Towards advanced social media metrics: understanding the diversity and characteristics of Twitter interactions around science

Fang, Z.

Towards Advanced Social Media Metrics: Understanding the Diversity and Characteristics of Twitter Interactions around Science

Zhichao Fang

Towards Advanced Social Media Metrics: Understanding the Diversity and Characteristics of Twitter Interactions around Science

Cover design and layout: Zhichao Fang

Print: ProefschriftMaken

ISBN: 978-94-6423-594-4

The research described in this dissertation was performed at the Centre for Science and Technology Studies (CWTS), Faculty of Social and Behavioural Sciences at Leiden University, in the Netherlands. The research was supported by the China Scholarship Council (Grant No. 201706060201).

Towards Advanced Social Media Metrics: Understanding the Diversity and Characteristics of Twitter Interactions around Science

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op dinsdag 21 december 2021
klokke 15:00 uur

door

Zhichao Fang

geboren te Yueyang, China
in 1992

Promotores en promotiecommissie

Promotor:

Prof.dr. P.F. Wouters

Co-promotor:

Dr. R. Costas Comesana

Leden Promotiecommissie:

Prof.dr. S. de Rijcke

Prof.dr. L.R. Waltman

Dr. E.C.M. Noijons

Prof.dr. M. Thelwall (University of Wolverhampton, UK)

Prof.dr. I. Peters (Kiel University, Germany)

Table of Contents

CHAPTER 1

General introduction

1.1 Introduction

Since the late 1990s, the popularity of social media has increasingly enriched and diversified the means of scholarly communication (Priem & Hemminger, 2010; Sugimoto, Work, et al., 2017) as well as public engagement with science (Howell et al., 2019; Kouper, 2010; Regenberg, 2010). Along with the growing enthusiasm of scientists (Collins et al., 2016; Van Eperen & Marincola, 2011), graduate students (Howell et al., 2019), and the general public (Brossard, 2013; Huber et al., 2019) for leveraging social media tools to communicate and engage with scientific developments, numerous digital traces of interactions around science have been emerging in the social media environment. These traces, such as blog citations, Twitter mentions, and Facebook mentions to scientific papers, alongside a family of indicators drawn upon them have been collectively incorporated under the umbrella term "altmetrics" (Priem et al., 2010) which was first coined by Jason Priem in a tweet posted in 2010.[1] With a more specific focus on the events around scholarly objects captured in online social media platforms, "social media metrics" has usually been preferred to underline the social media component of these new measures (Costas, 2017; Haustein, Bowman, & Costas, 2016; Wouters et al., 2019).

Amongst the various data sources forming the basis of social media metrics, Twitter has arguably been the most popular one (Haustein, 2019). Across different types of social media metric data of scientific papers, in general, scholarly tweets (i.e., tweets including URLs referring to scholarly outputs) exhibit considerable data volume and rather substantial data coverage (Meschede & Siebenlist, 2018; Robinson-Garcia et al., 2014). Furthermore, in a survey conducted by *Nature* about the uptake of social media by scientists (Van Noorden, 2014), Twitter was reported as the most "interactive" site, with more respondents stating that they use Twitter to follow scientific discussions and comment on relevant research in their fields. The interactive nature of Twitter is particularly embodied in a broad array of engagement functionalities available in its information ecosystem, for example, users can disseminate scientific information by retweeting, participate in scientific conversations by replying, and access other scholarly sources by clicking on the tweeted URLs (Didegah et al., 2018; Kahle et al., 2016). To sum up, enabled by the big data available and its interactive nature, Twitter opens the possibility to characterize, on a large scale, how scientific information is interacted with by broader audiences from both academic and non-academic environments.

The overarching aim of this PhD dissertation is to characterize diverse forms of Twitter interactions around science to approach more advanced Twitter-based metrics. This chapter

[1] Jason Priem tweeted that "I like the term #articlelevelmetrics, but it fails to imply *diversity* of measures. Lately, I'm liking #altmetrics." (https://twitter.com/jasonpriem/status/25844968813) (Accessed May 25, 2021).

presents a general introduction to conceptualize interactions between science and social media and review relevant literature. Specifically, section 1.2 proposes a conceptual framework of interactions within and between science and social media. Section 1.3 applies this conceptual framework to review literature related to interactions between science and Twitter in a systematic manner. Section 1.4 sets out the motivations and objectives of this PhD dissertation. Finally, section 1.5 outlines the structure and details the research questions to be addressed in the dissertation.

1.2 Conceptualization of interactions between science and social media

This section starts with an introduction to social media and its scholarly use. Then, the section conceptualizes and illustrates interactions within and between science and social media. Finally, based on the proposed conceptual framework, this section interprets the scope of social media metrics of science, and distinguishes between *primary social media metrics* and *secondary social media metrics*.

1.2.1 Social media and its scholarly use

The rapid development of social media has unarguably changed the ways in which people are connected and information is shared (Aral & Walker, 2012; Collins et al., 2016). As surveyed by the Pew Research Center (2021), in 2021 roughly 72% of American adults use at least one social media site, in contrast to 5% in 2005, indicating the fast-growing uptake of social media in the last two decades. Due to the multitude of stand-alone and built-in social media services and the ever-evolving landscape of online tools (Obar & Wildman, 2015), social media is a broadly used term without a mutually agreed-upon definition. There has been a variety of definitions of social media proffered in previous literature, for instance, social media has been referred to as:

- "User-generated content utilizing Internet-based publishing technologies, distinct from traditional print and broadcast media." (Terry, 2009)

- "A group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of User Generated Content." (Kaplan & Haenlein, 2010)

- "Social media employ mobile and web-based technologies to create highly interactive platforms via which individuals and communities share, co-create, discuss, and modify user-generated content." (Kietzmann et al., 2011)

- "Internet-based channels that allow users to opportunistically interact and selectively self-present, either in real-time or asynchronously, with both broad and narrow audiences who derive value from user-generated content and the perception of interaction with others." (Carr & Hayes, 2015)

Despite the differences, the aforementioned definitions of social media share three fundamental elements, including *Internet-based* (i.e., social media are online tools operating via the Internet), *user-generated content* (i.e., content created by end-users that is available to the public or a selected group of people), and *interactive features* (i.e., users are empowered to interact with other users or user-generated content). Within these general definitions, there are various social media tools identified and distinguished by researchers based on nuanced categorizations (Aichner & Jacob, 2015; Kaplan & Haenlein, 2010; Rowlands et al., 2011; Sugimoto, Work, et al., 2017; Tenopir et al., 2013). Overall, the landscape of social media comprises:

- Blogs (e.g., the Scholarly Kitchen, the LSE blogs, or the Leiden Madtrics to name a few)
- Social networking sites (e.g., Facebook, LinkedIn, or ResearchGate)
- Microblogs (e.g., Twitter, Sina Weibo, or Tumblr)
- Social bookmarking and online reference managers (e.g., Mendeley or Zotero)
- Wikis (e.g., Wikipedia or Baidu Baike)
- Media sharing services (e.g., YouTube, TikTok, or Instagram)
- Data and code sharing services (e.g., Figshare, Slideshare, or GitHub)
- Discussion forums (e.g., Reddit or Baidu Tieba)
- Question and answer sites (e.g., Quaro, Stack Exchange, or Zhihu)
- Social recommending, rating, and reviewing platforms (e.g., Faculty Opinions, PubPeer, or Publons)
- Virtual social games and social worlds (e.g., World of Warcraft or Second Life)

Over the last two decades, these various social media have exploded in popularity and become integrated into many aspects of people's daily lives (Osterrieder, 2013). Some of them have been widely adopted by scientists in scholarly communication (Sugimoto, Work, et al., 2017; Tenopir et al., 2013) and employed by the public in engaging with scientific information (Brossard & Scheufele, 2013; Huber et al., 2019). From the perspective of scholarly communication, although the attitude of scientists towards social media differs by individual, discipline, and country (Haustein, Peters, Bar-Ilan, et al., 2014; Holmberg & Thelwall, 2014; Mohammadi et al., 2015; Thelwall & Kousha, 2015; Weingart & Guenther, 2016), it has generally been found that scientists are increasingly embracing social media as an interactive way to discern research opportunities, improve research efficiency, communicate with colleagues, and advertise research findings (Bik & Goldstein, 2013; Didegah et al., 2018; Rowlands et al., 2011; Van Noorden, 2014). From the perspective of public engagement with science, according to a survey by the U.S. National Science Board (2020), in 2018 around 57% of Americans cited the Internet as their primary source of science and technology information, up from 9% in 2001. More specifically, a Pew Research Center (2017) survey found that most social media users in the U.S. reported seeing science-related posts. Social media have been observed as one of the predominant sources for the public to seek, comment on, and share specific information about scientific issues (Kahle et al., 2016).

The growing uptake of social media by both scientists and members of the public has narrowed the “gap” (H. P. Peters, 2013) between internal scholarly communication and public science communication, because social media provide an arena where scientists and the public can directly meet and interact with each other around science (Didegah et al., 2018). Different from the once quasi-monopoly of science journalism in the communication between science and the public before the 1990s (H. P. Peters, 2013), via social media nowadays, everyone is capable of acting as a science communicator to broadcast scientific information to broader audiences, thus helping bridge the gap between science and the public, and facilitating the circulation of scientific knowledge in online media environments. Furthermore, relying on the interactive nature of social media, users can engage in scientific discussions through a wide range of interaction behaviors (Costas et al., 2021), such as commenting on science blogs and replying to scholarly tweets. These interactions are potentially accompanied by cues about the accuracy, importance, and popularity of science stories embedded in the interacted scholarly objects, as described by Brossard and Scheufele (2013), thereby “adding meaning beyond what the author of the original story intended to convey”.

1.2.2 Conceptualizing interactions between (and within) science and social media

Figure 1 illustrates the conceptualization of various interactions existing within and between the science environment and the social media environment. The framework in Figure 1 is composed of different major conceptual components. We describe them below.

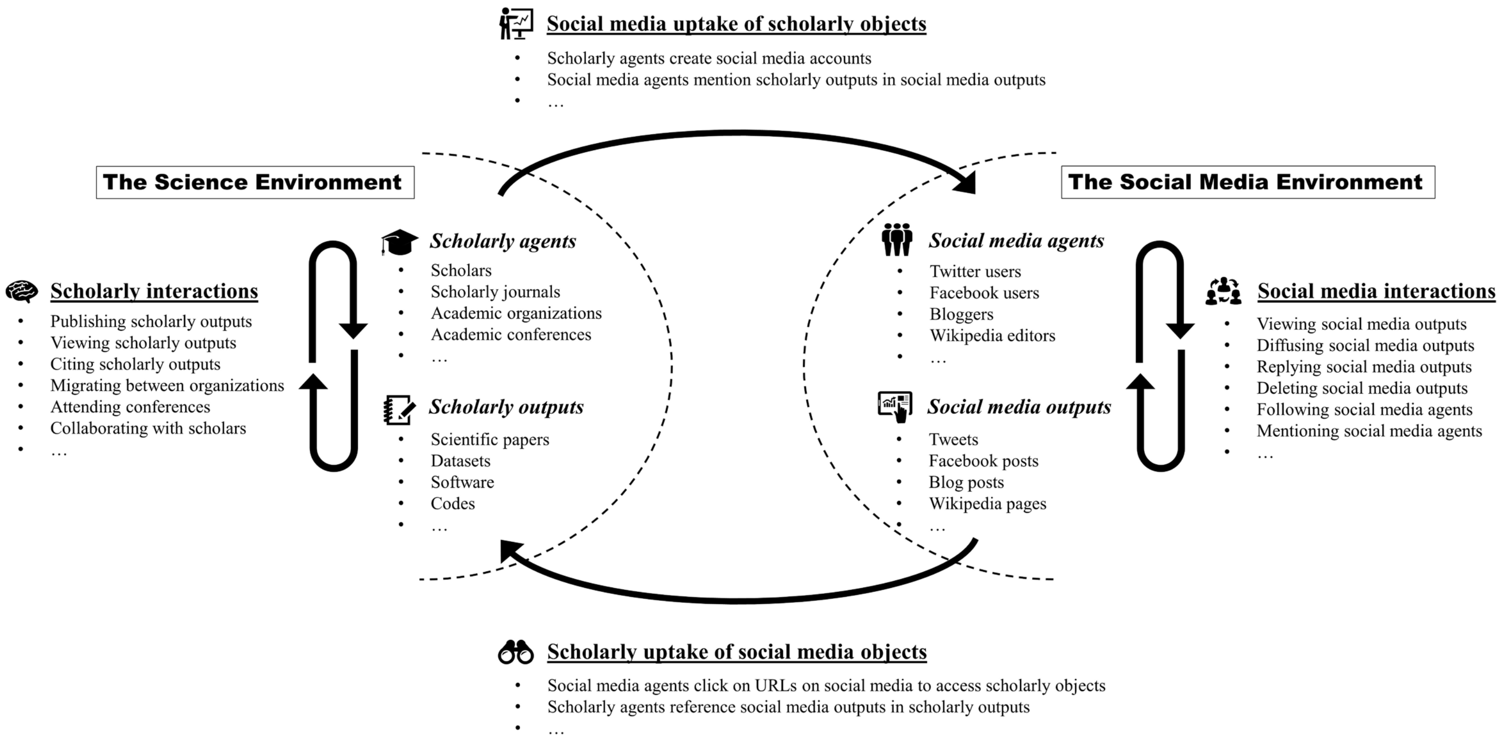

Figure 1. Conceptualized interactions and *information flows* within and between the science environment and the social media environment

- *Science environment* refers to the environment where scientists, scientific institutions, and other scholarly actors network and conduct scientific activities (e.g., publishing new knowledge, running scientific events, or conducting teaching activities). Within the science environment, there are two particular types of scholarly objects: *scholarly agents* as actors and *scholarly outputs* as outcomes of scientific activities (Haustein, Bowman, & Costas, 2016).
    - *Scholarly agents* refers to both individuals (e.g., individual scholars) and institutions (e.g., scholarly journals/publishers, funding agencies, and academic organizations).
    - *Scholarly outputs* include all kinds of research products coming from different scientific activities, including not only traditional scientific publications (e.g., journal articles or book chapters), but also datasets, software, and other types of outcomes resulting from research activities (Piwowar, 2013).
- *Social media environment* refers to the online environment where people post user-generated content and interact with others through diverse social media tools. Within the social media environment, there also exist two categories of social media objects: *social media agents* as actors and *social media outputs* as outcomes of social media events.
    - *Social media agents* are those individuals and organizations who use and carry out events and activities on online social media platforms.[1]
    - *Social media outputs* are user-generated content posted on social media applications, such as tweets posted on Twitter, or blog posts.

As conceptualized in Figure 1, between scholarly objects and social media objects, there exist heterogeneous interactions (i.e., reciprocal activities and influence between two objects). Correspondingly, interactions will result in directional information flows (i.e., transfer of information from an object to another object through a given process) as visualized with arrows in Figure 1. According to the environment where interactions take place and the

[1] Note that an individual can have a dual role as both a *scholarly agent* and a *social media agent* (e.g., a scientist using Twitter). It rests with the context in which the individual acts and the outcomes the individual's activity produces. For example, when a scientist conducts scientific activities and produces scholarly outputs, she acts as a scholarly agent, but when the scientist creates a social media account and engages in social media events, she is seen as a social media agent with an academic background. Same would apply for academic organizations (e.g., a university and its Twitter account).

direction of information flows, interactions within and between science and social media can be categorized into four types as follows:

(1) *Scholarly interactions* refers to interactions taking place amongst scholarly objects, which arouse information flows within the science environment. Scholarly interactions comprise not only interactions between scholarly agents and scholarly outputs, for instance, a scholar reads, publishes, or cites a scientific paper, but also those between scholarly agents or scholarly outputs per se, such as collaborations between scholars or scientific mobility of scholars across academic organizations (i.e., scholarly agent-scholarly agent interactions), and citation relationships between scientific papers (i.e., scholarly output-scholarly output interactions). Scholarly interactions are the main research objects of more traditional scientometric research, providing concrete evidence for studying, for example, scientific productivity (Cole & Phelan, 1999; Lotka, 1926), citations (Garfield, 1972; Wouters, 1999), scientific collaborations (Beaver & Rosen, 1978; S. Lee & Bozeman, 2005; Sonnenwald, 2007), and scientific mobility of scientists (Moed et al., 2013; Sugimoto, Robinson-Garcia, et al., 2017).

