What Makes People Watch Online Videos: An Exploratory Study

ART#

CHRISTOPH LAGGER, MATHIAS LUX
University of Klagenfurt, AUSTRIA

OGE MARQUES
Florida Atlantic University, USA

Searching and retrieving videos in a meaningful way on the web is still an open problem. The integration of a user's context and intentions into the search process is one of the most promising approaches to enhance current search interfaces and algorithms. In this paper we present the results of two exploratory studies on the topic of online video searching, retrieving, watching, and sharing: a qualitative study where 22 participants reported on situations when they retrieved and watched videos; and an online quantitative survey with more than 200 participants answering comparable questions. We provide a detailed analysis of the results from both studies and report on the insights that they provide in terms of video search, retrieval, watching, and sharing behavior. Our findings can be used to enhance current video retrieval systems, search interfaces, and algorithms, in order to improve the overall user satisfaction and experience. As an example of such improvements, we also propose a prototype that addresses the problem of taking the user's intentions into account when designing video retrieval interfaces.

Categories and Subject Descriptors: H.1.2 User/Machine Systems: Human Factors;
General Terms: Human Factors;
Additional Key Words and Phrases: video retrieval, exploratory survey, multimedia, user intentions, prototype;

1. INTRODUCTION

The massive growth of multimedia content – particularly video – on the Web, has changed the way we watch and share video contents online. Users can watch videos on various platforms, stream videos to their gaming consoles and TVs from different content providers, as well as download video content to their PCs via peer-to-peer (P2P) or other mechanisms. YouTube reports that one hour worth of video content is published every second on their site and that they have exceeded 4 billion views a day [1]. Cisco forecasts that by 2014 the sum of all forms of video (TV, video-on-demand (VoD), Internet (video platforms) and P2P) will

Author address: Christoph Lagger (clagger@edu.uni-klu.ac.at), Mathias Lux (mlux@itec.uni-klu.ac.at): University of Klagenfurt, Universitätsstraße 65, 9020 Klagenfurt, AUSTRIA; Oge Marques (omarques@fau.edu), Florida Atlantic University, 777 Glades Road, Boca Raton, FL 33431, USA.

DOI DOI: 10.1145/XXXXXX.XXXXXX. http://doi.acm.org/10.1145/ XXXXXX.XXXXXX

exceed 91% of all Internet traffic, out of which 57% would come from major online video platforms (e.g., YouTube, VEVO, etc.) [2].

Despite the tremendous growth in online video production, consumption, and sharing, there has not been a comparable increase in understanding the users' *intentions* as they search for, retrieve, watch, and share videos of interest. This lack of knowledge is reflected in the way currently used video search user interfaces are not mindful of the users' context and their intentions.

In general, an *intention* is an aim that guides a certain action[1]. Almost every action of a person is caused by an intention to accomplish a certain goal. Similarly, almost every search performed on the Internet, either on general search engines or more specialized (e.g., image and video) search engines, is triggered by an intention. It is reasonable to take a user's intention into account to deliver more precise and meaningful results.

We postulate that the integration of a user's context and intentions into the search process is one of the most promising approaches to enhance current video search interfaces and solve common problems, e.g., ambiguous queries. We also anticipate that it is possible to achieve better search results, i.e., higher precision relative to a user's context, based on intentions, e.g., by dynamically adapting relevance functions used by search algorithms. Furthermore, the paradigm of using intentions to derive information needs may lead to a better understanding of queries – including query reformulation, query extension, and relevance feedback mechanisms.

In order to better understand users and their behavior, we conducted two different exploratory studies between October 2010 and February 2011:

(1) Semi-structured interviews with 22 participants, where we asked questions about general platform usage, online video consuming and sharing behavior, as well as specific scenarios illustrating *why*, *where* and *when* people watched a video online.
(2) Online survey with 270 participants, in order to confirm the findings from the interviews and to gain further insight into user motivations for watching videos online.

These two types of studies led to an understanding of the basics of the video search, retrieval, watching, and sharing processes from the users' perspective in a more personal manner than conventional technical observation methods such as query log analysis or network traffic monitoring approaches.

[1] Dictionary: http://www.thefreedictionary.com/intentions (accessed February 12, 2011)

1.1 Research questions

Little is known about the reasons why people watch videos online, their intentions, motivation, and preferences. In our work we aim to fill the gap between multimedia information systems design and development and user context and behavior, especially user motivations and goals. We do not conduct research in a system-centric way, but rather in a holistic approach, integrating relevant actions before and after the video consumption on a specific platform. In this article we focus on the following research questions:

(1) *How* and *why* are people using video platforms?
(2) *What* is the actual user's behavior and motivation to retrieve videos?
(3) *How* do users share their video experience?
(4) Are *events* that lead to video consumption related to video *genres*?
(5) How can these findings be translated into code and incorporated in a prototype of a novel video retrieval interface?

1.2 Structure of this article

The remainder of this paper is organized as follows. Section 2 reviews related work, particularly articles that focus on user surveys in this field as well as literature focusing on the analysis of YouTube log files and metadata. In Section 3 we describe the design of our two studies in detail, whereas Section 4 highlights their most important results. Section 5 discusses our findings and Section 6 presents the basic aspects of a prototype of a novel, adaptive, video search and retrieval interface that takes into account the results and insights from both studies. Finally, Section 7 presents our concluding remarks, highlighting the contributions of this work and suggesting opportunities for future work.

2. RELATED WORK

There have been several user studies concerning mobile TV or mobile video usage in various countries. Miyauchi et al. [4] explored people’s different behavior on live mobile TV and mobile video usage in Japan by performing qualitative interviews with 11 participants. They compared usage on commuter vehicles, at home, and in experience-sharing via mobile TV, and found differences in attitudes concerning mobile TV and mobile video usage. They observed that mobile TV users on commuter transportation preferred light entertainment after work to relax, whereas mobile video users watched content which they prepared beforehand (e.g., the night before) because they wanted to use the time while commuting for a useful purpose (e.g., to watch how-to/education videos), although they also watched entertainment programs. When at home, both user groups did not like to watch content on

the small screens of their mobile devices, but sometimes mobile TV users sacrificed larger displays (TV, monitor) for the convenience of a portable viewing experience. Watching programs together (while commuting) was more convenient for mobile TV users because they did not prepare the content for a specific purpose and did not have to consider any privacy problems. Mobile video users, on the other hand, prepared the content to their needs and this content may be private to some extent.

Other studies related to mobile TV or mobile video usage where conducted in Finland [5][6], US and UK [7], Korea [8], and Australia [9]. There are significant differences in users' behavior among those studies, due to cultural and lifestyle reasons. The results of the study conducted in Finland showed that people prefer to watch short programs or pieces of longer programs while waiting for something or just to pass time. This was not restricted to a specific task or location – people did this, for example, while waiting for the bus in public or at home. Moreover, people in the US and the UK prefer to watch light entertainment (e.g., cartoons, comedy) instead of videos covering self-development or work topics. In Korea mobile TV devices are often used to avoid boredom, to stay up to date with current events, or to play games. People also use their mobile TV devices much more at home, even if a big TV screen is available.

Buchinger et al. [10] provide a comprehensive survey of the state of the art and open issues for mobile TV, focusing on the users' needs and experiences. The authors combined two research methods – Human Centered Design and Quality of Experience (QoE) – to analyze the results obtained by multiple studies on mobile TV from a users' perspective. The subject of research was the user itself, including his motivation, future trends and needs, as well as the mobile device and the context where, when and how long mobile TV is consumed. Furthermore they investigated which content is consumed, and the technical performance (regarding to audio, video and transmission) of the service. They summarized some open issues in the field, notably: (i) The predominant gender depends highly on the content and service; (ii) Users show an urge to share and modify content on their mobile devices; (iii) Open problems include the fees charged for different types of services as well as the technical performance; and (iv) Sometimes the quality is not good enough (especially for sport broadcasts) and the channel switching response time is too high.

Another research direction related to the topic of this paper is the analysis of query log files or datasets providing metadata of videos (rating, number of views, external links etc.) of various video platforms or content providers. In [11] the authors analyze online video search and sharing behavior on YouTube. They used a web crawler and the YouTube API to create a database with all the metadata and statistics provided by YouTube. Using this database, they analyzed people's search behavior and determine what makes some videos

more popular than others. Videos with longer titles, better description and more meaningful tags receive more views than videos with less metadata, therefore this information is an important factor to improve video popularity. A video receives the highest number of views and clicks in the first few days after the upload; after that period, the popularity generally decreases. The number of views of a video does not follow a Zipf distribution as observed in other types of web surfing paradigms.

Cha et al. [12] performed an extensive data-driven analysis on YouTube from the perspective of user-generated content (UGC). The subject of the study was to analyze the popularity distribution and evolution for various categories and videos and their relationship with video age. Additionally they provide insights into current UGC distribution systems, and propose more efficient designs, namely caching and peer-to-peer. In the case of popularity evolution, they showed that if a video does not receive enough requests during the first days it is unlikely that it will get many requests in future. Predicting near-future popularity would be useful for service providers to cache videos before they get popular to overcome bottlenecks. By comparing the first few days of video views with views after a certain time (5, 7, 90 days) the results showed that the second day record provides a good estimate about future popularity with a relatively high accuracy (correlation coefficient above 0.8) even for predicting distant future popularity (up to three weeks).

Another study [13] presents a systematic and in-depth measurement study on the statistics of YouTube videos. By examining 3 million YouTube videos the authors found differences to traditional video systems regarding length of videos, access patterns, popularity trends and active life span. They argue that the social network of YouTube videos substantially contributed to the success of this service.

These technical in-depth analyses of log files and video metadata are a different method to explore user behavior and cover other relevant aspects of video retrieval that cannot be covered by surveys or interviews, as done by our work, which effectively supplements their findings.

Human factors in video retrieval systems have not been extensively studied yet and constitute an open and promising research topic. More specifically, a broader user study on overall user behavior and preferences, especially in online video retrieval – without restricting the research to mobile TV – has not been done yet. Therefore we aim to fill this research gap with the exploratory study described in this paper.

3. DESIGN OF THE EXPLORATORY STUDIES

The exploratory studies reported in this paper aim at understanding *how* and *why* online videos are watched, and *what* are the user's actual behavior and motivation regarding video retrieval, search and sharing.

The first study (Section 3.1) was of qualitative nature and comprised interviews with 22 randomly selected people with different background, occupation and hobbies. Each participant was questioned about their experiences related to searching, retrieving, watching and sharing videos online. An important goal of the interview was to uncover *why* they watch videos online, i.e., what was their motivation in the first place. More specifically, we wanted to know which *event* caused the action to watch a video.

After the qualitative interviews were finished, we designed an online survey with SurveyMonkey[2]. We decided to reuse the questions from the qualitative interviews and provide predefined answers (inspired by the answers from the interviewees) that survey participants could choose from. Additionally, we introduced new questions that were of interest. We distributed the survey through various communication channels such as: email, instant messaging, and several social networks. By doing so, we reached 270 people from various countries, out of which 217 participants (80.4%) completed the whole survey. The main goals of this second study (Section 3.2) were to test whether our findings from the qualitative interviews would still hold in a larger and more diversified population, and to gain even further insight into user behavior in aspects associated with watching videos online.

Participants in both studies took part in the interview / survey on a completely voluntary basis, and were not given any money, vouchers or gifts as compensation for their time.

