Introducing Credibility Signals and Citations to Video-Sharing Platforms

Emelia Hughes, Renee Wang, Prerna Juneja, Tanushree Mitra, Amy X. Zhang

University of Washington

Seattle, Washington 98105

{*emhughes, prerna79, tmitra*}*@uw.edu*, {*renee17, axz*}*@cs.uw.edu*

Abstract

Misinformation is a growing problem in today's online information ecosystem. In particular, on major video-sharing platforms, there are few signals for users to assess the credibility of the videos they encounter, even as more people turn towards them for getting information. In this work, we propose a series of designs to indicate credibility on video-sharing platforms, including via signals on video search result pages, as well as citations on video pages, using the major platform YouTube. From an initial interview study on how users perceive YouTube credibility indicators, we found that video intent and context heavily shape how a user would apply credibility signals or citations. Along with an additional planned survey, these results inform design decisions for a browser extension we will build and deploy to test our designs.

Introduction

A common method for combating misinformation on social platforms involves employing content moderators and fact checkers to spot and remove misinformation for users (Potthast et al. 2016). More recently, researchers and platforms have examined additional methods that give users greater context about the content they see, so that they can spot instances of misinformation for themselves (Zhang et al. 2018; Morris et al. 2012; Smith 2018; Coleman 2021). These approaches are complementary to take-down forms of moderation because they give users the flexibility to apply their own credibility standards and trust measures to evaluate content that does not get caught by platform moderators.

Prior research has highlighted a wide range of possible indicators for evaluating content credibility such as visual appearance, tone, and representative citations (Zhang et al. 2018; Morris et al. 2012). In particular, Wikipedia and Wikidata have emerged as key players in the space of credibility signals, as many platforms have turned towards displaying links to or signals from Wikipedia. These indicators help users gain greater context about a topic by engaging in lateral reading practices, where readers cross-reference external materials while reading the original source. Lateral readers have been shown to gain a better sense as to whether
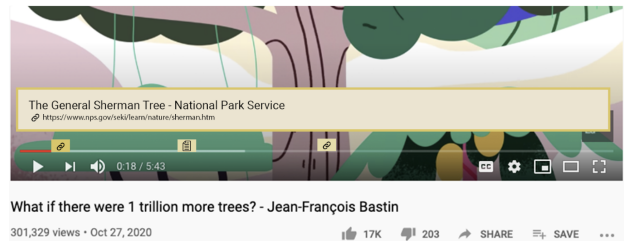
Figure 1: A design mock-up of how citations could look within a video playback page on YouTube. The citation bar pops up at the particular timestamp that references the information from the cited source.

to trust the facts and analysis presented to them (Caulfield 2017; Wineburg and McGrew 2017).

However, it is still unclear which types of credibility indicators are most useful to online information consumers and how they should be presented. In this work, we explore how credibility is conveyed on video-sharing platforms, due to the lack of standard credibility factors or citation displays in such spaces. For example, a search query made through a web search engine will result in common Knowledge Panels and informative Wikipedia pages that are missing when the same query is made on a video-sharing platform like YouTube. This lack of general credibility signals is concerning, considering that YouTube has been shown to be used as an informational source for many users (Smith, Toor, and Kessel 2018).

In addition, the wide base of content creators on common video-sharing platforms such as YouTube complicates credibility because it can magnify videos made by non-expert individuals over ones made by mainstream institutions. For instance, YouTube's streamlined video uploading process makes it possible for anyone to upload a video. Traditional search engine results lean towards mainstream sources because these institutions have the resources to make websites and other content favored by search engine algorithms. On YouTube, search results are affected by user watch history (Hussein, Juneja, and Mitra 2020), which can create a space for non-institutional content creators. This difference makes YouTube a compelling site for a study on credibility factors. For our work, we focus on two particular areas on the

YouTube Platform: 1) the YouTube search results page, and 2) the YouTube video playback page.