(2) *Social media uptake of scholarly objects* refers to interactions between science and social media, specifically capturing information flows originating from the science environment to the social media environment (i.e., science-social media). The social media uptake of scholarly objects is mainly accomplished by scholarly outputs being shared on social media or scholarly agents being active in social media platforms. The former captures specific social media attention towards scholarly outputs, while the latter reflects social media activities of scholarly agents. These science-social media interactions will cause information originally rooted in the science environment (whether about scholarly outputs or scholarly agents) to be visible and interactive in the social media environment. This is the type of interactions that has most often been studied in altmetric research (e.g., the number of tweets to scientific papers, the number of Wikipedia citations to scientific papers, or the number of Mendeley readers of research datasets) (Colavizza, 2020; I. Peters et al., 2016; Thelwall, Haustein, et al., 2013).

(3) *Social media interactions* stands for interactions within the social media environment, which lead to information flowing between social media objects. Social media interactions include not only interactions between social media agents and social media outputs, for instance, a Twitter user likes or retweets a tweet, but also interactions existing between social media agents or social media outputs themselves, for example, a Facebook user follows other users (i.e., social media agent-social media agent interactions) or a Wikipedia page contains hyperlinks connecting to other Wikipedia pages (i.e., social media output-social media output

interactions). Social media interactions come in a variety of forms because of the diverse interaction functionalities made available by different social media platforms, but they are essentially the same in terms of the internal information flows within the social media environment. In the context of social media metrics of science, this type of interactions has been less studied, although some examples include the study of the likes and retweets given by Twitter users engaging with science (Díaz-Faes et al., 2019), and the studies of the user engagement (e.g., likes, retweets, comments, or shares) with the content posted by some Twitter or Facebook accounts run by academic institutions (Bhattacharya et al., 2014, 2017; Kahle et al., 2016; H. Park et al., 2016).

(4) *Scholarly uptake of social media objects* refers to interactions between science and social media, but it triggers information flows originating from the social media environment back to the science environment (i.e., social media-science). For example, by clicking on scholarly URLs (e.g., links to scientific papers or homepages of academic organizations) embedded in social media outputs, social media agents will depart from the social media application in use and be directed to the webpages of scholarly objects, thus making for the information flows from social media to science. Besides, the information flows from social media back to science can also be observed through the lens of references to social media outputs in scholarly outputs, just like in this dissertation (scholarly objects) we referenced Jason Priem's tweet (social media objects) earlier to introduce the origin of the term altmetrics. This type of research is the least developed, with only a few related research discussing, for example, the citations in scientific papers to Wikipedia pages (Noruzi, 2009; T. K. Park, 2011), newspaper articles (Hicks & Wang, 2013), or blog posts (Late et al., 2019).

Some examples of the four types of conceptualized interactions are: (1) First, a scholar (scholarly agent) may read and cite a paper (scholarly output), this is a scholarly interaction because the interaction and the information flow stay within the science environment. (2) Then, the scholar may share the paper (scholarly output) by posting a tweet about it (social media output). This represents an uptake of scientific information in the Twitter universe. (3) After seeing this tweet (social media output) linking to a paper, some other Twitter users (social media agent) may retweet it, which is regarded as a social media interaction since it happens within the social media environment. (4) Finally, some Twitter users (social media agent) may click on the tweeted link to acquire more details of the paper, which eventually leads them to the webpage of the scientific paper (scholarly object). This is a kind of scholarly uptake of social media objects, which substantially makes an impact on scholarly objects (e.g., increases the visits of scientific papers).

1.2.3 Interpreting social media metrics of science: definition and developments

Definition of social media metrics of science

In the proposed conceptual framework, the three interactions involving social media make up the main objects of study for *social media metrics of science*. Social media metrics of science can be broadly defined as metrics aiming at capturing and characterizing interactions on social media platforms related to scholarly objects. Under this definition, social media metrics comprise not only the analysis of social media interactions with regard to scholarly objects (i.e., social media uptake of scholarly objects, social media interactions, and scholarly uptake of social media objects), but also the characterization of social media agents (e.g., demographics, identities, and behaviors of Twitter users, Facebook users, or Mendeley users) and social media outputs (e.g., content, sentiment, and feature use of tweets, Facebook posts, or blog posts) integrated in the social media interactions with science. Costas (2017) argued that it can also be referred to as *social media studies of science*, which generalizes the studies of "the relationships and interactions between social media and scholarly objects".

According to the differences in the objects of study, Díaz-Faes et al. (2019) classified social media metrics of science into *primary* and *secondary social media metrics*. Primary social media metrics refer to "metrics of the use and visibility" of scholarly objects on social media; while secondary social media metrics are "metrics about the social media users and their online activities", thus focusing on social media objects. Put in our conceptual framework, primary social media metrics mainly focus on the exploration of social media uptake of scholarly objects, while secondary social media metrics focus on the characterization of social media interactions and scholarly uptake of social media objects.

Interpreting primary social media metrics

As to primary social media metrics, related research so far has had a strong focal point on the uptake of scholarly outputs on social media. Investigations of the extent to which scholarly outputs (scientific papers in particular) are shared and discussed on distinct social media platforms constitute the main body of existing social media metrics literature (Fenner, 2013; Hammarfelt, 2014; Haustein, Peters, Bar-Ilan, et al., 2014; Robinson-Garcia et al., 2014; Zahedi et al., 2014) and form the basis for the article-level indicators developed by main altmetric aggregators like Altmetric.com and PlumX (Adie & Roe, 2013; Ortega, 2018a; Zahedi & Costas, 2018).

The particular research interest in the reception of scholarly outputs on social media parallels the initial expectation towards the potential of social media metrics in supplementing or even replacing traditional scholarly impact measures (e.g., peer-review, journal impact factors, and citations) for research evaluation (Priem et al., 2010; Priem & Hemminger, 2010). Many comparisons have been done to examine the correlations between traditional scholarly impact indicators and emerging social media metrics at the article level (Bardus et al., 2020; Costas et al., 2015a; Huang et al., 2018; Shema et al., 2014; Thelwall, Haustein, et al., 2013; Waltman & Costas, 2014). In general, the correlations with citations vary across different types of social media metrics. For example, it has been argued that Mendeley readership can capture a similar type of impact of scientific papers as citations do, on the basis of the positive and moderate to strong correlations found between these two measures (Thelwall & Sud, 2016; Thelwall & Wilson, 2016; Zahedi et al., 2017). Nevertheless, the majority of social media metrics derived from other sources like Twitter, Facebook, or blogs, have mostly weak or negligible correlations with citations (Costas et al., 2015a; Haustein, Peters, Sugimoto, et al., 2014), supporting the idea that science and social media are two fundamentally distinct environments. Therefore, rather than serving as a measure of scientific excellence, the uptake of scholarly outputs on social media has been deemed to open the possibilities to explore the diverse societal aspects of scientific developments, also specifically referred to as the *societal impact* of research (Bornmann, 2014a, 2015b; Bornmann et al., 2019; Noyons, 2019).

Towards secondary social media metrics

In contrast to traditional citation-based metrics, the context of social media metrics is more heterogeneous (Haustein, 2016). The heterogeneity of social media metrics is visible in several aspects: first, social media metrics entail heterogeneous metrics derived from different data sources with distinct purposes, functionalities, and user groups (Haustein, 2016); second, social media metrics are drawn upon digital traces left by heterogeneous actors from different backgrounds, including not merely scholars and practitioners, but also journalists or the general public (Yu, 2017); third, as conceptualized in our framework, social media metrics involve heterogeneous interaction behaviors driven by diverse user motivations. In short, behind the mere quantification of the social media uptake of scholarly objects measured by primary social media metrics lies abundant additional information regarding the context in which scholarly content is discussed, disseminated, and valued on social media. In order to delineate the context, many research efforts have been made, thus giving rise to the notion of secondary social media metrics.

Previous research related to secondary social media metrics mainly focused on (1) the characterization and clustering of social media agents based on their profiles and interaction relationships (Alperin et al., 2019; Díaz-Faes et al., 2019; Hassan et al., 2019; Mohammadi et al., 2015; Pearce et al., 2014; Said et al., 2019; Van Schalkwyk et al., 2020); (2) the characterization and networking of social media outputs based on their content and features use (Haustein, Bowman, Holmberg, Peters, et al., 2014; Holmberg et al., 2014; Pulido et al., 2018; Robinson-Garcia et al., 2017; Thelwall, Tsou, et al., 2013); and (3) the characterization and measurement of engagement behaviors of social media agents around social media outputs (Bhattacharya et al., 2017; Hoang et al., 2015; Sugimoto et al., 2013; Tsou et al., 2014; X. Wang et al., 2017; X. Wang, Fang, & Guo, 2016).

Although primary social media metrics capture the information flows from science to social media, they fail to account for the subsequent information flows within the social media environment. Secondary social media metrics make up for this gap by contextualizing how and by whom scholarly content is further interacted with. For example, two scientific papers may attract equivalent attention on Twitter because both of them have been tweeted the same number of times, however, the nature and degree of the Twitter attention may be further characterized through the lens of the Twitter users involved, the detailed tweet content, or the possible engagement triggered. Thus, it can be argued that secondary social media metrics play an important role in further understanding the mechanisms by which scientific information is processed and circulated on social media. Such understanding paves the way towards the development of more advanced indicators of the dissemination of scientific information in social media environments.

1.3 Scholarly Twitter metrics: Characterizing Twitter interactions around science

This dissertation focuses on Twitter to characterize diverse interactions around scientific information within and between the science environment and the Twitter environment. As a subset of social media metrics of science, studies on Twitter interactions related to scholarly objects have been termed as *scholarly Twitter metrics* by Haustein (2019). Following the same logic as social media metrics, scholarly Twitter metrics can be categorized into *primary Twitter metrics* and *secondary Twitter metrics*. This section begins with a literature review of primary Twitter metrics-related research. Then, the section also reviews research with respect to secondary Twitter metrics. Last, this section discusses the challenges that scholarly Twitter metrics currently confront, and the opportunities that studying Twitter interactions around science can bring.

1.3.1 Primary Twitter metrics: Twitter uptake of scholarly objects

Primary Twitter metrics focus on scholarly objects, offering insights into whether and how many times scholarly objects are present in the Twittersphere, and, in reverse, how scholarly objects can be connected based on their "heterogeneous couplings" (Costas et al., 2021) on Twitter. The Twitter uptake of both scholarly agents and scholarly outputs has widely been studied in previous literature to explore the extent to which information flows from science to Twitter.

Twitter uptake of scholarly agents

There are mainly two ways to investigate Twitter uptake of scholarly agents. One is to survey scholarly agents about their uptake of Twitter through questionnaires and interviews; the other is to identify scholarly agents on Twitter based on Twitter users' profiles and behaviors.

Multiple surveys have been done to inquire scholars' uptake of Twitter as well as their motivations for using it. In a 2010 international survey, Rowlands et al. (2011) found that 9.2% of the 2,414 scholars responded to the survey were active on microblogs (93% using Twitter). Later, in a global survey conducted by *Nature* in 2014 (Van Noorden, 2014), more than 3,500 scholars shared their attitudes towards popular social media tools. This survey reported that over 80% of the responses from scientists and engineers said that they were aware of Twitter, and 13% of them visited it regularly and used it in an interactive way to follow research-related discussions and comment on relevant research. In another survey to U.S. professors in 2014 (Bowman, 2015), a total of 613 out of 1,910 respondents (32%) reported having a Twitter account. Similarly, Haustein, Peters, Bar-Ilan, et al. (2014) surveyed a sample of 57 bibliometricians about their uptake of various social media tools. Their survey results showed that nearly half of the respondents had Twitter accounts and they used Twitter for both personal and professional purposes. In another survey that investigated the use of Twitter by scientists, Collins et al. (2016) found that scientists mostly tweeted about "research within their own field", followed by "science outreach and communication", "personal research", and then "personal life and experiences".

In addition to surveys, there has been a variety of methodologies developed to identify and characterize Twitter users who are scholarly agents, in particular scholars on Twitter. For example, most previous research identified scholars on Twitter by searching for the names of preselected scholars directly or searching for specific professional keywords (e.g., scientific occupations or academic conferences) in Twitter users' names or biographies, and then further expanded the preliminary samples obtained based on the identified scholars' Twitter followers and lists (Chretien et al., 2011; Holmberg & Thelwall, 2014; Ke et al., 2017; Lulic & Kovic, 2013; Pearce et al., 2014; Priem et al., 2011). In addition, more systematic methods have been developed to study the Twitter uptake of scholars on a larger scale. For instance, Costas et al. (2020) developed a scoring system which considers not only the names of scholars and Twitter users but also other matching elements like e-mail URLs, geolocations, and tweeted papers, based on which they identified and characterized 296,504 scholars active on Twitter by matching between over 25 million disambiguated scholars who have published Web of Science-indexed papers and Twitter users who have posted scholarly tweets recorded by Altmetric.com. In addition to scholars, Twitter uptake of other types of scholarly agents has also been explored in previous research, such as scholarly journals (Kamel Boulos & Anderson, 2014; Zheng et al., 2019), universities (Veletsianos et al., 2017), and scholarly publishers (Zedda & Barbaro, 2015).

Twitter uptake of scholarly outputs

As for Twitter uptake of scholarly outputs, most studies have looked into the extent to which scientific papers are mentioned on Twitter. Overall, the proportion of scientific papers receiving at least one scholarly tweet ranges from 9% up to 36% (Costas et al., 2015a; Haustein, Larivière, et al., 2014; Meschede & Siebenlist, 2018; Robinson-Garcia et al., 2014; Thelwall, Haustein, et al., 2013), making up the second largest social media metric data source of scientific papers, next only to Mendeley (Sugimoto, Work, et al., 2017). Twitter uptake of scientific papers varies by scientific discipline, document type, publication date of papers, and affiliated country of papers' authors. For example, papers from the fields of social sciences, life sciences, and medical and health sciences are more likely to be present on Twitter, in contrast to those related to natural sciences and engineering (Costas et al., 2015a; Didegah et al., 2018; Haustein, Costas, et al., 2015). In terms of document types, different from the advantage of research articles and reviews in accruing citations, news items show the highest possibility of getting Twitter mentions, followed by reviews and editorial materials (Haustein, Costas, et al., 2015). With respect to the publication dates of papers, Twitter uptake is higher for more recently published papers (Haustein, Larivière, et al., 2014). Twitter uptake of scientific papers also varies across the country of the authors, for instance, papers from China or Latin America tend to have a much lower Twitter uptake as compared to the U.S. and European countries (Alperin, 2015; Wang, Fang, Li, et al., 2016).

From the standpoint of research evaluation, the Twitter uptake of scientific papers has been assumed to capture broader forms of attention to science. Twitter metrics have been frequently compared with traditional citation counts to determine the potential that scholarly Twitter metrics could have in predicting future citations. Although some research based on papers from certain disciplines suggested that the level of Twitter uptake of scientific papers could serve as a predictive indicator of future citations (Eysenbach, 2011; Peoples et al., 2016; Shuai et al., 2012), the generalization of this conclusion has been refuted by the weak or even negligible correlations found between number of citations and Twitter mentions in many large-scale and cross-disciplinary analyses (Costas et al., 2015a; Haustein, Costas, et al., 2015; Jabaley et al., 2018; Rosenkrantz et al., 2017; Zahedi et al., 2014). These weak correlations point to the idea that the uptake of scholarly outputs in the science environment (e.g., as reflected by citations) and the uptake in the Twitter environment (e.g., in primary Twitter metrics, reflected by Twitter mentions) are fundamentally different, in terms of their audiences (e.g., scholarly audience in the science environment, more diverse audience in the Twitter environment) as well as in the norms followed in the engagements and decisions of these audiences in the Twitter environment. Simply put, the reasons why a scholar would choose a paper to cite are fundamentally different from the reasons why a Twitter user would choose a paper to tweet about. Therefore, the Twitter uptake of scholarly outputs has been argued more as a complement to citation-based indicators (Haustein, Costas, et al., 2015), rather than an alternative as it was set forth earlier on (Priem et al., 2010; Priem & Hemminger, 2010).

In addition to the potential application in research evaluation, the co-occurrence of scholarly objects on Twitter in different forms also opens up some novel approaches to network and further cluster scholarly objects. For example, in previous research heterogeneous "co-tweeted" networks have been constructed for scholarly objects (e.g., scientific papers, scholarly journals, and subject fields) tweeted by the same user account or tweeted in the same tweet post (Didegah & Thelwall, 2018; Hassan, Aljohani, Shabbir, et al., 2020; Jung et al., 2016; Robinson-Garcia et al., 2019), shedding light on a new way of unearthing the intrinsic relations amongst scholarly objects captured in the Twitter environment.

1.3.2 Secondary Twitter metrics: Twitter objects as the research objects

Secondary Twitter metrics are Twitter objects-focused, aiming at the characterization of who participates in the dissemination of scholarly content on Twitter, how scholarly content is shared and engaged with by Twitter users, and how Twitter objects associated with scholarly content are connected. As highlighted by Haustein (2019), "Twitter is less about what people tweet rather than how they are connected". Secondary Twitter metrics fit this argument by mapping the contexts in which Twitter interactions around scholarly objects occur, and depicting the details of Twitter interactions that primary Twitter metrics cannot deliver. Previous research related to secondary Twitter metrics mainly focused on the characterization of Twitter users, scholarly tweets, and Twitter engagement behaviors.

