
Efektivita kategorií videozáznamů na volně dostupných videoportálech

Efficiency of Video Categories on free video sharing websites

Ondřej Voců

Nadace Sirius

Ústav informačních studií a knihovnictví, Filozofická fakulta, Univerzita Karlova v Praze

Recenzenti:

Mgr. Kamil Matula

Mgr. Josef Schwarz

Abstrakt:

Cílem studie je stanovení míry efektivity kategorií videí volně dostupných videoportálů. Pro uvedení do problematiky poslouží výsledky analýzy kategorií volně dostupných videoportálů (YouTube, Dailymotion, Vimeo a Metacafe). Po stanovení jádrových kategorií studie představí výsledky kvantitativního výzkumu vzorku videozáznamů, majícího podobu základních statistických šetření na přítomnost různých jevů. Primárně bude sledováno správné zařazení videozáznamu v kategoriích volně dostupného videoportálu, doplňkově také stanovení relevance videozáznamů vzhledem k rešeršnímu dotazu, určení multiplicity videozáznamů a zjištění dohledatelnosti uživatelů vystavujících videozáznamy. Výzkumný vzorek videozáznamů byl stanoven na základě precizně formulovaných rešeršních dotazů, jež byly jednotné pro všechny volně dostupné videoportály. Studie bude rovněž obsahovat autorská doporučení a náměty nejen ke kategoriím volně dostupných videoportálů.

Klíčová slova: videa, videoportály, webové stránky, volný přístup, kategorie, YouTube, Vimeo, Dailymotion, Metacafe, analýzy, studie

Abstract:

Aim of this paper is to prescribe efficiency rate of video categories of free access video sharing sites. As introduction to the issue will serve analysis' results of categories of free access video sharing sites (YouTube, Dailymotion, Vimeo and Metacafe). After the setting of essential categories, the paper will introduce results of qualitative research of videos' sample, using basic statistic survey to prove different features. Primary focus is to determine correct category insertion on free access video sharing sites, supportive focus is to specify the relevance of videos regarding to search query, specify the multiplicity of videos and tracking survey of users, who posted the videos. Based on precisely formulated search queries, the research sample of videos was set, those queries were the same for all four free access video sharing sites. Paper will also include author's recommendations and suggestions not only to categories of free access video sharing sites.

Keywords: videos, video-sharing portals, websites, free access, categories, YouTube, Vimeo, Dailymotion, Metacafe, analyses, papers

1 Úvod

Volně dostupné videoportály jsou bezesporu fenoménem od počátku milénia až do dnešních dní. Přestože jejich primárními cíli je v duchu zásad webu 2.0 zprostředkovávání specifického obsahu a podnícení sociální interakce v pravdě masovém měřítku, lze na volně dostupné videoportály pohlížet také jako na informační zdroj s videozáznamy, přičemž některé z nich mají odborně a akademicky hodnotný obsah. V tomto smyslu je ale zapotřebí při zacházení s nimi určité vynalézavosti, všudypřítomné opatrnosti a je rovněž nutné si uvědomit, že principy tvorby a fungování většiny volně dostupných videoportálů, jako jsou Metacafe, YouTube, Dailymotion či Vimeo jsou odlišné od knihovnických či informačně-vědních odborných zdrojů. Vstupní zpracování videozáznamů, které je výlučně v rukou uživatele, s sebou přináší problémy procesu vyhledávání a využívání volně dostupných videoportálů samotných.

Není-li v textu uvedeno jinak, slova video, záznam a videozáznam jsou zcela zaměnitelná, stejně jako slovo portál a sousloví videoportál nebo volně dostupný videoportál.

1.1 Cíle a metody práce

Hlavním cílem této studie je zjistit, jak jsou kategorie videí na volně dostupných videoportálech efektivní, tedy do jaké míry jsou videa v kategoriích umístěna správně a díky tomu vyhledatelná. Doplňkovým cílem této studie je zmapování dalších parametrů, jako jsou multiplicita videozáznamů a dohledatelnost autorů videí.

První pracovní metodou této studie je analýza a komparace kategorií volně dostupných videoportálů YouTube, Dailymotion, Vimeo a Metacafe, následovaná vyhodnocením zjištěných skutečností.

Druhou a dominantní metodou studie je vyhodnocení dat z kvantitativního výzkumu. Kvantitativní výzkum má podobu základních statistických šetření na (ne)přítomnost monitorovaných ukazatelů v rámci výzkumných vzorků. Výzkumnými vzorky jsou dvě tematicky vymezené množiny videozáznamů. Každý videozáznam ve vzorcích má svou unikátní webovou adresu a lze se od něj přesunout i na kanál dotyčného uživatele. Výzkumné vzorky byly stanoveny na základě výsledků analýzy kategorií volně dostupných videoportálů a na základě precizně formulovaných vyhledávacích dotazů, které byly společné všem volně dostupným videoportálům. Více o výzkumných vzorcích ve třetí části této studie.

1.2 Současný stav poznání

Volně dostupné videoportály jsou předmětem mnoha studií v odborných periodikách. Tyto studie mají velmi rozmanitou délku, strukturu i použité výzkumné metody. Jejich společným rysem je fakt, že jediným zkoumaným volně dostupným videoportálem je portál YouTube, ostatní videoportály do studií zahrnuté nebyly.

Do první podstatné skupiny lze zařadit všechny studie, v nichž si autor či autorský kolektiv kladl výzkumnou otázku, zda na YouTube existují materiály relevantní tématu jejich výběru. Lze konstatovat, že se tyto studie soustřeďují na povahu relevantních videozáznamů, na charakter entit, které video na YouTube nahrály, a zkoumají sociální interakce u těchto videí. Mezi studie

tohoto typu lze zařadit např. Terantinovu studii [TERANTINO, 2011], v níž se autor soustřeďoval na problematiku výuky cizích jazyků za použití YouTube. Autor došel k závěru, že YouTube pomáhá ve výuce cizích jazyků vyučujícím i studujícím, zejména díky jeho sociálnímu potenciálu. Druhou stručně zmíněnou studií je Tianova studie [TIAN, 2010], v níž autor zkoumal tematiku dárcovství orgánů na YouTube. Výsledky této studie byly pozitivní a překvapující – drtivá většina relevantních videí se vůči dárcovství orgánů stavěla pozitivně a sociální interakce na tato videa byla rovněž pozitivní.

Do druhé skupiny lze zařadit takové studie, které zkoumají určitý jev, jenž je na YouTube prokazatelně vysledovatelný, leč YouTube do jiných vědních oborů přesahující. Mezi tyto studie lze zařadit studii Koushy a kol. [KOUSHA, 2012], v níž autoři zkoumali YouTube a jeho roli ve výzkumné komunikaci. Tato rozsáhlá a podrobná studie, pracující s videy v řádech tisíců jednotek, prokazuje, že do roku 2008 byl YouTube akademiky používán jako zdroj a jeho videa citována exponenciálně, od roku 2008 se nárůst zmírnil, až zastavil, a že ponejvíce YouTube akademici citují v exaktních vědách, medicíně a vědách o zdraví. Druhá stručně představená studie, která byla vypracována Rodriguesem a kol. [RODRIGUES, 2010], se zabývala sociálními reakcemi na duplicitní a multiplicitní videa na YouTube. Hlavními zjištěními studie bylo, že existuje přímá vazba mezi popularitou uživatele a popularitou duplikátů jeho videí a že existují dramatické rozdíly v metadatovém vybavení a obsahu těchto metadat u totožných videí.

2 Videá a jejich kategorie na webu

Kategorie videozáznamů existují na všech volně dostupných videoportálech celosvětového záběru. V odborných databázích jsou při vyhledávání využívány tezaury, předmětové hesláře, anebo lze výsledek vyhledávání omezit typem publikace. Naproti tomu ve světě volně dostupných videoportálů je hlavní důraz kladen na popis videa, jehož (ne)vyplnění nemá žádná striktní pravidla a podmínky. Velký zřetel je rovněž kladen na tagy, které jsou folksonomickým nástrojem k popisu webových zdrojů ze strany uživatelů., a jejichž vyplnění je opět věcí nahrávajícího.

