvalifikovanými lidskými zdroji a technologickým zajištěním. Data jsou

většinou získávána interně, z vlastních datových jezer, které mají provozovatelé VoD služeb zřízené. Zpracování a analýza se však realizuje často za použití externí technologie, ať už softwarové nebo hardwarové.

Jaký podíl mají na trhu VoD poskytovatelé, u kterých je předpoklad, že využívají big data, oproti těm, kteří této technologii nevyužívají?

Tržní podíl VoD poskytovatelů, kteří využívají big data, je podstatně větší než těch, kteří tohoto přístupu nevyužívají. Mezi úspěšné společnosti patří Netflix, HBO, Hulu, nebo Amazon Prime, které slouží jako příklad vhodného využití big data pro dosažení tržního úspěchu, a kterým se daří využívat big data pro udržení svojí konkurenceschopnosti. Síla a hodnota těchto velkých provozovatelů VoD služeb na trhu v USA, ale také na globálním trhu neustále roste.

6.2 Vyhodnocení hypotetických předpokladů

A. Současní klíčoví hráči v oblasti VoD platforem získají značnou převahu nad klasickými stávajícími kanály a výrobci AV obsahu, a to díky masově personalizovanému vlastnímu obsahu, vytvořeného na základě práce s big data.

Tuto hypotézu lze potvrdit. Při konzumaci audiovizuálního obsahu hrají stále důležitější roli jiná technologická zařízení, než klasická televize nebo rozhlas. Stále větší počet zákazníků spotřebovává audiovizuální obsah prostřednictvím počítačů, tabletů, chytrých telefonů apod. Platformy VoD se na tomto trendu velmi výrazně podílejí a aktivně reagují na vyšší poptávku po konzumaci audiovizuálního obsahu výše zmíněnými prostředky. Tím získávají výhodu nad klasickými poskytovateli televizního či rozhlasového vysílání. Zákazníci si mohou sami určit, co a kdy chtějí sledovat. Zároveň provozovatelům VoD služeb poskytují svoje data, která pak platforma využívá pro další zdokonalování svých služeb. Klasické televizní stanice takovou možnost nemají, což je staví do nevýhodné pozice. VoD poskytovatelé služeb

disponují větším množstvím způsobů, jak pochopit své zákazníky a jejich preference či potřeby. Tato situace nastává právě díky aktivnímu a efektivnímu využití dat. Tváří v tvář této výhodě hledají klasické televizní stanice jen složitě adekvátní reakci na novou konkurenci ze strany VoD poskytovatelů služeb. VoD poskytovatelům se daří získávat zákazníky těchto televizních stanic, či kabelového vysílání. Masivně personalizovaný vlastní obsah, který je vytvářen na základě práce s big daty, hraje výraznou roli v boji o diváka. Z tohoto důvodu lze s touto hypotézou souhlasit.

B. Big data jsou nafouknutou bublinou, která slouží zejména jako pojem v marketingových strategiích inovativních institucí, které se snaží tvářit technologicky vyspěle, avšak reálné využití podkladů vyhodnocených po sběru dat je zanedbatelné.

Tuto hypotézu lze vyvrátit. Využití big data umožňuje zlepšit výkonnost poskytovatelů VoD služeb. Argumentem jsou i vysoké investice tržních lídrů do využití big data. Pokud by žádný přínos neexistoval, a tedy šlo pouze o marketingovou bublinu, žádná z těchto společností by neinvestovala vysoké částky do jejich využití. Účelem existence těchto společností je tvorba a maximalizace zisku, a proto lze předpokládat, že investují skutečně účelně do činností, které jim zajistí zisk. Na příkladu seriálů House of Cards či Alpha House se prokázalo, že big data lze účinně využít při tvorbě audiovizuálního obsahu, ale také, že konečný úspěch ovlivňuje více faktorů než pouze využití big data. Z tohoto vyplývá, že big data nejsou automaticky předpokladem úspěchu, pokud však dojde k jejich využití kombinaci s dalšími prvky a faktory, mohou výrazně podpořit úspěch audiovizuálního díla.

Například kreativní ztvárnění seriálu House of Cards bylo primárně ovlivněno sběrem big data. Reálné využití podkladů z analýzy big data zde rozhodně nebylo zanedbatelné a seriál dosáhl výrazného úspěchu u svého publika, čímž přinesl společnosti Netflix podstatné tržby a nové zákazníky, kteří konzumují obsah na VoD platformě do dnešní doby. Použití big data umožnilo nalézt

vzorci zaručující úspěch audiovizuálního díla. Nelze souhlasit, že big data jsou jen nafouknutou bublinou, která slouží zejména jako pojem v marketingových strategiích inovativních institucí, jež se snaží tvářit technologicky vyspěle, avšak reálné využití podkladů vyhodnocených po sběru dat je zanedbatelné.

C. VOD platformy budou nadále doplňkovou službou k masové TV a kino distribuci a budou přinášet zajímavý doplňkový veřejně dostupný i placený prémiový obsah, který budou vyhodnocovat na základě využití podkladů získaných od uživatelů

Tuto hypotézu lze vyvrátit. VoD platformy mění paradigma fungování audiovizuálního trhu a budou do jeho podoby výrazně promlouvat i v budoucnosti. Klasickým televizním stanicím a kinům sice postupně přebírají zákaznicky díky většímu množství výhod v podobě personalizované nabídky a pohodlnosti sledování, avšak tradiční média potvrzují svojí roli na trhu již několik desetiletí v řadě. Dá se předpokládat, že zájem o kino distribuci zcela nevyumizí, neboť část diváků bude stále preferovat tento druh konzumace audiovizuálních děl z důvodu sdíleného zážitku.

