Are Trustworthy Health Videos Reachable on YouTube?
A Study of YouTube Ranking of Diabetes Health Videos

Randi Karlsen[1], José Enrique Borrás Morell[1] and Vicente Traver Salcedo[2]
[1]*Department of Computer Science, UiT The Arctic University of Norway, Tromsø, Norway*
[2]*ITACA - Health and Wellbeing Technologies, Universidad Politècnica de València, Valencia, Spain*

Keywords: Consumer Health Information, Information Retrieval, Social Networks, YouTube, Health Video Retrieval, Ranking Evaluation, Diabetes, Personal Health.

Abstract: While health consumers are increasingly searching health information on the Internet, information overload is a serious obstacle for finding relevant and good-quality information among inaccurate, obsolete or incorrect health information. While a lot of information exists, information from credible sources, such as hospitals and health organisations, may be difficult to find. The aim of this study is to analyse ranking of diabetes health videos on YouTube over a time period, to learn whether videos from credible sources are ranked sufficiently high to be reachable to users. 19 diabetes-related queries were issued to YouTube each day over a 1.5-month period, and in total 2584 videos from credible sources was detected and their ranking position tracked. We found that only a small number of the tracked videos were in practice available to the user, as most videos were given a persistent low ranking. Also, since ranking is fairly stable, users cannot expect to find many new videos (from credible sources) when issuing a query multiple times. We conclude that new tools are needed that enable health video retrieval based on requirements concerning not only relevance and popularity, but also credibility of the sources and trustworthiness of the videos.

1 INTRODUCTION

Internet has, during the last years, become a major source of health information (AlGhamdi and Moussa, 2012; Griffiths et al., 2012; Madathil et al., 2015). Users are typically searching for information about specific diseases or symptoms, treatment side effects, second opinions, complementary or alternative medicines, search for others with similar health concerns and follow personal health experiences through blogs (de Boer et al., 2007; Diaz et al., 2002; Fox, 2011b; Powell et al., 2011). Also, online health information is used, not only by health consumers to gain knowledge about some health issue, but also by physicians, for clinical decision support and for education purposes (Hughes et al., 2009).

However, a general problem when searching the Internet, is the information overload and difficulty of finding relevant information satisfying the information need. Adding to this problem, too many websites have inaccurate, missing, obsolete, incorrect, biased or misleading information, and it may be difficult to distinguish between trustworthy and specious information (Briones et al., 2012; Madathil et al., 2015; Pant et al., 2012; Shabbir et al., 2013; Singh et al., 2012; Steinberg et al., 2010). When people are relying on online health information to take medical decisions or handle their health issues, it is obviously of highest importance that the health information provided to users is not only relevant, but also correct and trustworthy. Existing search engines select and rank information based on relevance to a search query and popularity. Evaluating quality aspects, such as reliability and validity of information, is currently left to the user. Thus, the overwhelming amount of health information together with the mixed quality, makes it difficult for users to identify good-quality health information on the Internet, especially when users are not familiar with new technologies or when their health knowledge is limited. Also, certification approaches, such as the ethical HON code, are not solving the issue (Diaz et al., 2002).

Health information on the Internet comes from different sources, including hospitals, health organisations, government, educational institutions, for-profit actors and private persons reporting on personal experiences with some disease. User studies have shown that the credibility of an information source is

Karlsen R., BorrÃąs Morell J. and Traver Salcedo V.
Are Trustworthy Health Videos Reachable on YouTube? - A Study of YouTube Ranking of Diabetes Health Videos.
DOI: 10.5220/0006114000170025
In *Proceedings of the 10th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2017)*, pages 17-25
ISBN: 978-989-758-213-4

one of the most powerful factors affecting information credibility (Freeman and Spyridakis, 2009). Users are for example more likely to trust health information published or authored by physicians or major health institutions than information provided by other sources (Dutta-Bergman, 2003; Moturu et al., 2008; Bermúdez-Tamayo et al., 2013). Such studies show that users show greater interest in health information published by professional sources, such as hospitals and health organisations, since these are considered more credible than the average health information on the Internet.

In our study we focus on health information provided through videos on YouTube and investigate to what extent health videos from professional sources, such as hospitals and health organisations, are available to the user. YouTube is today the most important video-sharing website on the Internet (Cheng et al., 2008). It has over a billion users (almost one-third of all people on the Internet) and every day people watch hundreds of millions of hours on YouTube and generate billions of views (YouTube, 2016). YouTube social media tools allow users to easily upload, view and share videos, and enable interaction by letting users rate videos and post comments.