Characterization of Twitter users interacting with science

The identities and characteristics of the Twitter users are relevant to the interpretation of the nature of Twitter interactions. Previous research has paid much attention to who tweets about scientific information, analyzing the real-world identities of Twitter users. Users' profile descriptions provide direct evidence of how users describe themselves, so this information has been widely leveraged to analyze the portrait of Twitter users. Largely based on users' profile descriptions, Altmetric.com categorizes Twitter users with scholarly tweets posted into four types: researcher, practitioner, science communicator, and member of the public,[2] which has been adopted by several altmetrics researchers to compare the tweeting behaviors across different types of users (Hassan, Aljohani, Idrees, et al., 2020; Xia et al., 2016; Yu, 2017). Nevertheless, this coarse-grained classification has its limitations such as overestimating the amount of users classified as the general public (Didegah et al., 2018). Instead, some research manually coded relatively small samples to pinpoint the identities of

[2] See more information about how Altmetric.com categorizes Twitter users at: https://help.altmetric.com/support/solutions/articles/6000235926-twitter (Accessed May 25, 2021).

Twitter users interacting with science. By manually coding a random sample of 2,000 unique Twitter users who had tweeted scientific papers, Tsou et al. (2015) found that 76% of them were individual users and 23% were organizational users. Furthermore, of the identified individual users, nearly 12% were identified as students and 34% were identified as possessing a Ph.D., indicating that about half of the individual users had an academic background. Similarly, Yu et al. (2019) manually coded a sample of 1,468 Twitter users who had tweeted scientific papers and found that about 49% of them with an academic background (e.g., researchers, Ph.D. students, and universities), 38% identified as "general public" users, and 13% identified as science communicators (e.g., scholarly journals and publishers). From a disciplinary point of view, Didegah et al. (2018) found that the composition of the identity of Twitter users who tweeted scientific papers differed across subject fields. For example, for the tweeted papers related to social sciences and humanities, the involved Twitter users as members of the public and individual researchers accounted for 37% and 35%, respectively, whereas for the tweeted papers in the field of life and earth sciences, civil society organizations accounted for 48% of the involved Twitter users, followed by those identified as individual researchers (45%). Overall, previous research suggests that although Twitter users interacting with science do not reflect the general population of Twitter users in consideration of the existence of the large portion of users with an academic background (Tsou et al., 2015), Twitter interactions around science indeed include considerable numbers of members of the public and organizations outside academia (Didegah et al., 2018), thus offering opportunities to track the sharing of scientific knowledge within broader segments of society and capture the science-society interactions.

In addition to the characterization of users individually, previous research has also detected user communities to reveal the dissemination patterns of scientific information. Twitter users interacting with science have been networked based on diverse relations, such as the co-occurrence of words used in their profile descriptions (Díaz-Faes et al., 2019; Vainio & Holmberg, 2017), the user coupling relationship (Van Schalkwyk et al., 2020), and the user mentioning/retweeting/following relationships (Alperin et al., 2019; Araujo, 2020; Hassan et al., 2019; Said et al., 2019). Different from the manual coding method aiming at a selected sample of users, such networking approach can help detect groups of users sharing similar interests or backgrounds on a much larger scale.

Characterization of the content of scholarly tweets

Scholarly tweets are carriers of scientific information in the Twitter universe, delivering traces not only about how the authors of the tweets processed scientific information, but also how they managed to disseminate information and connect to other users. These traces are mainly embodied in tweet texts and used tweet features.

Tweet texts of scholarly tweets provide straightforward insights into how Twitter users introduce scientific information while posting scholarly tweets. However, in practice tweet texts of scholarly tweets have been found to be lacking original thought. By examining the content of the tweets to the top-10 most tweeted dental papers, Robinson-Garcia et al. (2017) found that the texts of the majority of the studied tweets were devoid of original thought but inundated with mechanical and duplicate content. Similarly, with a case study containing 270 scholarly tweets, Thelwall, Tsou, et al. (2013) reported that 42% of the analyzed tweets only echoed the title of the tweeted papers and 41% briefly summarized the key points of the papers. They speculated that one of the possible reasons for users summarizing the tweeted papers was to translate the scientific information to general audiences. Sentiment analyses of the texts of scholarly tweets pointed to a similar observation: only a limited share of scholarly tweets exhibited positive or negative sentiment expressed by users, while the majority were neutral in sentiment, implying that the texts of scholarly tweets are generally factual with emotional opinions rarely observed (Friedrich et al., 2015; Hassan, Aljohani, Idrees, et al., 2020; Thelwall, Tsou, et al., 2013; S. Xu et al., 2018).

There are several user-driven features made available by Twitter to help users increase the visibility of their tweets and interact with other users, with hashtags (keyword or phrase prefixed with #) and user mentions (user's handle name prefixed with @) being the most analyzed ones in previous research on scholarly Twitter metrics. Similar to the role of traditional metadata in tagging scientific documents, the use of hashtags may enhance the description and retrievability of tweets and thus facilitate the connections amongst users interested in the same topics (Haustein, 2019; Holmberg et al., 2014). Haunschild et al. (2019) found that 35% of their analyzed scholarly tweets referring to climate change papers contained at least one hashtag, which is comparable to the research by Haustein (2019) based on the 24.3 million scholarly tweets recorded by Altmetric.com until June 2016, in which she reported that 31% of the tweets contained at least one hashtag. Hashtags used in scholarly tweets pertaining to certain research fields have been used to capture the focuses of Twitter users on specific research topics (Haunschild et al., 2019; Lyu & Costas, 2020; S. Xu et al., 2018). Different from the role of hashtags in tagging and broadcasting tweets, user mentions function as a feature to target and address specific users (i.e., the mentioned users), presenting a conversational nature. Due to this nature, as mentioned earlier, user mentions in scholarly tweets are usually used as the clues of connections between users and thus are built upon to cluster users into communities (Hassan et al., 2019; Pearce et al., 2014).

Characterization of Twitter engagement behaviors around scholarly tweets

In contrast to the characterization of scholarly tweets and involved Twitter users, Twitter engagement behaviors around scholarly tweets have been less discussed in existing literature. Amongst all sorts of Twitter engagement behaviors, retweeting is the most studied one, largely due to the data availability enabled by main altmetric data aggregators (e.g., Altmetric.com, PlumX, and Crossref Event Data) tracking and treating retweets as scholarly tweets as well. In 2010, Priem and Costello (2010) found that retweets only accounted for 19% of a sample of scholarly tweets posted by 28 academic users. In more recent research, it was generally found that retweets accounted for close to or over half of scholarly tweets of papers (Alperin et al., 2019; Didegah et al., 2018; Haustein, 2019), implying that retweets play a large part in the indicator system seeing retweets as a category of scholarly tweets rather than a type of user engagement with original scholarly tweets.

There is also some research treating number of retweets as one of the engagement metrics to measure the degree of Twitter attention attracted by the tweets posted by some organizational user accounts like health agencies (Bhattacharya et al., 2014; H. Park et al., 2016) and research organizations (Kahle et al., 2016). Besides retweeting, other engagement behaviors may also happen around scholarly tweets. For example, users may quote a scholarly tweet with their own comments added, reply to a scholarly tweet to start or participate in a Twitter conversation, add a scholarly tweet to bookmarks, and click on the embedding URLs to access the original scholarly content. According to the survey results by Mohammadi et al. (2018), users engaging with scholarly tweets are driven by various motivations such as informing the authors that their tweets are interesting, disseminating the tweets, and saving the tweets for future access. Some of the engagement metrics drawn upon these engagement behaviors have also been picked up to reflect the degree to which the public engages with scholarly content posted on social media (Kahle et al., 2016), however, most of them have hardly been analyzed in a scholarly context.

1.3.3 Opportunities and challenges facing scholarly Twitter metrics

Opportunities opened by primary Twitter metrics

Along with the emerging impact agenda in science policy of valuing the relevance of scholarly outputs beyond the scientific institutions (Gunn & Mintrom, 2016; Wilsdon et al., 2015), scholarly Twitter metrics have gained much attention because of its presumed potential in demonstrating societal impact of scholarly outputs as a significant component of altmetrics (Bornmann, 2014a; Thelwall, 2020). The attention to scholarly Twitter metrics is not only limited to the scientometric community, but has proliferated in a broad array of research fields from which researchers are increasingly applying scholarly Twitter metrics to assess the extent to which scholarly outputs in their own areas are shared on Twitter (Jabaley et al., 2018; Kolahi & Khazaei, 2018; Maggio et al., 2017; Rosenkrantz et al., 2017). As a result, primary Twitter metrics, which offer a concise way to evaluate the Twitter uptake of scholarly outputs, dominate the existing altmetric literature including or centering on Twitter data. As the most commonly used form of scholarly Twitter metrics, primary Twitter metrics have opened up several opportunities:

(1) *Capturing social media attention towards scholarly outputs*: In a traditional evaluation system reigned by indicators measuring scholarly attention, like citations and journal impact factor, the advent of primary Twitter metrics, as well as other social media metrics of a similar nature, creates the opportunity to capture more diverse forms of attention that scholarly outputs may attract (Crotty, 2014). The degree of social media attention also indicates the visibility of scholarly outputs amongst people active on social media. In practice, the fundamental indicator stemmed from primary Twitter metrics – number of scholarly tweets – and its variants have generally been utilized by main altmetric data aggregators and incorporated into article-level metrics by many scholarly publishers like Elsevier, PLoS, and Springer Nature, enabling readers to easily access how much social media attention scholarly outputs have drawn.

(2) *Providing early evidence of interactions around scholarly outputs shortly after their publication*: In contrast to citation data which need relatively longer period to accumulate resulted from both publication delay and citation delay (Bollen & Van De Sompel, 2006), scholarly tweets can happen in a much shorter period of time after the publication of scholarly outputs, even within hours or minutes (Shuai et al., 2012). The high speed of Twitter reception of scholarly outputs allows for the observation of early interactions around scholarly outputs when citations are absent. Although primary Twitter metrics capture a different kind of attention from citations as mentioned above (Haustein, Costas, et al., 2015), it, to some extent, fills the gap existing in the very initial life cycle of scholarly outputs by opening a window to track the reactions from possibly both academic and non-academic users (Darling et al., 2013; Ortega, 2018b).

(3) *Identifying scientific developments of interest by a broader public*: Citation data have long been used to detect research topics or emerging trends of interest by academia (Chen, 2006; Small et al., 2014; Q. Wang, 2018). By looking into what scholarly outputs got more Twitter mentions, primary Twitter metrics bring the possibility to identify scientific developments of particular interest by a broader public. Especially given the exponential growth of the number of scholarly outputs (Larsen & Von Ins, 2010), Twitter uptake of scholarly outputs can help map the thematic focus of science in the eyes of Twitter users (Colavizza et al., 2021; Costas et al., 2015b; Robinson-Garcia et al., 2019), thereby serving as a tool for both researchers and the general public to tap into scientific developments that have driven extensive social or popular interest in online environments (Priem, Groth, et al., 2012).

Challenges facing primary Twitter metrics

Although primary Twitter metrics draw a quantitative and intelligible picture of the information flows from science to Twitter, here are some particular challenges facing primary Twitter metrics, including the difficulty in the interpretation of the nature of Twitter attention, the misuse of the composite indicators mixing up different tweet types, and the neglect of the data quality issues due to the volatility of tweets:

(1) *The difficulty in interpreting the nature of social media attention captured on Twitter*: Only based on the Twitter uptake of scholarly outputs tracked by primary Twitter metrics, it is difficult to interpret the nature of the captured social media attention for two main reasons: one is the heterogeneity of the identities of Twitter users contributing to the Twitter uptake of scholarly outputs; the other one is the heterogeneity of the motivations driving users to interact with science (Haustein, 2016). For example, academic users may take part in the Twitter interactions with their attention coming more from a scholarly nature, while non-academic users stand more for a kind of societal or popular attention. Even for the same type of Twitter users, they may pay attention to scholarly outputs motivated by different reasons such as viral jokes, scientific hoaxes, or even just tweeting robotically (Haustein, Bowman, Holmberg, et al., 2016; Sugimoto, 2015), which further exacerbates the difficulty in properly interpreting the nature of Twitter attention without exploring the contexts behind numbers.

(2) *The coarse-grained indicators compounded by individual indicators drawn upon different types of tweets*: Twitter has developed a series of functionalities for users to create and engage with tweets, which enable the creation of different types of tweets (e.g., original tweets, retweets, quote tweets, and reply tweets) and different engagement metrics for tweets (e.g., number of retweets, quotes, replies, likes, and clicks). [3] Largely because the most used altmetric data providers, such as Altmetric.com and PlumX, aggregate scholarly tweets into composite indicators regardless of the differences amongst tweet types and the engagement occurred around tweets,[4] the current indicator system of primary Twitter metrics, whether adopted in research or in practice, is generally sketchy and coarse-grained.

(3) *The neglect of the data quality issues due to the volatility of Twitter data*: In the context of primary Twitter metrics, data quality issues have been discussed with

[3] https://help.twitter.com/en/managing-your-account/using-the-tweet-activity-dashboard (Accessed May 25, 2021).

[4] Altmetric.com reports the total number of scholarly tweets received by scholarly outputs without distinguishing between tweet types. PlumX only distinguishes between "tweets" (aggregation of original tweets, reply tweets, and quote tweets) and "retweets".

caveats mainly in terms of the inconsistency of the Twitter uptake of scholarly outputs reported across altmetric data aggregators (Meschede & Siebenlist, 2018; Zahedi & Costas, 2018). As a challenge of central importance, data quality issues also exist in the metadata extracted from Twitter. Different from the accumulated scientometric data (e.g., citations, publications) which remain in theory persistent, Twitter data present a congenitally volatile nature: Twitter users can easily post tweets, likewise, they can easily delete tweets or hide tweets by protecting their user accounts, resulting in the disappearance or unavailability of tweets (Zubiaga, 2018). Through the lens of primary Twitter metrics, the volatility of Twitter data may reflect in the fluctuations of the number of scholarly tweets accrued by scholarly outputs over time. There are less efforts paid to explain and discuss the causes as well as its possible influence on Twitter-based metrics.

New opportunities brought about by secondary Twitter metrics

The aforementioned challenges that primary Twitter metrics are confronting underscore the importance of diving into the Twitter environment where scholarly tweets are produced and further interacted with. Against this background, the emergence of secondary Twitter metrics stimulates the transformation from mere counting the Twitter uptake of scholarly objects towards interpreting the processes of Twitter interactions, thereby creating more opportunities to complement and underpin primary Twitter metrics:

(1) *Understanding the nature of Twitter uptake of scholarly outputs*: In order to understand what scholarly tweets essentially measure, as suggested by Haustein (2019), it is necessary to solve the problems about "how, when, and by whom" scholarly tweets are posted. The nature of Twitter attention towards scholarly outputs can be elaborated on by looking into who the involved users are and how they tweet about scientific information, which fall into the scope of secondary Twitter metrics. For example, given that the proportion of academic users and non-academic users interacting with science on Twitter found to be almost half-and-half by many previous research independently (Mohammadi et al., 2018; Tsou et al., 2015; Yu et al., 2019), Thelwall (2020) concluded that, in general, scholarly tweets might reflect half academic, half non-academic attention. Similarly, by scrutinizing the content of scholarly tweets, one can easily tell whether the scholarly tweets are the results of interest (Thelwall, Tsou, et al., 2013), humor (Didegah et al., 2018), doubt (Haunschild & Bornmann, 2021), mechanical tweeting behavior (Robinson-Garcia et al., 2017) and so on. Therefore, by contextualizing scholarly tweets, secondary Twitter metrics help better understand the underlying mechanism through which Twitter attention emerges towards scholarly outputs, which mere primary Twitter metrics alone cannot convey.

(2) *Improving scholarly Twitter metrics with more advanced and fine-grained indicators*: Amongst all sorts of tweet types, original tweets, quote tweets, and reply tweets are the tweet types capable of functioning as *initiators* of bringing scientific information to Twitter by including URLs to scholarly outputs. After being posted, these types of scholarly tweets can be further engaged with via retweeting, quoting, replying, clicking and so on (Kahle et al., 2016; Kalia et al., 2018), leaving traces of the further Twitter attention triggered by the engaged scholarly tweets. The engagement behaviors around scholarly tweets represent deeper levels of Twitter reception of scientific information by broader audiences. Metrics of these engagement behaviors have always been left out or conflated with the indicators of primary Twitter metrics. For example, likes, as a widespread engagement metric of scholarly tweets, has been less analyzed despite its role in the assessment of the attention that scholarly tweets attract, whereas retweets, as another engagement metric of scholarly tweets, has been assigned with the same credit as the original scholarly tweets they rely on by Altmetric.com. Clarifying and distinguishing these different tweet types and incorporating engagement metrics will benefit the development of more advanced indicators for systematically assessing Twitter attention.