Klasické masové televizní stanice či kino distributoři nemají takové technologické možnosti k cílení produktu a musí spoléhat na jiné postupy, jak získat a oslovit zákazníky, tyto „klasické“ postupy však mohou být v dnešní době už neúčinné. Nelze souhlasit s tím, že VoD platformy budou pouze doplňkovou službou k masové TV a kino distribuci. Z vývoje na trzích ve Spojených státech, ve Velké Británii, a částečně i v České republice vyplývá, že jsou služby VoD platforem ve stále větší míře poptávané a žádané. V tuto chvíli sice nemůžeme zcela potvrdit jejich úplné budoucí převzetí trhu a tak i zánik klasických médií, ukazuje se však, že VoD platformy už nejsou pouze catch-up doplňkovou službou pro sledování odloženého nebo bonusového obsahu, nýbrž mají svou vlastní diváckou základnu a jsou nezávislé na stávajících mediálních domech nebo hráčích ve filmovém průmyslu, kteří distribuují obsah klasickým způsobem.

Závěr

Cílem práce bylo zhodnotit možnosti potenciálního využití big data v audiovizuální tvorbě v budoucnosti, a to zejména v oblasti VOD (video on demand tedy video na vyžádání).

V rámci práce se podařilo zjistit, že má využití big data v audiovizuální tvorbě velký potenciál, protože přináší nesporné výhody těm platformám, které s big data dokáží aktivně a efektivně pracovat. Big data v audiovizuální tvorbě vytvářejí podklady pro lepší rozhodování, které zase pomáhá k získání konkurenční výhody. Umožňují identifikovat realistické cíle a ukazují, jakým způsobem takové cíle naplnit. Tvůrci audiovizuálního obsahu pak mohou svoje filmy nebo seriály přizpůsobit konkrétním trhům a cílit na ně prostřednictvím přímé reklamy.

Další výhodou big data v audiovizuální tvorbě je zvýšená produktivita, která plyne ze schopnosti využít a interpretovat velké množství dat, a umožňuje tak rychlou reakci na poptávku zákazníků – prostřednictvím big data lze mimo jiné určit nejvhodnější den pro vydání filmu. Data v kontextu AV tvorby pomáhají tvůrcům audiovizuálního obsahu snižovat náklady, protože prostřednictvím analýzy big data dochází ke zvýšení provozní efektivity a snížení výrobních nákladů či marketingových nákladů. Analytika umožňuje vyhnout se nesprávnému rozhodnutí, které by vedlo ke ztrátě. Prostřednictvím big data tak mohou tvůrci audiovizuálního obsahu přinášet zákazníkům vylepšené služby. Všechna big data, pokud jsou správně interpretována, následně vedou k přizpůsobení se přáním a potřebám zákazníků.

Prostřednictvím analýzy big data mohou tvůrci audiovizuálního obsahu snáze měnit svoji strategii a taktiku, aby se jim podařilo udržovat svoji konkurenční výhodu a získat náskok před konkurencí. Neustálá inovace se stává předpokladem úspěchu na trhu, a právě big data toto umožňují. Jejich

analýza a interpretace vede k uvádění nových produktů a služeb na trh rychlejším tempem.

Práce popisuje nejen stav trhu v českém a globálním kontextu, ale přináší také stručnou historii vybraných platform. Dále rozebírá rozdíly mezi veřejnoprávní platformou a komerční platformou, a to z hlediska způsobu a motivace využití dat při vývoji, výrobě a distribuci.

Kromě výčtu pozitivních jevů, které data při správném použití mohou přinášet, je nutné zmínit, že jsou s nimi běžně spojené také velké náklady na investice nejen v podobě hardwarových a softwarových řešení, ale – a to především – v podobě odborného personálu. Struktury, které s daty pracují, vytvářejí IT odborníci, a data jsou hlavním nástrojem především pro datové analytiku, kteří je pomáhají správně interpretovat. Bez této interpretace mohou být sebelepší data bezcenná.

Pomyslná finální třetina práce je věnována problematice právní ochrany uživatelů, ale i možnostem právní ochrany podnikatelské investice VOD platform. V krátkosti předkládá také úvahy o etice práce s daty uživatelů.

O uživatelských datech se někdy hovoří také jako ropě 21. století, neboť jsou pomyslně těžena od uživatelů a jejich zpracování může mít hodnotu zlata. Práce však ukazuje, že samotná data na poli audiovizuální tvorby nestačí a vždy je důležité zachovávat nenahraditelný kreativní přístup autorů a intuici producentů.

Práce nahlíží na problematiku dat a VoD z mnoha úhlů pohledu, od uživatelského až po manažersko-provozní. Nabízí vhled do technických překážek, rizik i výhod, a může být předzvěstí dalšího výzkumu, který se bude věnovat analýze a sestavení funkčního modelu inovativního řešení nové VoD platformy na trhu.

Seznam použitých zdrojů

AKADEMIE VĚD ČESKÉ REPUBLIKY. Big Data. *Www.avcr.cz/* [online]. 2019 [cit. 2020-04-17]. Dostupné z: <http://www.avcr.cz/opencms/export/sites/avcr.cz/cs/veda-a-vyzkum/avex/files/1-2019-AVex-Big-Data.pdf>

AMRIT, Pal Singh a GURVINDER, Singh. Analysis of Amazon Product Reviews Using Big Data- Apache Pig Tool. *International Journal of Information Engineering and Electronic Business* [online]. 2019, vol. 12, no. 1, s. 11. ISSN 20749023.

AUST, Ondřej. Český filmový trh zažil rekordní rok, ale reklama je na dně. *Www.mediar.cz* [online]. 2019 [cit. 2020-05-07]. Dostupné z: <https://www.mediar.cz/cesky-filmovy-trh-zazil-rekordni-rok-ale-reklama-je-na-dne/>

Barvíř, T., Hampl, J., & Melišová, Š. (2011). *Základy práce s počítačem a kancelářskými programy* (První vydání. vyd.). Praha: Grada Publishing.