YouTube is increasingly being used to share health information offered by a variety of sources (channels), including hospitals, organisations, government, companies and private users (Bennett, 2011). However, it may be difficult to find videos from credible channels, since YouTube video ranking is known to favour content from popular channels. This may cause for instance hospital videos, where social interaction through likes/dislikes and comments are not so common, to appear low in the ranked list. Also, YouTube ranking does not focus on trustworthiness, and both misleading and incorrect videos may well be popular and may therefore be given a high ranking (Briones et al., 2012; Shabbir et al., 2013).

A considerable amount of literature has been published on YouTube data analysis, such as studying relations between video ratings and their comments (Yee et al., 2009) or focusing on the social networking aspect of YouTube and social features (Cheng et al., 2008; Chelaru et al., 2012). Studies of YouTube performance have mainly focused on YouTube in general, rather than on specific domains, such as health. However, there have recently been some studies evaluating YouTube health video content with respect to their quality of information for patient education and professional training (Gabarron et al., 2013; Topps et al., 2013). Such studies, focusing on different areas of medicine, include the work of (Briones et al., 2012; Singh et al., 2012; Steinberg et al., 2010; Butler et al., 2013; Schreiber et al., 2013; Murugiah et al., 2011; Fat et al., 2011; Azer et al., 2013). In these studies, reviewers evaluate the quality or content of selected videos, and assess their usefulness as information source within their respective area.

This paper reports on a study where we tracked diabetes health videos on YouTube over a period of 1.5 month, to gain knowledge on how videos from professional channels are ranked on YouTube. The study was intended to answer the following questions: *"Where are videos from hospitals and health organisations ranked on YouTube?" "Are these videos ranked in positions that make them reachable to users?"* To the best of our knowledge, there has previously not been conducted a study where the availability of YouTube health videos has been tracked over time, as was done in our work.

The structure of the paper is the following. The next section presents the methodology used in our study. Section 3 presents the results of the work, while findings are discussed in Section 4. Section 5 concludes.

2 METHOD

This study is based on health videos obtained from YouTube through textual search queries on diabetes-related issues. We set up a test environment, where 19 diabetes-related queries were issued to YouTube each day over a period of 1.5 months, from March until April 2013. During this period, we daily collected the top 500 YouTube results for each query. Videos from white-listed (presumably credible) sources were identified and tracked during each day of the study, and their ranking position registered.

We implemented a system that for each day automatically issued the 19 queries and extracted information about the top 500 YouTube results. In addition to ranking position, we collected information such as video name and identifier, channel identifier, number of likes, dislikes and comments to the video. All 19 queries included the term "diabetes" and were focused towards different aspects concerning the disease. We used queries such as "diabetes a1", "diabetes glucose", "diabetes hyperglycemia" and "diabetes lada", and issued them as regular search queries on the YouTube home page using an anonymous profile (to avoid any bias) and with language option set to English. Video ranking was obtained by parsing the html of the result page, while video and channel information were collected through YouTube API version 2.0. All search queries can be seen in Table 1.

Through our study of YouTube health videos, we

Table 1: List of You Tube search queries.

diabetes type 1	diabetes hyperglycemia	diabetes insulin
diabetes type 2	diabetes hypoglycemia	diabetes injection
diabetes a1c	diabetes complications	diabetes glucose
diabetes food	diabetes retinopathy	diabetes mellitus
diabetes diet	diabetes ketoacidosis	diabetes education
diabetes obese	diabetes insulin pump	
diabetes lada	diabetes monitoring	

identified a number of (assumed) credible health video sources, such as hospitals and health organisations. We organised these channels into a hospital white-list and a health organisations white-list, containing channel identifiers for hospitals and health organisations respectively. In the light of user-interests in peer-to-peer healthcare (Ziebland and Herxheimer, 2008; Fox, 2011a), we also generated a third white-list of channels, which includes users that are active and predominantly publishing diabetes videos. Our white-lists contained a total of 699 channels, where 651 were hospitals, 30 were organisations and 18 were active users. We used the Health Care Social Media List started by Ed Bennett (Bennett, 2011) as an initial white-list, and expanded with more channels that we identified during our studies (Karlsen et al., 2013; Morell et al., 2012).