3.1 Qualitative Study

We used qualitative interviews as a starting point for our research. In-depth interviews, such as the ones performed for this survey, are more flexible than other methods because they allow participants to freely talk about their video watching and sharing experiences, without being bound by predefined answers. Moreover, the interviewer can decide – based on the interviewee's response – which question to ask next, how the next question should be phrased, and whether a certain topic deserves deeper and more detailed exploration. The analytical objectives for qualitative interviews are: to describe and explain relationships, variation, individual experiences and, if possible, to describe group norms [14].

[2] SurveyMonkey: http://www.surveymonkey.com

Setup: We prepared a questionnaire with open-ended questions beforehand in order to cover all necessary topics during the interviews. The questionnaire contained a mix of quantitative questions with qualitative elements designed to find out and understand why people watch videos and uncover the triggering factors behind the decision to watch such videos. The questions were divided into four major blocks:

- *Demographics*: Consists of questions about job, gender, age and general interests and hobbies.
- *General usage*: We wanted to find out which online video platforms a user knows and actively uses, and how frequently. We were also interested in learning on which devices a user is consuming digital video content. Part of this block was an open question about the user's motivation to search for a video and watch it, and whether the quality of a video is a decision criterion for watching it. Additionally, participants were asked questions about the usage of a diverse set of features of video platforms (e.g., subscriptions, related videos, and suggestions on what to watch next).
- *Communication*: In this block, we wanted to find out how users share their video experience with other people in their social environment. It was of interest *with whom* and *why* one is sharing online videos. Interviewees were also asked about: the content of the video being shared, the intention behind sharing the (link to the) video, the frequency of sharing, and which sharing mechanisms were used.
- *Instances*: Instances are descriptions of specific scenarios (or situations) a user has experienced. Interviewees were asked to recall one or more such situations and report on: which *event* (e.g., search, recommendation from another person, or a platform feature such as "related videos") had triggered the action of watching that specific video, what was the content of the video, and – very important – what was the intention behind watching such video. Additionally, questions about the setting (place/time), device used, and the length of the video were asked, as well as whether the interviewee had shared the (link to the) video with someone else after watching it. Each interviewee could report up to three instances, with preference for recent video watching situations.

Interviews were performed on a semi-structured fashion, which allowed an engaging and interactive conversation while still covering the predefined set of questions. We approached students, relatives and friends in a spontaneous manner and asked them if they had time for an interview without mentioning what the survey was about. Each interview was conducted face-to-face and lasted up to 30 minutes. During the interview, we wrote down the answers from the interviewee on paper immediately and then later transcribed them to a digital format.

Participants: Within our study, 22 participants (19 males and 3 females) – all of which were German-speaking Austrians with different background, occupation, and hobbies – were interviewed. Their occupations included students (9) and professionals (13) and their hobbies varied on a broad scale. Participants reported from memory on one, two, or three video watching situations, particularly on the following aspects: *why* they watched a video online, *what* the video was about, and *where* they watched it. In the end, the 22 participants reported on 49 different situations, an average of 2.23 scenarios per participant. The age distribution of the participants can be seen in *Figure 1*. The vast majority of participants (~ 77%) were younger than 30 years of age.

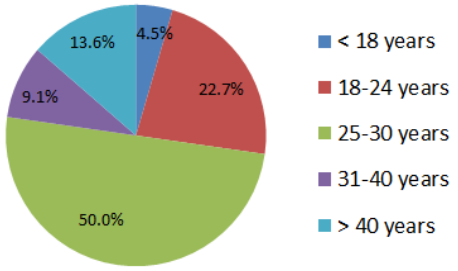

Figure 1: Age distribution of participants in the qualitative interviews.

3.2 Quantitative Survey

Based on the preliminary findings of the qualitative study, we decided to design and perform an online quantitative survey to test and confirm our findings with a larger and more demographically diverse audience.

Setup: We used SurveyMonkey to design a structured online questionnaire, which included questions with multiple-choice answers following an ordinal scale, where items were given a value from 1 to 5, which let us calculate the average rating, as seen in many of the upcoming tables and figures. We completely redesigned the "instances" part of the study to fit the quantitative interview because we did not want to rely on the memory of the participants. We introduced new questions such as: questions about the user's favorite genres (based on YouTube categories), most common events which triggered the decision to watch an online video, places and times where and when videos were watched, and other preferences such as: picture quality, content quality, sound quality, etc. Additionally, some of the questions offered the participant an open text box so that they could provide us with further information or a more detailed answer. The interview structure and questions were based on the insights and knowledge given in [3] for performing and evaluating quantitative surveys. The survey was distributed through different communication channels, namely: email,

instant messaging, Reddit, Twitter, Facebook, Slashdot, and two personal blogs. A SurveyMonkey mechanism that distinguishes among them was configured to identify how participants were referred to the survey.

Participants: 270 participants took part in the survey, out of which 217 (i.e., 80.4%) completed and answered all questions[3]. The response rate per communication channel from all 270 participants can be seen in Figure 2. Most of the respondents attended the survey through email (173), followed by the Reddit community (48), instant messaging (17), and Facebook (12).

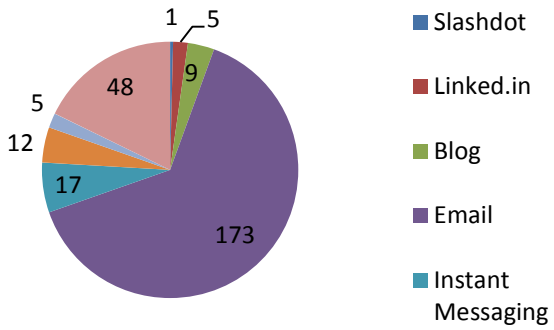

Figure 2: Response rate by distribution-channel in number of overall participants. Channels are listed according to their position in the pie chart (starting from the top, moving clockwise).

Among the 217 persons who finished the survey, 69.4% were male (150), 30.6% were female (66), and one person skipped the gender question. The age distribution (Figure 3) shows that the vast majority of respondents (~ 81%) are less than 30 years of age, similarly to the age distribution for the qualitative interviews (Figure 1). Their country of origin (determined via http://www.ip2country.cc) can be seen in Figure 4: the majority of participants came from Austria (136) followed by the US (50).

[3] For our evaluation, only data of the 217 completed surveys were used.

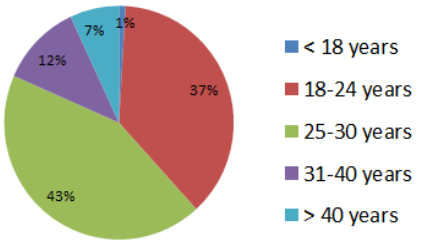

Figure 3: Age distribution of participants of quantitative study

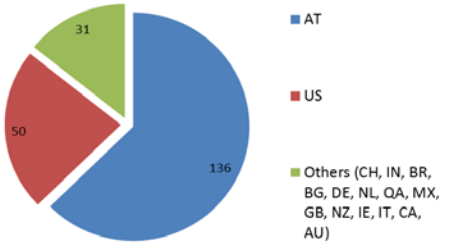

Figure 4: Participants per country of the quantitative study

A comprehensive overview and comparison from both studies in terms of demographics, education and age can be seen in Table 1.

		Qualitative Interviews	in %	Quantitative Survey	in %
Demographics	Overall participants	22	100.0%	217 [4]	100.0%
	Female	3	13.6%	66 [5]	30.4%
	Male	19	86.4%	150	69.1%
	Country of origin	AT: 22	-	AT: 136 US: 50 Others[6]: 31	-
Education	Graduate degree	3	13.6%	47	21.8%
	Bachelor, college or university degree	7	31.8%	85	39.4%
	High school	5	22.7%	68	31.5%
	Elementary/secondary school	6	27.3%	5	2.3%
	Others	1	4.5%	11	5.1%
Age	< 18	1	4.5%	2	0.9%
	18 – 24	5	22.7%	81	37.7%
	25-30	11	50.0%	94	43.3%
	31-40	2	9.1%	25	11.5%
	> 40	3	13.6%	15	6.9%

Table 1: Demographics overview

[4] Completed the survey
[5] One person skipped this question
[6] Countries of origin of other participants: CH (1), IN (1), BR (2), BG (2), DE (5), NL (3), QA (1), MX (1), GB (3), NZ (1), IE (1), IT (1), CA (2), AU (1)

4. RESULTS

In this section we summarize and compare the results from both studies. We also highlight interesting facts and figures from both surveys, as well as personal comments from the participants. This section is structured in three parts: in Section 4.1 we explore how participants of the studies used video platforms in general; in Section 4.2 we provide insights into their video sharing behavior; finally, in Section 4.3 we analyze the specific scenarios reported in qualitative interviews and attempt to validate those findings with the data from the quantitative survey by using statistical analysis. Each subsection has been structured using short paragraphs preceded by questions that cover selected topics of interest from our studies.

4.1 General usage of video platforms

Which platforms and how often do you use them?

Both questionnaires covered questions about the general usage of video platforms, including questions about most favorite platforms, usage patterns of platforms regarding frequency of visits, reasons for using the platform, etc.

Figure 5 shows a summary of the preferred video platforms used by participants on a regular basis. YouTube appears as the most-used and best-known platform, followed by Facebook. The high rating of Facebook is somewhat surprising, due to the fact that it is not primarily a video platform. Moreover, Facebook posts often contain links to other platforms such as YouTube and that could be the cause for its popularity. This finding provides insight on the fact that users do not seem to care much about which platform hosts or streams their videos, but instead which platform (or, in this case, social network) acts as a portal through which they learn about such videos. This is also true for most blogs or news sites, which don’t necessarily host their own content or use their own media player. Overall, the percentages reported in our qualitative study are higher than the quantitative counterpart, which is most likely due to the differences in demographics. A notable exception to that pattern is the fact that no one mentioned Netflix in the qualitative study, which is clearly due to the fact that the service is not yet available in Austria.

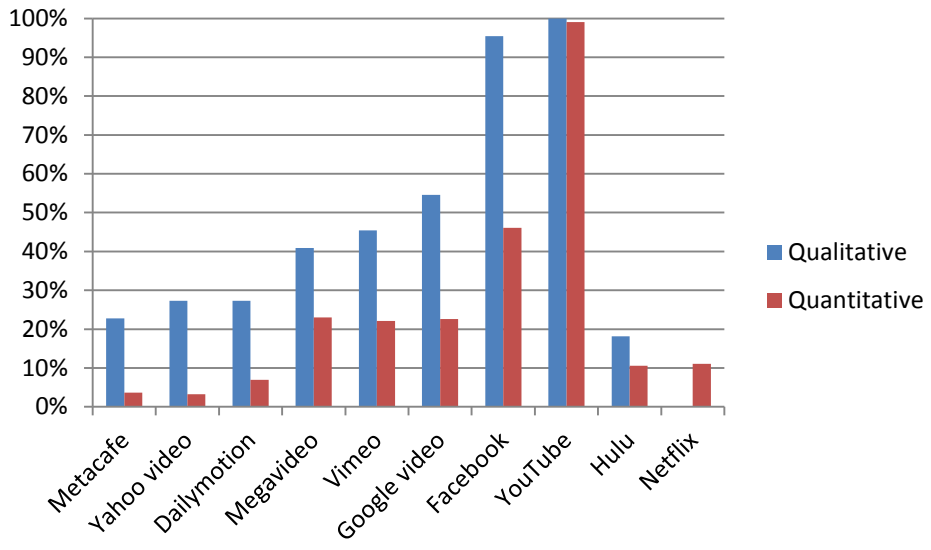

Figure 5: Most frequently used video platforms

Some users have (in some cases, completely) moved away from cable TV because they watch all of their video content online.