Odborné prameny k problematice pořádání videomateriálů lze rozdělit do dvou hlavních skupin. První skupinu tvoří odborné studie o kategorizaci všeobecné povahy, jako je např. studie Gnoliho a Meie [GNOLI, 2006]. Druhou skupinou studií jsou studie, které do popředí stavějí materiál, tedy videozáznam. Příkladem je studie Huanga, Fua a Chena [HUANG, 2010]. Avšak obě zmiňované studie shodně ve výsledku sdělují, že klasifikaci webových zdrojů, volně dostupných videoportálů též, nelze efektivně zkonstruovat za použití oborové klasifikace, resp. vytvoření klasifikace pomocí stanovení nejvíce zastoupených videomateriálů. Lepší cestou je dle Gnoliho je fasetová klasifikace, kdy pro její tvorbu mohou posloužit textová data, která uživatelé vyplňují při nahrávání videozáznamu na videoportál.

2.1 Pozadí a význam kategorií na volně dostupných videoportálech

Kategorie, přesněji obsahové kategorie na volně dostupných videoportálech byly vytvářeny s ohledem na uživatele. Tyto kategorie nemají více úrovní, neexistují tedy žádné podkategorie, stejně tak princip jejich tvorby není řízen žádným systematickým opatřením. A proto je nezbytné na tyto kategorie pohlížet spíše jako na pomocný, podpůrný nástroj, jímž jsou videoportály mimo dalších funkcí, v duchu zásad webu 2.0, vybaveny.

Je nutné od sebe oddělovat dvě skupiny vizuálních elementů, které lze spatřit pohromadě při pohybu na volně dostupných videoportálech. Tyto skupiny vizuálních elementů mohou být ze strany videoportálů nazývány kategoriemi, nicméně jejich princip je odlišný. Za (pravé, obsahové) kategorie se označují skupiny obsahově příbuzných videí, která jsou při brouzdání

na portále pohromadě pod zastřešením jedné kategorie. Tato kategorie může být u každého z videí zmíněna. Za (nepravé, neobsahové) kategorie, byť tak bývají ze strany volně dostupných videoportálů nazývány, lze označit skupiny videí, která byla seřazena na základě neobsahového kritéria při vstupu na portál nebo při vyhledávání na něm, např. nejlépe hodnocená videa, nejnavštěvovanější videa, nejvíce komentovaná videa, videa či množiny videí sestavené na základě předchozí uživatelské aktivity a další [JACOB, 2004].

Původ kategorií volně dostupných videoportálů lze vystopovat ve dvou oblastech – kategorie televizních pořadů a stanic a materiálů na webu.

Při pohledu i do archivního výtisku televizního programu nebo na tematický profil libovolné vysílací společnosti a jejich televizních kanálů si nelze nepovšimnout, že jednotlivé pořady v rámci televizního programu mívají kategorii, o jaký typ pořadu se jedná, stejně tak televizní společnosti profilují svoje kanály obdobným způsobem, např. sport, zpravodajství, zábava, vzdělání, romantika, krimi, vysílání pro děti, věda, technika atd [HUANG, 2010]. Podle R. Glushka lze kategorie na volně dostupných videoportálech zařadit do skupiny „individuálních kategorií“, které jsou vytvářeny výlučně pro potřebu daného zdroje či systému [GLUSHKO, 2013].

Druhým a neméně důležitým faktorem vzniku některých kategorií na volně dostupných videoportálech je nepsané členění webových zdrojů např. pro sdílení souborů, obrazových materiálů, hudby aj., v duchu zdrojů webu 2.0. Odtud pocházejí kategorie, jako jsou počítačové hry, videoblogy, webkamery, některé firmy mají část www stránek vyhrazenou pro neziskový sektor atd.

Kategorie na volně dostupných videoportálech plní zejména tyto funkce:

- 1) Funkce pořádací – i když to není nejdůležitější úkol, kategorie na volně dostupných videoportálech plní pořádací úlohu, neboť je cílem videa podobných témat zastřešit pod jednu kategorii.
- 2) Funkce vyhledávací – jedním z hlavních úkolů, které kategorie volně dostupného videoportálu plní, je její použití při vyhledávání. Kategorie může být jako hledisko zahrnuta do rešerše anebo může posloužit jako kritérium omezující výsledky rešerše.
- 3) Funkce orientační – ne všichni zavítají na volně dostupný videoportál s určitým cílem, někteří uživatelé na portále videa procházejí v dané kategorii (tzv. browsing). Kategorie jsou v tomto případě hlavním pomocníkem při navigaci na volně dostupných videoportálech.
- 4) Funkce propagační – s kategoriemi, stejně jako např. se statisticky proměnlivými výběry (nejnavštěvovanější videa, nejlépe hodnocená videa), je nakládáno tak, aby byly vidět, uživatelé si jich všimli a na portále setrvali co nejdéle.

Videoportály YouTube, Metacafe i Dailymotion mají společný způsob, jak a kdy videa do kategorií zařadit. Uživatel, který video nahrává, video přiřazuje jednu (YouTube, Dailymotion) nebo dvě kategorie (Metacafe), v níž / nichž bude video dostupné pro ostatní návštěvníky, pakliže se rozhodnou brouzdat v rámci kategorií volně dostupného videoportálu. Tento způsob, společný pro všechny výše zmíněné videoportály, má dva zásadní problémy -

- 1) Uživatel, který zařazení do kategorie nevěnuje pozornost anebo záměrně video zařadí do špatné kategorie, celý systém kategorií devaluje. Neexistuje ani žádný dozor ze strany provozovatele volně dostupných videoportálů, který by hlídal „čistotu“ kategorií.
- 2) Video na portálech Dailymotion a YouTube může být zařazeno právě v jedné z kategorií. Toto omezení se ukazuje jako nevyhovující, pokud nahrané video svou povahou nepřísluší pouze do jediné kategorie. V tom případě je video, které by z logiky věci zasluhovalo být ve dvou kategoriích, přítomno pouze v kategorii jedné a kategorie druhá jej nezahrnuje.

Partneři YouTube, což jsou registrovaní a placení uživatelé, mohou k vlastním videím přiřadit více než jednu kategorii. Tato množina uživatelů je vzhledem k celkovému počtu uživatelů zanedbatelná, a tudíž není ve výzkumné části této studie zohledněna.

Volně dostupný videoportál Vimeo pracuje s kategoriemi videa na jiném principu. Uživatel, který video nahrává, je vyzván, aby věnoval náležitou pozornost názvovým údajům videa, tagům,

kterými video opatří, a aby video zařadil do nekategoriálních struktur portálu. Na základě těchto skutečností pak bude video na základě algoritmu automaticky zařazeno do jedné či dvou kategorií, které odpovídají parametrům, jejichž vyplnění či zařazení měl v kompetenci nahrávající uživatel.

2.2 Výsledky srovnání kategorií volně dostupných videoportálů

Všechny videoportály zmíněné v této studii disponují obsahovými kategoriemi, do nichž jsou videa zařazována. Při porovnání těchto kategorií mezi všemi portály vystoupí do popředí pětice kategorií, která je zastoupena na všech portálech.

Tento překryv v trojici kategorií „Comedy“ + „Entertainment“ + „Funny“, a dále v kategoriích „Music“, „Tech(nology)“, „Sports“ a „Travel“ posloužil ke stanovení jádrových kategorií. Oba výzkumné vzorky videozáznamů byly stanoveny tak, aby obsahově spadaly do těchto kategorií, videa s písní Aleshy Dixon „The boy does nothing“ do kategorie „Music“, zatímco videa s loutkou břichomluvce Jeffa Dunhama, Achmed, mrtvý terorista, do kategorií „Comedy“ a „Funny“ či „Entertainment“ u uměleckých vystoupení a do kategorie „Music“ u videí obsahující pouze písně této loutky. Soupisy kategorií, aktuálních k 15. 9. 2013, vč. překryvů, viz *Tabulka č. 1*.

