\chapter{Introduction}

\label{chap01}

本章では，本研究の背景，目的，貢献及び構成について述べる．

This chapter describes the background, objectives, contributions, and structure of this study.

Introduction

In the modern era, to cope with the increasing complexity and scale of scientific research problems, multiple researchers are now working together, and sharing research problems and results has become essential. For this purpose, various means of scholarly communication have been proposed and utilized. For example, academics established academic journals and conferences as standard scholarly communication methods in the past few decades. The number of media and events for such communication has grown to a vast number and scale. In this climate, academic research has been shared and disseminated mainly within the academic community.

Recently, however, there has been much talk about rethinking the value of scientific research. In the past, the research value was assessed by discovering and creating unknown knowledge by the researcher and the utilization of the research results by the direct stakeholders to create secondary value. The value of research given in this way reflects only the perspective of a narrow community of stakeholders, which may diverge from the socially accepted notion of research value. Such a phenomenon can interfere with empathy between the academic community and society and induce obstacles to scientific research's smooth planning and progress. In this reason, many have advocated the need to expand the scope of research evaluators and incorporate the perspectives of various members of society into the evaluation process. In response to this challenge, people are now focusing on various communication activities that reflect the social interest in research and the channels through which they appear. For example, in the mass media, scientific research has been communicated to the general public through newspaper articles and television broadcasts of news and documentaries. On the other hand, recently, with the activation of social networking services, stakeholders in research and non-specialists in research have been referring to research on various new platforms such as blogs, Facebook, and Twitter, accelerating the dissemination and discussion of scientific research in society\cite{57}.

There have been various attempts to estimate the effects of such broad science communication surrounding research. As academic publishing and communication moved online, Usage metrics based on downloads and views, and Webometrics based on web links were proposed. Since 2010, the usefulness of Altmetrics, a new metric based on communication activities and interactions on various social media, has been proposed \cite{3}. The word "Altmetrics" is a combination of the words "alternative" and "metrics", although, in general, the issues and approaches addressed by Altmetrics are different from those of bibliometrics, and are not alternative concepts\cite{2}. Altmetrics are often considered more transparent than bibliometrics and webometrics. They provide a quicker and more real-time view of impact, are accessible to all, cover non-academic audiences, and can encompass a wider variety of research outputs and sources\cite{4}. One of the widely accepted Altmetrics, Altmetric Attention Score (AAS), provided by altmetric.com, offers a broad measure of an article's value by comprehensively considering the number of times an academic article is mentioned in various popular communication sources including social media. The AAS provides a broad measure of the value of an article.

On the other hand, as the Web and the media develop, various communication formats have emerged in academia. Also, in the early 2000s, electronic scholarly communication was proactively used in informal communication, cooperation, publication and dissemination of results, and relationship building among academics\cite{6}. The Web provides an opportunity to distribute and share a wide range of atypical content in science. The use of non-standard electronic media such as diagrams, presentations, and videos is increasing in research and education. Among them, communication using online academic videos, a relatively new method, has been recognized as useful because of its ease of dynamic expression compared to other media. For example, online video could be useful for academic purpose in the arts and humanities, such as dance and film, where human motion is essential\cite{6}. Video recordings of elaborate scientific research demonstrations, scientific documentaries, and lectures may be more effective than text for explaining the scientific experience. For example, online videos can effectively communicate scientific methodologies, protocols, research results, and market educational and volunteer activities\cite{7}. In response to reports of the effectiveness of such online video-based scientific communication, the use of online video is expanding in academia as well. For example, Nature and Cell, the well-known journals, proactively encourage accepted papers to submit videos that introduce the research topic and results, which are published in a dedicated section for online videos on the journal's website. The Journal of Visualized Experiments, a journal for videos of experiments mainly in life sciences, has pioneered and established a new journal format, the online video journal, while some academic conferences, mainly in the media field, hold presentations focusing on videos.