Blumenthal, E., Katzmaier, D., & Pendlebury, T. (12. Červen 2020). *HBO Max vs. HBO Now vs. HBO Go: What are the differences and how do you upgrade?* Získáno 12. Červen 2020, z Cnet.com: <https://www.cnet.com/news/hbo-max-vs-hbo-now-vs-hbo-go-what-are-differences-how-do-you-upgrade-friends-elmo-roku-amazon-fire-tv/>

Bém, R. (29. Prosinec 2011). *ČT letos spustila nové webové stránky i ceněnou aplikaci iVysílání.* Získáno 11. Červen 2020, z Česká televize: <https://ct24.ceskatelevize.cz/media/1201162-ct-letos-spustila-nove-webove-stranky-i-cenenou-aplikaci-ivysilani>

BERGER, Jiří. Co jsou BIG DATA a co nám mohou přinést. *Www.businessinfo.cz/* [online]. 2016 [cit. 2020-04-15]. Dostupné z: <https://www.businessinfo.cz/clanky/co-jsou-big-data-a-co-nam-mohou-prinest>

BIFET, Albert. Mining big data in real time. *Informatica*, 2013, 37.1.

BOULTON-WALLACE, Steve. *Analytics Strategy to Launch HBO Max Streaming* [konference]. Orange, USA: Filmmaking by the Numbers: How Data is Revolutionizing the Media. 15. 11. 2019

BUGHIN, Jacques. Big data, Big bang?. *Journal of Big Data*, 2016, 3.1: 2.

BULYGO, Zach. How Netflix Uses Analytics To Select Movies, Create Content, and Make Multimillion Dollar Decisions. *Neilpatel.com* [online]. 2019 [cit. 2020-05-06]. Dostupné z: <https://neilpatel.com/blog/how-netflix-uses-analytics/>

BURIAN, Pavel. *Internet inteligentních aktivit*. Praha: Grada, 2014. Průvodce (Grada). ISBN 978-80-247-5137-5.

BRANDON, John. How 'Big Data' could help TV networks make better shows. *Www.foxnews.com/* [online]. 2014 [cit. 2020-05-06]. Dostupné z: <https://www.foxnews.com/tech/how-big-data-could-help-tv-networks-make-better-shows>

BRYCHTA, Jan. Evropská unie zveřejnila nová pravidla pro VOD služby i televizní reklamu. *Www.lupa.cz* [online]. 2018 [cit. 2020-04-17]. Dostupné z: <https://www.lupa.cz/aktuality/evropska-unie-zverejnila-nova-pravidla-pro-vod-sluzby-i-televizni-reklamu/>

Čansado, J. M. (6. Červen 2017). *WILL INTERNET TV KILL IPTV?* Získáno 2. Červen 2020, z Disruption Matters: <http://disruptionmatters.com/2008/10/13/will-internet-tv-kill-iptv/>

Couric, K. (Květen. 31 2017). *Katie Couric and Beau Willimon (Rebroadcast)*. Získáno 1. Červen 2020, z The Aspen Institute: <https://stage2017.aspeninstitute.org/podcasts/katie-couric-beau-willimon-rebroadcast/>

Čapek, M. (12. Červen 2019). Právní aspekty big data. Praha, Česká republika: Univerzita Karlova.

Česká televize. (6. Červen 2020). *iVysílání*. Získáno 6. Červen 2020, z Česká Televize: <https://www.ceskatelevize.cz/ivysilani>

Česká televize. (Leden 2020). *Základní principy vývoje originálního online obsahu v CT*. Získáno 12. Červen 2020, z Česká televize: <https://www.ceskatelevize.cz/vse-o-ct/podavani-nametu-a-projektu/pro-autory/>

Čížek, J. (21. Duben 2016). *Seznam zveřejňuje analytická data v aplikaci Reporter. Podívejte se, jak používáme třeba Mapy.cz Více na:* <https://www.zive.cz/bleskovky/seznam-zverejnuje-analyticka-data-v-aplikaci-reporter-podivejte-se-jak-pouzivame-treba-mapycz/sc-4-a-182152/default.asp>. Získáno 5. Červen 2020, z Živě.cz: <https://www.zive.cz/bleskovky/seznam-zverejnuje-analyticka-data-v-aplikaci-reporter-podivejte-se-jak-pouzivame-treba-mapycz/sc-4-a-182152/default.aspx>

ČESKÝ STATISTICKÝ ÚŘAD. KULTURNÍ PRŮMYSLY V ČR: AUDIOVIZUÁLNÍ A MEDIÁLNÍ SEKTOR 2018. *Www.czso.cz* [online]. 2019 [cit. 2020-03-12]. Dostupné z: <https://www.czso.cz/documents/10180/90577085/090004-19.pdf/19665b09-cf3b-42be-a746-8b901d365034?version=1.1>

ČESKÝ STATISTICKÝ ÚŘAD. Využívání informačních a komunikačních technologií v domácnostech a mezi jednotlivci – 2019. *Www.czso.cz* [online]. 2019 [cit. 2020-03-12]. Dostupné z: <https://www.czso.cz/documents/10180/90577057/06200419.pdf/6f465ecb-ec36-492e-bcc8-fd136d154cd1?version=1.1>

DAI, Xin. Alpha House vs. House of Card: How Data Analysis Helped Netflix But Not Amazon. *Medium.com/* [online]. 2017 [cit. 2020-05-07]. Dostupné z: <https://medium.com/@cxdai99/alpha-house-vs-house-of-card-how-data-analysis-helped-netflix-but-not-amazon-92c71eb55b3e>

DANIEL, Thomas. How video-on-demand tech is changing the TV landscape. *Www.bbc.com* [online]. 2015 [cit. 2020-05-07]. Dostupné z: <https://www.bbc.com/news/business-34050865>

DANZIGER, Caleb. How Amazon Used Big Data to Rule E-Commerce. *Insidebigdata.com* [online]. 2019 [cit. 2020-04-15]. Dostupné z: <https://insidebigdata.com/2019/11/30/how-amazon-used-big-data-to-rule-e-commerce/>

DAVIES, Aran. Pros and Cons of Using Big Data in Movie Making. *Sofy.tv/* [online]. 2019 [cit. 2020-05-07]. Dostupné z: <https://sofy.tv/blog/pros-cons-using-big-data-movie-making/>

DELL TECHNOLOGIES. Filmmaking by the Numbers: How Data is Revolutionizing the Media Industry – 2 day Entertainment Analytics Summit. 2019.