YouTube	Dailymotion	Vimeo	Metacafe
Autos & Vehicles Comedy Education Entertainment Film & Animation Gaming Howto & Style Music News & Politics Nonprofits & Activism People & Blogs Pets & Animals Science & Technology Sports Travel & Events	Animals Arts Auto-Moto College Funny Gaming Lifestyle Movies News Music Sexy Sports Tech Travel Webcam & Vlogs <i>Celeb</i> <i>Film & TV</i> <i>Gay & Lesbian</i> <i>Latino</i> <i>News & Politics</i> <i>People & Family</i> <i>Tech & Science</i>	Activism & Non Profits Animation & Motion Graphics Art Comedy Education & DIY Everyday life Experimental Films HD Music Nature Products & Equipment Science & Tech Sports Travel & Events Vimeo Projects Web series	Art & Animation Comedy Cool Commercials Entertainment How to Music & Dance News & Events People & Stories Pets & Animals Science & Tech Sports Travel & Outdoors Video Games Wheels & Wings Other 18+

Tabulka č. 1 – Srovnání kategorií videoportálů a stanovení jejich překryvu

Až na volně dostupný videoportál Dailymotion geografické ani jazykové mutace neovlivňují počet a název kategorií. Na portále Dailymotion se počet kategorií vzhledem ke zvolené mutaci jazyka či země liší, což je v *Tabulce č.1* vyjádřeno odlišně zbarvenou buňkou „nekmenových“ kategorií.

3 Výzkumné vzorky

Kolekce videozáznamů na volně dostupných videoportálech YouTube, Dailymotion a Vimeo jsou kolekce originální, tedy videa byla nahrána registrovanými uživateli a nebyla odnikud převzata. Volně dostupný videoportál Metacafe přebírá v nemalé míře videozáznamy z portálu YouTube.

3.1 Výzkumný vzorek č. 1: Alesha Dixon – The boy does nothing

Prvním vzorkem videozáznamů byla videa s písní „The boy does nothing“ od zpěvačky Aleshy Dixon.

Vyhledávací řetězec - „Alesha Dixon – The boy does nothing“

Počet videozáznamů ve výzkumném vzorku - 345

Relevantní videozáznamy

Jako kritérium pro odfiltrování nerelevantních videozáznamů bylo stanoveno, že video bude obsahovat statický či pohyblivý obraz interpretky písně, Aleshy Dixon nebo text písně. Zvuková stopa bude obsahovat živou či studiovou nahrávku této písně. Videozáznamy, kdy byla píseň pouze použita k jinému, než výše uvedenému obrazovému dění, byly považovány za nerelevantní, pro ilustraci - amatérský či profesionální tanec na tuto píseň, statické koláže či záběry nesouvisející s písní a interpretkou aj.

Unikátní / Duplicitní videozáznamy

U tohoto vzorku, tedy u relevantních videozáznamů s písní „The boy does nothing“, byla duplicita, či spíše unikátnost definována poněkud přísně. Pokud byla dvě videa identická, jde logicky o duplicitu, nicméně za duplicitní byly považovány i videozáznamy, které se lišily třeba vloženou úvodní či koncovou sekvencí. Za duplicitní, resp. neunikátní, což je přesnější, byla považována i videa se stejným obsahem, leč odlišným grafickým vyjádřením, konkrétně u videozáznamů, kde byl v průběhu písně zobrazován text na obrazovce. Lhostejno, že v jednom případě bylo pozadí červené a písmena bílá a v jiných případech byla barevnost a typ písma změněny.

Identita nahrávajícího uživatele

Přestože lze, v relaci registrovaný uživatel – registrovaný uživatel na volně dostupných videoportálech zahájit korespondenci a snažit se touto cestou zjistit skutečnou identitu protistrany, může jako lepší vodítko posloužit, pokud je k dispozici, profil na sociální síti, nebo jiná informační stopa mimo volně dostupných videoportálů. Stejně jako v případě videozáznamů, i v případě údajů o uživateli, který zveřejnil video, je tato informace migrována z portálu YouTube na portál Metacafe. Proto jsou odhady dohledatelnosti uživatelů videoportálů limitovány na relevantní neduplicitní záznamy, správné zařazení v rámci kategorií rozhodující pro tuto část výzkumu není.

Zařazení v kategorii

Videozáznamy s písní Aleshy Dixon „The boy does nothing“ logicky spadají do kategorie Hudba. Tato kategorie je přítomna na všech videoportálech, které byly zahrnuty do této studie.

3.2 Výzkumný vzorek č. 2: *Achmed, the dead terrorist*

Druhým vzorkem videozáznamů, logicky spadajícím do kategorie Hudba a do kategorií Komédie či Zábava (v závislosti na názvu a přítomnosti těchto kategorií), byla videa s loutkou břichomluvce Jeffa Dunhama „Achmed, mrtvý terorista“.

Vyhledávací řetězec - „Achmed, the dead terrorist“

Počet videozáznamů ve výzkumném vzorku - 466

Relevantní videozáznamy

Jako kritérium pro odfiltrování nerelevantních videozáznamů bylo stanoveno, že video bude obsahovat videozáznam z živého či předtočeného představení Jeffa Dunhama, v němž bude (mimo jiných) vystupovat loutka „Achmed, mrtvý terorista“. Za nerelevantní videozáznamy byla považována např. hudební díla vzniklá na základě „verbálního projevu“ loutky Achmeda, nebo předvedení vyzvánění či upomínkových předmětů s touto loutkou.

Unikátní / Duplicitní videozáznamy

U druhého vzorku, tedy u relevantních videozáznamů s loutkou „Achmed, mrtvý terorista“ byla duplicita definována na základě zjištění počtu předtočených celovečerních vystoupení Jeffa Dunhama, jichž bylo 5 a reklamních spotů či záznamů z živých představení. Za duplicitní byly považovány všechny videozáznamy, které byly jinak dlouhou či sestříhanou částí již dříve excerpovaných záznamů vystoupení J. Dunhama.

Identita nahrávajícího uživatele

Přestože lze, v relaci registrovaný uživatel – registrovaný uživatel na volně dostupných videoportálech zahájit korespondenci a snažit se touto cestou zjistit skutečnou identitu protistrany, může jako lepší vodičko posloužit, pokud je k dispozici, profil na sociální síti, nebo jiná informační stopa mimo volně dostupných videoportálů. V případě videozáznamů, i v případě údajů o uživateli, který zveřejnil video, je tato informace migrována z portálu YouTube na portál Metacafe. Proto jsou odhady dohledatelnosti uživatelů volně dostupných portálů limitovány na relevantní neduplicitní záznamy, správné zařazení v rámci kategorií rozhodující pro tuto část výzkumu není.

Zařazení v kategorii

Vystoupení břichomluvce Jeffa Dunhama a loutky „Achmed, the dead terrorist“ spadají do kategorií Hudba, pokud jde o pěveckou část vystoupení. Dále tato vystoupení spadají do kategorií Komédie a Zábava, za předpokladu, že dominantní složkou vystoupení je mluvené slovo.

3.3 Společné poznámky k výzkumným vzorkům

Pro lepší orientaci a snadnější analýzy byly videoportály YouTube, Metacafe a Vimeo nastaveny jazykově na angličtinu, resp. americkou angličtinu – English (US). Dailymotion byl nastaven na mutaci „International“, kde je předdefinovaným jazykem rovněž angličtina.

Oba výzkumné vzorky byly stanoveny tak, aby všechny volně dostupné portály obsahovaly nenulový počet relevantních videozáznamů. Dva výzkumné vzorky byly vybrány proto, aby lépe vynikly odlišnosti mezi videoportály zahrnutými v této studii. Oba výzkumné vzorky videozáznamů byly stanoveny tak, aby obsahově spadaly do kategorií, v nichž vznikl překryv podle analýzy v podkapitole 2.2.

Výsledky výzkumných vzorků byly odstupňovány takto – 100 % - 81 % - vynikající, 80 % - 61 % - nadprůměrný, 60 % - 41 % - průměrný, 40 % - 21 % - podprůměrný, 20 % - 0 % - neuspokojivý.