> *"I do not have cable TV, so all of my video watching is in lieu of television"* (male, US, 31-40)

> *"No cable TV, no bunny ears. All video is strictly streaming from Hulu, Netflix, or other leading services"* (male, US, 25-30)

Regarding frequency of usage, in both studies the majority of participants (80% or more) reported watching videos online several times a week or more than once a day (Figure 6), reinforcing the notion that Internet-based video platforms are becoming a part of everyday life.

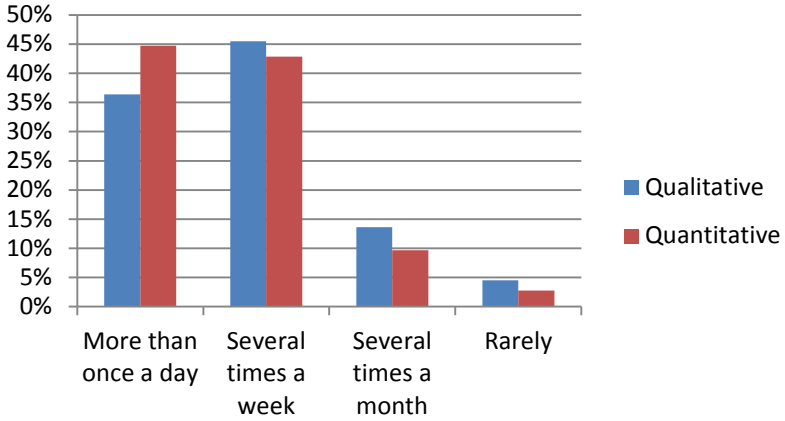

Figure 6: Frequency of platform usage

Which devices do you use to watch video contents?

We are also interested in learning about the devices used or preferred among the users to consume online video content. In the qualitative interviews we found that all 22 participants used their PC/laptop/netbook to watch videos online, nine reported to watch videos on a mobile device, and five also watched videos on a TV. One interviewee even watched online videos on his car PC. In the quantitative study we asked about the most preferred devices of participants, giving them the option to quantify the usage for each device by selecting a ordinal scale item between "Never (1)" and "Always (5)". This resulted in an average rating (Figure 7) that shows that most of the participants used laptop/notebook/netbook devices to watch online video content followed by the classical PC and smartphones. As new TV sets and gaming consoles have the ability to stream online content and access online video services, the high rating of TV is not unusual. Tablets are still not used very often, but this may be due to the fact that the market for such devices is still very young and growing.

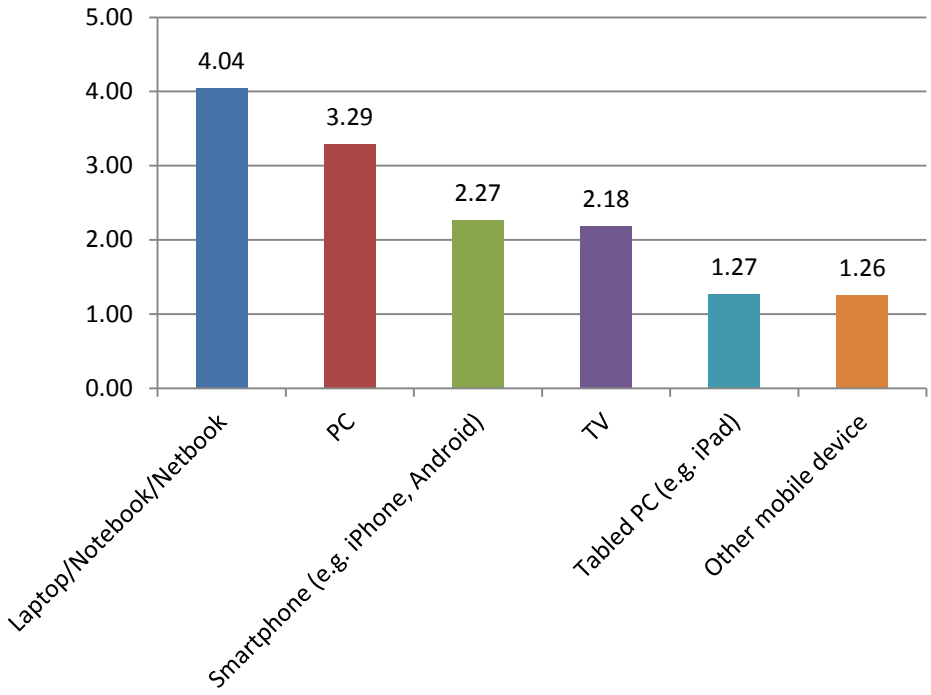

Figure 7: Average rating of device usage
(based on ordinal scale items from Never (1) to Always (5))

Some people also mentioned other devices and ways to consume online video content:

> *"I watch videos on the web less frequently after I could stream Netflix on my gaming consoles"* (female, US, 25-30)

> "*I have a WDLive Drive. Basically, you set it up with a TV and can access Netflix, YouTube, your stored videos, etc. Makes YouTube sharing much easier"* (male, US, 25-30)

Professional or private usage?

The growing availability of video contents for educational and professional use may also lead to a change in attitude in the mindset of employers and employees. Educational video content is cheap (or even free) and easy to access and the information presented in a video is often easier to understand than reading about the same topic. In the qualitative interview four participants reported to use video content for professional reasons. Three of the four were students who answered "yes" because they watched video tutorials and lectures on specific topics they needed to acquire knowledge about. In the quantitative survey, participants reported to use online video content for professional use once a day or more in 3.8%, once a week or more in 14.2%, once a month or more in 21.2% of all cases. The majority reported to use videos for professional reasons rarely (37.7%) or never (23.1%). These numbers seem low, however they indicate an upcoming trend as more and more people recognize the opportunity to use video as a convenient tool for learning.

What about quality?

The quality of a video encompasses various dimensions such as picture, sound and content quality. However, the quality requirements vary from one person to the next and also depend on other factors, e.g., the device someone is currently using, which video one is actually watching, or the reason for watching it. In the qualitative interviews 15 people reported that quality is a very important factor for them to watch a video. They reported that if the video or audio quality does not match their expectations they either close the video immediately or they search for another one with the same content but in a better quality. The requirements for picture and sound quality also heavily depend on the content of the video. Users reported that picture quality is very important for video tutorials and how-to videos (because often you need to read or follow a text or the like), and sports/action videos but not significantly important for most entertainment or music videos. For music videos the sound quality is often more relevant than the video quality because music videos are often used just for background entertainment like radio. In general, users make exceptions if the task to watch a specific video is very important and no alternative video that covers the same content is

available. Other technical factors that impact the decision to watch a video are loading and buffering time and (lack of) sync between video and audio tracks.

> *"When a video is repeatedly interrupted or paused due to slow download, I always press exit or delete because I lose patience and interest quickly. As this happens quite often, I rarely watch videos at all"* (female, AT, > 40)

Since there is also a growing number of people using high end equipment at home, they also stated that they often explicitly search for HD videos to watch it either on their TV or in full screen mode on their computing devices which support HD.

> *"Because of 30Mbps download speeds and all of my monitors and internet connected TV's are 1920x1080 resolutions, I prefer highest quality videos. However, while traveling and using iPhone or other slower connected products, I prefer medium to low quality so buffering is not an issue"* (male, US, 31-40)

We also asked similar but more specific questions in the quantitative survey, whose results are summarized in Table 2. Here, once again, the majority of users prefer high picture and sound quality, but consider the content of the video the most important factor to impact their decision to watch it.

	Strongly disagree (1)	Disagree (2)	Neutral (3)	Agree (4)	Strongly agree (5)	Rating average
I prefer videos with high visual quality	1.9%	1.9%	18.1%	**42.1%**	36.1%	4.09
I prefer videos with good sound quality	0.9%	0.9%	20.1%	**47.2%**	30.8%	4.06
My decision to watch a video is based on the content	1.9%	1.4%	13.9%	34.3%	**48.6%**	4.26

Table 2: Quality preferences among users in quantitative survey.

Why and how are you watching videos?

We also asked the participants of the qualitative interview why in general they searched for videos. Certainly and as expected, participants often told us that they use such platforms when they have time, are bored and need to kill time or have nothing better to do and want some entertainment. In this case they watch all kinds of video ranging from entertainment and fun videos to videos related to their hobbies and interests (such as sports videos). Another interesting fact is that some people use YouTube solely as a radio station. They load up a queue of their favorite music videos and just play it in the background, while giving no

attention to the video at all. They state that it is the easiest way of listening to their favorite music.

Besides entertainment, videos are used to learn different things and get information about certain topics. Using video tutorials and how-to videos to learn new or rehearse known content is very common nowadays. For example, two students used videos that explained certain math paradigms to prepare for an upcoming exam. Others wanted to improve at their hobby/sport so they watched videos about special technique training for rock climbing and others watched videos to gain knowledge on how something works (a specific microcontroller, a drilling machine, our solar system, etc.). They all agreed that videos are an easy way to gain new knowledge, learn about unknown things, or to refresh knowledge about a certain topic when needed.

Some users reported having used online video platforms to see content that they missed on normal broadcast TV programs like summaries of sport events (e.g., missed goals of a soccer game) or the latest local news. It is also an easy way to see how some things look like (e.g., a place that one is planning to visit). One participant watched videos about a specific cruise ship because he planned to do a cruise vacation in future and he wanted to see what the accommodation looks like and what the ship itself has to offer. He also does this with hotels where he is planning to stay while on vacation just to get an impression beforehand. Others used online videos to watch reviews about a product that they wanted to buy because it is much less stressful than reading a product review online.

Participants from both studies also reported to use video platforms to watch videos not accessible in their country or to watch old TV shows from their youth, because it is the only way to see that content again.

> *"I belong to the 2nd generation of TV viewing, and I use YouTube to reminisce my childhood viewing, and to watch original, undubbed, non-censored, or content not available in my country"* (male, US, > 40)

> *"I especially enjoy watching videos on the net that couldn't be accessed otherwise for whatever reasons"* (female, AT, 25-30)

Certain users also reported to watch movie trailers or trailers/advertisement for upcoming events (e.g., concerts, movies, musical premieres, etc.) online. A few people also just watch all sorts of videos to have a background noise (i.e., using it as a radio station as mentioned before); they don't pay attention to the video itself that much because they are busy with more important work.

"I have dual-monitors on my main computer. I'm more likely to watch videos if I can continue doing my current activity on one screen and have the video play on the other" (male, US, 18-24)

"I'm working from home, alone, but after a while I need something in the background, sometimes its music, sometimes it is a Movie. When I need to work and I choose a movie, it is always a movie I've already seen 3 or 4 times before, so I don't get distracted. So I guess in average I watch a lot of videos, by not watching them" (male, AT, 31-40)

To be more exact and summarize the answers from our participants, the most frequent responses from the qualitative interviews can be categorized as follows:

- *Information search* (16 out of 22), e.g., watch missed broadcast news, watch product reviews, see how a place looks like, research and follow-up on a certain topic.
- *Educational purposes* (8 out of 22), e.g., how-to videos, learn math, tactics for video games, and video tutorials on diverse topics.
- *Entertainment* (19 out of 22), e.g., sports, comedy, listening to music or watch music videos, movie trailers, discontinued/old TV material.