(3) *Verifying the data quality for scholarly Twitter metrics*: Digging deep into the metadata behind counts is an important step towards the verification of the data quality of metrics (Ortega, 2019b; Yu et al., 2021). Particularly in consideration of the aforementioned volatile nature of Twitter data, it is of importance to keep an eye on the availability of scholarly tweets to ensure the stability and reliability of Twitter metrics. The volatility of Twitter data is deep-seated, because it is a kind of inherent data quality issue caused by legitimate user behaviors (e.g., deleting tweets and protecting accounts) and Twitter regulation (e.g., suspending accounts behaving against Twitter policy) instead of technical problems. Adopting a dynamic perspective, primary Twitter metrics can observe the fluctuations of the number of scholarly tweets to scholarly outputs, but cannot determine the reasons behind this phenomenon. In contrast, secondary Twitter metrics can provide compelling explanations to such fluctuations by looking into the interactions in the Twitter environment which lead to the unavailability of scholarly tweets.

1.4 Motivations and objectives of this PhD dissertation

1.4.1 Motivations

The need to consider more advanced social media metrics arises from a series of issues surrounding currently prevalent social media indicators, as well as the challenges these issues pose to the rationality, diversity, and reliability of social media metrics of science. For scholarly Twitter metrics which is the focus of this dissertation, the main issues discussed in the dissertation include (1) the misappropriation of hybrid Twitter-based indicators; (2) the neglect of the role of secondary Twitter metrics; and (3) the insufficient consideration of the instability of Twitter data.

More specifically, the misappropriation of hybrid Twitter-based indicators is mainly caused by the lack of clear distinction amongst diverse Twitter interactions. For example, retweets, one of the typical engagement metrics for scholarly tweets, are generally not distinguished from original tweets, but rather compounded into hybrid indicators to evaluate the Twitter uptake of scientific papers. The misuse of hybrid Twitter-based indicators conflates different Twitter interactions regardless of their discrepancies, thus making it more difficult to provide a robust interpretation of what these indicators are intended to measure. The neglect of the role of secondary Twitter metrics excludes a range of meaningful Twitter interactions with scholarly tweets from the measurement of Twitter attention. The exclusion of secondary Twitter metrics from the system of Twitter-based indicators hinders the possibility of capturing the attention of wider audiences, and assessing the effectiveness of scholarly tweets themselves in disseminating scientific information. The insufficient consideration of the instability of Twitter data also leads to the exclusion of the volatile nature of Twitter data from the measurement of Twitter attention, partly causing the inconsistency of observations across time points.

Although we scrutinize these issues in the Twitter context, similar challenges also exist in other social media sources. In order to optimize the system of social media indicators, it is necessary to face and address these issues, thus approaching more advanced social media metrics by which the attention to scholarly objects and social media objects relevant to scholarship can be characterized and measured in a more systematic and responsible manner.

1.4.2 Objectives

The main objective of this PhD dissertation is to characterize diverse Twitter interactions around science to understand in greater depth the Twitter uptake of scientific information, and contribute to improve Twitter-based metrics for research evaluation. Put specifically in the proposed conceptual framework, this dissertation aims to capture and characterize the information flows between science and Twitter caused by different Twitter interaction behaviors, covering research questions concerning both primary Twitter metrics and secondary Twitter metrics.

In terms of primary Twitter metrics, this dissertation presents state-of-the-art analyses of the volume and speed of information flows from science to Twitter by exploring the extent to which scientific papers are mentioned on Twitter and when they are mentioned on Twitter after publication.

In terms of secondary Twitter metrics, on the one hand, this dissertation studies, within the Twitter environment, how Twitter users engage with scholarly tweets through diverse types of engagement behaviors (i.e., retweeting, quoting, liking, and replying) and how the stability of Twitter metrics might be affected by some interaction behaviors which will cause the unavailability of scholarly tweets (e.g., tweet deleting and account protecting); on the other hand, this dissertation focuses on the clicking behavior around URLs in scholarly tweets to explore the extent to which scholarly tweets lead Twitter users to access scientific papers, which results in the information flows from Twitter back to science.

By examining these Twitter interactions around science, this dissertation aims to pave the way for a better understanding of the diversity and characteristics of Twitter interactions, and more importantly, approach a more fine-grained indicator system of scholarly Twitter metrics by taking into consideration the diverse interactions around scholarly tweets and clarifying their differences in the measurement of Twitter attention.

1.5 Structure and research questions of this PhD dissertation

To achieve the main objective, this dissertation investigates three main types of information flows within and between science and Twitter and the diverse forms of interactions behind them. Chapter 2 and chapter 3 belong to the scope of primary Twitter metrics. These two chapters study the Twitter uptake of scientific papers from the aspects of volume and speed, respectively, thus unravelling how broad and how fast scientific information flows to the Twitter environment. Chapter 4, chapter 5, and chapter 6 fall into the realm of secondary Twitter metrics, with chapter 4 and chapter 5 focusing on the interactions around scholarly tweets within the Twitter environment (i.e., information flows within the Twitter environment) and chapter 6 specifically focusing on the URL clicking behavior which brings about information flows from Twitter back to science. Specifically, chapters 2 to 6 set out to answer the following five research questions (RQs):

RQ1. *To what extent are scientific papers mentioned on Twitter? In particular, which subject fields and research topics are more likely to have related scientific papers mentioned on Twitter?*

Chapter 2 answers this research question by presenting a large-scale analysis of the presence of Twitter mention data as well as other eleven types of altmetric data and citation data for a total of 12.3 million Web of Science-indexed (WoS) papers published between 2012 and 2018. The eleven types of altmetric data include Mendeley readers, Facebook mentions, news mentions, blog citations, Wikipedia citations, policy document citations, Reddit mentions, Faculty Opinions (formerly F1000Prime) recommendations, video mentions, peer review comments, and Q&A mentions. Their presence amongst scientific papers is analyzed as a reference to grasp the broadness of the Twitter uptake of scientific papers. Combining with the bibliometric information of scientific papers, chapter 2 also investigates which publication years, which subject fields, and which research topics show relatively higher possibility to have scientific papers present on Twitter, unveiling the Twitter users' biases towards certain types of scholarly outputs. Because sufficient data presence is one of the key preconditions for applying metrics in practice, findings of this chapter demonstrate the potential of generalizing scholarly Twitter metrics as a means of evaluating scientific papers on a large scale.

RQ2. *When are scientific papers mentioned on Twitter after publication? In other words, how fast do Twitter mentions of scientific papers accumulate after papers are published*?

Chapter 3 answers this research question by studying the accumulation velocity of Twitter mention data as well as other eleven types of altmetric data for a total set of 2.4 million WoS papers at the day level. Based on the DOI created date recorded by Crossref as the proxy of publication date of scientific papers, and the post date recorded by Altmetric.com for all kinds of altmetric events, chapter 3 calculates the time intervals between these two time points for each altmetric event to depict how fast different altmetric data accumulate after the publication of scientific papers. To gain an overall picture, the accumulation velocity of Twitter data is compared with Facebook mentions, news mentions, Google+ mentions, blog citations, Wikipedia citations, policy document citations, Reddit mentions, Faculty Opinions (formerly F1000Prime) recommendations, video mentions, peer review comments, and Q&A mentions. Similarly, combining with the bibliometric information of scientific papers, chapter 3 compares the velocity of Twitter uptake of scientific papers by document type, subject field, and research topic. Findings of this chapter provide empirical evidence of the accumulation velocity of Twitter data, and highlight the importance of noticing the different accumulation velocity of different sources of altmetric data while selecting different time windows in practice.

RQ3. *To what extent do scholarly tweets get engaged with through different engagement behaviors (i.e., liking, retweeting, replying, and quoting)?*

Chapter 4 answers this research question with an extensive analysis of the user engagement metrics of a total set of 7 million original scholarly tweets mentioning scientific papers. Four types of user engagement metrics (i.e., likes, retweets, replies, and quotes) are collected by using the Twitter API to investigate the extent to which scholarly tweets are further engaged with by Twitter users. Chapter 4 also compares the subject field differences of the user engagement around scholarly tweets. Furthermore, to study the relationships between user engagement metrics and other factors related to scholarly tweets, chapter 4 performs both correlation analysis and regression analysis for user engagement metrics and a wide range of science-based indicators (e.g., number of citations and number of Mendeley readers of tweeted papers) and Twitter-based indicators (e.g., number of hashtags, number of mentioned users in tweets, and number of followers and friends of Twitter users). Findings of this chapter shed light on the possibility to apply user engagement metrics in measuring deeper levels of Twitter reception of scientific information.

RQ4. *To what extent and for what reasons do scholarly tweets become unavailable as time goes by? What is the potential effect that the unavailability of tweets may make on the stability of Twitter metrics?*

Chapter 5 answers these research questions through a case study consisting of over 2.6 million scholarly tweets received by the 1,154 most tweeted scientific papers recorded by Altmetric.com up to October 2017. The (un)availability of the tweets are rechecked in April 2019 with the Twitter API. For the unavailable tweets identified during the recheck, the error codes responded by the Twitter API are collected as well to figure out the specific unavailable reasons. In addition to reporting the overall proportion of unavailable tweets (i.e., unavailability rate), chapter 5 also explores what kinds of scientific papers are facing greater risk of unstable Twitter metrics according to their Twitter dissemination structures. To this end, two indicators – *Degree of Originality* (DO) and *Degree of Concentration* (DC) – are proposed to delineate papers' Twitter dissemination structures and examine the potential influence of Twitter dissemination structures on the stability of Twitter metrics. Findings of this chapter confirm the volatile nature of tweets partly caused by some post-interaction behaviors that users conduct to their tweets and accounts (e.g., deleting tweets and protecting accounts). More importantly, findings of this chapter put emphasis on the necessity of paying attention to such nature which may dramatically affect the stability of Twitter metrics.

RQ5. *To what extent do scholarly URLs to scientific papers embedded in scholarly tweets get clicked?*

Chapter 6 answers this research question on the basis of the click metric data provided by Bitly – a link shortening service platform which records how its generated short links are clicked from different sources and on different dates. By analyzing the click metric data of over 1.1 million Bitly short links referring to scientific papers cited in scholarly tweets, chapter 6 provides an insight into how frequently short links embedded in scholarly tweets are clicked by Twitter users to visit the original webpages of scientific papers. Besides, chapter 6 presents the patterns of clicking behavior with regard to scientific information in terms of both occurrence speed and subject field preference. Finally, chapter 6 explores how the number of Twitter clicks of short links correlates with the scholarly attention that the scientific papers attract (i.e., number of citations and number of Mendeley readers) and the Twitter attention that the scholarly tweets attract (i.e., number of likes and number of retweets). Findings of this chapter open up a novel opportunity to assess deeper levels of Twitter reception of scientific information which has direct effect on the consumption of scholarly content.

At last, chapter 7, as the discussion and conclusion part, summarizes and contextualizes the main findings presented in chapters 2 to 6. Based on the research findings, chapter 7 further elaborates on the implications of the findings for improving scholarly Twitter metrics, and puts forward prospects for future research.

CHAPTER 2

An extensive analysis of the presence of altmetric data for Web of Science papers across subject fields and research topics[1]

[1] **This chapter is based on**:

Fang, Z., Costas, R., Tian, W., Wang, X., & Wouters, P. (2020). An extensive analysis of the presence of altmetric data for web of science publications across subject fields and research topics. *Scientometrics*, *124*(3), 2519-2549. https://doi.org/10.1007/s11192-020-03564-9

Author contributions:

Fang, Z. (Conceptualization, Methodology, Formal analysis, Investigation, Visualization, Data Curation, Writing - Original Draft, Writing - Review & Editing)

Costas, R. (Conceptualization, Methodology, Investigation, Supervision, Writing - Review & Editing)

Tian, W. (Visualization, Writing - Review & Editing)

Wang, X. (Conceptualization, Writing - Review & Editing)

Wouters, P. (Conceptualization, Methodology, Supervision, Writing - Review & Editing)

Abstract

Sufficient data presence is one of the key preconditions for applying metrics in practice. Based on both Altmetric.com data and Mendeley data collected up to 2019, this paper presents a state-of-the-art analysis of the presence of 12 kinds of altmetric events for nearly 12.3 million Web of Science papers published between 2012 and 2018. Results show that even though an upward trend of data presence can be observed over time, except for Mendeley readers and Twitter mentions, the overall presence of most altmetric data is still low. The majority of altmetric events go to papers in the fields of *Biomedical and Health Sciences*, *Social Sciences and Humanities*, and *Life and Earth Sciences*. As to research topics, the level of attention received by research topics varies across altmetric data, and specific altmetric data show different preferences for research topics, on the basis of which a framework for identifying *hot* research topics is proposed and applied to detect research topics with higher levels of attention garnered on certain altmetric data source. Twitter mentions and policy document citations were selected as two examples to identify hot research topics of interest of Twitter users and policy-makers, respectively, shedding light on the potential of altmetric data in monitoring research trends of specific social attention.

Keywords

Altmetrics, social media metrics, data coverage, data intensity, hot topics, social attention

2.1 Introduction

Ever since the term "altmetrics" was coined in Jason Priem's tweet in 2010,[1] a range of theoretical and practical investigations have been taking place in this emerging area (Sugimoto, Work, et al., 2017). Given that many types of altmetric data outperform traditional citation counts with regard to the accumulation speed after publication (Fang & Costas, 2020), initially, altmetrics were expected to serve as faster and more fine-grained alternatives to measure scholarly impact of research outputs (Priem et al., 2010; Priem, Groth, et al., 2012). Nevertheless, except for Mendeley readership which was found to be moderately correlated with citations (Zahedi et al., 2014; Zahedi & Haustein, 2018), a series of studies have confirmed the negligible or weak correlations between citations and most altmetric indicators at the paper level (Bornmann, 2015a; Costas et al., 2015a; de Winter, 2015; Zahedi et al., 2014), indicating that altmetrics might capture diverse forms of impact of scholarship which are different from citation impact (Wouters & Costas, 2012).

The diversity of impact beyond science reflected by altmetrics, which is summarized as "broadness" by Bornmann (2014a) as one of the important characteristics of altmetrics, relies on diverse kinds of altmetric data sources. Altmetrics do not only include events on social and mainstream media platforms related to scholarly content or scholars, but also incorporate data sources outside the social and mainstream media ecosystem such as policy documents and peer review platforms (Haustein, Bowman, & Costas, 2016). The expansive landscape of altmetrics and their fundamental differences highlight the importance of keeping them as separate entities without mixing, and selecting datasets carefully when making generalizable claims about altmetrics (Alperin, 2015; Wouters et al., 2019). In this sense, data presence, as one of the significant preconditions for applying metrics in research evaluation, also needs to be analyzed separately for various altmetric data sources.

2.1.1 Presence of altmetric data for scientific papers

Bornmann (2016) regarded altmetrics as one of the hot topics in the field of Scientometrics for several reasons, being one of them that there are large altmetric datasets available to be empirically analyzed for studying the impact of scientific papers. However, according to existing studies, there are important differences of data coverage across diverse altmetric data. In one of the first, Thelwall, Haustein, et al. (2013) conducted a comparison of the correlations between citations and 11 categories of altmetric indicators finding that, except for Twitter mentions, the coverage of all selected altmetric data of PubMed articles was substantially low. This observation was reinforced by other following studies, which provided more evidence about the exact coverage for Web of Science (WoS) papers. Based

[1] On September 29, 2010, Jason Priem posted a tweet with the hashtag "altmetrics". See more details about this tweet at: https://twitter.com/jasonpriem/status/25844968813 (Accessed May 3, 2020).

on altmetric data retrieved from ImpactStory (IS), Zahedi et al. (2014) reported the coverage of four types of altmetric data for a sample of WoS papers: Mendeley readers (62.6%), Twitter mentions (1.6%), Wikipedia citations (1.4%), and Delicious bookmarks (0.3%). In a follow-up study using altmetric data from Altmetric.com, Costas et al. (2015a) studied the coverage of five altmetric data for WoS papers: Twitter mentions (13.3%), Facebook mentions (2.5%), blogs citations (1.9%), Google+ mentions (0.6%), and news mentions (0.5%). They also found that research outputs in the fields of Biomedical and Health Sciences and Social Sciences and Humanities showed the highest altmetric data coverage in terms of these five altmetric data. Similarly, it was reported by Haustein, Costas, et al. (2015) that the coverage of five social and mainstream media data for WoS papers varied as follows: Twitter mentions (21.5%), Facebook mentions (4.7%), blogs citations (1.9%), Google + mentions (0.8%), and news mentions (0.7%).