Furthermore, online academic videos are not limited to the academic community but are also being actively published on popular platforms, such as YouTube, founded in 2005 and the second most visited website in the world after Google\cite{8}. It is the second most visited website in the world after Google. While people mainly use YouTube for entertainment purposes such as music and comedy, some academics have been distributing their academic work online since the foundation of YouTube. For example, videos of renowned scientist Stephen Hawking's lectures on the universe (youtube.com/watch?v=xjBIsp8mS-c) and a physics lecture recorded at MIT (youtube.com/watch?v=sJG-rXBbmCc), both released online almost as soon as YouTube was founded, have been viewed about 6.5 million times and 5 million times respectively. For another example, Science, a well-known academic journal, has been running a YouTube channel since 2008 and has released videos introducing some of the research published in the journal. Such efforts enable the journal to produce and distribute videos to explain and socially disseminate complex research to those involved.

With these reports on the use and effectiveness of online academic videos, it is believed that video production and publication are becoming a worthy investment of research resources for academics. For videos that mention the URL of Science journals on YouTube, Fig\ref{fig1-1} shows the estimated number of new videos that mention Science journal articles and other new videos. The number of videos mentioning articles is the number of results obtained by searching YouTube queries in the literature browsing domain (science.sciencemag.org/content), and the number of other videos is the balance between the number of results searched in the main page domain (science.sciencemag.org) and the number of videos mentioning the article. The figure implies that the numbers of videos mentioning articles and other videos related to Science on YouTube have been increasing trendily since the foundation of YouTube until the most recent year, 2019.

\begin{figure}[tbp].

\begin{center}

\includegraphics[width=10cm]{pics/fig1-1.jpg}

\end{center}

\caption{Estimated number of new YouTube videos mentioning Science Magazine}

\label{fig1-1}

\end{figure}

However, there is a significant lack of research on the impact of online academic videos on the value of research, which should be discussed in the process of investing in video production. Suppose it is acknowledged that online academic videos have a significant effect on research value, i.e., citation count and the level of public attention. In that case, it could provide a rationale for investing in video production and publication of research resources. On the other hand, if it could be confirmed that there are differences in the impact of academic video depending on the characteristics of the research, such as the field of research, researchers, and publication sources, this can provide support for understanding what kind of research to which video can effectively contribute. Thus, understanding the interaction between research and online academic videos is crucial in designing effective communication through academic videos. Furthermore, if it is possible to classify online academic videos, which are an atypical communication method, by extracting their characteristics, making proposals on the content of videos for effective communication will be possible.

In this study, we propose a method to assess academic videos' impact on citation counts and Altmetrics of academic papers. Specifically, we prepare a large dataset of academic articles in specific research fields and collect YouTube videos that mention the articles in the title or description as article-mentioning-videos. As a control group for the articles with mentioning-videos, we sample the articles without mentioning-videos with the same quality as the articles with mentioning-videos using our original homogenization method. We then test for the difference in the populations of origin for the citation count and AAS distributions of the two article groups to verify article-mentioning-videos' effectiveness. Next, based on the qualitative analysis of the large number of academic videos collected, we define the taxonomy of "the purpose of article mention" for the video method and identify the effective video method by comparing the distribution of citation counts and AAS among the articles with mentioning-videos divided along with the label. Finally, we propose the "YouTube score," an article index representing the article's popularity on YouTube and verify that the YouTube score is a leading indicator that saturates earlier than citation count. We then perform a regression analysis between the YouTube scores and citation counts for article groups mentioned by each video group divided by the video method, followed by identifying the video group that shows a significant correlation, estimating its label as an effective video method for predicting the future citation counts using the YouTube scores. This method allows researchers to predict future citation count from the YouTube scores of articles and evaluate the academic impact of articles in the early stage of publication.

To evaluate the proposed method, we conducted experiments using academic paper datasets in two broad fields: mathematical and computer sciences, and life and earth sciences. We sampled the papers without videos against the papers with videos and tested that the citation counts and AAS distributions between the two groups originated from different populations, finding a significant effect of article-mentioning-video. Also, categorizing videos by labelling the purpose of article mention, we confirmed the differences in citation counts and AAS distribution by the video method and identified the effective video methods for each indicator. Furthermore, we verified that the YouTube score saturates earlier than citation count, and extracted some video methods with a significant correlation between the YouTube score and the citation count, suggesting the possibility of predicting future citation counts based on the YouTube score of early-stage papers.