If we compare this result to our quantitative survey where we asked the same question ("Why do you search for videos and watch them?") and gave the participants predefined ordinal scale items ("Never (1)" to "Always (5)") to choose from, the detailed results are shown in Table 3.

	Never (1)	Rarely (2)	Sometimes (3)	Often (4)	Always (5)	Rating Average	Rank
For leisure/entertainment reasons	2.3%	4.2%	14.9%	**59.5%**	19.1%	3.89	1
Out of boredom	13.9%	21.3%	26.2%	**33.7%**	5.0%	2.95	6
To listen to music or to watch a music video	3.8%	10.8%	24.4%	**44.6%**	16.4%	3.59	2
To get information about certain topics	4.3%	14.2%	**40.3%**	8.1%	8.1%	2.27	4
To get product reviews	22.2%	**31.4%**	30.0%	14.0%	2.4%	2.43	8
To see missed news-broadcasts	22.2%	**32.9%**	29.0%	14.0%	1.9%	2.41	9
To learn something new/to gain new knowledge	4.3%	14.3%	36.7%	**38.1%**	6.7%	3.29	3
To see how something looks like (e.g. place or product)	11.2%	**28.8%**	27.8%	27.8%	4.4%	2.85	7
Too see how something works	5.4%	15.8%	**41.9%**	32.0%	4.9%	3.15	5

Table 3: Why do people watch videos?

A close inspection of the 'Rank' column in Table 3 shows that most people use video platforms for entertainment in their leisure time. Entertainment in this case could be whatever a person enjoys, e.g., watching a video about their favorite sport, TV show, comedy or the like. Secondly, people often use video platforms to listen to music because music video clips are easily accessible and mostly of good quality because music companies or the artist themselves promote their new songs over such platforms. The third-highest-ranked reason mentioned by participants was to use video platforms to learn something new or improve their knowledge on a certain topic.

Do people like platform features?

When asked about which on-site (platform-dependent) features they use, and how often, participants of the qualitative interviews indicated a frequent use of features such as "video suggestions" or "related video". They often use it if they have enough time and want to get further information about the topic they are currently looking at. In some cases they also used it if the first video for their search was not the expected one: instead of refining the query terms of their search, they reported looking first at the suggested or related videos section to see if they could find the intended video there. In the quantitative survey we also evaluated the usage of that feature and confirmed that most participants used it (only 3.2% reported having never used it) (Figure 8).

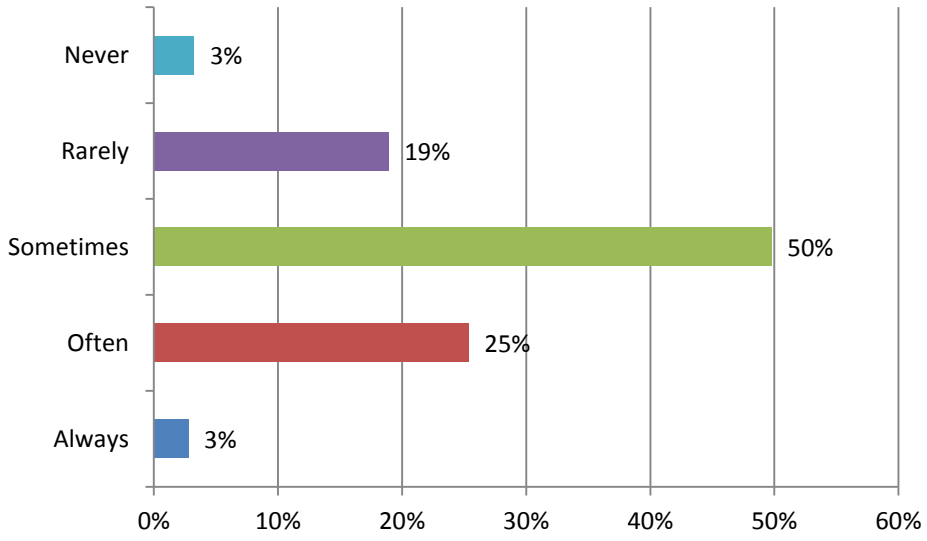

Figure 8: How often do people follow recommendations?

Table 4 summarizes answers to questions related to on-site features, such as "channel subscriptions" and "navigation through categories" from the quantitative study. During the

qualitative interviews, by contrast, only two participants reported having used "Channel subscriptions", whereas the "navigation through categories" option was not used at all because it is more convenient and efficient to search by terms and select a video from the result list.

	Never (1)	Rarely (2)	Once a month or more (3)	Once a week or more (4)	Once a day or more (5)	Rating Average
I watch video recommendations coming from on-site features (channel subscription, related videos, video suggestions)	9.3%	**30.8%**	29.4%	25.7%	4.7%	2.86
I browse through categories and select videos from there	17.8%	**32.9%**	24.4%	18.8%	6.1%	2.62

Table 4: General feature usage

Even if most of the participants used these two features rarely, the overall percentage of respondents who use them is quite significant (82.2% or higher), which suggests that these features are used more often than expected before and that they are somewhat important for the user's video experience.

In the quantitative survey, participants also reported on features that they believe should be available on video platforms, chief among them the ability to download a video in case it gets removed or just to have an offline copy for further usage.

> *"I prefer to check out sites that allow easy download of said videos. I hate streaming sites (buffering)"* (female, MX, age 25-30)

> *"… For YouTube I use youtube-dl. Vimeo offers downloads for registered users for some videos, but for some time I had trouble logging in at their webpage, so no more Vimeo. Luckily youtube-dl works most of the time and some people host their videos on their own sites so you can simply download them. …"* (male, AT, 25-30)

> *"… On another note, I tend to download online videos if they are important enough to me and if they are likely to be removed from the respective hosting site in the future. ..."* (male, AT, 18-24)

What do people like to watch?

In our qualitative interviews we analyzed the 49 scenarios reported by the participants and matched each scenario to a YouTube genre. The highest fraction of watched videos belonged to the "Music" genre with 34.7%, followed by he "Sports" genre (28.6%). Other genres watched were "Film & Animation" and "Science & Technology" with 8.2% of all cases, "Autos & Vehicles", "Comedy, "Education", "Gaming" with 4.1%, and "News & Politics" and "Travel & Events" with 2.0%.

In the quantitative survey, we asked participants to express their interest in various genres (we adopted the 15 YouTube categories as predefined genres) by selecting one of the following ordinal scale items: "Not interested (1)", "Rarely interested (2)", "Neutral (3)", "More interested (4)", "Very interested (5)". The average rating based on the values can be seen in Figure 9.

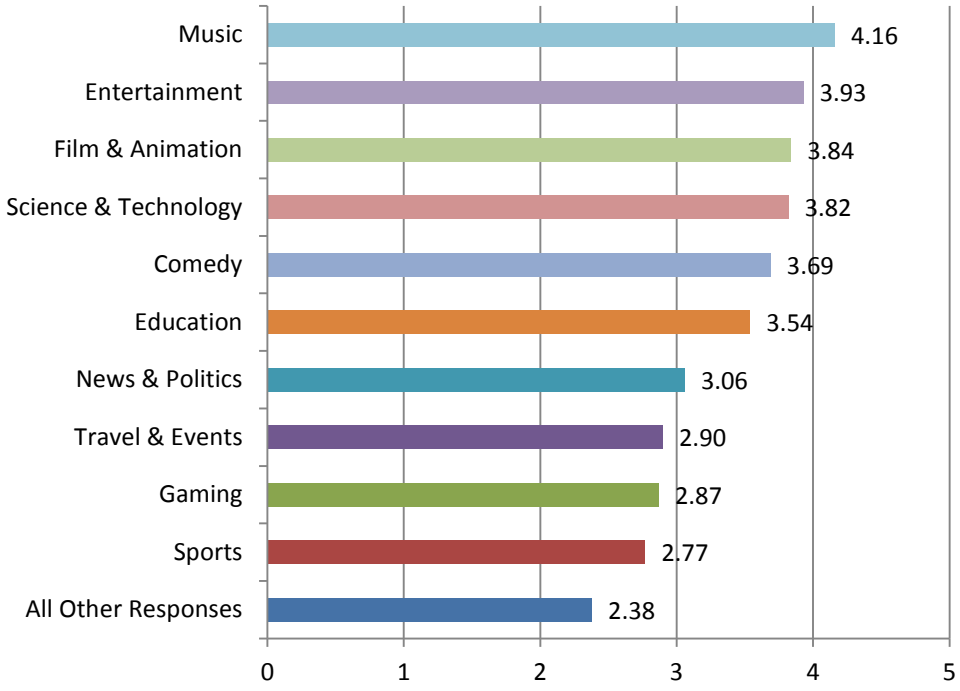

Figure 9: Favorite genres of the overall population

Once again, the "Music" genre is the most preferred one, but followed by other genres – not "Sports" as our result from the qualitative interviews suggested. Based on the comparison of both studies, we can easily argue that the rating of the preferred genres is heavily based on the survey population. But in order to understand how preferred genres change when restricted to certain demographic groups or usage patterns, we further analyzed age groups

and frequency of usage regarding genre preference. Table 5 shows the first five most preferred genres based on the age of the participants.

Rank	< 18 years	18-24 years	25-30 years	31-40 Years	> 40 years
1	Entertainment	Music	Music	Science & Technology	Science & Technology
2	Comedy	Entertainment	Science & Technology	Music	Comedy
3	Music	Film & Animation	Film & Animation	Entertainment	Entertainment
4	Film & Animation	Comedy	Entertainment	Film & Animation	Film & Animation
5	News & Politics	Science & Technology	Education	Comedy	Music

Table 5: Genre interest per age group

An investigation of genre preferences based on platform usage frequency did not show much variation or any specific preferences.

Table 6 shows the first five most preferred genres based on the frequency of platform usage. No matter how often a user visits video platforms, the “Music”, “Entertainment” and “Film & Animation” genres seem to be popular among all users except for people who rarely visit such platforms, which seem to be more interested in videos from the “Education” genre.

Rank	Rarely	Several times a month	Several times a week	More than once a day
1	Education	Music	Music	Music
2	Music	Entertainment	Entertainment	Entertainment
3	Comedy	Film & Animation	Science & Technology	Film & Animation
4	Science & Technology	Science & Technology	Film & Animation	Comedy
5	Entertainment	Comedy	Comedy	Science & Technology

Table 6: Genre preference based on frequency of usage

Annoyances of online video

Online video content is not always satisfying for every user; there are always drawbacks. Since many videos include advertisements – either during the actual video or before it –, we should briefly discuss this matter. Our quantitative study did not cover this question *per se*, but participants frequently reported in the open text box that the growing trend of advertisements in or before videos often annoys them, especially when the advertisement is placed in such a way that covers important parts of the video.

> *"I don't think this is interesting per se but I hate when a site puts ads that obstruct the video I am trying to watch"*, female/US (age > 40)

> *"Advertisement links, banner or movies which are shown before the real movie or clip starts is very annoying"*, female/AT (age 25-30)

Users (particularly those outside the US) found it particularly annoying that some video content providers (e.g., Hulu and Netflix) limit access to their resources to specific countries or regions.