DELOITTE. Filmy a hry: Zábava v Česku z pohledu ekonomiky. *Www2.deloitte.com/* [online]. 2020 [cit. 2020-04-12]. Dostupné z: <https://www2.deloitte.com/cz/cs/pages/about-deloitte/articles/film-hry-zabava-ekonomika.html>

Demerjian, D. (15. Březen 2007). *Rise of the Netflix Hackers*. Načteno z Wired: <https://www.wired.com/2007/03/rise-of-the-netflix-hackers/>

Dredge, S. (8. Duben 2019). *MIPTV Wrap 1: France Country of Honour keynotes, TV trends, 4K and more*. Získáno 13. Červen 2020, z *miptrends*: <https://mipblog.com/2019/04/miptv-wrap-1-france-country-of-honour-keynotes-tv-trends-4k-and-more/>

DESOUZA, Kevin C.; SMITH, Kendra L. Big data for social innovation. *Stanford Social Innovation Review*, 2014, 12.3: 38-43.

DUKE, Jennifer. The Netflix effect: TV networks' shifting video on demand strategies. *Www.smh.com.au* [online]. 2018 [cit. 2020-05-07]. Dostupné z: <https://www.smh.com.au/business/companies/the-netflix-effect-tv-networks-shifting-video-on-demand-strategies-20180218-p4z0r6.html>

DYE, Melody. Stanford at Netflix: From Script to Screen. In: *Youtube* [online]. Stanford (California): Netflix UI Engineering, 2019, 19.11.2019 [cit. 2020-05-31]. Dostupné z: <https://www.youtube.com/watch?v=cR86ebSymXc>

EUROPEAN COMMISSION. *ON-DEMAND AUDIOVISUAL MARKETS IN THE EUROPEAN UNION*. EUROPEAN UNION: Brusel, 2014. ISBN 978-92-79-38425-7.

FERNÁNDEZ-MANZANO, Eva-Patricia; NEIRA, Elena; CLARES-GAVILÁN, Judith. Data management in audiovisual business: Netflix as a case study. *El profesional de la información (EPI)*, 2016, 25.4: 568-576.

FLORIDI, Luciano. *Čtvrtá revoluce: jak infosféra mění tvář lidské reality*. Přeložil Čestmír PELIKÁN. Praha: Univerzita Karlova, nakladatelství Karolinum, 2019. Studia nových médií. ISBN 978-80-246-3803-4.

FRITZ, Ben, *The Big Picture: The Fight for the Future of Movies*. USA: Eamon Dolan, 2018. ISBN 978-05447889760.

Gaylor, B. (Režisér). (2015). *Do Not Track* [Film]. Arte.

Gilbert, B., & Webb, K. (27. Květen 2020). *HBO Max vs. HBO Go vs. HBO Now: The key differences explained*. Získáno 2. Červen 2020, z Business Insider: <https://www.businessinsider.com/hbo-go-vs-hbo-now-vs-hbo-max>

Gregersen, E. (19. Květen 2020). *HBO*. Získáno 12. Červen 2020, z Encyclopædia Britannica: <https://www.britannica.com/topic/HBO>

Grunt, D. (12. Listopad 2019). *HbbTV - klíčová platforma televizního VOD?* Získáno 23. květen 2020, z Innovation Day 2019: <https://i-day.cz/admin/fileGet.aspx?f=islnkpzya>

GEMIGNANI, Zach, Chris GEMIGNANI, Richard GALENTINO a Patrick Jude SCHUERMANN. *Efektivní analýza a využití dat*. Přeložil Jiří HUF. Brno: Computer Press, 2015. ISBN 978-80-251-4571-5.

GOLEMAN, Daniel. *Pozornost: skrytá cesta k dokonalosti*. V Brně: Jan Melvil, 2014. Pod povrchem. ISBN 978-80-87270-94-3.

GOODWORKLABS. How Big Data is Revolutionizing the Entertainment Industry. *Www.goodworklabs.com/* [online]. 2020 [cit. 2020-04-16]. Dostupné z: <https://www.goodworklabs.com/big-data-in-entertainment/>

GOYAL, Ritu. Stanford at Netflix: From Script to Screen. In: *Youtube* [online]. Stanford (California): Netflix UI Engineering, 2019, 19.11.2019 [cit. 2020-05-31]. Dostupné z: <https://www.youtube.com/watch?v=cR86ebSymXc>

HALADA, Jan a Barbora OSVALDOVÁ, ed. *Slovník žurnalistiky: výklad pojmů a teorie oboru*. Praha: Univerzita Karlova, nakladatelství Karolinum, 2017. ISBN 978-80-246-3752-5.

HAVENS, Timothy. Media programming in an era of big data. *Media Industries Journal*, 2014, 1.2.

HAZELTON, John. Big film, big data: how analytics is shaping the business. *Www.screendaily.com* [online]. 2016 [cit. 2020-05-07]. Dostupné z: <https://www.screendaily.com/features/big-film-big-data-how-analytics-is-shaping-the-business/5104922.article>

HBO. Filmmaking by the Numbers: How Data is Revolutionizing the Media Industry – 2 day Entertainment Analytics Summit. 2019.