Rešeršní dotazy pro zjištění videozáznamů do výzkumných vzorků byly do vyhledávacího pole (až na portál Vimeo) zapisovány s uvozovkami „“, aby bylo nalezeno co nejvíce relevantních

videozáznamů i ve středu a na konci výsledků vyhledávání. Na portále Vimeo uvozovky pro malý počet nalezených výsledků použity nebyly. Vyhledávání proběhlo v rozmezí 15.7 – 15.8. 2013.

Na tomto místě je nutné uvést, že duplicita identických či téměř identických videozáznamů je jen v základní míře hlídána na videoportálech YouTube a Metacafe, na videoportálech Vimeo a Dailymotion hlídána není vůbec. Duplicita na videoportále YouTube a Metacafe je hlídána skrze kontrolní součty nahrávaných souborů s těmi, která jsou již na videoportálech vystavena. Navíc, na žádném ze zkoumaných videoportálů není obsah zvukové či obrazové stopy zkoumán z hlediska obsahové duplicity, pouze z hlediska závadného obsahu či obsahu, který podléhá ochraně autorských práv.

Přestože na videoportálech YouTube a Metacafe je, i když slabá, kontrola duplicity zavedena, je nutno zmínit, že tento slabý mechanismus kontroly duplicity selže, pokud totéž video bude kódováno různými kodéry obrazu a zvuku a výsledná velikost souboru bude odlišná od toho, který již na YouTube je vystaven [RODRIGUES, 2010].

Videozáznamy do výzkumných vzorků, dovolil-li to jejich počet, byly vybírány z prvních stránek seznamu výsledků, z posledních stránek seznamu výsledků a ze stránek, které se pohybovaly aritmetickým průměrem uprostřed seznamu výsledků vyhledávání. Poměr zastoupení videozáznamů ze začátku, středu a konce seznamu výsledků vyhledávání byl stanoven poměrem 3:4:3, např. při výběru 200 videí z portálu YouTube bylo vybráno prvních 60, středních 80 a posledních 60 videí. Z jednoho portálu bylo do výzkumného vzorku zařazeno maximálně 200 videí.

Údaje o videozáznamech z výzkumných vzorků byly shromažďovány a zpracovány v tabulkovém editoru Open Office Calc. Každý řádek obsahuje sledované ukazatele a základní údaje o videozáznamu, viz *Tabulka č. 2*.

Prefix URL	URL suffix	Název videa
http://www.youtube.com/watch?v=	b6VMQCFYZJ0	WeihnachtsSpecialAchmed 2 (deutsche Untertitel)
http://www.youtube.com/watch?v=	S7B-cHSy8F0	Ahmed - The Dead Terrorist "Jingle Bombs" (polskie napisy)
http://www.youtube.com/watch?v=	IJD9xn00ZW8	Achmed The Dead Terrorist remix 08/11
http://www.youtube.com/watch?v=	ETv3UvGqyPA	comedy achmed the dead terrorist ventriloquist
http://www.youtube.com/watch?v=	wAGawYPtmk0	Achmed the Dead Terrorist
http://www.youtube.com/watch?v=	B10bccJuvi	Jeff Dunham & Peanut Washington Part1 VOSTFR

Délka videa	Relevantní	Unikátní	Vystavovatel	Kategorie
10:07	1	0	A	1
04:53	1	1	A	1
02:06	0	X	X	X
01:18	1	0	A	1
10:00	1	1	A	1
06:04	0	X	X	X

Tabulka č. 2 – Náhled zpracování výzkumných vzorků

3.4 Odhady výsledků výzkumných vzorků

Protože se autor studie na volně dostupných videoportálech pohybuje denně, a to vzhledem k akademickým i dalším aktivitám, odvážil se odhadnout dílčí výsledky obou výzkumných vzorků.

Výzkumný vzorek č. 1: *Alesha Dixon – The boy does nothing*

Předpokládaný podíl relevantních videozáznamů z celku nalezené množiny byl odhadnut pouze na **60 %**. Tento odhad byl stanoven na základě toho, že píseň je často pouhým podkresem pro domácí či umělecké video, avšak tyto materiály nejsou z hlediska této studie považovány za relevantní.

Předpokládaný podíl multiplicitních relevantních videozáznamů byl odhadnut na **25 % až 33 %**, tedy že podíl unikátních relevantních videozáznamů bude činit **66 % až 75 %**. Tento odhad pracuje s fakty, že mnohé videozáznamy budou z živých koncertů či zaznamenaných televizních pořadů a nesmí být opomenuta ani videa, která mají jako obrazovou složku videoklip, nicméně mohou mít titulky či překlad písně v různých jazycích.

U relevantních videozáznamů s písní „The boy does nothing“ bylo odhadováno **20 % až 25 %** dohledatelných uživatelů díky externím odkazům. Toto číslo bylo odhadnuto takto nízké proto, že jde o autorsky chráněný materiál, jehož volné použití podléhá autorskoprávní legislativě.

Videozáznamy s písní „The boy does nothing“ v kategorii Hudba byly odhadnuty v rozmezí **66 % až 75 %**, tedy že dvě třetiny až tři čtvrtiny relevantních videozáznamů budou v kategorii Hudba. Toto číslo bylo stanoveno na základě zjištění, že některá předvedení této písně byla např. na živých koncertech, v rámci nějakého zábavného pořadu, a tudíž vystavovatel videa zařadil video do jiné kategorie, sledujíc širší kontext. Rozmezí 66 % až 75 % bylo stanoveno také s ohledem na uživatele, kteří zařazení do kategorie nevěnovali pozornost.

*Výzkumný vzorek č. 2: **Achmed, the dead terrorist***

Odhad relevantních videozáznamů na celku byl stanoven na **70 % až 75 %**. Tento odhad se zakládá na faktu, že „Achmed, mrtvý terorista“ byl a je doslova fenoménem volně dostupných videoportálů.

Odhadovaný podíl duplicitních záznamů byl odhadnut na **80 % až 85 %**, tedy podíl unikátních videozáznamů byl odhadnut na **15 % - 20 %**. Tento odhad pracuje s faktem, že mnohé videozáznamy jsou části týchž, zejména předtočených vystoupení, jen jsou různě rozstříhané. Rovněž je ale nutné (jako unikátní videozáznamy) zohlednit ty z nich, které byť mají totožnou obrazovou stopu, liší se v jazyku titulků.

U relevantních videozáznamů s loutkou „Achmed, mrtvý terorista“ bylo odhadováno **20 % až 25 %** dohledatelných uživatelů díky externím odkazům. Toto číslo bylo odhadnuto takto nízké proto, že jde o materiál, jehož volné užití podléhá autorskoprávní legislativě.

Předpokládaným výsledkem, tedy přítomností videozáznamů vystoupení břichomluvice Jeffa Dunhama a jeho loutky „Achmed, the dead terrorist“, v kategoriích Hudba, Komédie a Zábava (podle povahy videozáznamů a propozic jednotlivých volně dostupných videoportálů), je výsledek v rozmezí **75 % - 80 %**, tedy že zhruba tři čtvrtiny relevantních videozáznamů a jejich kolekcí budou v kategoriích usazeny správně. Problém v zařazení do správných kategorií bude v momentě, kdy součástí vystoupení je kromě dialogů také píseň anebo tato píseň byla vystřížena a stalo se z ní samostatné video.