Citations are one method we explore as a way to increase context of information on YouTube. Some of the most recent work conducted by the Wikimedia Foundation has examined the impact of citations on trust (Morgan et al. 2019), though this focused on citations within Wikipedia as opposed to citations elsewhere on the web that link to Wikipedia. Through initial interviews, we have found that the intent and context of a YouTube video heavily shape how a user would apply credibility signals or citations. To gain a greater understanding of how the context of a video can change the effectiveness of credibility signals, we will deploy a survey expanding upon our interviews. These results will inform design decisions for a browser extension we will build and deploy to test our designs.

Study Overview

1. Inserting Credibility Signals into YouTube Search Results

The YouTube search results page is formatted similarly to that of web search engines, suggesting that similar credibility signals could exist on the YouTube platform. Besides implementing signals on the search result page like a Google Knowledge Panel, we would also like to design individual indicators associated with each search result. Some of these individualized indicators could later also transfer to the video playback page.

This part of the study looks at credibility signals for both the video source channel as well as the video itself. For the channel specifically, we will study the displayed information that could indicate a channel's overall credibility. For example, YouTube currently displays a link to a Wikipedia article if a channel is owned by a government-funded news source. We will look into how displaying Wikipedia links and other third-party information could be extended to other channel categories.

Some of the questions we explore with this part of the study are:

- How does channel verification contribute to credibility?

- What are the best information points to display about a channel for a user to determine its credibility?

- What is the best way to integrate Wikipedia-based information into the video experience?

2. Inserting Citations into YouTube Videos

We also explore how credibility can be displayed on the video playback page, particularly looking into what a standardized citation within a video should look like, and what types of videos could benefit from citations. As a starting place for this part of the study, we take inspiration from citation norms on Wikipedia. Looking at the format of Wikipedia also elicits the question of whether citations could be added by the viewers and YouTube community. The possibility of a crowd-sourced approach to fact-checking videos could be another exciting mode of interaction for a community that already collaborates on translated captioning and
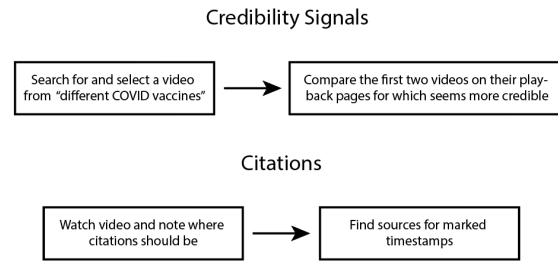

Figure 2: Interviewee tasks during the user study. In the first part, participants react to credibility signals, first from the search-results page by choosing the video they would watch from a search query, and then from the video playback page by comparing two video-playback pages for which they believe to be more credible. In the second part, interviewees watch a provided video and mark timestamps where they believe there should be a citation. The final task is to find appropriate sources for their marked spots.

interact in comment sections. Additionally, citations could also help creators make more credible informational videos, similar to how citations give Wikipedia articles credibility.

Some of the questions we explore with this part of the study are:

- How would a standardized citation within a video look?

- How can citations and other signals associated with a video be presented in places where that video is embedded, such as in social media or articles?

- Who should be adding citations to a video—could they be generated automatically, inserted by the creator, or added by third party?

Methods

We are currently going through an iterative design and study process to find effective UX choices that communicate credibility clearly. First, we collected existing user opinions and practices of YouTube credibility indicators. Using this information, we are developing a Chrome browser extension that allows users to introduce citations and other credibility signals onto both the video-playback and search results pages. We will conclude this project by studying the efficacy of the design through a final user study, and consider future steps we could take with that information.

We begin this process with user research involving interviews and contextual inquiry in order to gain an understanding of how users currently address credibility on YouTube. This study includes methods to understand both how the study can capitalize on current YouTube credibility signals and how users may possibly interact and create citations for videos (Figure 1). For instance, one exercise the interviewee performs is to locate timestamps in an anti-vaccination video where they believe a citation is needed. The interviewee is then asked to find an appropriate source for their citation.

How often do you "fact check" the information you see in a Youtube video (outside of Youtube)?