In addition to the aforementioned large-scale research on WoS papers, there have been also studies focusing on the coverage of altmetric data for research outputs from a certain subject field or publisher. For example, on the basis of the selected journal articles in the field of Humanities, Hammarfelt (2014) investigated the coverage of five kinds of altmetric data, including Mendeley readers (61.3%), Twitter mentions (20.6%), CiteULike readers (5.2%), Facebook mentions (2.9%), and blogs citations (2.2%). Waltman and Costas (2014) found that just about 2% of the biomedical literature received at least one F1000Prime recommendation. For papers published in the Public Library of Science (PLoS) journals, Bornmann (2015b) reported the coverage of a group of altmetric data sources tracked by PLoS's Article-Level Metrics (ALM). Since the data coverage is a value usually computed for most altmetric studies, similar coverage levels are found scattered across many other studies as well (Alperin, 2015; Fenner, 2013; Robinson-Garcia et al., 2014). By summing up the total number of papers and those covered by altmetric data in 25 related studies, Erdt et al. (2016) calculated the aggregated percentage of coverage for 11 altmetric data. Their aggregated results showed that Mendeley readers covered the highest share of papers (59.2%), followed by Twitter mentions (24.3%) and CiteULike readers (10.6%), while other altmetric data showed relatively low coverage in general (below 10%).

2.1.2 Identification of hot research topics using altmetric data

The distributions of publications and article-level metrics across research topics are often uneven, which has been observed through the lens of text-based (Gan & Wang, 2015), citation-based (Shibata et al., 2008), usage-based (X. Wang et al., 2013), and altmetric-based (Noyons, 2019) approaches, making it possible to identify research topics of interest in different contexts, namely, the identification of *hot research topics*. By combining the concept made by Tseng et al. (2009), hot research topics are defined as topics that are of particular interest to certain communities such as researchers, Twitter users, Wikipedia editors, or policy-makers. Thus, *hot* is defined as the description of a relatively high level of attention that research topics have received on different altmetric data sources. *Attention* here is understood as the amount of interactions that different communities have generated around research topics, therefore those topics with high levels of attention can be identified and characterized as hot research topics from an altmetric point of view.

Traditionally, several text-based and citation-based methodologies have been widely developed and employed in detecting research topics of particular interest to researchers, like co-word analysis (Ding & Chen, 2014; W. H. Lee, 2008), direct citation and co-citation analysis (Chen, 2006; Small, 2006; Small et al., 2014), and the "core documents" based on bibliographic coupling (Glänzel & Czerwon, 1996; Glänzel & Thijs, 2012). Besides, usage metrics, which are generated by broader sets of users through various behaviors such as viewing, downloading, or clicking, have been also used to track and identify hot research topics. For example, based on the usage count data provided by Web of Science, X. Wang and Fang (2016) detected hot research topics in the field of Computational Neuroscience, which were listed as the keywords of the most frequently used papers. By monitoring the downloads of papers in *Scientometrics*, X. Wang et al. (2013) identified hot research topics in the field of Scientometrics, operationalized as the most downloaded papers in the field.

From the point of view that altmetrics can capture the attention around scholarly objects from the broader public (Crotty, 2014; Sugimoto, 2015), some altmetric data were also used to characterize research topics based on the interest exhibited by different altmetric and social media users. For example, Robinson-Garcia et al. (2019) studied the field of Microbiology to map research topics which are highly mentioned within news media outlets, policy briefs, and tweets over time. Zahedi and Van Eck (2018) presented an overview of specific topics of interest of different types of Mendeley users, like professors, students, and librarians, and found that they showed different preferences in reading papers from different topics. Fang and Costas (2020) identified research topics of papers that are faster to be mentioned by Twitter users or cited by Wikipedia page editors, respectively. By comparing the term network based on author keywords of climate change research papers, the term network of author keywords of those tweeted papers, and the network of "hashtags" attached to related tweets, Haunschild et al. (2019) concluded that Twitter users were more interested in topics about the consequences of climate change to humans, especially those papers forecasting effects of a changing climate on the environment.

2.1.3 Objectives

Although there are multiple previous studies discussing the coverage of different altmetric data, after nearly 10 years of altmetric research, we find that a renewed large-scale empirical analysis of the up-to-date presence of altmetric data for WoS papers is highly relevant. Particularly, since amongst previous studies, there still exist several types of altmetric data sources that have not been quantitatively analyzed. Moreover, although the correlations between citations and altmetric indicators have been widely analyzed at the paper level in the past, the correlations of their presence at the research topic level are still unknown. To fill these research gaps, this paper presents a renovated analysis of the presence of various altmetric data for scientific papers, together with a more focused discussion about the presence of altmetric data across broad subject fields and smaller research topics.

The main objective of this study is two-fold: (1) to reveal the development and current situation of the presence of altmetric data across papers and subject fields, and (2) to explore the potential application of altmetric data in identifying and tracking research trends that are of interest to certain communities such as Twitter users and policy-makers. The following specific research questions are put forward:

RQ1. Compared to previous studies, how the presence of different altmetric data for WoS papers has developed until now? What is the difference of altmetric data presence across WoS papers published in different years?

RQ2. How is the presence of different altmetric data across subject fields of science? For each type of altmetric data, which subject fields show higher levels of data prevalence?

RQ3. How are the relationships among various altmetric and citation data in covering different research topics? Based on specific altmetric data, in each subject field which research topics received higher levels of altmetric attention?

2.2 Data and methods

2.2.1 Dataset

A total of 12,271,991 WoS papers published between 2012 and 2018 were retrieved from the CWTS in-house database. Since identifiers are necessary for matching papers with their altmetric data, only papers with a Digital Object Identifier (DOI) or a PubMed Identifier (PubMed ID) recorded in WoS were considered.

Using the two identifiers, WoS papers were matched with 12 types of altmetric data from Altmetric.com and Mendeley readership as listed in Table 1. The data from Altmetric.com were extracted from a research snapshot file with data collected up to October 2019. Mendeley readership data were separately collected through the Mendeley API in July 2019.[1] Altmetric.com provides two counting methods of altmetric performance for papers, including

[1] This is to avoid the limitation in the Mendeley data reported by Altmetric.com, which is restricted to only papers with other metrics in Altmetric.com (Haustein, Costas, et al., 2015).

the number of each altmetric event that mentioned the paper and the number of unique users who mentioned the paper. To keep a parallelism with Mendeley readership, which is counted at the user level, the number of unique users was selected as the indicator for counting altmetric events in this study. For the selected papers, the total number of events they accumulated on each altmetric data source are provided in Table 1 as well.

Table 1. Descriptive statistics of 12 types of altmetric data analyzed in this study[1]

Data source	Concept measured with regard to research outputs	NP	NE
Mendeley	Mendeley readers with the output in their Library.	10,959,393	293,922,534
Twitter	Twitter mentions, including original tweets, reply tweets, quote tweets, and retweets.	4,173,353	36,092,805
Facebook	Facebook mentions, including posts on a curated list of public pages only.	1,052,235	2,388,875
News	News media mentions on a list of news sources tracked by Altmetric.com, which contains over 5,000 English and non-English global news outlets.	491,855	2,803,824
Blogs	Blog citations on a list of blogs tracked by Altmetric.com, which contains over 15,000 academic and non-academic blogs.	448,663	767,381
Wikipedia	Wikipedia citations on English Wikipedia pages only.	165,170	239,686
Policy documents	Policy document citations on a wide range of public policy documents tracked by Altmetric.com, including policy, guidance, or guidelines documents from a governmental or non-governmental organization.	137,326	156,813
Reddit	Reddit mentions on all sub-reddits, including original posts only.	69,356	90,758
F1000Prime	F1000Prime recommendations.	69,180	69,197
Video	Video mentions on YouTube.	48,561	71,191
Peer review	Post-publication peer review comments collected from two forums: PubPeer and Publons.	32,154	32,217
Q&A	Q&A mentions on Stack Overflow.	7,005	8,021

Note: NP refers to the number of papers with corresponding altmetric data, NE refers to the total number of corresponding altmetric events. As of October 2019, Altmetric.com has stopped collecting data from CiteULike, Sina Weibo, LinkedIn, Pinterest, and Google+. Syllabus data only posted in 2015 were provided by Altmetric.com and almost all publications mentioned by Syllabus are not indexed by Web of Science. Therefore, these data sources have not been included in this study.

Besides, we collected the WoS citation counts in October 2019 for the selected papers. Citations serves as a benchmark for a better discussion and understanding of the presence

[1] See more information about the different data sources tracked by Altmetric.com at: https://help.altmetric.com/support/solutions/articles/6000060968-what-outputs-and-sources-does-altmetric-track- (Accessed February 26, 2020).

and distribution of altmetric data. To keep the consistency with altmetric data, a variable citation time window from the year of publication to 2019 was utilized and self-citations were considered for our dataset of papers.

2.2.2 CWTS publication-level classification system

To study subject fields and research topics, we employed the CWTS classification system (also knowns as the *Leiden Ranking classification*). Waltman and Van Eck (2012) developed this publication-level classification system mainly for citable WoS publications (Article, Review, Letter) based on their citation relations. In its 2019 version, papers are clustered into 4535 micro-level fields of science with similar research topics (here and after known as *micro-topics*) as shown in Figure 1 with VOSviewer. For each micro-topic, the top five most characteristic terms are extracted from the titles of the papers in order to label the different micro-topics. Furthermore, these micro-topics are assigned to five main subject fields of science algorithmically obtained, including *Social Sciences and Humanities* (SSH), *Biomedical and Health Sciences* (BHS), *Physical Sciences and Engineering* (PES), *Life and Earth Sciences* (LES), and *Mathematics and Computer Science* (MCS).[1] The CWTS classification system has been applied not only in the Leiden Ranking, but also in many different previous studies related with subject field analyses (Costas et al., 2015a; Didegah & Thelwall, 2018; Zahedi & Van Eck, 2018).

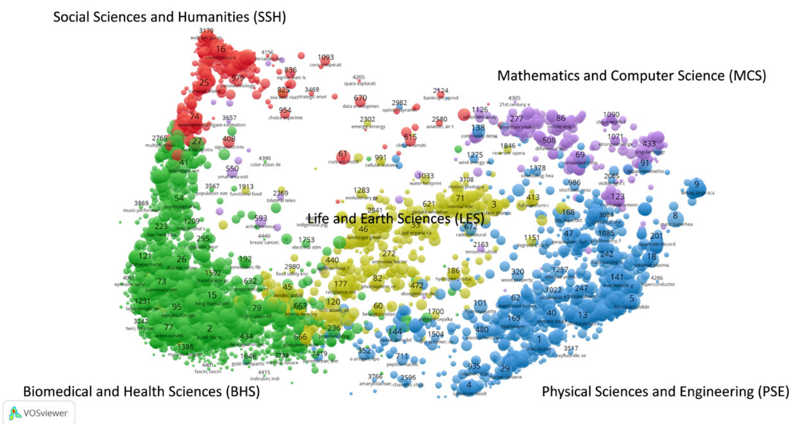

Figure 1. Five main subject fields of science of the CWTS classification system. Each circle represents a micro-level field (micro-topics) of clustered papers based on direct citation relations

[1] See more details about CWTS classification system at: https://www.leidenranking.com/information/fields (Accessed May 3, 2020).

A total of 10,615,881 of the initially selected papers (accounting for 86.5%) have CWTS classification information. This set of papers was drawn as a subset for the comparison of altmetric data presence across subject fields and research topics. Table 2 presents the number of selected papers in each main subject field.

Table 2. Number of papers in each subject field

Subject field	Abbr.	Number of papers	Percentage
Social Sciences and Humanities	SSH	910,011	8.57%
Biomedical and Health Sciences	BHS	4,272,079	40.24%
Physical Sciences and Engineering	PSE	3,075,125	28.97%
Life and Earth Sciences	LES	1,555,443	14.65%
Mathematics and Computer Science	MCS	803,223	7.57%

2.2.3 Indicators and analytical approaches

In order to measure the presence of different kinds of altmetric data or citation data across different sets of papers, we employed the three indicators proposed by Haustein, Costas, et al. (2015): *Coverage*, *Density*, and *Intensity*. For a specific set of papers, these three indicators are defined and calculated as follows:

- Coverage (C) indicates the percentage of papers with at least one altmetric event (or one citation) recorded in the set of papers. Therefore, the value of coverage ranges from 0 to 100%. The higher the coverage, the higher the share of papers with altmetric event data (or citation counts).

- Density (D) is the average number of altmetric events (or citations) of the set of papers. Both papers with altmetric events (or citations) and those without any altmetric events (or citations) are considered in the calculation of density, so it is heavily influenced by the coverage and zero values.[1] The higher the value of density, the more altmetric events (or citations) received by the set of papers on average.

- Intensity (I) is defined as the average number of altmetric events (or citations) of papers with at least one altmetric event (or citation) recorded. Different from D, the calculation of I only takes papers with non-zero values in each altmetric event (or citation event) into consideration, so the value must be higher or equal to one. Only in those cases of groups of papers without any altmetric events (or citations), the

[1] Papers without altmetric events or citations are assumed to have zero values.

intensity is set to zero by default. The higher the value of intensity, the more altmetric events (or citations) that have occurred around the papers with altmetric/citation data on average.

In order to reveal the relationships among these three indicators at the research topic level, as well as the relationships of preferences for research topics among different data, the Spearman correlation analysis was performed with IBM SPSS Statistics 25.

2.3 Results

This section consists of four parts: the first one presents the overall presence of altmetric data for the whole set of WoS papers (in contrast with previous studies) and the evolution of altmetric data presence over the publication years. The second part compares the altmetric data presence of papers across five main subject fields of science. The third part focuses on the differences of preferences of altmetric data for research topics. In the fourth part, Twitter mentions and policy document citations are selected as two examples for identifying hot research topics with higher levels of altmetric attention received.

2.3.1 Overall presence of altmetric data over the publication years

Coverage, density, and intensity of the 12 sources of altmetric data and citations were calculated for the nearly 12.3 million sample WoS papers to reveal their overall presence. Table 3 presents not only the results based on our dataset, but also, for comparability purposes, the findings of data coverage (C_ref) reported by some previous altmetric empirical studies that also used Altmetric.com (and Mendeley API for Mendeley readership) as the altmetric data source, and WoS as the database for scientific papers; and also without applying restrictions of certain discipline, country, or publisher. As these previous studies analyzed datasets with size, publication years (PY), and data collection years (DY) different from ours, we present them as references for discussing the retrospective historical development of altmetric data prevalence.

Table 3. The overall presence of 12 types of altmetric data and citation data

Data	C	D	I	Reference	PY	DY	C_ref
Mendeley readers	89.30%	23.951	26.819	Haustein, Larivière, et al. (2014)	2010-2012	-	66.20%
				Mohammadi et al. (2015)	2008	-	45.60%
				Bornmann and Haunschild (2017)	2014	Jul. 2016	89.27%
				D'Angelo and Di Russo (2019)	2009-2016	Feb. 2018	96.10%
Twitter mentions	34.01%	2.941	8.648	Robinson-García et al. (2014)	2011-2013	Jan. 2014	16.10%
				Haustein, Larivière, et al. (2014)	2010-2012	Dec. 2012	9.40%
				Costas et al. (2015a)	Jul.-Dec. 2011	Oct. 2013	13.30%
				Haustein, Costas, et al. (2015)	2012	Oct. 2013	21.50%
				Meschede and Siebenlist (2018)	2015	-	35.78%
Facebook mentions	8.57%	0.195	2.270	Robinson-García et al. (2014)	2011-2013	Jan. 2014	3.70%
				Costas et al. (2015a)	Jul.-Dec. 2011	Oct. 2013	2.50%
				Haustein, Costas, et al. (2015)	2012	Oct. 2013	4.70%
				Meschede and Siebenlist (2018)	2015	-	8.46%
News mentions	4.01%	0.229	5.701	Costas et al. (2015a)	Jul.-Dec. 2011	Oct. 2013	0.50%
				Haustein, Costas, et al. (2015)	2012	Oct. 2013	0.70%
				Meschede and Siebenlist (2018)	2015	-	4.42%
Blog citations	3.66%	0.063	1.710	Robinson-García et al. (2014)	2011-2013	Jan. 2014	1.80%
				Costas et al. (2015a)	Jul.-Dec. 2011	Oct. 2013	1.90%
				Haustein, Costas, et al. (2015)	2012	Oct. 2013	1.90%
				Meschede and Siebenlist (2018)	2015	-	2.56%
Wikipedia citations	1.35%	0.020	1.451	Meschede and Siebenlist (2018)	2015	-	0.70%
Policy document citations	1.12%	0.013	1.142	Haunschild and Bornmann (2017)	2000-2014	Dec. 2015	0.32%
Reddit mentions	0.57%	0.007	1.309	Meschede and Siebenlist (2018)	2015	-	1.16%
F1000Prime recommendations	0.56%	0.006	1.000	-	-	-	-
Video mentions	0.40%	0.006	1.466	-	-	-	-
Peer review comments	0.26%	0.003	1.002	-	-	-	-
Q&A mentions	0.06%	0.001	1.145	-	-	-	-
WoS citations	77.43%	9.681	12.502	-	-	-	-

2

According to the results, the presence of different altmetric data varies greatly. Mendeley readership provides the largest values of coverage (89.30%), density (23.95), and intensity (26.82), even higher than citations. As to other altmetric data, their presence is much lower than Mendeley readers and citations. Twitter mentions holds the second largest values among all other altmetric data, with 34.01% of papers mentioned by Twitter users and those mentioned papers accrued about 8.65 Twitter mentions on average. It is followed by several social and mainstream media data, like Facebook mentions, news mentions, and blogs citations. About 8.57% of papers have been mentioned by Facebook, 4.01% have been mentioned by news outlets, and 3.66% have been cited by blog posts. But among these three data sources, papers mentioned by news outlets accumulated more intensive attention in consideration of its higher value of intensity (5.70), which means that mentioned papers got more news mentions on average. In contrast, even though there are more papers mentioned by Facebook, they received fewer mentions at the individual paper level (with the intensity value of 2.27). For the remaining altmetric data, their data coverage values are extremely low. Wikipedia citations and policy document citations only covered 1.35% and 1.12% of the sample papers, respectively, while the coverage of Reddit mentions, F1000Prime recommendations, video mentions, peer review comments, and Q&A mentions are lower than 1%. In terms of these data, the altmetric data of papers are seriously zero-inflated.