3 RESULTS

Using the 19 search terms shown in Table 1, we tracked the rank position of a total of 2584 YouTube health videos from white-listed channels during the test period. The videos were uploaded from 73 hospital channels, 30 organisation channels and 18 user channels. Among these, 2372 videos were uploaded to YouTube before the study began, whereas 212 videos were uploaded while the study was performed.

For each day of the study, our system detected a number of new videos from white-listed channels (for which tracking started and continued to the end of the study). The number of new videos was large in the first days of the study, and after some days stabilised at around 10 new videos each day.

3.1 Ranking of Videos from White-listed Channels

A goal of this study is to identify the number of videos from hospitals, health organisations and active users that are in practice available to users. When a YouTube search returns over 600.000 ranked videos (which is the case for the "diabetes type 1" search), it is obvious that the lowest ranked videos are not very available. A question is: "How far down in the ranked list of videos is a user willing to browse in order to find a relevant video?" The answer may to some extent be a matter of how patient the user is, but testing several hundred videos are beyond what can be expected from an average user.

To characterise videos w.r.t availability, we have grouped the tracked videos using ranking position intervals that were chosen based on our perception of how available videos in the different groups are. We consider videos ranked in position 1-40 as *highly available*, position 41-100 as *available*, position 101-200 as *not very available* and position 201-500 as *almost unavailable*. In this work, we assume that videos ranked lower than position 500, are in practice unavailable, and we have therefore tracked only videos appearing in the top-500 ranking.

To learn where videos from hospitals, health organisations and active users were ranked, we examined, for each day, the rank positions for all videos from our white lists, and determined the number of videos that were ranked in position intervals (1-40), (41-100), (101-200), and (201-500). Based on this study, we found that only a small number of videos from white-listed channels were in practice available to the user. When examining the top-40 ranked videos, we found that on average, only 3.2% were from hospitals, 10.4% from health organisations and 3.6% from active users. This means that we on average will retrieve approximately 7 videos from white-listed channels among the top-40 ranked videos. In the next position interval (41-100), the average number of videos from white-listed channels will be approximately 6. The results for all rank intervals are seen in Figure 1. The numbers for the top-500 videos (not given in Figure 1) were 2.3% from hospitals, 6.4% from health organisations and 1.8% from active users.

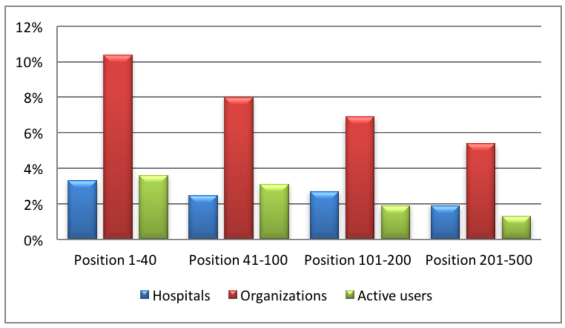

Figure 1: Average number (in percentage) of videos from white-listed channels within different rank intervals.

As our white-lists obviously do not contain every hospital or health organisation available, we took, for all 19 queries, the top-100 YouTube hits from one

Table 2: Classification of the tracked videos from white-listed channels, grouped according to their most frequent ranking position during the test period. The number of videos in each rank-position group are given (percentage between parentheses).

	Group A videos (pos. 1-40)	Group B videos (pos. 41-100)	Group C videos (pos. 101-200)	Group D videos (pos. 201-500)	Group E videos (pos. > 500)	Total no. of videos
Hospitals	26 (1.0%)	27 (1.1%)	52 (2.0%)	99 (3.8%)	334 (12.9%)	538 (20.8%)
Organizations	80 (3.1%)	89 (3.4%)	127 (4.9%)	279(10.8%)	1017 (39.4%)	1592(61.6%)
Active users	30 (1.2%)	33 (1.3%)	33 (1.3%)	55 (2.1%)	303 (11.7%)	454 (17.6%)
Total	136(5.3%)	149(5.8%)	212(8.2%)	433(16.7%)	1654(64.0%)	2584

day's search and manually checked the channel of each video. We found 15 new videos from hospitals not included in the white-list. This addition represents a modest 0.8% (of the 1900 videos), and does not significantly improve the amount of hospital videos given a high ranking.