> *"I hate it with passion when I'm not able to watch video because I'm in the "wrong" country and there is a copyright issue (YouTube, etc.). How something like this is still possible in this time and place is a mystery to me."* (male, DE, age 25-30)

> *"I watched YouTube since the very early start and enjoyed it all the time. The only thing which starts to annoy me are adds in front of spots which you cannot skip or that I am unable to use some YouTube main sites like Vevo out of my country. Same is true for sites like Hulu or FOX video which scan your IP to determine your location and block you"* (male, AT, 25-30)

4.2 Sharing behavior

Within our studies we also explored the sharing behavior of participants. What is their main motivation to share video links with other people, who are the people they share the links with and why? 16 participants (73%) of our qualitative interview reported to share video links in general. The motivations why they shared a link varied, but the main reasons or sharing a video can be summarized as follows:

- To allow others to laugh and be amused with some entertainment video.
- To show a new or interesting video to a selected group of persons (friends, family or colleagues) that share the same hobbies/interests or movie/music taste.
- To inform people about breaking news, upcoming events or products (product reviews).
- To show others what one has accomplished, or what one has experienced. Such videos are mostly filmed by them and serve as self-representation.

If we compare this to the results of the quantitative study where we asked the same question (based on ordinal scale items), we can clearly see some similarities. People are often sharing entertainment and fun videos (rating average of 4.08) as expected, but still an interesting fact is that most people don't forward links without any special reasons, they think about it more consciously, consider which other persons could be interested in the video, and only forward it to a certain group.

	Never (1)	Rarely (2)	Sometimes (3)	Often (4)	Always (5)	Rating Average
I want others to learn about things, transport knowledge	3.5%	20.6%	**36.2%**	31.2%	8.5%	3.21
I want others to be entertained, have fun watching the video	1.4%	1.4%	8.5%	**65.5%**	23.2%	4.08
I assume that friends I forward the link are interested in this video	1.4%	2.1%	19.7%	**58.5%**	18.3%	3.90
To share information about products or events	10.0%	33.6%	**35.0%**	17.1%	4.3%	2.72
To share an experience, to share something I have experienced	12.9%	28.1%	**31.7%**	23.7%	3.6%	2.77
To share emotions (like anger, disgust, beauty, love etc.)	21.9%	27.0%	**30.7%**	15.3%	5.1%	2.55
Someone asked me to share this video	**36.5%**	30.7%	22.6%	10.2%	0.0%	2.07
No special reason	**36.8%**	28.8%	27.2%	5.6%	1.6%	2.06

Table 7: Why are people sharing links?

People do share video links with various groups, mostly friends, family, colleagues and the public. The participants of the qualitative interviews shared video links with the groups seen in Figure 10. The most common group that people share links with is friends (55.56 %), followed by an identical share of family members and colleagues (18.52% each).

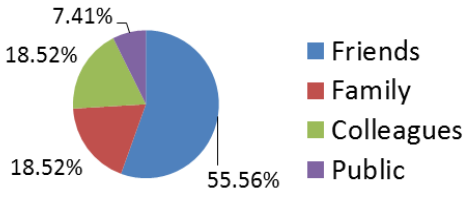

Figure 10: Sharing groups of qualitative interviews

This also strongly relates to our findings in the quantitative survey, where we asked the same question. Of the 217 participants, 143 (65.9%) shared video links in general with the following groups:

	Never (1)	Rarely (2)	Sometimes (3)	Often (4)	Always (5)	Rating average
Friends	0.7%	0.0%	7.8%	**59.6%**	31.9%	4.22
Family members	9.6%	22.1%	**31.6%**	30.9%	5.9%	3.01
Colleagues	6.6%	27.7%	25.5%	**35.0%**	5.1%	3.04

Table 8: With whom are people sharing links

The quantitative study also confirmed that "friends" are the most preferred group to share links with, resulting in an average rating of 4.22, followed by an almost identical average rating of family members with 3.01 and colleagues with 3.04. This correlates very well to our findings in the qualitative interviews.

In our exploratory study we also wanted to find out how often and by which means video links get shared with other people. Results from our quantitative and qualitative study showed that most people share video links several times per month (*Figure 11*). Only in rare cases people reported to share video links more than once a day.

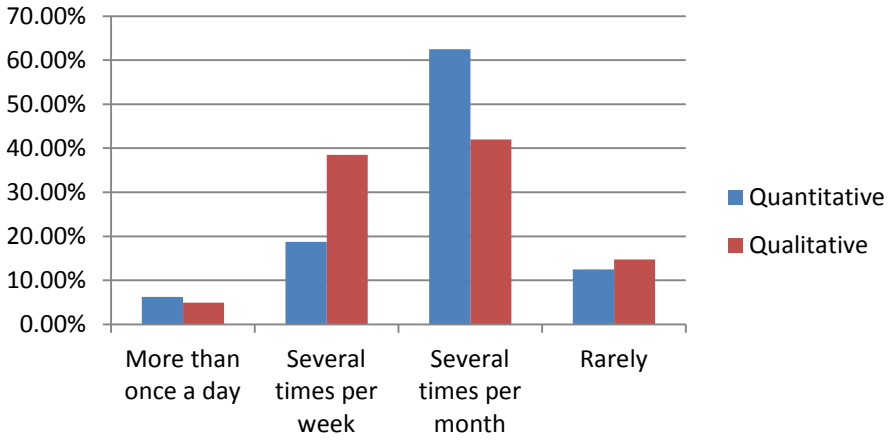

Figure 11: Link sharing frequency

Our qualitative interviews showed that the most common mechanism to share video links is still email, followed by instant messaging (*Figure 12*). Social media or on-site mechanisms (e.g., the "Share" button on YouTube) are not yet very popular. Some people

also just share a “link” or give a video recommendation while talking to another person. They simply explain where to find it or just give hints for search terms to find the video, which can also be considered a kind of sharing.

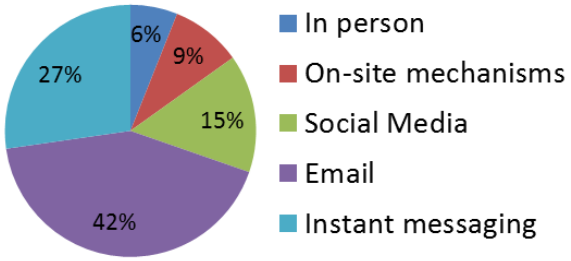

Figure 12: Link sharing mechanisms

The quantitative survey shows some differences in the users’ preferences. Participants favored sharing links over social media the most, followed by email and instant messaging (Table 9). We also validated if this result may be biased through the great response from the Reddit community, but this is not the case. When selecting all responses except those coming from Reddit, we got nearly the same results.

	Never (1)	Rarely (2)	Sometimes (3)	Often (4)	Always (5)	Rating average
In person by showing them on a computer screen or mobile device	9.5%	**36.5%**	35.8%	16.8%	1.5%	2.64
On-site mechanisms (“share” button)	25.9%	17.8%	22.2%	**27.4%**	6.7%	2.71
Social Media (Twitter, Facebook, Reddit, Digg etc.)	18.0%	5.8%	19.4%	**43.2%**	13.7%	3.29
Email	14.3%	25.7%	22.1%	**31.4%**	6.4%	2.90
Instant Messaging	20.6%	17.6%	**28.7%**	23.5%	9.6%	2.84

Table 9: Which mechanisms do you use to share video links?

What do people do when they receive a video link?

Of all 217 participants, people watch videos from recommendations (e.g. personal, via email, etc.) quite often. The majority of people (93) watch videos from recommendations once a week or more, followed by 67 people who watch them once a month or more (25 once a day or more, 26 rarely, 3 never).

In our quantitative survey, we asked how people react when receiving a video link: the average rating is shown in Figure 13. In most cases people try to watch the video right away and avoid delaying it or they watch it as soon as possible. Very few people ignore it and watch it sometime later or make a note or bookmark for that video to watch it when they have time for it.

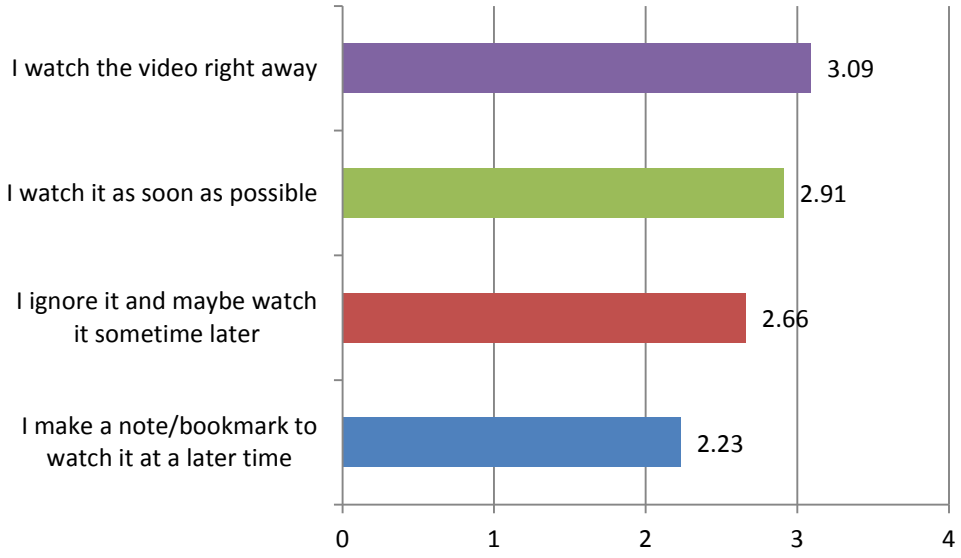

Figure 13: Average rating for different behavior patterns when receiving a video link.

4.3 Online video consumption behavior

In this section we look at *scenarios*, or *situations*, associated with video watching. A scenario, or situation, is defined as the set of conditions under which video consumption takes place. It encompasses the event that triggered the intention to watch the video content as well as the time, place and device in which it was watched.

Based on our qualitative interviews, each user reported on up to three scenarios when they consumed video content. Overall, we collected 49 specific scenarios – i.e., an average of 2.2 scenarios per user. In this section we summarize our findings and report on the most interesting situations experienced by our participants. We then also compare our findings with the quantitative survey in order to confirm that some of the behavior is common throughout video retrieval. The most important difference in methodology between the two studies is that in the qualitative interviews users were asked open questions and relied on memory of past events, whereas in the quantitative survey, they were asked more

specific questions regarding their viewing preferences, interests, events which caused them to watch a video, as well as preferred place and time to watch videos.

In summary, active search by a user was performed in 67% of the reported scenarios and that in the remaining 33% of the scenarios a direct link to a video was received (via Facebook, email, instant messaging, or face-to-face conversations).

In the qualitative studies, more than half of the videos associated with the 49 reported scenarios were watched in the evening or at night, followed by the afternoon and morning (Figure 14).