4 Výsledky a jejich zhodnocení

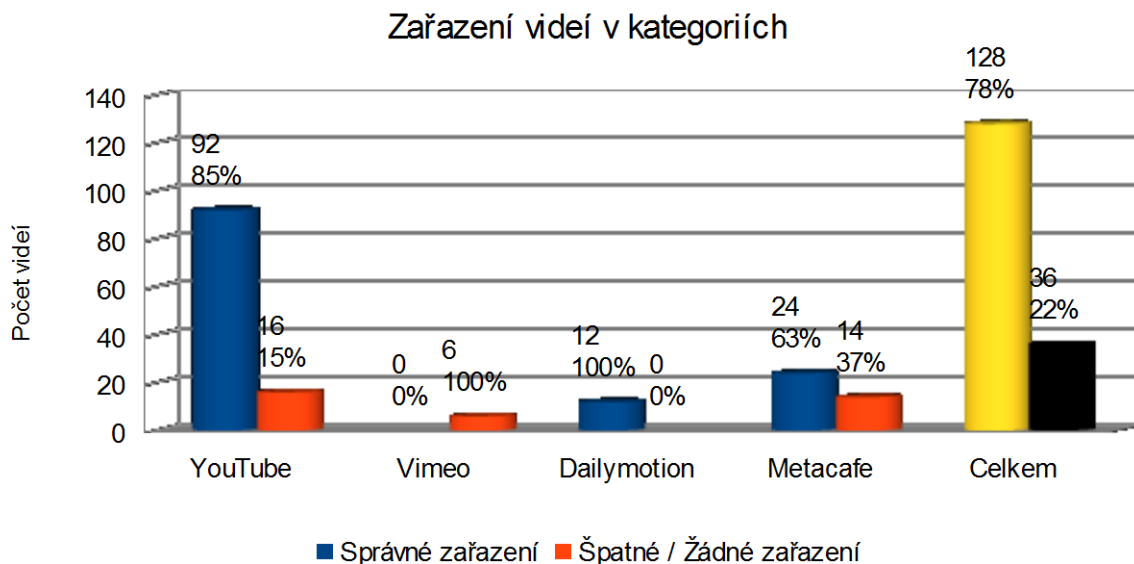
Jako uživatel, který se ve světě volně dostupných videoportálů pohybuje od roku 2007, si autor studie troul odhadnout, viz kapitoly o výzkumných vzorcích, jak která část výzkumu dopadne, zejména proto, že se na volně dostupných videoportálech pohybuje denně a denně na nich vyhledává [VOCŮ, 2012].

Přestože to ve studii již zaznělo, je na tomto místě nutné znovu zopakovat, že portál Metacafe přebírá videa a údaje o nahrávajícím uživateli z portálu YouTube. Tento jev se odráží ve výsledcích výzkumných vzorků v rámci ukazatele „Unikátní / Duplicitní videozáznamy“ tím, že mnoho videí z Metacafe bylo označeno za „duplicitní“. Stejně tak ukazatel „Identita nahrávajícího uživatele“ byla u videí převzatých z YouTube (vy)hodnocena podle zdroje, nikoliv podle videoportálu Metacafe.

4.1 Výzkumný vzorek č. 1: Alesha Dixon – The boy does nothing

Prvním vzorkem videozáznamů, logicky spadajícím do kategorie Hudba, byla videa s písní „The boy does nothing“ od zpěvačky Aleshy Dixon.

Zařazení v kategorii - Odhad – 66 % - 75 %, skutečnost – 78 %, viz Graf č. 1 – Zařazení videí v kategoriích.

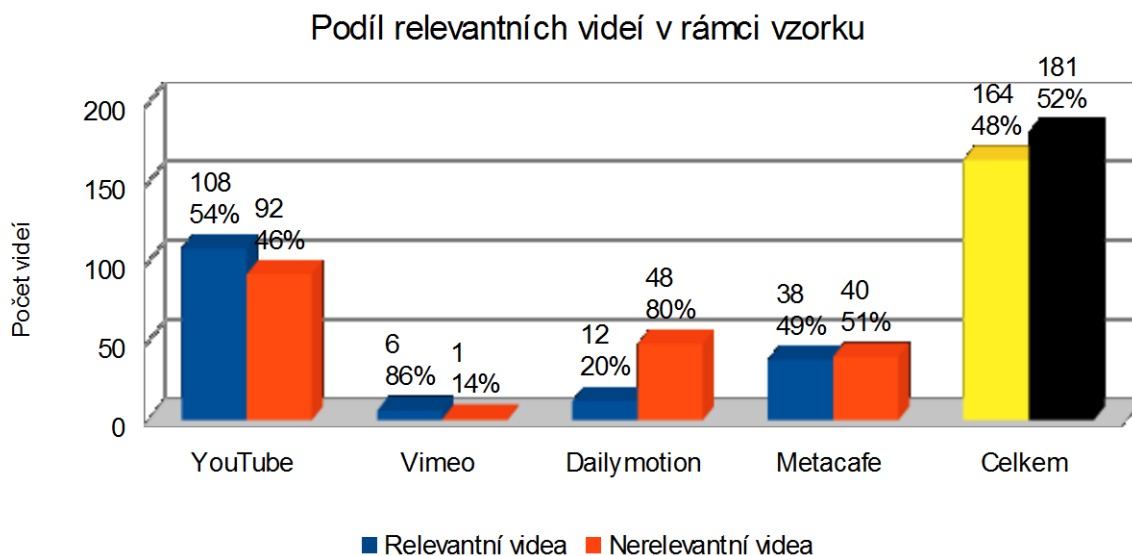


Graf č. 1 – Zařazení videí v kategoriích

Skutečnost mírně předčila horní hranici očekávaného výsledku, za povšimnutí stojí výsledek portálu Dailymotion, kde byla všechna relevantní videa zařazena správně, zatímco na portále Vimeo nebylo zařazeno žádné video vůbec. Tuto skutečnost si lze vysvětlit tím, že automatické

zařazování videí do kategorií probíhá od videí s nejlepším uživatelským aparátem, napomáhajícím při automatickém zařazování a také od videí s největším penzom sociální interakce.

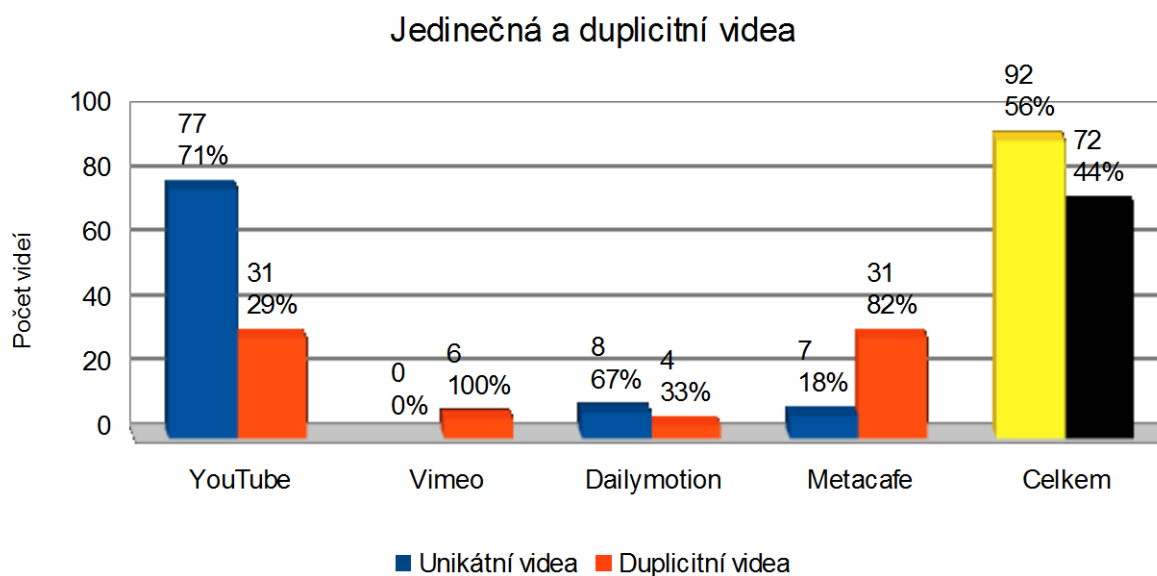
Relevantní videozáznamy – Odhad – **60 %**, skutečnost – **48 %**, viz Graf č. 2 – Podíl relevantních videí v rámci vzorku.



Graf č. 2 – Podíl relevantních videí v rámci vzorku

Oproti předpokladu 60 % relevantních videí je souhrnná skutečnost na úrovni 48 %, vysvětlení lze nalézt ve skutečnosti, že v mnohých nerelevantních videích byla píseň „The boy does nothing“ použita jako hudební podkres pro nejrůznější videa obrazově nesouvisející s písní. Portál Vimeo, ač jeho návratnost činila pouhých 7 záznamů, měl nejvyšší podíl relevantních záznamů na celku.