3.2 Changes in Ranking Position

To investigate variations in video ranking, we first classified the tracked videos into five rank-position groups (Group A-E) according to their most frequent rank position during the test period. Occasional changes in ranking position were registered by counting the number of videos from each group that one or more days had a ranking position associated with a different group. We also calculated mean position and standard deviation for videos that were ranked within a top-500 position the whole period.

Among the 2584 videos from white-listed channels, we found 136 videos (5.3%) (26 hospital, 80 organisation and 30 user videos) that most frequently appeared in the top-40 results, while 1654 videos (64%) only occasionally appeared among the top-500 results. These were classified as Group A and Group E videos, respectively. Table 2 shows the number of videos within each of the five rank-position groups.

Occasional change in ranking is presented in Table 3, showing the proportion of videos classified as Group A-E that occasionally (i.e. one or more days) appeared in a different rank-position interval. For example, 38.2% of Group A videos appeared occasionally in position 41-100, while 14% occasionally appeared in position 101-200. We found that only a small proportion of videos from Group C, D and E were occasionally given a high-ranked position within 1 - 40 (11.8%, 2.5% and 4.4% respectively).

These results indicate that most of the videos appearing in low rank positions, are stuck in low positions, and will consequently remain (almost) out of reach to the user.

Mean position and standard deviation (SD) were calculated for the 175 videos that were ranked within top-500 the whole test period. Table 4 shows mean position and standard deviation for videos, grouped according to their most frequent rank position (i.e. rank-position group).

Table 3: Proportion of Group A-E videos (classified according to their most frequent ranking position) that occasionally changed ranking position to a different position interval.

	Group A	Group B	Group C	Group D	Group E
Occ. pos 1-40	-	35.6%	11.8%	2.5%	4.4%
Occ. pos 41-100	38.2%	-	48.6%	11.3%	9.6%
Occ. pos 101-200	14.0%	54.4%	-	39.7%	19.6%
Occ. pos 201-500	6.6%	26.8%	59.0%	-	81.4%
Occ. pos > 500	27.9%	63.1%	76.9%	82.7%	-

Table 4: Mean position and standard deviation for videos in different rank-position groups. Including the 175 videos that were ranked within top-500 the whole test period.

	No. of videos	Mean position	Standard deviation
Group A	71	15.5	7.0
Group B	43	71.0	20.9
Group C	37	145.9	38.2
Group D	24	272.3	48.0

We found that the highest ranked videos (Group A) had the lowest standard deviation, i.e. 7.0. These videos seemed to be established in a top-ranked position, and had in general less variation in rank-position than videos from other groups. In fact, stability in rank position seemed to be the case for all groups, even though the standard deviation for Group B, C and D is higher.

SD-values indicate that changes in rank position in general do not make Group D videos more accessible to users, while Group C and B videos may occasionally be given a more accessible ranking. As an example, take Group B videos having a mean position of 71 and an SD value of 20.9. This means that most videos (about 68%, assuming a normal distribution) were ranked within position 50-92. Approximately 15% of the videos occasionally had a position within top-40, while approximately 15% were occasionally not included in the top-100. For Group C videos, less than 0.5% of the videos would occasionally have a position within the top-40, while approximately 10% would occasionally have a position within top-100.

The rank stability observed through these numbers, indicates that highly ranked videos remain available to users, while low ranked videos will almost always remain out of reach for the user.

3.3 Relevance of Videos

One could suspect that videos given a low-ranked position were not relevant to the query. To investigate this, we selected two queries ("diabetes hyperglycemia" and "diabetes retinopathy") and determined relevance of each tracked video by manually comparing keywords in the query to video title and description, and by watching the video to compare video content to query.

Over the test period, the system tracked 130 videos for the "diabetes hyperglycemia" query and 64 videos for the "diabetes retinopathy" query. Table 5 shows the number of videos that were i) relevant to the query, ii) relevant to diabetes in general and iii) not relevant to diabetes. For example, for Group E videos of the "diabetes hyperglycemia" query, we found that 50% were relevant to the query, an additional 47% were relevant to diabetes, while only 3% were not relevant to diabetes. For the "diabetes retinopathy" query, 55% of Group E videos were relevant to the query, an additional 18% were relevant to diabetes, while 27% were not relevant. For Group A and B videos (of both queries), every video was relevant to the query. In conclusion, we found that a large number of low-ranked videos were relevant to the query, implying that lack of relevance could not be the reason for their low ranking.