Hlaváčová, V. (23. Březen 2018). *Dopad kauzy Facebook? 55 procent Čechů podle průzkumu změni své chování, na síti zůstanou*. Získáno 11. Červen 2020, z iRozhlas: https://www.irozhlas.cz/zpravy-domov/cesi-na-facebooku-chovani-na-socialnich-sitich-pruzkum-median-cesky-rozhlas_1803281130_haf

HOLUBOVÁ, Irena et al. *Big Data a NoSQL databáze*. Praha: Grada, 2015. Profesionál. ISBN 978-80-247-5466-6.

HOSCH, William L. Hulu. *Www.britannica.com* [online]. 2020 [cit. 2020-05-08]. Dostupné z: <https://www.britannica.com/topic/Hulu>

INSIDEBIGDATA. How Netflix Uses Big Data to Drive Success. *Insidebigdata.com* [online]. 2018 [cit. 2020-05-06]. Dostupné z: <https://insidebigdata.com/2018/01/20/netflix-uses-big-data-drive-success/>

JANČA, Jan. Big data: Malé chyby, velký problém. *Connect.zive.cz* [online]. 2015 [cit. 2020-04-15]. Dostupné z: <https://connect.zive.cz/clanky/big-data-male-chyby-velky-problem/sc-320-a-176646/default.aspx>

JEŽEK, Mojmir. Big data a pokročilá analytika datových souborů. *Www.ecovislegal.cz/* [online]. 2018 [cit. 2020-04-15]. Dostupné z: <https://www.ecovislegal.cz/aktuality/big-data-a-pokrocila-analytika-datovych-souboru/>

JOHNSON, Ben. Big data – the recipe for movie success? *Www.gruvi.tv* [online]. 2016 [cit. 2020-05-06]. Dostupné z: <https://www.gruvi.tv/post/big-data-movies>

JOHNSON, Derek. Reluctant Retailing: Amazon Prime Video and the Non-Merchandising of Kids' Television. *Film Criticism* [online]. 2018, vol. 42, no. 2. ISSN 01635069.

Jain, A. (17. Září 2016). *The 5 V's of big data*. Získáno 12. Duben 2020, z IBM: <https://www.ibm.com/blogs/watson-health/the-5-vs-of-big-data/>

Jetmar, J. (15. Květen 2020). *Vydavatelé s tvorbou vlastních seriálů skončili*. Získáno 3. Červen 2020, z Mediář: https://www.mediar.cz/vydavatele-s-tvorbou-vlastnich-serialu-skoncili/?fbclid=IwAR3CLOCaqN0q7dGIcPvlup8qF5wh5qdi4sYy1nw-oK0sm0ud3E_YDWxphcw

Kafka, P., & Molla, R. (8. Duben 2019). *Millions of Netflix users are using someone else's password, but that's not bad for Netflix*. Získáno 9. Červen 2020, z Vox: <https://www.vox.com/2019/4/8/18300256/netflix-password-sharing-14-percent-survey-content-experience-moffettnathanson>

Kapustina, L. (12. Září 2015). *VOD RECOMMENDATION FOR OTT VIDEO PLATFORMS*. Načteno z AI Ukraine: https://aiukraine.com/2015/presentations/2015/09/Lyubov_Kapustina.pdf

Kartner, M. (15. Srpen 2015). *Právní regulace cookies v České republice*. Získáno 4. Červen 2020, z epravo.cz: <https://www.epravo.cz/top/clanky/pravni-regulace-cookies-v-ceske-republice-98406.html>

Kimmel, R., Bronstein, M., & Bronstein, A. (27. března 2010). *The Video Genome*. Získáno 24. Květen 2020, z arXiv® Cornell University: <https://arxiv.org/pdf/1003.5320v1>

Kollárová, D. (Prosinec 2019). *Informace o činnosti úseku Nová média v roce 2019*. Získáno 7. Červen 2020, z Česká televize: <https://img.ceskatelevize.cz/boss/document/1578.pdf?v=1>

KLOUDOVÁ, Jitka. *Kreativní ekonomika: Trendy, výzvy, příležitosti*. Praha: Grada Publishing, 2010, 224 s. ISBN 9788024773452.

KOLAJO, Taiwo, Olawande DARAMOLA a Ayodele ADEBIYI. Big data stream analysis: a systematic literature review. *Journal of Big Data* [online]. 2019, vol. 6, no. 1, s. 1-30.

KUMAR, Ritwik. Data Science and the Art of Producing Entertainment at Netflix. *Netflixtechblog.com* [online]. 2018 [cit. 2020-05-06]. Dostupné z: <https://netflixtechblog.com/studio-production-data-science-646ee2cc21a1>

KŘÍŽ, Jonáš. Trump i brexit. Úspěšnou kampaň vedla umělá inteligence. *Www.euro.cz* [online]. 2017 [cit. 2020-04-15]. Dostupné z: <https://www.euro.cz/politika/trump-i-brexit-uspesnou-kampan-vedla-umela-inteligence-1336517>

LAMONT, Ian. Creating hit programs at Netflix and Amazon: It's not just about data. *Leanmedia.org/* [online]. 2016 [cit. 2020-05-07]. Dostupné z: <http://leanmedia.org/creating-hit-programs-at-netflix-and-amazon-its-not-just-about-data/>

LAZZARO, Sage. Netflix Purposely Designed 'House of Cards' to Be a Major Hit—Here's How They Did It. *Observer.com* [online]. 2016 [cit. 2020-05-07]. Dostupné z: <https://observer.com/2016/01/can-we-use-big-data-to-create-hit-tv-shows-as-addictive-as-breaking-bad/>

MACEK, Jakub. Traditional and Convergent Domestic Audiences: Towards a Typology of the Transforming Czech Viewership of Films and TV Series. *Illuminace*, 2017, 29.2.