Very Frequently

Frequently

Occasionally

Rarely

Very Rarely

Never

What do you think could help in your determination of whether a video has true information?

Figure 3: An example of questions included in our survey-based study. This survey will help us to gain a more generalized understanding of user credibility practices both on YouTube and elsewhere on the web.

While this interview process is still in progress, some interesting results have already emerged. A consistent pattern among interviewees so far is that participants all noted that the context of the video plays a significant role in deciding how they would apply citations. For example, one participant noted that they were less likely to think about creating citations for entertainment videos or videos giving personal advice, saying:

> "*It only really matters when they're trying to present something that is informational.*"

Another interviewee commented that they were more likely to use citations as an extension to information presented in news videos, rather than as a fact-checking tool for ambiguously credible content.

Emerging from these interviews is the question of how users want to be involved in citation creation. Part of the user study asked the user to create citations, and while many participants identified credible sources for their citations, they also reported that the labor of adding citations themselves would be more effort than they are willing to contribute on a normal basis. One participant noted that:

> "*If I would feel comfortable contributing to closed captions or translations, I would also be comfortable adding citations.*"

However, the ability to flag information that needed citations has been appealing to the majority of participants.

Alongside this interview process, we are planning a survey-based study (Figure 3) to gain a more generalized understanding of user credibility practices. From there, we will be finalizing the browser extension. In Figure 1, we present a mockup example of how a citation tool could be presented on a YouTube video. Once finalized, we will have an additional user study to evaluate the effectiveness of our Chrome extension for both credibility signals on the search-result page and citations on the video playback page.

Conclusion

Our work aims to strengthen credibility on YouTube by introducing a standardized citation process along with new credibility signals. While citations are standard elsewhere on the web, there is no current standard for video-sharing platforms like YouTube. We hope to design a Chrome extension that allows users to generate citations, based on our studies of how users interact with credibility signals on YouTube result and playback pages. By offering such a tool, we provide YouTube users another way to get more context about the information they see, and ultimately, be better able to identify misinformation for themselves.

Acknowledgements

References

Caulfield, M. 2017. *Web-Literacy for Student Fact-Checkers*. Pressbooks.

Coleman, K. 2021. Introducing birdwatch, a community-based approach to misinformation. Twitter. https://blog.twitter.com/en_us/topics/product/2021/introducing-birdwatch-a-community-based-approach-to-misinformation.html.

Hussein, E.; Juneja, P.; and Mitra, T. 2020. Measuring misinformation in video search platforms: An audit study on youtube. volume 4. New York, NY, USA: Association for Computing Machinery.

Morgan, J.; Forte, A.; Mimouni, H. E.; and Johnson, I. 2019. The role of citations in how readers evaluate wikipedia articles. Wikimedia.

Morris, M. R.; Counts, S.; Roseway, A.; Hoff, A.; and Schwarz, J. 2012. Tweeting is believing? understanding microblog credibility perceptions. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, CSCW '12, 441–450. New York, NY, USA: Association for Computing Machinery.

Potthast, M.; Köpsel, S.; Stein, B.; and Hagen, M. 2016. Clickbait detection. In *European Conference on Information Retrieval*, 810–817. Springer.

Smith, A.; Toor, S.; and Kessel, P. V. 2018. Many turn to youtube for children's content, news, how-to lessons. Pew Research Center.

Smith, J. 2018. Designing new ways to give context to news stories. Facebook. https://medium.com/facebook-design/designing-new-ways-to-give-context-to-news-stories-f6c13604f450.

Wineburg, S., and McGrew, S. 2017. Lateral reading: Reading less and learning more when evaluating digital information. Stanford history education group working paper.

Zhang, A.; Ranganathan, A.; Metz, S. E.; Appling, S.; Sehat, C. M.; Gilmore, N.; Adams, N. B.; Vincent, E.; Lee, J. .; Robbins, M.; Bice, E.; Hawke, S.; Karger, D.; and Mina,

A. X. 2018. A structured response to misinformation: Defining and annotating credibility indicators in news articles. International World Wide Web Conference (WebConf) Companion.