3.4 Video Properties

To detect possible correlations between video properties and ranking position, we compared video title and query terms, investigated social interaction by counting for each video the number of likes, dislikes, comments and views, and subsequently compared against the video's ranking position.

Having a *match between query terms and video title* is obviously an important criterion for considering the video relevant to the query. We found for Group A videos that 88% (120 of 136 videos) had a perfect match between video title and query (meaning that all terms in the query were found in the video title). The proportion of videos with a perfect query-title match was lower in the other groups, but there were still a large number of lower ranked videos that had a perfect query-title match. This shows that such a match is not sufficient for a high-ranked position.

Table 5: The number of videos relevant to i) the search query and ii) diabetes in general, for the two queries "diabetes hyperglycemia" and "diabetes retinopathy".

Rank position	relevant to	Diabetes hyperglicemia	Diabetes retinopathy
Group A videos	query	100%	100%
	diabetes	-	-
Group B videos	query	100%	100%
	diabetes	-	-
Group C videos	query	45%	100%
	diabetes	55%	-
Group D videos	query	41%	72%
	diabetes	55%	11%
	not relev	4%	17%
Group E videos	query	50%	55%
	diabetes	47%	18%
	not relev	3%	27%

The average number of *likes/dislikes and comments* for Group A-E videos are displayed in Figure 2. The general trend was that the highest ranked videos had the highest number of social interactions. This coincides well with previous studies, which found that very few videos get the users' attention. This can be explained through the Yule process (or rich-get-richer principle), as the videos that appear in the first page are more likely to be viewed and interacted (Chelaru et al., 2012; Cha et al., 2009).

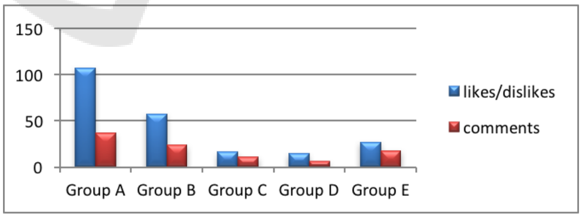

Figure 2: Average number of likes/dislikes and comments on videos.

However, when studying individual videos, we observed huge differences in the number of user interactions. We found for example that a number of videos without likes/dislikes and comments were highly ranked despite the lack of user activity. Table 6 shows the percentage of videos, within each rank-position group, that had zero likes, dislikes and comments. We see for example that 20% of Group A videos had no such social interaction.

Table 6: Number of videos without user interaction through likes/dislikes and comments.

	Videos tracked all period (175 videos)	All videos (2584 videos)
Group A	21%	20%
Group B	35%	29%
Group C	43%	33%
Group D	33%	39%
Group E	-	28%

When examining the number of *views* for individual videos, we found a close correlation between views and ranking (see Figure 3). This seems obvious since users can easily find and access highly ranked videos, which then get a higher number of views compared to low ranked videos. However, there were also a few exceptions. For instance, one Group A video had only 18 views, zero likes, dislikes and comments.

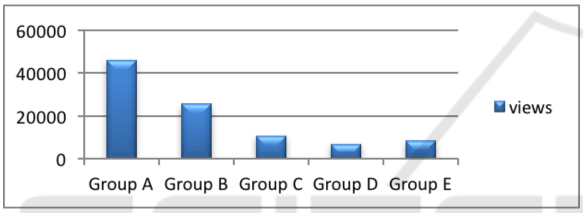

Figure 3: Average number of times a video has been viewed.

3.5 Videos From Non-white-listed Channels

To get an impression of the type of videos not tracked in our study, we also examined properties of videos published by non-white-listed channels. Because of the large number of videos, we restricted this investigation to the top-50 results of two queries: "diabetes type 1" and "diabetes injection". For all top-50 videos we manually examined relevance to query and channel type.

Relevance and channel type for the examined videos are shown in Table 7 and Table 8, respectively. Videos from both white-listed (WL) and non-white-listed channels are included in the tables. Among the non-white-listed videos, most came from private users (47 videos), while only 3 videos were from health organisations/centres. We further observed that white-listed channels had the highest proportion of relevant videos.

Table 7: Relevance of videos published by non-white-listed and white-listed channels.