MACNAB, Geoffrey. What are the pros and cons of using Big Data in the film and TV business? *Www.screendaily.com* [online]. 2018 [cit. 2020-04-17]. Dostupné z: <https://www.screendaily.com/news/what-are-the-pros-and-cons-of-using-big-data-in-the-film-and-tv-business/5133082.article>

MADDODI, Srivatsa, et al. Netflix Bigdata Analytics-The Emergence of Data Driven Recommendation. 2019.

MALATINSKÝ, Pavel. Nestrukturovaná data a GDPR compliance. *Www.pravniprostor.cz* [online]. 2018 [cit. 2020-04-15]. Dostupné z: <https://www.pravniprostor.cz/clanky/ostatni-pravo/nestrukturovana-data-a-gdpr-compliance>

MARR, Bernard. *Big data in practice: how 45 successful companies used big data analytics to deliver extraordinary*. West Sussex, United Kingdom: John Wiley and Sons, 2016. ISBN 9781119231387.

MARR, Bernard. Netflix Used Big Data To Identify The Movies That Are Too Scary To Finish. *Www.forbes.com* [online]. 2018 [cit. 2020-05-06].

Dostupné z:

<https://www.forbes.com/sites/bernardmarr/2018/04/18/netflix-used-big-data-to-identify-the-movies-that-are-too-scary-to-finish/>

McAlone, N. (15. Červen 2016). *Why Netflix thinks its personalised recommendation engine is worth \$1 billion per year*. Získáno 12. Červen 2020, z Business Insider Australie:

<https://www.businessinsider.com.au/netflix-recommendation-engine-worth-1-billion-per-year-2016-6>

MacDonald, C. (9. Prosinec 2015). *America really DOES love to Netflix and chill: Site now accounts for 37% of all US broadband traffic and video takes 70% overall*. Získáno 6. Červen 2020, z Daily Mail:

<https://www.dailymail.co.uk/sciencetech/article-3351849/America-really-DOES-love-Netflix-chill-Site-accounts-37-broadband-traffic-video-takes-70-overall.html>

MallTV. (Říjen 2019). VIDEO & DISPLAY PURCHASE UPLIFT SOLUTION. Praha.

Matoušek, V. (12. Červen 2020). Rozhovor o sběru a využití dat se zaměřením na Seznam, Mall a CNC. (J. Rálek, Tazatel) Praha, Česká republika.

Mesenberg, K. (1. Listopad 2019). The Internet of Life - Media for after tomorrow. (J. Rálek, Tazatel)

Miller, B. (18. Listopad 2019). *What Kind of Hits Does Google Analytics Track?* Získáno 3. Červen 2020, z Bryttdesigns.com:

<https://bryttdesigns.com/hits-google-analytics-track/>

MICHL, Petr. Jak funguje Cambridge Analytica: příběh zneužívání dat uživatelů a jejich strachu. *Www.focus-age.cz* [online]. 2018 [cit. 2020-04-15]. Dostupné z: https://www.focus-age.cz/m-journal/internet/jak-funguje-cambridge-analytica--pribeh-zneuzivani-dat-uzivatelu-a-jejich-strachu__s281x13594.html

MIKEŠOVÁ, Markéta. Lidé chtěli vědět, jak Netflix u interaktivních filmů ctí GDPR. Výsledek nepotěší. *Www.zive.cz* [online]. 2019 [cit. 2020-05-07]. Dostupné z: <https://www.zive.cz/clanky/lide-chteli-vedet-jak-netflix-u-interaktivnich-filmu-cti-gdpr-vysledek-nepotesi/sc-3-a-197215/default.aspx>

MORDOR INTELLIGENCE. Video on Demand Market - Growth, Trends, and Forecast (2020 - 2025). *Www.mordorintelligence.com* [online]. 2020 [cit. 2020-05-07]. Dostupné z: <https://www.mordorintelligence.com/industry-reports/video-on-demand-market>

MUSO. A Culture of Analytics: How Big Data is Shaking Up the Film & TV. *Www.muso.com* [online]. 2020 [cit. 2020-04-16]. Dostupné z: <https://www.muso.com/magazine/big-data-netflix-amazon>

MURGIA, M. Netflix, YouTube, Amazon and Apple accused of GDPR breach. *Www.ft.com* [online]. [cit. 2020-05-07]. Dostupné z: <https://www.ft.com/content/4a977276-1b3e-11e9-b93e-f4351a53f1c3>

Netflix. (18. 9 2009). *Netflix Prize: Forum*. Získáno 31. 5 2020, z Netflix Prize: <http://www.netflixprize.com/community/viewtopic.php?id=1537>

Netflix, Inc. (2020). *2020 Quarterly Earnings*. Získáno 20. Květen 2020, z Netflix Investors: <https://www.netflixinvestor.com/financials/quarterly-earnings/default.aspx>

Netflix, Inc. (2020). *Interaktivní obsah na Netflixu*. Získáno 11. Červen 2020, z Netflix - Centrum zákaznické podpory: <https://help.netflix.com/cs/node/62526>

NetMonitor. (3. Červen 2020). *Online data (OLA)*. Získáno 3. Červen 2020, z NetMonitor: <https://www.netmonitor.cz/online-data-ola>

ŠKORNIČKOVÁ, Eva. Narazí umělá inteligence u nařízení GDPR? *Www.gdpr.cz/* [online]. 2018 [cit. 2020-04-15]. Dostupné z: <https://www.gdpr.cz/blog/umela-inteligence/>

SADEH, Gabrielle. How Netflix uses big data to create content and enhance user experience. *Www.clickz.com* [online]. 2019 [cit. 2020-05-06]. Dostupné z: <https://www.clickz.com/how-netflix-uses-big-data-content/228201/>

SIDDIQA, Aisha, et al. A survey of big data management: Taxonomy and state-of-the-art. *Journal of Network and Computer Applications*, 2016, 71: 151-166.