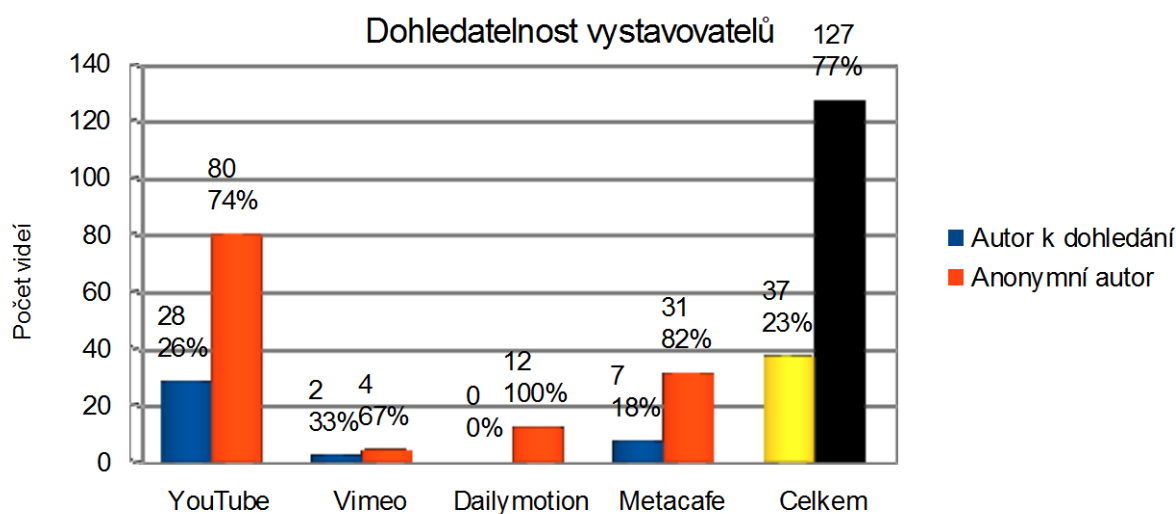
Unikátní / Duplicitní videozáznamy – Odhad – **66 % - 75 %**, skutečnost – **56 %**, viz Graf č. 3 – Jedinečná a duplicitní videa.



Graf č. 3 – Jedinečná a duplicitní videa

Předpoklad nebyl potvrzen, a podíl unikátních videí na celku je pod dolní odhadovanou hranicí. Nejčastěji se opakujícím relevantním videem s písní „The boy does nothing“ byl oficiální videoklip, který byl paradoxně na portále Vimeo zastoupen 6x ze 7 případů. Překvapující byl výsledek portálu Dailymotion, který měl při relevantním počtu 12 videozáznamů 8 unikátních s touto písní. Vysvětlení lze spatřit v tom, že Dailymotion je velmi navštěvován frankofonní komunitou a ta namísto nahrávání videoklipu raději na Dailymotion vystavila živé předvedení této písně. Portál Metacafe má nízký podíl unikátních videí proto, že mnoho z nich je převzato z videoportálu YouTube.

Identita nahrávajícího uživatele - Odhad – **20 % - 25 %**, skutečnost – **23 %**, viz Graf č. 4 – Dohledatelnost vystavovatelů.



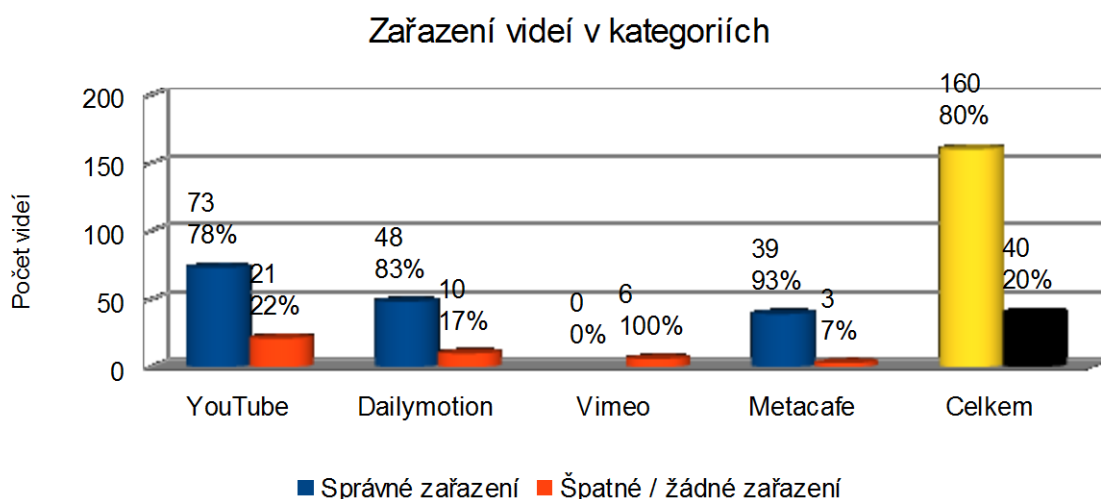
Graf č. 4 – Dohledatelnost vystavovatelů

Skutečnost se vešla do odhadu, na portále Vimeo je překvapující výsledek 33 %, pravdou je, že pokud uživatel „férově přizná“ cizí obsah, nemusí jeho video být odstraněno, stejně jako když pod video korektně uvede, že nevlastní žádná autorská práva ke zveřejňovanému materiálu a vystavuje jej pouze za účelem kritiky. Portál Metacafe dosáhl 18 % dohledatelných uživatelů právě díky YouTube a videí vč. základního informačního aparátu od něj převzatých.

4.2 Výzkumný vzorek č. 2: Achmed, the dead terrorist

Druhým vzorkem videozáznamů, logicky spadajícím do kategorií Zábava / Komédie (pro komediální představení) a Hudba (pro videa pouze s písněmi), byla vystoupení břichomluvců Jeffa Dunhama s jeho loutkou „Achmed, mrtvý terorista“.

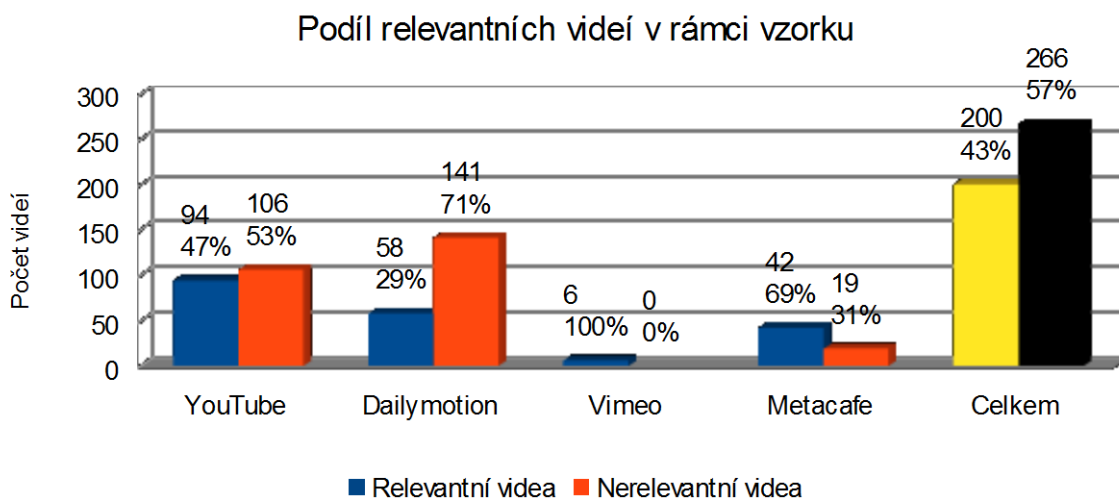
Zařazení v kategoriích - Odhad – **75 % - 80 %**, skutečnost – **80 %**, viz Graf č. 5 – Zařazení videí v kategoriích.