Relevant to	Diabetes type 1		Diabetes injection	
	non-WL-channels 35 videos	WL-channels 13 videos	non-WL-channels 38 videos	WL-channels 11 videos
query	32 (91%)	13 (100%)	19 (50%)	10 (91%)
diabetes	3 (9%)	0	9 (24%)	1 (9%)
not relev.	0	0	10 (26%)	0

Table 8: Channels of videos published by non-white-listed and white-listed channels.

	Diabetes type 1		Diabetes injection	
	non-WL-channels 35 videos	WL-channels 13 videos	non-WL-channels 38 videos	WL-channels 11 videos
Hospital	0	3 (23%)	0	0
Health organization	1 (3%)	5 (38.5%)	1 (2%)	8 (73%)
Active users	0	5 (38.5%)	0	3 (27%)
Health center	1 (3%)	-	0	-
Company	8 (23%)	-	5 (13%)	-
Private users	24 (68%)	-	23 (61%)	-
Others	1 (3%)	-	9 (24%)	-

4 DISCUSSION

The goal of this study was to identify the number of videos from hospitals, health organisations and active user channels that are in practice available to users. On the positive side, the study shows that for each query, videos from white-listed channels are in fact available among the top-500 ranked videos. A problem, however, is that these videos represent a small proportion of the total number of retrieved videos, and that many of them are found in low ranked positions that make them in practice beyond reach for the user issuing the query. Thus, precision for videos from white-listed channels is not very good. Among the top-100 ranked videos for a diabetes related query, one can on average expect 15% to be from white-listed channels (2.8% from hospitals, 8.9% from health organisations and 3.3% from active users).

Of the 2584 tracked videos, 64% were most frequently ranked in a position lower than 500, only occasionally appearing within the top-500 results. This shows that many relevant videos from credible channels are in practice unreachable for users. Also, standard deviation values and observed ranking variations for individual videos show that the ranking of videos is fairly stable. This implies that only a small percentage of low ranked videos improved their ranking position sufficiently to be available to users and that users hardly obtain any new videos (from white-listed channels) by issuing a query multiple times. On the other hand, ranking stability also guarantees that top-ranked videos from white-listed channels are available to users over a period of time. This benefits new users that will have access to a few popular and potentially good quality health videos from credible channels.

One conclusion from our study is therefore that relevant diabetes-related health videos are available on YouTube, but too few are given a ranking that make them reachable for the user.

The YouTube ranking algorithm is based on video popularity among users. Previously the algorithm was based on view count of the video, while the current version (since 2012) is based on Watch Time, which is the amount of time on aggregate that viewers watch a video (Robertson, 2014). Even though Watch Time is a better measure of success for a video (since it rewards engaging videos that keep viewers watching), it still requires videos to be available to users in order to get sufficient attention and improve Watch Time. Also, there is no guarantee that an engaging, much watched video is trustworthy with respect to the health information it provides.

In our study, the investigation of correlation between ranking position and user attention in the form of social interactions, gave mixed results. There were on average a higher number of social interactions (i.e. likes/dislikes, comments and views) for the highest ranked videos, but we also saw many examples of videos that had a high-ranked position with no social interaction and very few views.

A critical factor in identifying relevant videos based on a textual query, is the accuracy of the metadata with respect to video content. When examining the correlation between video title and query terms, we found that a majority (88%) of the highest ranked videos (Group A videos), but also a large number of low ranked videos, had a perfect match between video title and query terms. However, by inspection, we also found many video descriptions that were very short and of such a general nature that they did not describe the video content. Video titles were also in many cases inaccurate with respect to video content.

An implication of these findings is that video publishers should make an effort in providing precise textual description of videos, where video title and description matches the video content as accurately as possible. This is a simple way of improving the likelihood for being selected as relevant and possibly ranked sufficiently high to be reachable. Allowing and even encouraging social interaction on videos may also help visibility of the video.

However, an accurate video title/description is only a step in the right direction for improving video rank position and precision. We believe there is a need for new video retrieval tools that not only focus on relevance and popularity as it is done today, but also retrieves health information based on requirements for credibility of the sources and trustworthiness of the videos. This provides topics for future research.