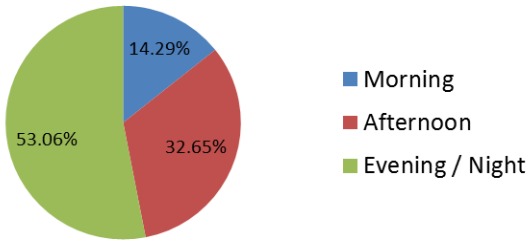

Figure 14: Time of day people prefer to watch online videos (qualitative)

We also found similar results in our quantitative study, but we introduced separate choices for evening and night (Figure 15). These findings are not very unusual due to the fact that most participants are at work or at school during the morning and most of the afternoon.

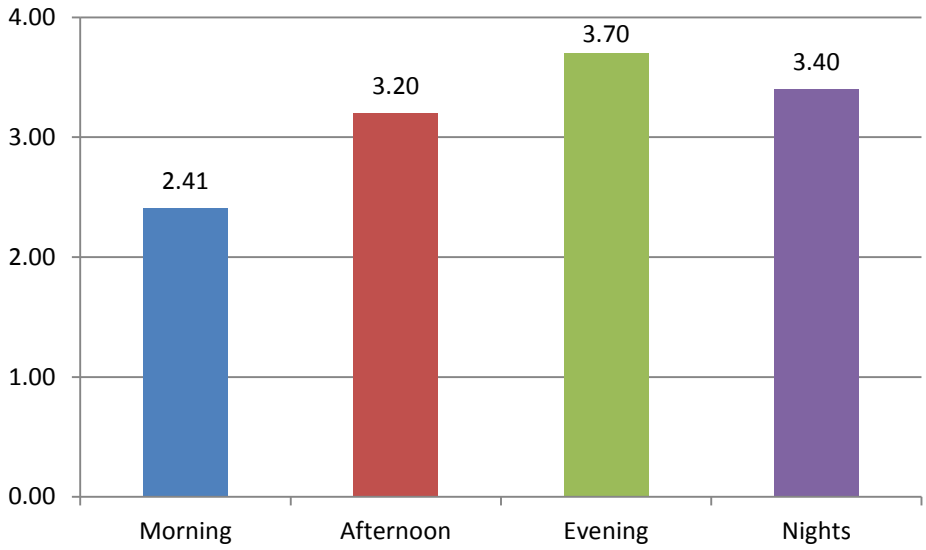

Figure 15: Average rating of preferred time to watch online videos (quantitative)

Additionally we investigated where videos are being watched. Using the same 49 situations in the qualitative interviews, we learned that in 40 cases the videos were watched at home, four videos were watched at the University and only three videos where watched at work. In two situations the video was watched at a friend's house while sitting together. One can argue that the number of videos watched at work might in reality be higher than reported due to the fact that people may not want to share that experience if it is not work related. Within our quantitative questionnaire, we asked where people prefer to watch videos by giving them ordinal scale items ranging from "Never (1)" to "Once a day or more (5)". The rating average can be seen in Figure 16. Clearly watching videos at home is the most preferred one, while all other choices scored significantly lower. An interesting fact is that watching videos "on the go" is not yet very popular, despite the growing usage of smartphones or tablets with online video capability.

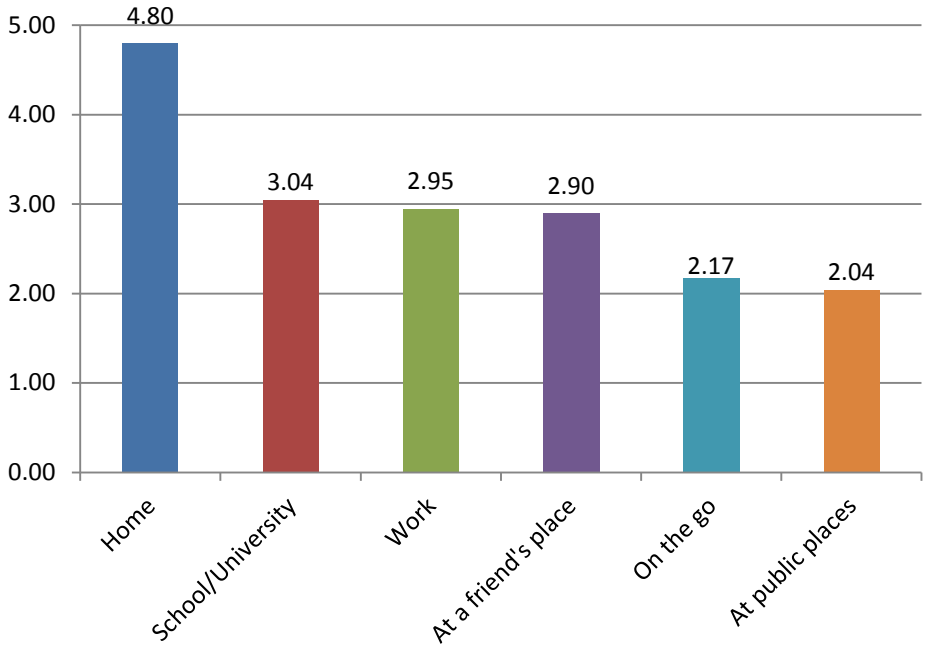

Figure 16: Where do people prefer to watch videos

In the qualitative interviews, for all the 49 reported scenarios the device used in 43 of the cases was either a PC/laptop or notebook, 5 times participants watched the video on a TV and just in one case a mobile device was used.

In the qualitative study, most of the videos watched in the reported situations were of short length: 45% are only 1 to 4 minutes long and 35% are 5-10 minutes long (*Figure 17*).

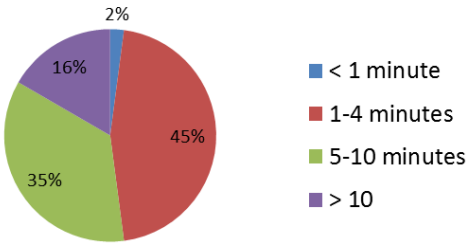

Figure 17: Average video lengths

During the quantitative surveys, in order to test the hypothesis that most of the videos watched are of short length, we specifically asked the question if participants prefer short videos (up to 5 minutes) by letting them express their level of agreement. Out of the 217 participants, 1.4% strongly disagreed, 2.8% disagreed, 30.0% were neutral where 37.1% agreed and 28.6% strongly agreed to this statement.

Matching events to genre/category

While analyzing and evaluating the interview transcripts from the qualitative questionnaire, we chose to assign each scenario to one of the 15 predefined categories by YouTube. To accomplish this, we tried to infer which video the user would have watched by analyzing the interview transcripts. It was possible for all scenarios to assign a clear and non-ambiguous category. Furthermore, we characterized seven events that caused the video consumption in the first place. These seven events are described as follows:

- Leisure: Videos in this category were watched in leisure time, mostly out of boredom, for entertainment, or to kill time. Active search by the user is performed and a video is selected from the results.
- Pull services: Video links pulled by the users, e.g., instant message, email. No search was needed because it was a direct link to a specific video.
- External events: Follow-up active search by users triggered by an event on TV, radio or newspaper, etc.
- Face-to-face communication: Users talked to another person about a certain topic or about a specific video already. At a later time, they then researched or followed up on that topic by actively searching.
- Task: Users had to solve a specific task and a video is retrieved in course of solving the task.

- Facebook/RSS/Feature usage: Facebook, RSS and any feature usage can be seen as a push service. Users get the information and click on a video link if it sounds interesting for them. The user performs no active search in this case.

The results of this matching between events and genres can be seen in Table 10. The table shows only 10 columns because none of the situations matched any of the following YouTube categories: Entertainment, Howto & Style, Nonprofits & Activism, People & Blogs, Pets & Animals.

	Autos & Vehicles	Comedy	Education	Film & Animation	Gaming	Music	News & Politics	Science & Technology	Sports	Travel & Events		
Leisure				2	1	**10**			3	1	17	34.7%
Push services									4		4	8.2%
External events	1	2		2		3	1	1	3		13	26.5%
Face-to-face communication						1		1			2	4.1%
Task	1		2		1			2	1		7	14.3%
Pull services						3			3		6	12.2%
	2	2	2	4	2	17	1	4	14	1	49	
	4.1%	4.1%	4.1%	8.2%	4.1%	34.7%	2.0%	8.2%	28.6%	2.0%		

Table 10: Matching between events (rows) and YouTube categories (columns).

We can see from the results that the most watched videos belong to the genre "Music" (34.7%) followed by "Sports" (28.6%). The most frequent trigger event was "leisure" with 17 occurrences, closely followed by "external events" with 13 instances. For example, the "Entertainment" category was not assigned to any of the videos reported by the interviewees and yet one can claim that a big goal of viewers spending their leisure time is to entertain themselves. The trigger event having the biggest spread in categories (seven categories out of 10) is "external events", marked blue. This can be easily explained with the broad semantics of the "external event". More specifically, we do not know about the semantics, topic or content of the external event that triggered the retrieval process. The same goes for task solving, which spreads across six genres and shows no peaks in absolute numbers.

Even more interesting is the inverse investigation: how do genres relate to trigger events. Once again, music gives a good example due to its dominance in the reported instances, marked grey. There is an indication that retrieving music videos is very much related to the leisure cluster. The "Sports" genre, on the other hand, features nearly all of the six trigger events.

In our quantitative questionnaire we also wanted to find out more about events that triggered video consumption, therefore we asked how often users follow any of the following statements. The results can be seen in Table 11.

	Never (1)	**Rarely (2)**	**Once a month or more (3)**	**Once a week or more (4)**	**Once a day or more (5)**	**Rating Average**
I watch videos when I am bored	6.9%	21.7%	18.9%	**30.4%**	22.1%	3.39
I watch videos (or follow video links) I get in emails	12.6%	**30.7%**	21.9%	30.2%	4.7%	2.84
I watch videos (or follow video links) I get in instant messages	20.8%	**27.4%**	19.8%	25.5%	6.6%	2.70
I search for videos to follow up something on TV/radio/newspaper/magazine	9.8%	24.7%	**37.2%**	24.7%	3.7%	2.88
I search for videos to follow up something after I talked to another person about it	3.3%	22.2%	**42.9%**	27.4%	4.2%	3.07
I search for videos to get help on what I'm going to do (e.g. how to)	6.6%	23.5%	**38.0%**	27.2%	4.7%	3.00
I watch videos published on RSS streams/feeds/blogs	**28.9%**	27.0%	24.2%	16.6%	3.3%	2.38
I watch videos posted on Facebook (MySpace, Twitter, etc.)	21.6%	17.4%	14.6%	**32.4%**	14.1%	3.00
I search for videos to learn how something works / looks like	4.2%	21.5%	**41.1%**	27.1%	6.1%	3.09

Table 11: Events that lead to watching videos online

Additionally we performed Spearman's rank correlation analysis and our quantitative survey data to further support the hypotheses that users' different intentions to watch videos correlate well with the genre of the watched videos. Table 12 shows the nine most significant events triggering video consumption (one per row) and the 15 YouTube categories used for video genres in our surveys (one per column). The Spearman's correlation coefficient r_s for each combination of event and genre/category and its two-sided significance p are also shown. High values of r_s, together with a high statistical significance ($p < 0.05$) indicate a relation between two variables. The three strongest relations of events and genres for each event are marked green, and the three weakest relations are marked red in Table 12. Based on those correlations we postulate that if users use the video platform with a specific intention, some categories will likely contain a larger number of relevant results than other categories. Besides the ranking of categories, we want to point out other insights derived from the numbers in Table 12, which include:

- The strongest correlation (r_s =0.404, p=0.000) was found between the event "I search for videos to get help on what I'm going to do" and the "Howto & Style" category.