SZCZEPANIK, Petr, Pavel Zahrádka (eds.), Radim Polčák, Jakub Macek, Ivan David, Dominika Galajdová, Johana Kotišová, Rudolf Leška, Petr Lupač, Tomáš Karger a Štěpán Žádník. *Mapa audiovizuálního pole v České republice z hlediska digitalizace a strategie pro jednotný digitální trh*. 1. vyd. Olomouc: Univerzita Palackého v Olomouci, 2018. 93 s. ISBN 978-80-244-5450-4.

SMITH, Michael D. a Rahul Telang. *Streaming, sharing, stealing: big data and the future of entertainment*. Cambridge, Massachusetts: MIT Press, 2017. ISBN 978-0-262-53452-9.

SOBERS, Rob. The Average Reading Level of a Privacy Policy. *Www.varonis.com* [online]. 2020 [cit. 2020-05-07]. Dostupné z: <https://www.varonis.com/blog/gdpr-privacy-policy/>

SØRENSEN, Jannick Kirk; HUTCHINSON, Jonathon. Algorithms and public service media. *Public Service Media in the Networked Society RIPE*, 2017, 91-106.

SPANGLER, Todd. Hulu Acquires Video Genome Project to Enhance Recommendations for Live TV, VOD. *Variety.com/* [online]. 2016 [cit. 2020-05-08]. Dostupné z: <https://variety.com/2016/digital/news/hulu-acquires-video-genome-project-vgp-1201918435/>

SUBÍAS, Mónica Herrero, Mercedes Medina LAVERÓN a Alicia María Urgellés MOLINA. Online recommendation systems in the Spanish Audiovisual market: comparative analysis between Atresmedia, Movistar+ and Netflix. *Universia Business Review* [online]. 2018, no. 60, s. 54-89. ISSN 16985117.

MURSCHETZ, Paul a Daniela SCHLÜTZ. Big Data and Television Broadcasting. A Critical Reflection on Big Data's Surge to Be-come a New Techno-Economic Paradigm and its Impacts on the Concept of the «Addressable Audience». *Fonseca* [online]. 2018, no. 17, s. 23-38.

TAJBR, Jaroslav. KOMENTÁŘ: Evropská unie srovnala podmínky audiovize. *Www.lidovky.cz* [online]. 2019 [cit. 2020-04-17]. Dostupné z: https://www.lidovky.cz/byznys/pravo-a-justice/komentar-evropska-unie-srovnala-podminky-audiovize.A190127_154228_In_byznys_pravo_ssu

TED. How Netflix changed entertainment -- and where it's headed | Reed Hastings. In: *YouTube* [Video]. USA: YouTube, 2018, 12.7.2018 [cit. 2020-05-31]. Dostupné z: <https://www.youtube.com/watch?v=LsAN-TEJfN0>

TRADING.CO.UK. 9 amazing ways big data influences our everyday lives. *Medium.com/* [online]. 2017 [cit. 2020-05-06]. Dostupné z: <https://medium.com/@tradingwebsite/9-amazing-ways-big-data-influences-our-everyday-lives-8b4deb7177c2>

Pelánek, R. (5. Květen 2011). *Doporučující systémy*. Získáno 14. Duben 2020, z Vesmír: <https://vesmir.cz/cz/casopis/archiv-casopisu/2011/cislo-5/doporuucujici-systemy.html>

Plummer, L. (22. Srpen 2017). *This is how Netflix's top-secret recommendation system works*. Získáno 10. Červen 2020, z Wired: <https://www.wired.co.uk/article/how-do-netflixs-algorithms-work-machine-learning-helps-to-predict-what-viewers-will-like>

PPF. (6. Červen 2020). *Mall Group je největší e-commerce platforma ve střední a východní Evropě, s obratem přes 500 milionů eur ročně*. Načteno z PPF: <https://www.ppf.eu/cs/companies/mall-group>

Rožánek, F. (30. Březen 2020). *Spokojenost s pořady ČT stoupla na 83 procent, její videa vedou na webu*. Získáno 12. Červen 2020, z Lupa.cz: <https://www.lupa.cz/clanky/spokojenost-s-porady-ct-stoupla-na-83-procent-jeji-video-vedou-na-webu/>

Sedlák, J. (19. Březen 2020). *Křetínský vstupuje do „albánského Seznamu“ Gjirafa. CNC se spojuje s Mall TV*. Získáno 28. Květen 2020, z Lupa.cz: <https://www.lupa.cz/aktuality/kretinsky-vstupuje-do-albanskeho-seznamu-gjirafa-cnc-se-spojuje-s-mall-tv/>

Singel, R. (17. Prosinec 2019). *Netflix Spilled Your Brokeback Mountain Secret, Lawsuit Claims*. Načteno z Wired: <https://www.wired.com/2009/12/netflix-privacy-lawsuit/>

Slížek, D. (15. Březen 2018). *Lukáš Záhoř (Mall Group): Neděláme kopii Streamu. A míříme i za hranice Česka*. Získáno 18. Květen 2020, z Lupa.cz: <https://www.lupa.cz/clanky/lukas-zahor-mall-group-nedelame-kopii-streamu-a-mirime-i-za-hranice-ceska/>

Smith, M. (2. Listopad 2018). *Is Big Data Killing Creativity? | Michael Smith | TEDxHarvardCollege*. Získáno 3. Červen 2020, z YouTube: <https://www.youtube.com/watch?v=A1XibEzp6K0>

Snížek, M. (5. Květen 2011). *A/B testování – kompletní průvodce*. Získáno 10. Květen 2020, z optimics: <https://www.optimics.cz/ab-testovani-kompletni-pruvodce/>

Sprangler, T. (10. Září 2019). *Netflix Bandwidth Consumption Eclipsed by Web Media Streaming Applications*. Získáno 6. Červen 2020, z Variety: <https://variety.com/2019/digital/news/netflix-loses-title-top-downstream-bandwidth-application-1203330313/>

Szczepanik, P., & Zahrádka (eds.), P. (2018). ÚVOD: AUDIOVIZUÁLNÍ POLE A DIGITALIZACE. V P. Szczepanik, & P. Zahrádka (eds.), *Mapa audiovizuálního pole v České republice z hlediska digitalizace a strategie pro jednotný digitální trh*. (stránky 4-14). Olomouc: Univerzita Palackého v Olomouci.