Some limitations to our work should be noted. Firstly, even though our white-lists of hospitals and health organisations include a large number of channels, they cannot include every hospital and health organisation available. The focus was not to track every relevant and trustworthy video in the result set from YouTube, but rather to track videos from specific channel types that are assumed to be of interest to health consumers. Also, it should be noted that the quality of each video was not assessed in this study. We base the study on the assumption that videos from hospitals, health organisations and also active users are of interest and therefore worthwhile investigating. We are fully aware that videos from other channels (not tracked in our study) may provide useful and trustworthy information. Furthermore, for each query we only examined the top-500 ranked videos from YouTube. When some queries return over 600.000 videos, this is a small number. However, we believe that a position over 500 is not significant in terms of availability to users.

5 CONCLUSION

To gain knowledge about how health videos are ranked on YouTube, we have tracked diabetes health videos on YouTube every day over a period of 1.5 month. We focused on videos published by credible channels, such as hospitals, health organisations and users actively publishing diabetes-related videos. Our findings show that most videos from these channels are given a persistent low ranking that makes them in practice unavailable to users. Additionally, since ranking position of videos is fairly stable, users receive the same videos over and over again if issuing a query multiple times. Thus, users may find it difficult to obtain new information from YouTube. A conclusion from this work is that research is needed to provide users with new tools that enable health video retrieval based on requirements concerning not only relevance and popularity, but also credibility of the sources and trustworthiness of the videos. Mechanisms for alternative ranking or less stable ranking could also be useful for making a larger number of relevant videos available to the user.

ACKNOWLEDGEMENTS

The authors appreciate support from UiT The Arctic University of Troms through funding from Troms Research Foundation, and the important contributions of the ITACA-TSB Universitat Politècnica de València group.

REFERENCES

AlGhamdi, K. M. and Moussa, N. A. (2012). Internet use by the public to search for health-related information. *International journal of medical informatics*, 81(6):363–373.

Azer, S. A., AlGrain, H. A., AlKhelaif, R. A., and AlEshaiwi, S. M. (2013). Evaluation of the educational value of youtube videos about physical examination of the cardiovascular and respiratory systems. *Journal of medical Internet research*, 15(11):e241.

Bennett, E. (2011). Social media and hospitals: From trendy to essential. *Futurescan 2001: Healthcare Trends and Implications 2011-2016*, pages 43–48.

Bermúdez-Tamayo, C., Alba-Ruiz, R., Jiménez-Pernett, J., García Gutiérrez, J.-F., Traver-Salcedo, V., and Yubraham-Sánchez, D. (2013). Use of social media by spanish hospitals: perceptions, difficulties, and success factors. *Telemedicine and e-Health*, 19(2):137–145.

Briones, R., Nan, X., Madden, K., and Waks, L. (2012). When vaccines go viral: an analysis of hpv vaccine coverage on youtube. *Health communication*, 27(5):478–485.

Butler, D. P., Perry, F., Shah, Z., and Leon-Villapalos, J. (2013). The quality of video information on burn first aid available on youtube. *Burns*, 39(5):856–859.

Cha, M., Kwak, H., Rodriguez, P., Ahn, Y.-Y., and Moon, S. (2009). Analyzing the video popularity characteristics of large-scale user generated content systems. *IEEE/ACM Transactions on Networking (TON)*, 17(5):1357–1370.

Chelaru, S. V., Orellana-Rodriguez, C., and Altingovde, I. S. (2012). Can social features help learning to rank youtube videos? In *International Conference on Web Information Systems Engineering*, pages 552–566. Springer.

Cheng, X., Dale, C., and Liu, J. (2008). Statistics and social network of youtube videos. In *Quality of Service, 2008. IWQoS 2008. 16th International Workshop on*, pages 229–238. IEEE.

de Boer, M. J., Versteegen, G. J., and van Wijhe, M. (2007). Patients use of the internet for pain-related medical information. *Patient education and counseling*, 68(1):86–97.

Diaz, J. A., Griffith, R. A., Ng, J. J., Reinert, S. E., Friedmann, P. D., and Moulton, A. W. (2002). Patients' use of the internet for medical information. *Journal of general internal medicine*, 17(3):180–185.

Dutta-Bergman, M. (2003). Trusted online sources of health information: differences in demographics, health beliefs, and health-information orientation. *Journal of medical Internet research*, 5(3):e21.

Fat, M. J. L., Doja, A., Barrowman, N., and Sell, E. (2011). Youtube videos as a teaching tool and patient resource for infantile spasms. *Journal of Child Neurology*, 26(7):804–809.