- The “Music” category has very poor statistical significance with $p > 0.094$ across all events in *active search*. One explanation can be found in statements given by our survey participants. Often they choose to just listen to music on YouTube without paying attention to the video at all while working on other, more important tasks.
- The “Education”, “Film & Animation”, and “Non-profits & Activism” categories did not make it to the “top 3” list for any event either.
- All negative correlations appeared with poor values for statistical significance ($p > 0.139$).
- The event “I search for videos to follow up something on TV/radio/newspaper/magazine” has the biggest number of significant positive correlations, namely 11 (out of 15).
- The event “I watch videos (or follow video links) I get in emails” has the lowest number of significant correlations, namely 6 (out of 15).

		Autos & Vehicles	Comedy	Education	Entertainment	Film & Animation	Gaming	Howto & Style	Music	News & Politics	Nonprofits & Activism	People & Blogs	Pets & Animals	Science & Technology	Sports	Travel & Events
Active search																
I watch videos when I am bored	r_s	**-0.078**	**0.362**	0.159	**0.232**	0.217	0.227	0.114	0.111	0.197	0.143	**0.262**	0.221	**0.039**	0.187	**0.092**
	p	**0.257**	**0.000**	0.020	**0.001**	0.001	0.001	0.098	0.108	0.004	0.038	**0.000**	0.001	**0.569**	0.006	**0.180**
I search for videos to follow up something on TV/radio/newspaper/magazine	r_s	0.139	**0.040**	0.180	0.219	0.229	**0.096**	0.212	0.115	**0.282**	0.209	**0.326**	0.202	**0.086**	0.206	**0.253**
	p	0.044	**0.561**	0.009	0.001	0.001	**0.165**	0.002	0.094	**0.000**	0.002	**0.000**	0.003	**0.212**	0.003	**0.000**
I search for videos to follow up something after I talked to another person about it	r_s	0.141	0.177	0.172	0.169	0.222	0.122	0.159	**0.089**	**0.055**	0.134	0.157	**0.232**	**-0.004**	**0.257**	**0.251**
	p	0.041	0.010	0.012	0.013	0.001	0.078	0.021	**0.179**	**0.431**	0.053	0.024	**0.001**	**0.950**	**0.000**	**0.000**
I search for videos to get help on what I'm going to do (e.g. How-to)	r_s	**0.237**	**0.037**	0.154	0.053	**0.022**	0.203	0.404	0.025	**0.033**	0.084	**0.222**	0.174	0.189	**0.267**	0.132
	p	**0.001**	**0.593**	0.026	0.447	**0.754**	0.003	0.000	0.713	**0.635**	0.228	**0.001**	0.110	0.006	**0.000**	0.057
I search for videos to learn how something works / looks like	r_s	0.276	**0.042**	0.279	0.071	**-0.023**	0.200	**0.400**	**-0.067**	0.056	0.208	**0.290**	0.211	**0.331**	0.146	0.174
	p	0.000	**0.544**	0.000	0.302	**0.734**	0.004	**0.000**	**0.336**	0.421	0.002	**0.000**	0.002	**0.000**	0.034	0.120
Direct Link Access																
I watch videos (or follow video links) I get in emails	r_s	0.124	0.116	0.064	0.094	**-0.010**	**-0.010**	**0.216**	0.147	**-0.048**	0.048	**0.211**	0.142	0.036	**0.258**	0.185
	p	0.072	0.092	0.351	0.170	**0.886**	**0.882**	**0.002**	0.032	**0.488**	0.493	**0.002**	0.040	0.602	**0.000**	0.007
I watch videos (or follow video links) I get in instant messages	r_s	0.110	**0.223**	**0.103**	0.171	0.157	**0.224**	0.136	0.189	**-0.102**	**0.024**	0.140	0.105	0.074	**0.334**	0.175
	p	0.095	**0.001**	**0.136**	0.013	0.023	**0.001**	0.049	0.006	**0.139**	**0.725**	0.044	0.128	0.286	**0.000**	0.011
I watch videos published on RSS streams/feeds/blogs	r_s	0.172	0.130	0.154	0.065	**0.018**	0.155	0.215	**-0.007**	0.199	0.235	**0.372**	0.222	**0.100**	**0.238**	**0.274**
	p	0.013	0.061	0.026	0.354	**0.793**	0.027	0.002	**0.919**	0.004	0.001	**0.000**	0.001	**0.148**	**0.001**	**0.000**
I watch videos posted on Facebook (MySpace, Twitter, etc.)	r_s	**0.041**	0.187	0.216	0.184	0.125	0.108	0.140	**0.289**	**0.024**	0.077	**0.303**	0.118	**-0.029**	**0.309**	0.251
	p	**0.556**	0.006	0.002	0.007	0.069	0.120	0.043	**0.000**	**0.725**	0.268	**0.000**	0.088	**0.677**	**0.000**	0.000

Table 12: Event-genre correlation coefficients and statistical significance.

5. DISCUSSION

The qualitative interview and quantitative survey reported in this paper aimed at collecting insightful data on online video watching behaviors and how they relate to users' intentions. These studies addressed a number of aspects, from the triggering event that caused the video to be watched in the first place, to the time/place/device in which it was watched, to the way the video was shared with others.

First, we wanted to find out *how* and *why* people are using video platforms. Most participants reported using video platforms almost on a daily basis, with YouTube as the most popular platform. The majority of participants still preferred to consume online video on laptop or PC devices, but it is likely that in the future more and more content will be watched on a TV (possibly through set-top boxes such as Roku[7] or Boxee[8], or videogame consoles such as Xbox or Wii) or on various mobile devices such as smartphones and tablets. In both of our studies most of the participants watched video content at home during evening and night times. We also noticed a shift towards using online videos for educational and/or professional purposes.

We further addressed the actual user's behavior and motivation to retrieve videos. Various aspects influence a user's motivation to watch videos online. The most obvious reasons for video retrieval are of informational or educational nature or have some kind of entertainment background. Depending on the content of the video, different quality factors are important, namely picture, sound and content quality. Investigating overall favorite genres of participants, we found out that mostly entertainment-related categories – such as "Music", "Entertainment", and "Film & Animation" – are preferred.

We also addressed the question of *how* (and *by which means*) users share their video experience. Videos are either shared with a closed group of known people through various media (e.g., instant message, Facebook, etc.) or with an open community (e.g., Digg, Reddit). The most frequent group to share videos with is friends, followed by family, colleagues and the public at large. The reasons to share videos reported by participants included: to amuse other people, to inform them of relevant videos they could be interested in, to inform them about news, events or the like, or to share an experience. Based on our quantitative survey, users on average send video links at least several times per month mostly using social media functionality, email or instant messaging.

Last, we investigated relationships between video genres and events. Our qualitative results show that there is a relation between genre and triggering event. The biggest peak in absolute numbers can be observed in instances when people watched music videos for

[7] Roku Streaming Player, http://www.roku.com/ (accessed on February 10, 2012)
[8] Boxee, http://www.boxee.tv/ (accessed on February 10, 2012)

leisure, with 10 such instances (see Table 10). Moreover, only five out of ten genre categories were reported for the "leisure" trigger event. This indicates that for users who want to entertain themselves, a subset of categories (which could be grouped under the label "fun") might be more relevant than others (which could be grouped under the label "serious"), e.g., politics, science or education. In order to investigate this matter, we performed statistical analysis (Spearman's Rho) on our quantitative survey data (Table 12) to demonstrate that genres correlate well with our predefined events. One of the merits of finding correlations between triggering events and genres is that such correlations can also be used to circumvent the fact that YouTube categories (which were used as proxies for genres in this study) overlap and do not constitute a neatly organized taxonomy.

Notwithstanding the quality of the insight brought about by results from both studies, there are a few shortcomings that should be addressed in subsequent work. Among many possible improvements to the methodology, we would like to suggest: introducing quotas on demographics (to avoid too many participants from very few countries), supplementing the interview/survey answers with information about user behavior through log file analysis, conducting an even larger study to provide more solid statistical support to the correlations between genres and trigger events, fine-tuning the granularity and selection of genres, and revisiting the classification of trigger events in terms of granularity, coverage, and possible overlaps. We also believe that the development of additional similar studies – focused on specific needs and demographics that were not covered by the present work – could be beneficial to the advancement of the field.

In summary, the work reported in this paper – the first study of this kind – represents a first step in the direction of understanding why people retrieve and watch online videos. We postulate that if a video search and retrieval system could be designed with the ability to infer the possible trigger events based on the viewing behavior of a user, this knowledge could be employed to provide more relevant search results to users and/or filter irrelevant results and recommendations.

6. PROTOTYPE OF AN ADAPTIVE VIDEO RETRIEVAL SYSTEM

In this section, we present the basic aspects of a prototype of an adaptive video retrieval system that leverages the knowledge of users' intentions (and their relationship to video genres / categories) uncovered by our studies to provide better search results and a user interface adapted to the intentions and needs of its users. Our intention with the development of this prototype is to provide a better solution to the problem of including the user's context into video retrieval than the one offered by baseline video retrieval interfaces (e.g., YouTube). From a software engineering point of view, the results and insights obtained from

both studies reported in this paper have helped "*determine the correctness of the final software product with respect to the user needs and requirements*", i.e., the validation process of the software development cycle.

Our prototype provides a way to demonstrate how our most promising findings can be used to enhance the user experience in video retrieval. Our ultimate goal is to show the impact and benefit of the integration of user intentions in search and retrieval processes in multimedia information systems.

The first step in the development of the prototype has been focused on improving the precision of search results by exploiting the correlation between intentions and genres. The basic block diagram for the prototype can be seen in Figure 18 and consists of three main blocks:

1. The user interface (top), which includes a traditional text input box for typing the query and four graphically pleasant buttons which the user could press to communicate their intention to the system, namely: (i) to learn something; (ii) to be entertained; (iii) to get informed; or (iv) to solve a task.
2. A set of ad-hoc, hard-coded rules derived from the results of the studies reported in this paper (such as the correlation values reported in Table 12), so that videos of certain categories with higher correlation coefficients are ranked higher in the search results.
3. A collection of sources of video contents (bottom), e.g., Vimeo, YouTube, or Khan Academy.

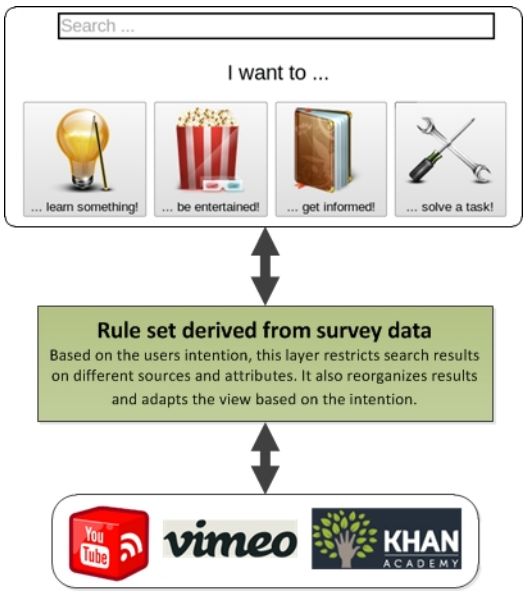

Figure 18: Block diagram of the prototype for an adaptive video retrieval system.