Graf č. 5 – Zařazení videí v kategoriích

Výsledek 80 % je na horní očekávané hranici, portál Vimeo opět neměl žádné relevantní video automaticky zařazeno do patřičné kategorie, ze stejných důvodů, jako u výzkumného vzorku č. 1 – žádné z videí s loutkou Achmed, mrtvý terorista nebylo divácky ani z hlediska sociálních interakcí dost významné, aby mu byly přiřazeny (až dvě) kategorie. Ostatní portály dosáhly velmi dobrých výsledků, vč. Metacafe, jehož videozáznamy byly opět v nemalé míře převzaty z YouTube.

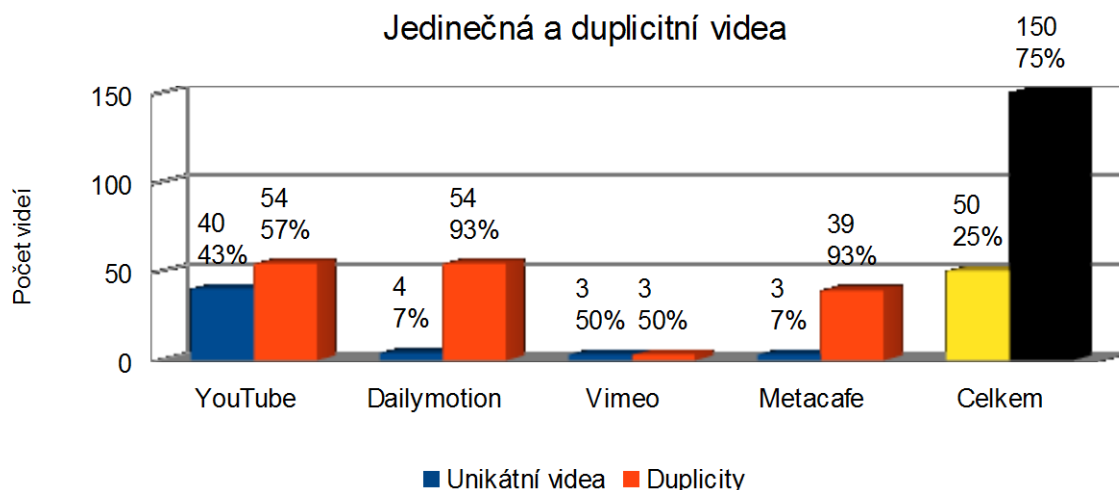
Relevantní videozáznamy – Odhad – **60 %**, skutečnost – **43 %**, viz Graf č. 6 – Podíl relevantních videí v rámci vzorku.



Graf č. 6 – Podíl relevantních videí v rámci vzorku

Podíl relevantních videozáznamů pouze na úrovni 43 % byl nepříjemným překvapením, a na vině tomu jsou nerelevantní videa na portálech YouTube a Dailymotion, která obsahovala zpravodajské šoty o útocích teroristů nebo přednášky či vystoupení odborníků či duchovních arabského etnika.

Unikátní / Duplicitní videozáznamy – Odhad – **15 % - 20 %**, skutečnost – **25 %**, viz Graf č. 7 – Jedinečná a duplicitní videa.

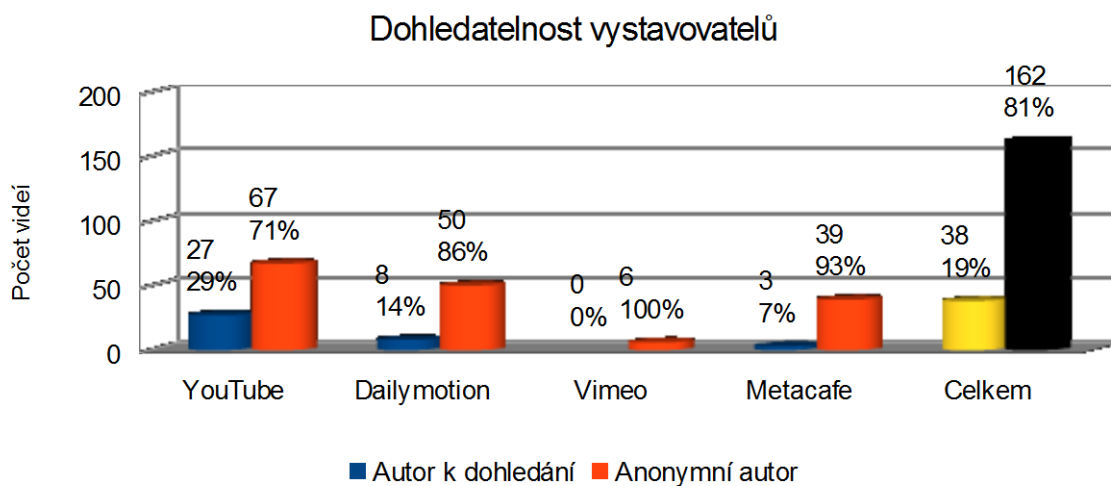


Graf č. 7 – Jedinečná a duplicitní videa

Byť to nemůže být ihned patrné, největším překvapením této části výzkumu nebyl fakt, že unikátních videí bylo nad očekávání 25 %, ale sice skromný výsledek portálu Vimeo, kde byla 3 unikátní videa z 6 nalezených. Portál Metacafe má nízký podíl unikátních videí proto, že mnoho z nich je převzato z YouTube.

Identita nahrávajícího uživatele - Odhad – **20 % - 25 %**, skutečnost – **19 %**, viz Graf č. 8 – Dohledatelnost vystavovatelů.

Pouze 19 % dohledatelných uživatelů, kteří vystavili relevantní video s loutkou Achmed, mrtvý terorista lze opět přičíst faktu, že jde o autorsky chráněnou postavku i vystoupení. Portál Metacafe dosáhl 7 % dohledatelných uživatelů jedine díky videím z YouTube, sám o sobě by skončil, stejně jako portál Vimeo, s 0 % na kontě.



Graf č. 8 – Dohledatelnost vystavovatelů

4.3 Závěry, komentáře a doporučení

Tato studie si nekladla za cíl být vyčerpávajícím materiálem, zevrubně zkoumajícím všechny volně dostupné videoportály, cílem studie bylo provést kvalitativní výzkum na tematicky vymezených, byť nevelkých, vzorcích videí a na jeho výsledcích demonstrovat určité fenomény. Svět volně dostupných videoportálů je nesmírně proměnlivý a bylo by určitě zajímavé provést obdobný výzkum na jiném vzorku anebo provést výzkum se stejnými parametry třeba za 3 roky a porovnat výsledky.

Co se týká zařazení videí v kategoriích, lze zkonstatovat, že u obou výzkumných vzorků byla videa na portálech YouTube a Dailymotion zařazena ve většině případů, v rozmezí 78 % až 100 %. Na portále Metacafe byl výzkumný vzorek č. 1 zařazen správně do kategorií v 63 %, což lze hodnotit jako nadprůměrné. Výzkumný vzorek č. 2 byl na portále Metacafe zařazen správně v 93 % případů, což lze hodnotit jako vynikající. Naopak portál Vimeo neměl zařazeno žádné video, tudíž zcela selhal.

Jelikož je portál Vimeo v současnosti nastaven tak, aby videa byly přiřazeny automaticky dvě kategorie, lze navrhnout, aby jednu z kategorií volil uživatel, který video vystavuje, a druhá kategorie aby byla přiřazena automaticky. Portál Vimeo tím může získat větší počet videí, které budou mít alespoň jednu kategorii a i kdyby od uživatelů nebyla správně přiřazena, může být automaticky posléze doplněna kategorií vhodnější.

Na portále YouTube může registrovaný uživatel (ale ne Partner YouTube) přiřadit videu jednu kategorii, zatímco Partner YouTube až tři kategorie. Neuškodilo by proto povolit registrovaným uživatelům přiřadit videu první kategorii při nahrávání a druhou a třetí po dosažení nějakého milníku, třeba určitého počtu zhlédnutí videa, nebo počtu komentářů pod videem, atd. Partneři YouTube tím neztrácejí ihned svou výhodu více kategorií, jen bude pomalu vzrůstat počet videí s dvěma či třemi kategoriemi.