Fox, S. (2011a). *Peer-to-peer healthcare*. Pew Internet & American Life Project.

Fox, S. (2011b). *The social life of health information, 2011*. Pew Internet & American Life Project Washington, DC.

Freeman, K. S. and Spyridakis, J. H. (2009). Effect of contact information on the credibility of online health information. *IEEE Transactions on Professional Communication*, 52(2):152–166.

Gabarron, E., Fernandez-Luque, L., Armayones, M., and Lau, A. Y. (2013). Identifying measures used for assessing quality of youtube videos with patient health information: a review of current literature. *Interactive Journal of Medical Research*, 2(1):e6.

Griffiths, F., Cave, J., Boardman, F., Ren, J., Pawlikowska, T., Ball, R., Clarke, A., and Cohen, A. (2012). Social networks–the future for health care delivery. *Social science & medicine*, 75(12):2233–2241.

Hughes, B., Joshi, I., Lemonde, H., and Wareham, J. (2009). Junior physician's use of web 2.0 for information seeking and medical education: a qualitative study. *International journal of medical informatics*, 78(10):645–655.

Karlsen, R., Morell, J. E. B., Fernández-Luque, L., and Salcedo, V. T. (2013). A domain-based approach for retrieving trustworthy health videos from youtube. In *MedInfo*, page 1008.

Madathil, K. C., Rivera-Rodriguez, A. J., Greenstein, J. S., and Gramopadhye, A. K. (2015). Healthcare information on youtube: a systematic review. *Health informatics journal*, 21(3):173–194.

Morell, J. E. B., Karlsen, R., Salcedo, V. T., and Luque, L. F. (2012). Analysis of hospital youtube videos as health reliable sources. *IVWorkshop on on Technology for Healthcare and Healthy Lifestyle*.

Moturu, S. T., Liu, H., and Johnson, W. G. (2008). Trust evaluation in health information on the world wide web. In *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1525–1528. IEEE.

Murugiah, K., Vallakati, A., Rajput, K., Sood, A., and Challa, N. R. (2011). Youtube as a source of information on cardiopulmonary resuscitation. *Resuscitation*, 82(3):332–334.

Pant, S., Deshmukh, A., Murugiah, K., Kumar, G., Sachdeva, R., and Mehta, J. L. (2012). Assessing the credibility of the youtube approach to health information on acute myocardial infarction. *Clinical cardiology*, 35(5):281–285.

Powell, J., Inglis, N., Ronnie, J., and Large, S. (2011). The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. *Journal of Medical Internet Research*, 13(1):e20.

Robertson, M. R. (2014). Watch time: A guide to youtubes top search ranking factor. *Internet: http://www. reelseo. com/youtube-watch-time/, Einsicht*, 8:2015.

Schreiber, J. J., Warren, R. F., Hotchkiss, R. N., and Daluiski, A. (2013). An online video investigation into the mechanism of elbow dislocation. *The Journal of hand surgery*, 38(3):488–494.

Shabbir, S. A., Fernandez-Luque, L., Jian, W.-S., Li, Y.-C., Crain, S., Hsu, M. H., Wang, Y. C., Khandregzen, D., Chuluunbaatar, E., Nguyen, P. A., et al. (2013). Misleading health-related information promoted through video-based social media. *Journal of Medical Internet Research*, 15(2).

Singh, A. G., Singh, S., and Singh, P. P. (2012). Youtube for information on rheumatoid arthritisa wakeup call? *The Journal of rheumatology*, 39(5):899–903.

Steinberg, P. L., Wason, S., Stern, J. M., Deters, L., Kowal, B., and Seigne, J. (2010). Youtube as source of prostate cancer information. *Urology*, 75(3):619–622.

Topps, D., Helmer, J., and Ellaway, R. (2013). Youtube as a platform for publishing clinical skills training videos. *Academic Medicine*, 88(2):192–197.

Yee, W. G., Yates, A., Liu, S., and Frieder, O. (2009). Are web user comments useful for search. *Proc. LSDS-IR*, pages 63–70.

YouTube (2016). *Statistics*. https://www.youtube.com/yt/press/statistics.html, Accessed: 2016-09-26.

Ziebland, S. and Herxheimer, A. (2008). How patients experiences contribute to decision making: illustrations from dipex (personal experiences of health and illness). *Journal of nursing management*, 16(4):433–439.