For better understanding, we outline a simple example scenario: let us assume that a user wants to learn about a specific topic and she types the query “Moonwalk”. Since the user has a *learning intention* in mind, she clicks on the *'I want to learn something'* button. Our pattern-based rule set then optimizes the search to certain sources and categories. In this example case, YouTube and the Khan Academy will be used as video sources and the videos will be ranked by categories from the strongest to the weakest correlation (according to the results in Table 12), leading to a result screen whose screenshot appears in Figure 19. Notice how three clips from the “How-to & Style” category (all of which are related to the “moonwalk” dance popularized by Michael Jackson) appear at the top of the screen, with the second-highest-ranked category (“Science & Technology”) appearing in a second block, mostly containing NASA footage from historical Apollo-era moon explorations. By prioritizing the categories that are most strongly correlated to the user’s intentions and adopting a visually pleasant layout that shows them in easily distinguishable blocks, we circumvent the ambiguity caused by the query term and provide an intuitive way to navigate to the desired result.

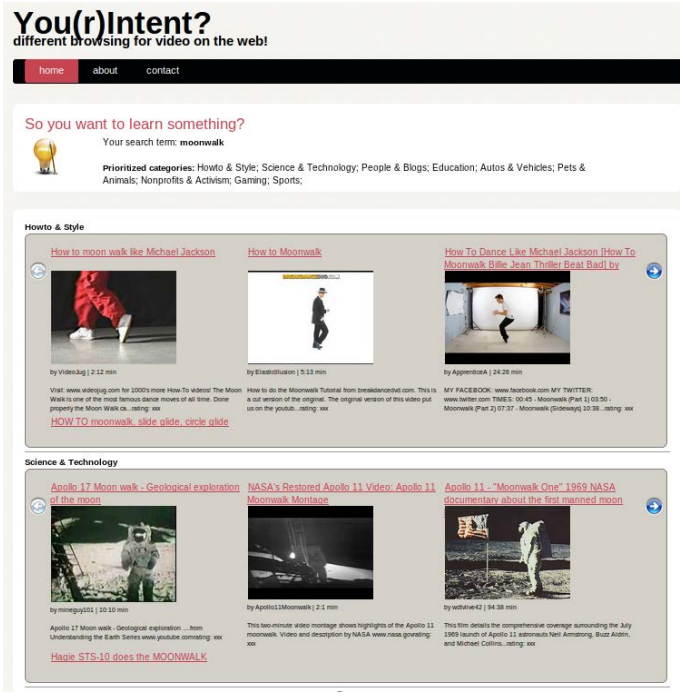

Figure 19: Result screen of the prototype when the user queries for videos containing the ‘moonwalk’ keyword and expresses the intention to learn something.

The adaptive video retrieval system presented in prototype form in this section is being expanded and improved to include additional options and features based on our survey results, such as:

- the ability to factor in the video quality of the resulting clips to match the user's intention: e.g., participants of our survey reported to prefer HD video quality when they want to watch a video to learn something;
- view adaptation based on the users' intention: e.g., survey participants reported that not every option and metadata information needs to be displayed at all times for every intention.

7. CONCLUDING REMARKS

In this section we highlight the contributions of this work and suggest opportunities for future work on the topic of "user intentions in video retrieval".

7.1 Contributions

How can we improve video retrieval and the precision of search results at a time where digital video content on the Internet is constantly growing? We have to start by understanding how users behave on video platforms and learn as mush as we can about their intentions. Based on that knowledge we can develop better video retrieval models and systems. This paper introduces user intentions in the field of video retrieval on the web. The qualitative and quantitative studies reported in this paper aimed at collecting insightful data on online video watching behavior and how they relate to users' intentions. They addressed a number of aspects, from the triggering event that caused the video to be watched in the first place, to the time/place/device in which it was watched, to the way the video was shared with others.

We further addressed the actual user's behavior and motivation to retrieve videos. Various aspects influence a user's motivation to watch videos online. The most obvious reasons for video retrieval are either entertainment-related or informational / educational in nature. Depending on the content of the video, different quality factors are important, namely picture, sound and content quality. Investigating overall favorite genres of participants, we found that their preference falls mostly on entertainment categories such as *"Music"* and *"Film & Animation"*.

We also addressed the question of how users share their video experience and by which means. Videos are either shared with a closed group of known people through various media (e.g., instant message, *Facebook*, etc.) or with an open community (e.g., *Digg*, *Reddit*). The most frequent group to share videos with is friends, followed by family, colleagues and the

public at large. Reasons to share videos include: to amuse other people, to send relevant videos they could be interested in, to inform people about news, events or the like; or to share a personal experience. Based on our quantitative survey, users on average send video links at least several times per month mostly using social media functionality, email or instant messaging.

Last, we investigated relationships between video genres and events. The classification of situations into video genre and trigger event (which ultimately maps to the user's intention) allowed for an exploratory analysis of relations between genres and event classes that actually lead to video retrieval. Our qualitative results have shown that there is a relation between genre and trigger event. In order to further investigate this matter, we performed statistical analysis (Spearman's Rho) on our quantitative survey data to show by numbers that genres correlate well with our predefined events. This knowledge was eventually used to build rules (of the type "if the user's intention is X, then videos are more likely to be of category Y") that could be used to translate the results of our studies into a building block of an actual video retrieval system.

Based on our results, we designed and implemented a prototype of an adaptive video retrieval interface that resembles a basic video retrieval platform. The big differences of our approach were:

- It offers the user four predefined intentions which could be selected when performing a video search; and
- It presents the search results following a different ordering and grouping, based on the selected intention (and a rule set that maps intentions / trigger events to genres / categories).

7.2 Future Work

The work described in this paper is of exploratory nature and, consequently, opens the door for many promising research directions and opportunities for future work, such as the development of additional studies, focused on specific needs (e.g., a more detailed analysis of granularity and selection of genres) and demographics not covered by the present work. The statistical analysis of the numerical results of such studies may uncover stronger and more interesting correlations in video search and retrieval behavior, which should help extend multimedia information retrieval systems towards intelligent systems capable of taking user context like intentions, tasks, and goals into account.

We are currently finishing the development of the prototype of the adaptive video retrieval system based on the results and insights obtained from the two studies reported in this paper. After deploying the prototype, we plan to perform user acceptance and satisfaction tests, including questions on quality of video search results and ease of use compared to popular video portals, notably YouTube. Evaluation methods will include analysis of mouse tracking heat maps and semi-structured interviews with the study participants.

ABOUT THE AUTHORS

Christoph Lagger: Christoph Lagger is a student at Klagenfurt University in Austria. He is currently working on his Master's thesis "User intent classification in video retrieval".

Mathias Lux: Dr. Mathias Lux is an Assistant Professor at Klagenfurt University in Austria. He received his PhD from Graz University of Technology. His research focuses on multimedia information systems, more specifically on visual information retrieval and user context. He has (co-) authored more than 50 peer-reviewed articles in conference proceedings, journals and books and he is a founding member of the Multimedia Metadata Community.

Oge Marques: Dr. Oge Marques is an Associate Professor in the Department of Computer & Electrical Engineering and Computer Science at Florida Atlantic University (Boca Raton, FL). His research interests include: image processing, analysis, annotation, search, and retrieval; human and computer vision; and video processing and analysis. He is the (co-) author of 4 (four) books in these topics, including the textbook "Practical Image and Video Processing Using MATLAB" (Wiley, 2011). He has also published several book chapters and more than 50 refereed journal and conference papers in these fields. He is a senior member of the IEEE and a senior member of the ACM.

ACKNOWLEDGMENTS

The authors would like to thank all participants from the qualitative interviews and the quantitative survey for taking the time to answer our questions.

REFERENCES

[1] The Official YouTube Blog. *Holy Nyans! 60 hours per minute and 4 billion views a day on YouTube*, http://youtube-global.blogspot.com/2012/01/holy-nyans-60-hours-per-minute-and-4.html, January 23, 2012 (accessed February 05, 2012).

[2] Cisco Systems Inc. Cisco Visual Networking Index: Forecast and Methodology, 2009 - 2014. Technical report, Cisco, 06 2010.

[3] Bortz, J., Döring, N. 2006. *Forschungsmethoden und Evaluation: für Human- und Sozialwissenschaftler, 4te Auflage*, Springer, Berlin

[4] Koji Miyauchi, Taro Sugahara, and Hiromi Oda. 2009. *Relax or study? A qualitative user study on the usage of live mobile TV and mobile video*. Computer in Entertainment, Vol. 7, No. 3, Article 43

[5] Södergard, C. 2003. *Mobile television – technology and user experiences. Report on the Mobile-TV project*. Report VTT Publications 506, VTT Information technology

[6] Petteri Repo, Kaarina Hyvönen, Mika Pantzar, and Päivi Timonen. 2004. *Users Inventing Ways to Enjoy New Mobile Services -- The Case of Watching Mobile Videos*. In Proceedings of the Proceedings of the 37th Annual Hawaii International Conference on System Sciences (HICSS'04) - Track 4 - Volume 4 (HICSS '04), Vol. 4. IEEE Computer Society, Washington, DC, USA, 40096.3-.

[7] W. Alex Vorbau, April Slayden Mitchell, and Kenton O'Hara. 2007. *"My iPod is my Pacifier": An Investigation on the Everyday Practices of Mobile Video Consumption.* In Proceedings of the Eighth IEEE Workshop on Mobile Computing Systems and Applications (HOTMOBILE '07). IEEE Computer Society, Washington, DC, USA, 29-33.

[8] Yanqing Cui, Jan Chipchase, and Younghee Jung. 2007. *Personal TV: a qualitative study of mobile TV users*. In Proceedings of the 5th European conference on Interactive TV: a shared experience (EuroITV'07), Pablo Cesar, Konstantinos Chorianopoulos, and Jens F. Jensen (Eds.). Springer-Verlag, Berlin, Heidelberg, 195-204.

[9] Wei Song and Dian Tjondronegoro. 2010. *A survey on usage of mobile video in Australia*. In Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction (OZCHI '10). ACM, New York, NY, USA, 5-8.

[10] Shelley Buchinger, Simone Kriglstein, and Helmut Hlavacs. 2009. *A comprehensive view on user studies: survey and open issues for mobile TV*. In Proceedings of the seventh european conference on European interactive television conference (EuroITV '09). ACM, New York, NY, USA, 179-188.

[11] Martin J. Halvey and Mark T. Keane. 2007. *Analysis of online video search and sharing*. In Proceedings of the eighteenth conference on Hypertext and hypermedia (HT '07). ACM, New York, NY, USA, 217-226.

[12] Meeyoung Cha, Haewoon Kwak, Pablo Rodriguez, Yong-Yeol Ahn, and Sue Moon. 2007. *I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system*. In Proceedings of the 7th ACM SIGCOMM conference on Internet measurement (IMC '07). ACM, New York, NY, USA, 1-14.

[13] Xu Cheng; Dale, C.; Jiangchuan Liu. 2008. *Statistics and Social Network of YouTube Videos*. *Quality of Service, 2008. IWQoS 2008.* pp.229-238, 2-4 June 2008

[14] Mack N., Woodsong C., MacQueen K., Guest G., Namey E. 2005. *Qualitative Research Methods: A Data Collector's Field Guide, 2008. Family Health International (FHI)*, NC, USA