Compared to the coverage reported by previous studies, an increasing trend of altmetric data presence can be observed as time goes by. Mendeley, Twitter, Facebook, news, and blogs are the most studied altmetric data sources. On the whole, the more recent the studies, the higher the values of coverage they report. Our results show one of the highest data presence for most altmetric data. Although the coverage of Twitter mentions, news mentions, and Reddit mentions reported by Meschede and Siebenlist (2018) is slightly higher than ours, it should be noted that they used a random sample consisting of 5000 WoS papers published in 2015, and as shown in Figure 2, there exist biases toward publication years when investigating data presence for altmetrics.

After calculating the three indicators for research outputs in each publication year, Figure 2 shows the change trends of the presence of altmetric data. Overall there are two types of tendencies for all altmetric data, which are in correspondence with the accumulation velocity patterns identified in the research conducted by Fang and Costas (2020). Thus, for altmetric data with higher speed in data accumulating, such as Twitter mentions, Facebook mentions, news mentions, blog citations, and Reddit mentions, newly published papers have higher coverage levels. In contrast, those altmetric data taking a longer time to accumulate (i.e., the slow sources defined by Fang and Costas (2020)), they tend to accumulate more prominently for older papers. Wikipedia citations, policy document citations, F1000Prime recommendations, video mentions, peer review comments, and Q&A mentions fall into this “slower” category. As a matter of fact, their temporal distribution patterns resemble more that of citations counts. Regarding Mendeley readers, although it keeps quite high coverage in every publication year, it shows a downward trend as citations too, indicating a kind of readership delay, by which newly published papers have to take time to accumulate Mendeley readers (Haustein, Larivière, et al., 2014; Thelwall, 2017; Zahedi et al., 2017).

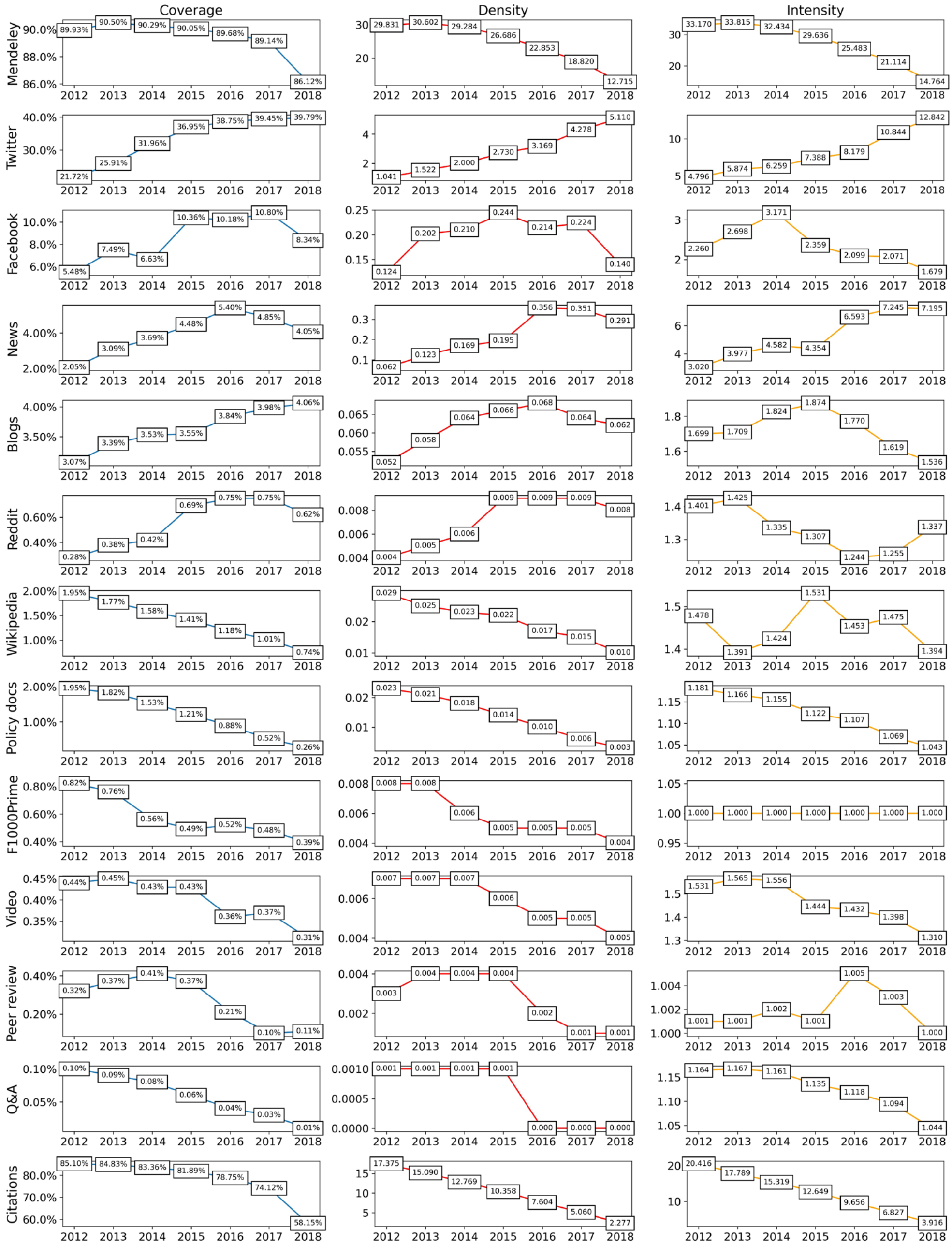

Figure 2. The presence of altmetric data and citations over the publication years

2.3.2 Presence of altmetric data across subject fields

In general, papers in the fields of natural sciences and medical and health sciences received more citations (Marx & Bornmann, 2015), but for altmetric data, the distribution across subject fields shows another picture. As shown in Figure 3, on the basis of our dataset, it is confirmed that papers in the subject fields of BHS, PSE, and LES hold the highest presence of citation data, and papers in the fields of SSH and MCS accumulated obviously fewer citation counts. However, as observed by Costas et al. (2015a) for Twitter mentions, Facebook mentions, news mentions, blog citations, and Google+ mentions, most altmetric data in Figure 3 are more likely to concentrate on papers from the fields of BHS, SSH, and LES, while PSE papers lose the advantage of attracting attention as they show in terms of citations, thereby performing weakly in altmetric data presence as MCS papers do.

Amongst altmetric data, there are some showing special patterns of presence. For example, PSE papers reach the coverage of Mendeley readers as high as papers in BHS, SSH, and LES, but from the perspectives of density and intensity, PSE papers drop down, showing the lowest values of density and intensity of Mendeley readers only second to MCS papers. Since F1000Prime (now Faculty Opinions https://facultyopinions.com) is a platform mainly focusing on the research outputs in the fields of life sciences and medical sciences, BHS papers show a considerably higher presence of F1000Prime recommendations over other subject fields. In terms of peer review comments, SSH papers hold a higher coverage level. This result differs from what has been observed in Ortega (2019a)'s study on the coverage of Publons data, in which Publons data were found to be biased to papers in life sciences and health sciences. It should be noted that the peer review comment data provided by Altmetric.com is an aggregation of two platforms: Publons (https://publons.com) and PubPeer (https://pubpeer.com). In our dataset, there are 31,132 distinct papers with altmetric peer review data for the analysis of data presence across subject fields, 8,337 of them (accounting for 26.8%) having peer review comments from Publons and 22,851 of them (accounting for 73.4%) having peer review comments from PubPeer (56 papers have been commented by both). If we only consider the papers with Publons data, BHS papers and LES papers contribute the most (accounting for 53.4% and 17.2%, respectively), which is in line with Ortega (2019a)'s results about Publons on the whole. Nevertheless, PubPeer data, which covers more papers recorded by Altmetric.com, is biased towards SSH papers. SSH papers make up as high as 49.9% of all papers with PubPeer data, followed by BHS papers (accounting for 43.4%), besides the relatively small quantity of WoS papers in the field of SSH, thereby leading to the overall high coverage of peer review comments of SSH papers.

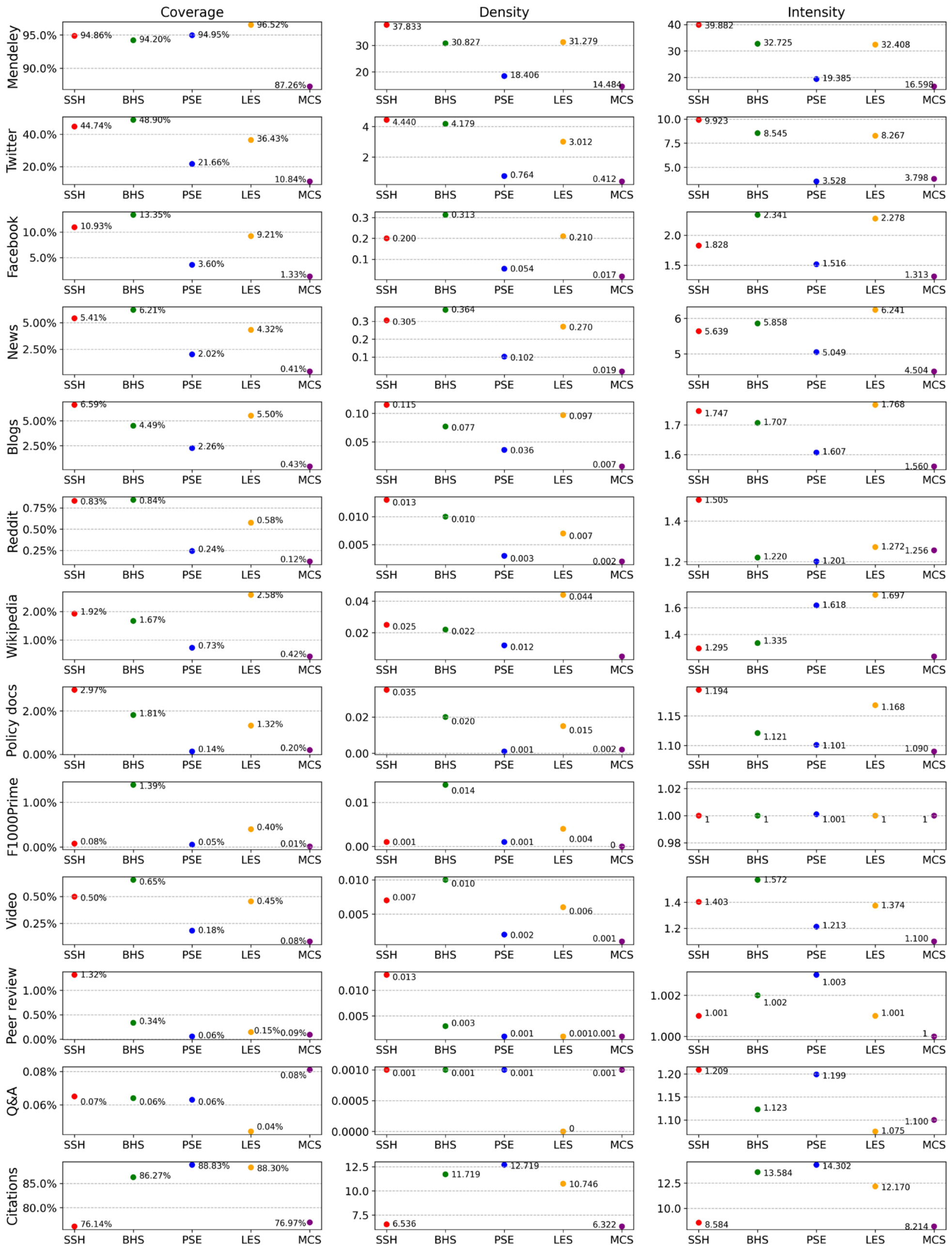

Figure 3. The presence of altmetric data and citations of scientific papers across five subject fields

Moreover, given the fact that the distributions of altmetric data are highly skewed, with the majority of papers only receiving very few altmetric events (see Figure 8 in Appendix), particularly for altmetric data with relatively small data volume, their density and intensity are very close across subject fields. But in terms of intensity, there exist some remarkable subject field differences for some altmetric data. For example, on Reddit, SSH papers received more intensive attention than other subject fields in consideration of their higher value of intensity. By comparison, those LES and PSE papers cited by Wikipedia pages accumulated more intensive attention, even though the coverage of Wikipedia citations of PSE papers is rather low, suggesting that although PSE papers have a lower coverage in Wikipedia, they are more repeatedly cited.

2.3.3 Presence of altmetric data across research topics

Due to the influence of highly skewed distribution of altmetric data (see Figure 8 in Appendix) on the calculation of coverage and density, these two indicators at the micro-topic level are strongly correlated for all kinds of altmetric data (see Figure 9 in Appendix). In comparison, the correlation between coverage and intensity is rather weaker. Moreover, in an explicit way, coverage tells how many papers around a micro-topic have been mentioned or cited at least once, and intensity describes how frequently those papers with altmetric data or citation data have been mentioned or cited. Consequently, for a specific micro-topic, these two indicators can reflect the degree of broadness (coverage) and degree of deepness (intensity) of its received attention. Therefore, we employed coverage and intensity to investigate the presence of altmetric data at the micro-topic level and identify research topics with higher levels of attention received on different data sources.

Coverage and intensity values were calculated and appended to micro-topics based on different types of altmetric and citation data, then the Spearman correlation analyses were performed at the micro-topic level between each pair of data respectively. Figure 4 illustrates the Spearman correlations of coverage amongst citations and 12 types of altmetric data at the micro-topic level, as well as those of intensity. The higher the correlation coefficient, the more similar the presence patterns across micro-topics between two types of data. Discrepancies in the correlations can be understood as differences in the relevance of every pair of data for micro-topics, therefore some pairs of data with stronger correlations may have a more similar preference for the same micro-topics, while those with relatively weaker correlations focus on more dissimilar micro-topics. Through the lens of data coverage, Mendeley readers is the only altmetric indicator that is moderately correlated with citations at the micro-topic level, being in line with the previous conclusions about the moderate correlation between Mendeley readership counts and citations at the publication level (Zahedi et al., 2014). In contrast, because of the different distribution patterns between citations and most altmetric data across subject fields we found in Figure 3, it is not surprising that the correlations of coverage between citations and other altmetric data are relatively weak, suggesting that most altmetric data cover research topics different than citations. Among altmetric data, Twitter mentions, Facebook mentions, news mentions, and blog citations are strongly correlated with each other, indicating that these social media data cover similar research topics. Most remaining altmetric data also present moderate correlations with the above social media data, however, Q&A mentions, as the only altmetric data showing the highest coverage of papers in the field of MCS, is weakly correlated with other altmetric data at the micro-topic level.