Portál Dailymotion by zasluhoval zrevidovat kategorie a ustálit je na fixním počtu pro všechny jazykové i geografické mutace portálu s tím, že z kategorií, které svůj status ztratily, se stanou prosté tagy a uživatelé, jejichž videa byla kategorie zbavena či přeřazena, o tom budou zpraveni.

Portál Metacafe nemá problémy s kategoriemi, ale s přebíráním videí z YouTube, která když jsou na YouTube smazána, tak jsou přesto ve výsledcích vyhledávání přítomna. Jelikož jsou na portále Metacafe videa z YouTube jednoduše identifikovatelná díky url, v němž je řetězec „yt-“, lze doporučit k těmto videím přiřadit zřetelné tlačítko s funkcí „Ohlásit neplatný odkaz“ (Report broken link), aby tato již neexistující videa byla co nejrychleji z výsledků odstraněna.

Výsledky vyhledávání na portále YouTube jsou zobrazeny max. do str. 25 při zobrazení 20 videí na stránku, tedy lze zobrazit videa 1-500. Na portále Dailymotion je toto omezení zavedeno na výsledky od str. 101, tedy lze zobrazit videa 1-2 000. Tento fakt je neuspokojivý zejména v momentě, kdy uživatel nezná unikátní webovou adresu videa a přitom hledá video, které může mít nespécifický název, chudý informační aparát a je zařazeno do špatné kategorie. Navíc, pro omezení výsledků vyhledávání slouží (by měla sloužit) omezující kritéria, která budou uplatněna na vyhledávání nebo množinu výsledků [VOCŮ, 2013].

Překvapující skutečností je, že na portále YouTube, na jeho poslední zobrazitelné straně s výsledky vyhledávání, což bývá strana 25 s videy na 481. - 500. místě, se u ani jednoho vzorku neobjevilo relevantní video. Pakliže byl odhad u obou výzkumných vzorků 16-18 000 videí, je podezřelé, že by videa okolo pětisté pozice a dále byla už všechna nerelevantní.

Stejným překvapením, jako o odstavec výše, byla skutečnost, že na posledních zobrazitelných stranách vyhledávání na portále Dailymotion, tedy na stranách 98-100 byl, byť jediný, relevantní videozáznam.

Přestože to lze považovat za ne zcela korektní, při rešerši na volně dostupných portálech YouTube a Dailymotion je množina zobrazitelných výsledků zredukována. Při pohledu na výsledky vyhledávání se může uživatel orientačně dozvědět, kolik videí odpovídá jeho rešeršnímu dotazu.

Na videoportálu Metacafe není uživatel o množině nalezených videozáznamů informován vůbec. Po uskutečnění vyhledávacího procesu je uživateli nabídnuta pouze první stránka s výsledky a doplňková třídící kritéria.

Příkladným videoportálem z hlediska zobrazení přibližného počtu výsledků a možnost jejich úplného procházení od prvního po poslední videozáznam, je portál Vimeo. I při vyhledávání, jehož výsledkem jsou statisíce záznamů, má uživatel možnost procházet výsledky od prvního po poslední nalezený videozáznam.

Doporučení vzešlá z této studie budou komunikována příslušným videoportálům, pakliže to dovolu informační aparát webových stránek těchto videoportálů.

Literatura:

1. Dailymotion (firma), 2013. *Dailymotion – Watch, publish, share videos* [online]. Paris (Francie): Dailymotion, 2005-2013 [cit. 2013-09-15]. Dostupné z: <http://www.dailymotion.com/>.
2. GNOLI, Claudio, MEI, Hong, 2006. Freely faceted classification for Web-based information retrieval. *New Review of Hypermedia and Multimedia*. June 2006, vol. 12, no. 1, s. 63-81. Special issue: Knowledge organization Systems and Services. ISSN 1361-4568 (Print). ISSN 1740-7842 (Online). Dostupné komerčně z Taylor & Francis Online: <http://dx.doi.org/10.1080/13614560600758944>.
3. GLUSHKO, Robert J., 2013 *The Disciple of Organizing*. 1st ed. Cambridge: MIT Press, 2013. ISBN 978-0262518505. Dostupné též komerčně z: <http://disciplineoforganizing.org/>.
4. HUANG, Chunneng, FU, Tianjun, CHEN, Hsinchun, 2010. Text-based Video Content Classification for Online Video-sharing Sites. *Journal of the American Society for Information Science and Technology*. May 2010, vol. 61, issue 5, s. 891-906. ISSN 1532-2882 (Print). ISSN 1532-2890 (Online). Dostupné komerčně z Wiley Online Library: <http://onlinelibrary.wiley.com/doi/10.1002/asi.21291/abstract>.
5. JACOB, Elin K, 2004. Classification and Categorization: a Difference that Makes a Difference. *Library Trends*. Winter 2004, vol. 52, issue 3, s. 515-540. ISSN 0024-2594. Dostupné též z: <https://www.ideals.illinois.edu/handle/2142/1686>.
6. KOUSHA, Kayvan, THELWALL, Mike, ABDOLI, Mahshid, 2012. The role of online videos in research communication: A content analysis of YouTube videos cited in academic publications. *Journal of the American Society for Information Science and Technology*. September 2012, vol. 63, issue 9, s. 1710-1727. ISSN 1532-2882 (Print). ISSN 1532-2890 (Online). Dostupné komerčně z Wiley Online Library: <http://dx.doi.org/10.1007/s10516-007-9022-z>.
7. Metacafe (firma), 2013. *Metacafe – Online Video Entertainment – Free video clips for you enjoyment* [online]. San Francisco (Kalifornie): Metacafe Inc., 2013 [cit. 2013-09-15]. Dostupné z: <http://www.metacafe.com/>.
8. RODRIGUES, Tiago, BENEVENUTO, Fabrício, ALMEIDA, Virgílio, ALMEIDA, Jussara, GONCALVES, Marcos, 2010. Equal but different: a contextual analysis of duplicated videos on YouTube. *Journal of the Brazilian Computer Society*. September 2010, vol. 16, issue 3, s. 201-214. ISSN 0104-6500 (Print). ISSN 1678-4804 (Online). Dostupné komerčně ze SpringerLink: <http://dx.doi.org/10.1007/s13173-010-0019-x>.
9. TERANTINO, Joseph, 2011. M. YouTube for Foreign Languages: You Have to See This Video. *Language Learning & Technology*. February 2011, vol. 15, no. 1, s. 10-16. ISSN 1094-3501. Dostupné též volně z: <http://llt.msu.edu/issues/february2011/emerging.pdf>.

10. TIAN, Yan, 2010. Organ donation on Web 2.0: content and audience analysis of organ donation videos on YouTube. *Health Communication*. April 2010, vol. 25, no. 3, s. 238-246. ISSN 1041-0236 (Print). ISSN 1532-7027 (Online). Dostupné komerčně z: Taylor & Francis Online:
<http://www.tandfonline.com/doi/full/10.1080/10410231003698911#.UZNgSNimW6Q>.
11. Vimeo (firma), 2013. *Vimeo: your videos belongs here* [online]. New York (New York): Vimeo LLC, 2013 [cit. 2013-09-15]. Dostupné z: <https://vimeo.com/>.
12. VOCŮ, Ondřej, 2012. Vyhledávání hudbou a jeho vazby na portály se sdíleným videem. *Knihovna*. 2012, roč. 23, č. 1, s. 63-83. ISSN 1801-3252 (Print). ISSN 1802-8772 (Online). Dostupné z: http://knihovna.nkp.cz/knihovna121/12_163.htm.
13. VOCŮ, Ondřej, 2013. Vyhledávání a video. *Ikaros* [online]. 2013, roč. 17, č. 2 [cit. 2013-05-15]. ISSN 1212-5075. Dostupné z: <http://www.ikaros.cz/node/7853>.
14. YouTube (firma), 2013. *YouTube: broadcast Yourself* [online]. San Bruno (Kalifornie)