PARK, Sungwook; KWON, Youngsun. *Research on the Relationship between the Growth of OTT Service Market and the Change in the Structure of the Pay-TV Market*. 2019.

PRESSNELL, Josh. Better Video Recommendations Require AI and the Smart Use of Data. *Www.streamingmedia.com* [online]. 2019 [cit. 2020-05-08]. Dostupné z: https://www.streamingmedia.com/Articles/Editorial/Featured-Articles/Better-Video-Recommendations-Require-AI-and-the-Smart-Use-of-Data-129942.aspx?utm_source=related_articles&utm_medium=gutenberg&utm_campaign=editors_selection

ROYAL TELEVISION SOCIETY. Why big data is changing TV. *Rts.org.uk* [online]. 2016 [cit. 2020-05-06]. Dostupné z: <https://rts.org.uk/article/why-big-data-changing-tv>

Šiler, J. (29. květen 2020). Rozhovor o datové analytice, mediálním trhu, Mall.TV, CNC, Seznamu a dalších službách v ČR. (J. Rálek, Tazatel)

Uphoff, K. (9. Únor 2018). *Supporting decisions with ML*. Získáno 14. Květen 2020, z Netflix: <https://www.slideshare.net/MeganNeider/supporting-decisions-with-ml>

Upian. (4. Červen 2020). *YOUR DATA*. Načteno z Do Not Track: <https://donottrack-doc.com/en/your-datas/>

Usabilla. (14. Březen 2017). *All A/Bout Testing: How Netflix Knows Which Images You'll Click*. Získáno 17. Květen 2020, z Medium.com: <https://medium.com/@usabilla/all-a-bout-testing-how-netflix-knows-which-images-youll-click-4ba673142cbd>

UserTesting. (2018). *Streaming Media CX Index - A COMPETITIVE BENCHMARKING RESEARCH REPORT, 2018*. Získáno 20. Květen 2020, z UserTesting: <https://info.usertesting.com/rs/709-WMS-542/images/2018-Streaming-MediaCX-Index.pdf>

YLIJOKI, Ossi a Jari PORRAS. Perspectives to Definition of Big Data: A Mapping Study and Discussion. *Journal of Innovation Management* [online]. 2016, vol. 4, no. 1, s. 69-91.

VOGELS, Werner. How Amazon is solving big-data challenges with data lakes. *Siliconangle.com* [online]. 2019 [cit. 2020-04-15]. Dostupné z: <https://siliconangle.com/2020/01/30/amazon-solving-big-data-challenges-data-lakes/>

Vašíčková, D., & Zukal, J. (12. Červen 2020). Rozhovor k tématu připravované videoplatformy ČT. (J. Rálek, Tazatel) Praha, Česká republika.

Vojtěchovská, M. (28. Listopad 2018). *Gebauer: VOD služby začínají nahrazovat placené TV kanály*. Získáno 11. Červen 2020, z Mediaguru: <https://www.mediaguru.cz/clanky/2018/11/gebauer-vod-sluzby-zacinaji-nahrazovat-placene-tv-kanaly/>

Vojtěchovská, M. (12. Červen 2018). *MallTV odstartuje na podzim, asi kolem října*. Získáno 27. Květen 2020, z Mediaguru.cz: <https://www.mediaguru.cz/clanky/2018/06/mall-tv-odstartuje-na-podzim-asi-kolem-rijna/>

Vojtěchovská, M. (18. Únor 2019). *Průzkum: Na čele VOD v Česku Stream, Netflix má téměř milion měsíčně*. Získáno 5. Červen 2020, z MediaGuru.cz: <https://www.mediaguru.cz/clanky/2019/02/pruzkum-na-cele-vod-v-cesku-stream-netflix-ma-temer-milion-mesicne/>

Vranica, S. (4. Červen 2015). *Comcast Has Agreed To Acquire Ad Tech Firm Visible World*. Načteno z The Wall Street Journal: <https://blogs.wsj.com/cmo/2015/06/04/comcast-has-agreed-to-acquire-ad-tech-firm-visible-world/>

Welsh, J. (4. Květen 2010). *Sequencing the Video Genome*. Získáno 27. Květen 2020, z Wired: <https://www.wired.com/2010/04/video-genome/>

What's on Netflix. (15. Květen 2020). *What's on Netflix*. Získáno 19. Květen 2020, z Netflix Library A-Z: <https://www.whats-on-netflix.com/library/>

WEISS, Martin. Velký datový tunel Cambridge Analytica. *Echo24.cz* [online]. 2018 [cit. 2020-04-15]. Dostupné z: <https://echo24.cz/a/Sy9fd/velky-datovy-tunel-cambridge-analytica>

WEST, Emily. Amazon: Surveillance as a Service. *Surveillance & Society* [online]. 2019, vol. 17, no. 1, s. 27-33.

WU, Eddy. Stanford at Netflix: From Script to Screen. In: *Youtube* [online]. Stanford (California): Netflix UI Engineering, 2019, 19.11.2019 [cit. 2020-05-31]. Dostupné z: <https://www.youtube.com/watch?v=cR86ebSymXc>