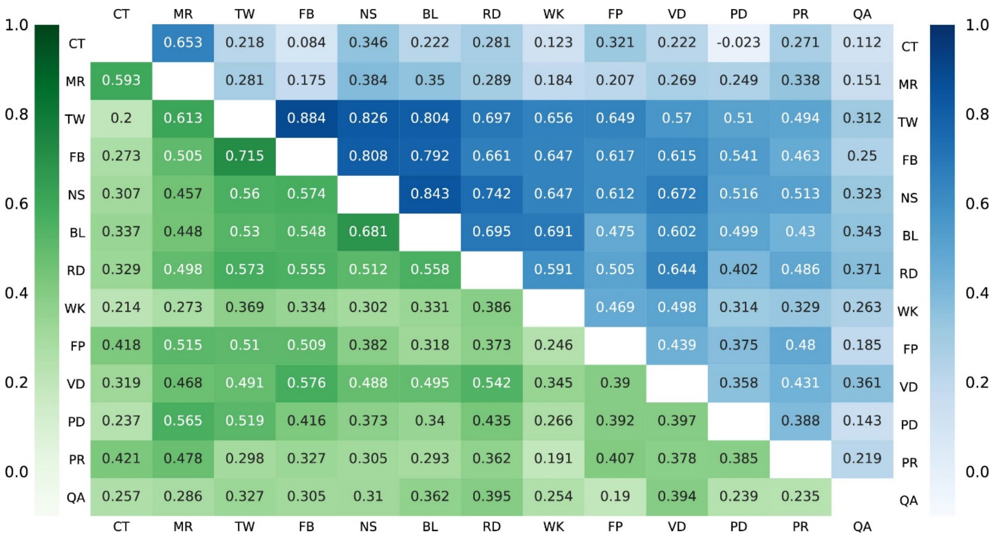

Figure 4. Spearman correlation analyses of coverage (upper-right triangle) and intensity (bottom-left triangle) among citations and 12 types of altmetric data at the micro-topic level. WoS citations (CT), Mendeley readers (MR), Twitter mentions (TW), Facebook mentions (FB), news mentions (NS), blog citations (BL), Reddit mentions (RD), Wikipedia citations (WK), F1000Prime recommendations (FP), video mentions (VD), policy document citations (PD), peer review comments (PR), Q&A mentions (QA)

Nevertheless, from the perspective of intensity, most altmetric data show different attention levels towards research topics, because the values of intensity of different data are generally weakly or moderately correlated. Twitter mentions and Facebook mentions, news mentions and blog citations, are the two pairs of altmetric data showing the strongest correlations from both coverage and intensity perspectives, thus supporting the idea that these two pairs of altmetric data do not only respectively cover very similar research topics, but also focus on similar research topics.

There exists a certain share of micro-topics in which their papers have not been mentioned at all by some specific altmetric data. In order to test the effect of those mutual zero-value micro-topics between each pair of data, the correlations have been performed also excluding them (see Figure 10 in Appendix). It is observed that particularly for those pairs of altmetric data with low overall data presence across papers (e.g., Q&A mentions and peer review comments, Q&A mentions and policy document citations), their correlation coefficients are even lower when mutual zero-value micro-topics are excluded, although the overall correlation patterns across different data types at the micro-topic level are consistent with what we observed in Figure 4.

2.3.4 Identification of hot research topics with altmetric data

On the basis of coverage and intensity, it is possible to compare the altmetric data presence across research topics and to further identify topics that received higher levels of attention. As shown in Figure 5, groups of papers with similar research topics (micro-topics) can be classified into four categories according to the levels of coverage and intensity of attention received. In this framework, *hot research topics* are those topics with a high coverage level of their papers, and at the same time they have also accumulated a relatively high intensive average attention (i.e., their papers exhibit high coverage and high intensity values). Differently, those research topics in which only few papers have received relatively high intensive attention can be regarded as *star-papers topics* (i.e., low coverage and high intensity values), since the attention they attracted has not expanded to a large number of papers within the same research topic. Thus, in star-papers topics the attention is mostly concentrated around a relatively reduced set of papers, namely, those star-papers with lots of attention accrued, while most of the other papers in the same research topic do not receive attention. Following this line of reasoning, there are also research topics with a relatively large share of papers covered by a specific altmetric data, but those covered papers do not show a high average intensity of attention (i.e., high coverage and low intensity values), these research topics are defined as *popular research topics* with mile-wide and inch-deep attention accrued. Finally, *unpopular research topics* indicate those topics with few papers covered by a specific altmetric data source, and the average of data accumulated by the covered papers is also relatively small (i.e., low coverage and low intensity values); these research topics have not attracted too much attention, thereby arguably remaining in an altmetric unpopular status. It should be noted that as time goes on and with newly altmetric activity generated, the status of a research topic might switch across the above four categories.

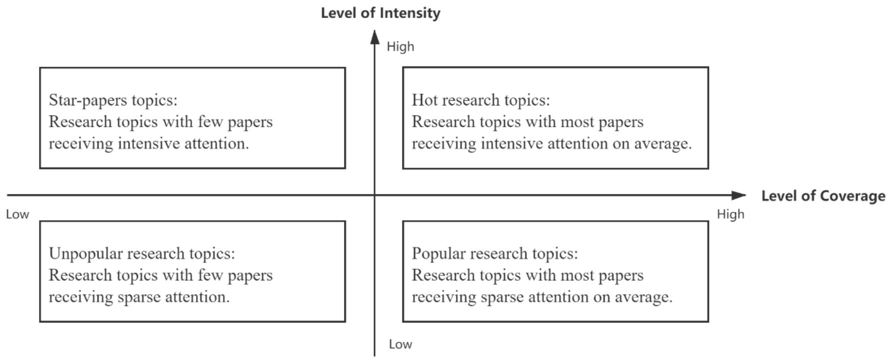

Figure 5. Two-dimensional system for classifying research topics with different levels of attention

Following the framework proposed in Figure 5, we took Twitter mention data as an example to empirically identify hot research topics in different subject fields. A total of 4531 micro-topics with at least one Twitter mention in Figure 1 were plotted into a two-dimensional system according to the levels of coverage and intensity they achieved (Figure 6**A**). Micro-topics are ranked based on their coverage and intensity at first, respectively. The higher the ranking a micro-topic achieves, the higher the level of its coverage or intensity. Size of micro-topics is determined by their total number of papers. In order to identify representative hot research topics on Twitter, here we selected the top 10% as the criterion for both levels of coverage and intensity (two dashed lines in Figure 6**A**) to partition micro-topics into four parts, which are in correspondence with Figure 5. As a result, micro-topics with higher levels of coverage and intensity are classified as hot research topics that received broader and more intensive attention from Twitter users (locate at the upper right corner of Figure 6**A**). Because papers in the fields of SSH, BHS, and LES have much higher coverage and intensity of Twitter data, micro-topics from these three subject fields are more likely to distribute at the upper right part. In contrast, micro-topics in PSE and MCS concentrate at the lower left part. In consideration of the biased presence of Twitter data across five main subject fields, we plotted micro-topics in each subject field by the same method as Figure 6**A**, respectively, and then zoomed in and only presented the part of hot research topics for each subject field in Figure 6**B**-**F** to show their identified hot research topics on Twitter. For clear visualization, one of the extracted terms by CWTS classification system was used as the label for each micro-topic.

In the field of SSH, there are 488 micro-topics considered, and 23 (5%) of them rank in top 10% from both coverage and intensity perspectives (Figure 6**B**). In this subject field, hot research topics tend to be about social issues, including topics related to gender and sex (e.g., "sexual orientation", "gender role conflict", "sexual harassment"), education (e.g., "teacher quality", "education", "undergraduate research experience"), climate ("global warming"), as well as psychological problems (e.g., "stereotype threat", "internet addiction", "stress reduction").

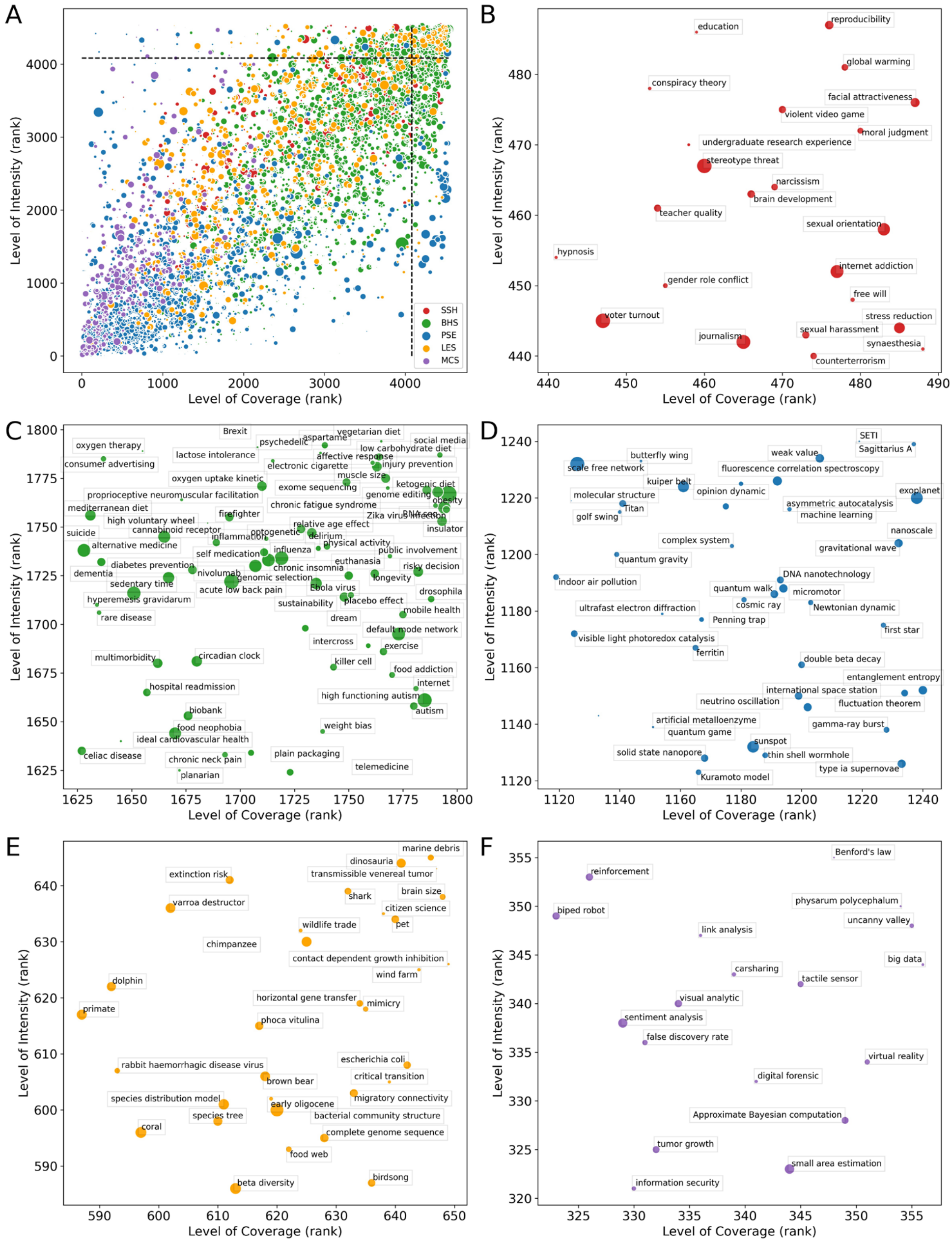

Figure 6. **A** The distribution of micro-topics with different levels of attention received on Twitter; and hot research topics mentioned on Twitter in **B** SSH; **C** BHS; **D** PSE; **E** LES; **F** MCS

BHS is the biggest field with both the most research outputs and the most Twitter mentions, so there are 1796 micro-topics considered, and 75 (4%) of them were detected as hot research topics in Figure 6**C**. Research topics about daily health keeping (e.g., "injury prevention", "low carbohydrate diet", "longevity"), worldwide infectious diseases (e.g., "Zika virus infection", "Ebola virus", "influenza"), lifestyle diseases (e.g., "obesity", "chronic neck pain"), and emerging biomedical technologies (e.g., "genome editing", "telemedicine", "mobile health") received more attention on Twitter. Moreover, problems and revolutions in the medical system caused by some social activities such as "Brexit" and "public involvement" are also brought into focus.

In the field of PSE, 42 (3%) out of 1241 micro-topics were identified as hot research topics in Figure 6**D**. As a field with less Twitter mentions accumulated, although most research topics are left out by Twitter users, those about the universe and astronomy (e.g., "gravitational wave", "exoplanet", "sunspot") and quantum (e.g., "quantum walk", "quantum game", "quantum gravity") received relatively higher levels of attention. In addition, there are also some hot research topics standing out from complexity sciences, such as "scale free network", "complex system", and "fluctuation theorem".

In the field of LES, there are 650 micro-topics in total, and Figure 6**E** shows 32 (5%) hot research topics in this field. These hot research topics are mainly about animals (e.g., "dinosauria", "shark", "dolphin") and natural environment problems (e.g., "extinction risk", "wildlife trade", "marine debris").

Finally, as the smallest subject field, MCS has 18 (5%) out of 356 micro-topics identified as hot research topics (Figure 6**F**), which are mainly about emerging information technologies (e.g., "big data", "virtual reality", "carsharing") and robotics (e.g., "biped robot", "uncanny valley").

To reflect the differences of hot research topics through the lens of different altmetric data sources, policy document citation data was selected as another example. Figure 7 shows the overall distribution of 3134 micro-topics with at least one policy document citation and the identified hot research topics in the five main subject fields. The methodology of visualization is same as Figure 6 based on Twitter data. However, due to the smaller data volume of policy document citations, there are 1868 micro-topics sharing the same intensity of 1. In this case, total number of policy document citations of each micro-topic was introduced as a benchmark to make distinctions. For micro-topics with the same intensity, the higher the total number of policy document citations accrued, the higher the level of attention in the dimension of intensity. After this, if micro-topics still share the same ranking, they are tied for the same place with the next equivalent rankings skipped. In general, these paralleling rankings of micro-topics with relatively low level of attention do not affect the identification of hot research topics.

2

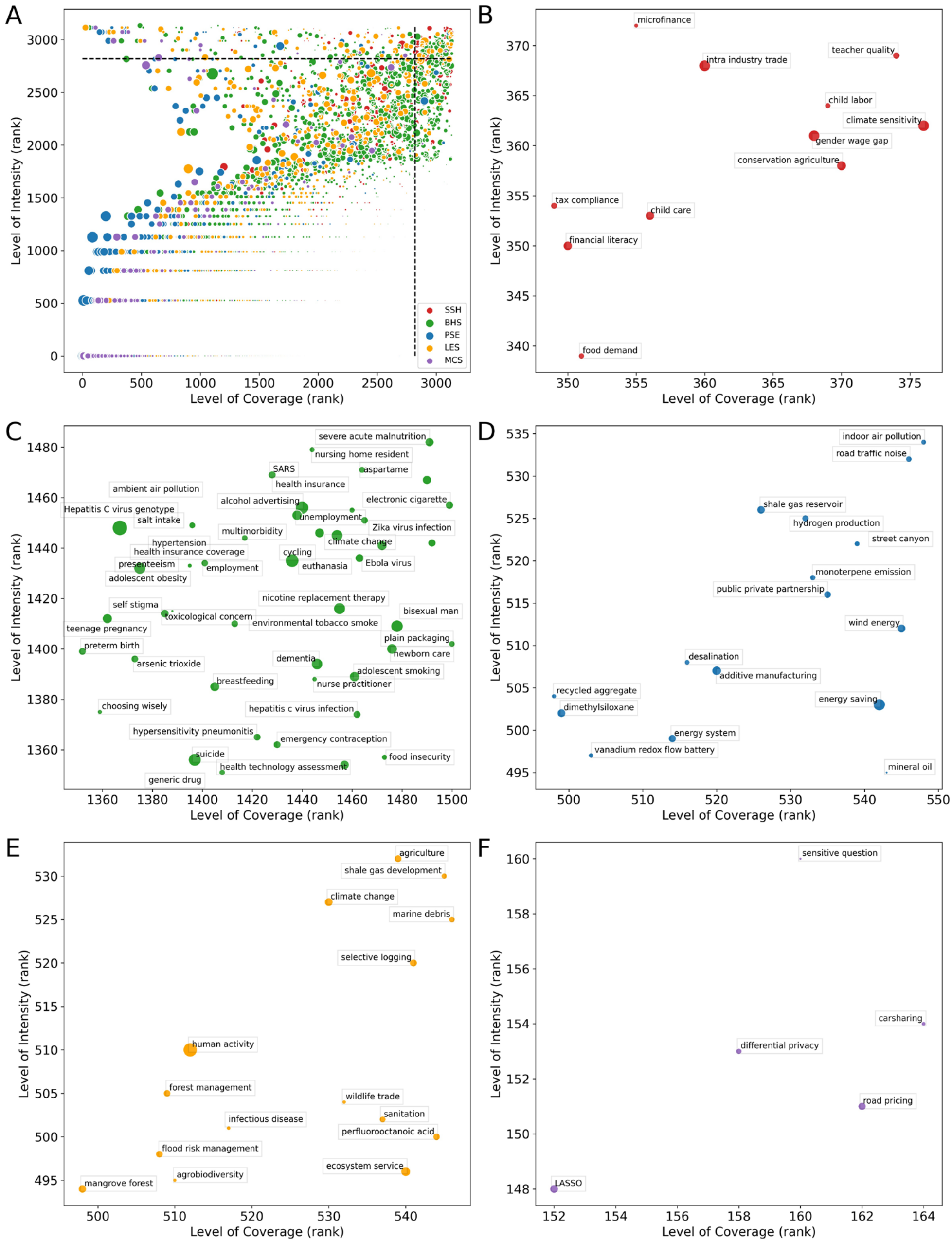

Figure 7. **A** The distribution of micro-topics with different levels of attention received in policy documents; and hot research topics cited by policy documents in **B** SSH; **C** BHS; **D** PSE; **E** LES; **F** MCS

Through the lens of policy document citations, identified hot research topics differ from those in the eyes of Twitter uses to some extents. In the field of SSH, 11 (3%) out of 376 micro-topics were classified as hot research topics (Figure 7**B**). These research topics mainly focus on industry and finance (e.g., “microfinance”, “tax compliance”, “intra industry trade”), as well as child and education (e.g., “child care”, “child labor”, “teacher quality”). Besides, “gender wage gap” is also a remarkable research topic appeared in policy documents.