Using this method, researchers can allocate limited research resources to video production and video design based on quantitative evidence. Furthermore, it is possible to predict citation counts in the future based on the YouTube score by releasing YouTube videos in the methods with significant prediction accuracy for papers in the early stages of publication, which will accelerate the evaluation of research stakeholders and help build a research portfolio.

\section{Contributions}

本研究の成果から以下のことが可能となる．

まず，本研究では学術文献を言及するユーチューブ動画が，文献の被引用数及びオルトメトリクスに有意な影響を与えることを検証し，計量的に分析することができた．これによって，ユーチューブにおける科学コミュニーケーションは研究の学術的・社会的注目度に寄与することが検証され，その効果を定量的に評価することが可能になった．

次に，ユーチューブ上の論文言及動画に対して論文言及目的に関する分類法を提案し，動画の分類ラベルによって分割された各論文群の被引用数及びオルトメトリクス分布を定量的に分析及び比較することができた．これにより，論文言及動画の被引用数及びオルトメトリクスへの寄与に関して，より効果的な動画方式の選択に関する定量的な根拠を提供することが可能になった．

最後に，ユーチューブ上での論文の人気度を表す指標として提案したユーチューブスコアは，論文の被引用数より早期に飽和する先行指標であることが検証された上に，当該指標と被引用数の回帰分析の結果，有意な相関を示す論文言及動画のラベルを部分的に抽出することができた．これにより，出版初期の論文に対し，ユーチューブスコアを用いた将来の被引用数の予測が有効と考えられる動画方式の推定が可能になった．

以上のことから，本研究の貢献は以下のように集約できる．

\begin{itemize}

\item 学術論文を言及するユーチューブ動画が，論文の被引用数及びオルトメトリクスに有意な影響を与えることを検証し，定量的に分析することができた．

\item 動画の論文言及目的による被引用数及びオルトメトリクスへの影響の相違を確認し，効果的な動画の論文言及目的を特定することができた．

\item ユーチューブ上での論文の人気度を表す指標を提案し，出版初期の論文の当該指標を用いた将来の被引用数の予測が有効な動画方式を部分的に推定することができた．

\end{itemize}

The results of this study allow followings.

First, this study was able to verify and quantitatively analyze that YouTube videos mentioning academic literature have a significant impact on citation counts and Altmetrics of the literature. The results verified that science communication on YouTube contributes to the research's academic and social attention and made it possible to evaluate the effect of such communication quantitatively.

Next, we proposed a taxonomy of the purpose of mentioning on YouTube and quantitatively analyzed and compared distributions of citation counts and Altmetrics for each article group divided by the videos' labels. The results allow us to provide a quantitative ground for selecting video methods that could more effectively contribute to the article's citation count and Altmetrics.

Finally, we verified that the YouTube score, a proposed measure of the popularity of an article on YouTube, is a leading indicator that saturates earlier than citation count. As a result of the regression analysis between two indicators, we identified the labels that showed a significant correlation between them. The results enable us to estimate the video method that effectively predicts future citation counts using the YouTube score for articles in the early stages of publication.

Based on the above, the contribution of this study can be summarized as follows.

\begin{itemize}

\item We verified and quantitatively analyzed that YouTube videos mentioning academic papers have a significant impact on citation count and Altmetrics of the papers.

\item We confirmed the differences in the video's impact on citation count and Altmetrics by the purpose of mentioning and identified effective purposes of mentioning.

\item We proposed an indicator for the popularity of papers on YouTube and partially estimated the video method that effectively predicts the future citation counts using the said indicator for papers in the early stage of publication.

\end{itemize}

\section{Structure}

Chapter 2 provides the related works and significance of this study. Chapter 3 presents our method for verifying and measuring YouTube videos' impact on citation counts and Altmetrics of papers and predicting citation counts using the videos. Chapter 4 describes the dataset used in the experiment and show the experimental results of the method proposed in the previous chapter. In Chapter 5, based on the results, we discuss the trends of article-mentioning videos on YouTube and some points that should be considered when using our proposed method. Chapter 6 presents the conclusions and future works of this study.