In the field of BHS, there are 1500 micro-topics have been cited by policy documents at least once, and 44 (3%) of them were classified as hot research topics (Figure 7**C**). Worldwide infectious diseases are typically concerned by policy-makers, consequently, there is no doubt that they were identified as hot research topics, such as “SARS”, “Ebola virus”, “Zika virus infection”, and “Hepatitis C virus genotype”. In addition, healthcare (e.g., “health insurance”, “nursing home resident”, “newborn care”), social issues (e.g., “suicide”, “teenage pregnancy”, “food insecurity”, “adolescent smoking”), and potential health-threatening environment problems (e.g., “ambient air pollution”, “environmental tobacco smoke”, “climate change”) drew high levels of attention from policy-makers too.

Different from the focus of attention on astronomy of Twitter users, in the field of PSE (Figure 7**D**), the 16 (3%) hot research topics out of 548 micro-topics that concerned by policy-makers are mainly around energy and resources, like “energy saving”, “wind energy”, “hydrogen production”, “shale gas reservoir”, “mineral oil”, and “recycled aggregate”.

In the field of LES, Figure 7**E** shows the 15 (3%) hot research topics identified out from 546 micro-topics. From the perspective of policy documents, environmental protection (e.g., “marine debris”, “forest management”, “sanitation”) and sustainable development (e.g., “selective logging”, “human activity”, “agrobiodiversity”) are hot research topics.

At last, in the field of MCS (Figure 7**F**), publications are hardly cited by policy documents, thus there are only 5 (3%) topics out of 164 micro-topics identified as hot research topics. In this field, policy-makers paid more attention to information security (“differential privacy”, “sensitive question”) and traffic economy (“road pricing”, “carsharing”).

2.4 Discussion

2.4.1 Increasing presence of altmetric data

Data presence is essential for the application of altmetrics in research evaluation and other potential areas. The heterogeneity of altmetrics makes it difficult to establish a common conceptual framework and to draw a unified conclusion (Haustein, 2016), thus in most cases it is necessary to separate altmetrics to look into their own performance. This paper investigated 12 types of altmetric data respectively based on a large-scale and up-to-date dataset, results show that various altmetric data vary a lot in the presence for WoS papers.

Data presence of several altmetric data has been widely discussed and explored in previous studies. There are also some reviews summarizing the previous observations of the coverage of altmetric data (Erdt et al., 2016; Ortega, 2020). Generally speaking, our results confirmed the overall situations of the data presence in those studies. For instance, Mendeley readership keeps showing a very high data coverage across scientific papers and provides the most metrics among all altmetric data, followed by Twitter mentions and Facebook mentions. However, there exist huge gaps among these altmetric data. Regarding the data coverage, 89.3% of sample papers have attracted at least one Mendeley reader, while for Twitter mentions and Facebook mentions, the value is 34.0% and 8.6%, respectively. Moreover, for those altmetric data which are hardly surveyed with the same dataset of WoS papers before, like Reddit mentions, F1000Prime recommendations, video mentions, peer review comments, and Q&A mentions, their data coverage is substantially lower than 1%, showing an extremely weak data presence across research outputs.

Comparing with previous observations of altmetric data coverage reported in earlier altmetric studies, it can be concluded that the presence of altmetric data is clearly increasing, and our results are generally higher than those previous studies using the same types of datasets. There are two possible reasons for the increasing presence of altmetric data across papers. One is the progress made by altmetric data aggregators (particularly Altmetric.com), by improving their publication detection techniques and by enlarging tracked data sources. For example, Altmetric.com redeveloped their news tracking system in December 2015 (Altmetric, 2020), which partially explains the rise of news coverage in 2016 (see Figure 2). The second reason for the increasing presence of some altmetric data is the rising uptake of social media by the public, researchers, and scholarly journals (Nugroho et al., 2020; Van Noorden, 2014; Zheng et al., 2019). Against this background, scientific papers are more likely to be disseminated on social media, thereby stimulating the accumulation of altmetric data. The fact that more papers with corresponding altmetric data accrued and detected is beneficial to consolidate the data foundation, thus promoting the development and possible application of altmetrics.

In the meantime, we emphasized the biases of altmetric data towards different publication years. Costas et al. (2015a) highlighted the “recent bias” they found in the overall altmetric scores, which refers to the dominance of most recent published papers in garnering altmetric data. Nevertheless, we found that the “recent bias” is not exhibited by all types of altmetric data. For altmetric data with relatively high speed in data accumulation after publication, like Twitter mentions, Facebook mentions, news mentions, blog citations, and Reddit mentions (Fang & Costas, 2020), it is demonstrated that their temporal distribution conforms to a “recent bias”. However, a “past bias” is found for altmetric data that take a relatively longer time to accumulate, such as Wikipedia citations, policy document citations, F1000Prime recommendations, video mentions, peer review comments, and Q&A mentions (Fang & Costas, 2020). Due to the slower pace of these altmetric events, they are more concentrated on relatively old papers. Even for Mendeley readers, its data presence across recent papers is obviously lower.

Overall, although an upward tendency of data presence has been observed over time, most altmetric data still keep an extremely low data presence, with the only exceptions of Mendeley readers and Twitter mentions. As suggested by Thelwall, Haustein, et al. (2013), until now most altmetric data may only be applicable to identify the occasional exceptional or above average articles rather than as universal sources of impact evidence. In addition, the distinguishing presence of altmetric data reinforces the necessity of keeping altmetrics separate in future analyses or research assessments.

2.4.2 Different presence of altmetric data across subject fields and research topics

With the information of subject fields and micro-topics assigned by the CWTS publication-level classification system, we further compared the presence of 12 types of altmetric data across subject fields and their inclinations to different research topics. Most altmetric data have a stronger focus on papers in the fields of SSH, BHS, and LES. In contrast, altmetric data presence in the fields of PSE and MCS are generally lower. This kind of data distribution differs from what has been observed based on citations, in what SSH are underrepresented while PSE stands out as the subject field with higher levels of citations. This finding supports the idea that altmetrics might have more added values for Social Sciences and Humanities when citations are absent (Costas et al., 2015a).

In this study, it is demonstrated that even within the same subject field, altmetric data show different levels of data presence across research topics. Amongst altmetric data, their correlations at the research topic level are similar with the correlations at the paper level (Costas et al., 2015a; Zahedi et al., 2014), with Mendeley readers the only altmetric data moderately correlated with citations, and Twitter mentions and Facebook mentions, news mentions and blog citations, the two pairs showing the strongest correlations. There might exist some underlying connections within these two pairs of strongly correlated altmetric data, such as the possible synchronous updating by users who utilize multiple platforms to share scientific information, which can be further investigated in future research. For the remaining altmetric data, although many of them achieved moderate to strong correlations with each other from the aspect of coverage because they have similar patterns of data coverage across subject fields, the correlations of data intensity are weaker, implying that research topics garnered different levels of attention across altmetric data (Robinson-Garcia et al., 2019).

In view of the uneven distribution of specific altmetric data across research topics, it is possible to identify hot research topics which received higher levels of attention from certain communities such as Twitter users and policy-makers. Based on two indicators for measuring data presence: coverage and intensity, we developed a framework to identify hot research topics operationalized as micro-topics that fall in the first decile in terms of the ranking distribution of both coverage and intensity. This means that hot research topics are those with large shares of the papers receiving intensive average attention. We have demonstrated the application of this approach in detecting hot research topics mentioned on Twitter and cited in policy documents. Since the subject field differences are so pronounced that they might hamper generalization (Mund & Neuhäusler, 2015), the identification of hot research topics was conducted for each subject field severally. Hot research topics on Twitter reflect the interest shown by Twitter users, while those in policy documents serve as the mirror of policy-makers' focuses on science, and these two groups of identified hot research topics are diverse and hardly overlapped. This result proves that different communities are keeping an eye on different scholarly topics driven by dissimilar motivations.

The methodology of identifying hot research topics sheds light on an innovative application of altmetric data in tracking research trends with particular levels of social attention. By taking the advantage of the clustered publication sets (i.e., micro-topics) algorithmically generated by the CWTS classification system, the methodology proposed measures how wide and intensive the altmetric attention to the research outputs of specific research topics is. This approach provides a new option to monitor the focus of attention on science, thus representing an important difference with prior studies about the application of altmetric data in identifying topics of interest, which mostly were based on co-occurrence networks of topics with specific altmetric data accrued (Haunschild et al., 2019; Robinson-Garcia et al., 2019). The methodology proposed employs a two-dimensional framework to classify research topics into four main categories according to the levels of the specific altmetric attention they received. As such, the framework represents a more simplified approach to study and characterize different types of attention received by individual research topics. In our proposal for the identification of hot research topics, the influence of individual papers with extremely intensive attention received is to some extent diminished, relying the assessment of the whole topic on the overall attention of the papers around the topic, although of course those topics characterized by singularized papers with high levels of attention are also considered as "star-papers topics". It should be acknowledged that the results of this approach give an overview of the attention situations of generalized research topics, however, to get more detailed pictures of specific micro-level research fields, other complementary methods based on the detailed text information of the papers should be employed to go deep into micro-topics. Moreover, in this study, the identification of hot research topics is based on the whole dataset, in future studies, through introducing the factors of publication time of research outputs and the released time of altmetric events, it is suggested to monitor those hot research topics in real time in order to reflect the dynamic of social attention to science.

2.4.3 Limitations

There are some limitations in this study. First, the dataset of papers is restricted to papers with DOIs or PubMed IDs. The strong reliance on these identifiers is also seen as one of the challenges of altmetrics (Haustein, 2016). Second, although all types of documents are included in the overall analysis of data presence, only Article, Review, and Letter are assigned with main subject fields of science and micro-topics by the CWTS publication-level classification system, so only these three document types are considered in the following analysis of data presence across subject fields and research topics. But these three types account for 87.5% of sample papers (see Table 4 in Appendix), they can be used to reveal relatively common phenomena. Lastly, the CWTS classification system is a coarse-grained system of disciplines in consideration of that some different fields are clustered into an integral whole, like social sciences and humanities, making it difficult to present more fine-grained results. But the advantages of this system lie in that it solves the problem caused by multi-disciplinary journals, and individual papers with similar research topics are clustered into micro-level fields, namely, micro-topics, providing us with the possibility of comparing the distribution of altmetric data at the research topic level, and identifying hot research topics based on data presence.

2.5 Conclusions

This study investigated the state-of-the-art presence of 12 types of altmetric data for nearly 12.3 million Web of Science papers across subject fields and research topics. Except for Mendeley readers and Twitter mentions, the presence of most altmetric data is still very low, even though it is increasing over time. Altmetric data with high speed of data accumulation are biased to newly published papers, while those with lower speed bias to relatively old papers. The majority of altmetric data concentrate on papers from the fields of Biomedical and Health Sciences, Social Sciences and Humanities, and Life and Earth Sciences. These findings underline the importance of applying different altmetric data with suitable time windows and fields of science considered. Within a specific subject field, altmetric data show different preferences for research topics, thus research topics attracted different levels of attention across altmetric data sources, making it possible to identify hot research topics with higher levels of attention received in different altmetric contexts. Based on the data presence at the research topic level, a framework for identifying hot research topics with specific altmetric data was developed and applied, shedding light onto the potential of altmetric data in tracking research trends with a particular social attention focus.

2

2.6 Appendix

It is reported that the distributions of citation counts (Seglen, 1992), usage counts (X. Wang, Fang, & Sun, 2016), and Twitter mentions (Fang, Dudek, et al., 2020) are highly skewed. Results in Figure 8 show that the same situation happens to other altmetric data as well. Even though the data volume differs greatly, the distributions of all kinds of altmetric data are highly skewed, suggesting that most scientific papers only accrued few corresponding events and very few of them received high levels of attention.

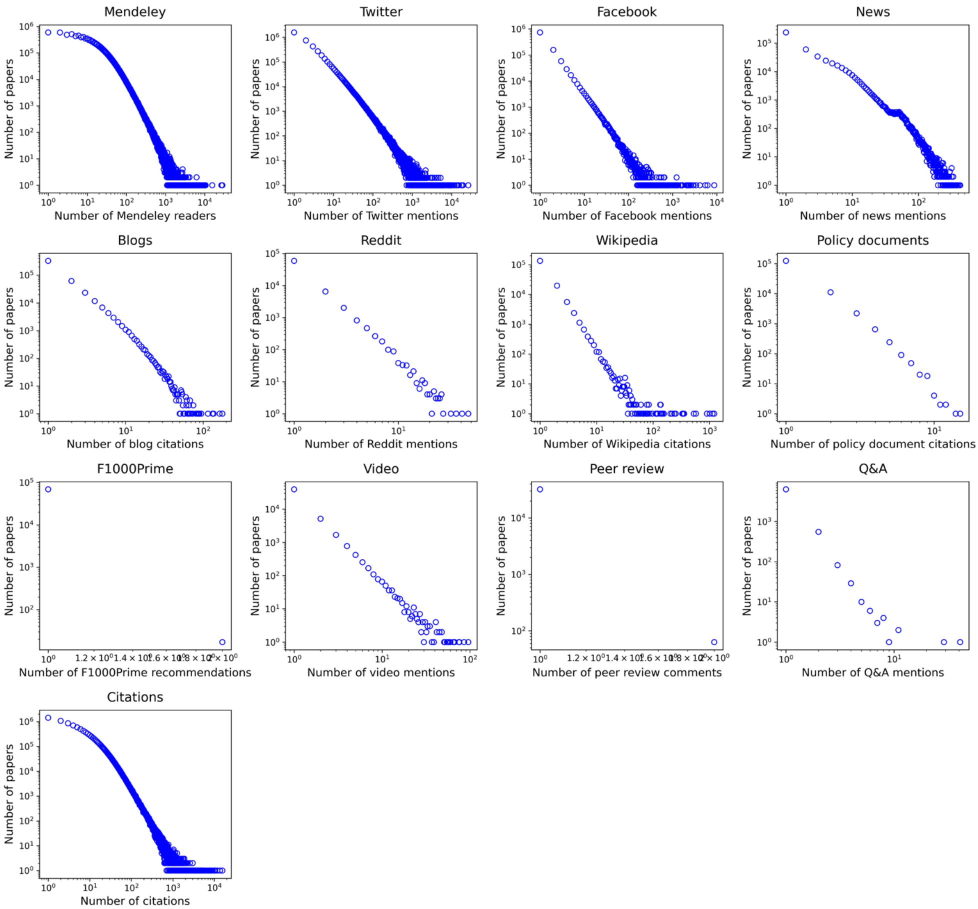

Figure 8. Distribution of 12 types of altmetric data and citations of sample papers

Spearman correlation analyses among coverage, density, and intensity of micro-topics were conducted for each altmetric data and citations, and the results are shown in Figure 9. Because of the highly skewed distribution of all kinds of altmetric data, the calculation of coverage and density are prone to get similar results, especially for altmetric data with smaller data volume. Therefore, the correlation between coverage and density is quite strong for every altmetric data. For most altmetric data, density and intensity are moderately or strongly correlated, and their correlations are always slightly stronger than that between coverage and intensity.

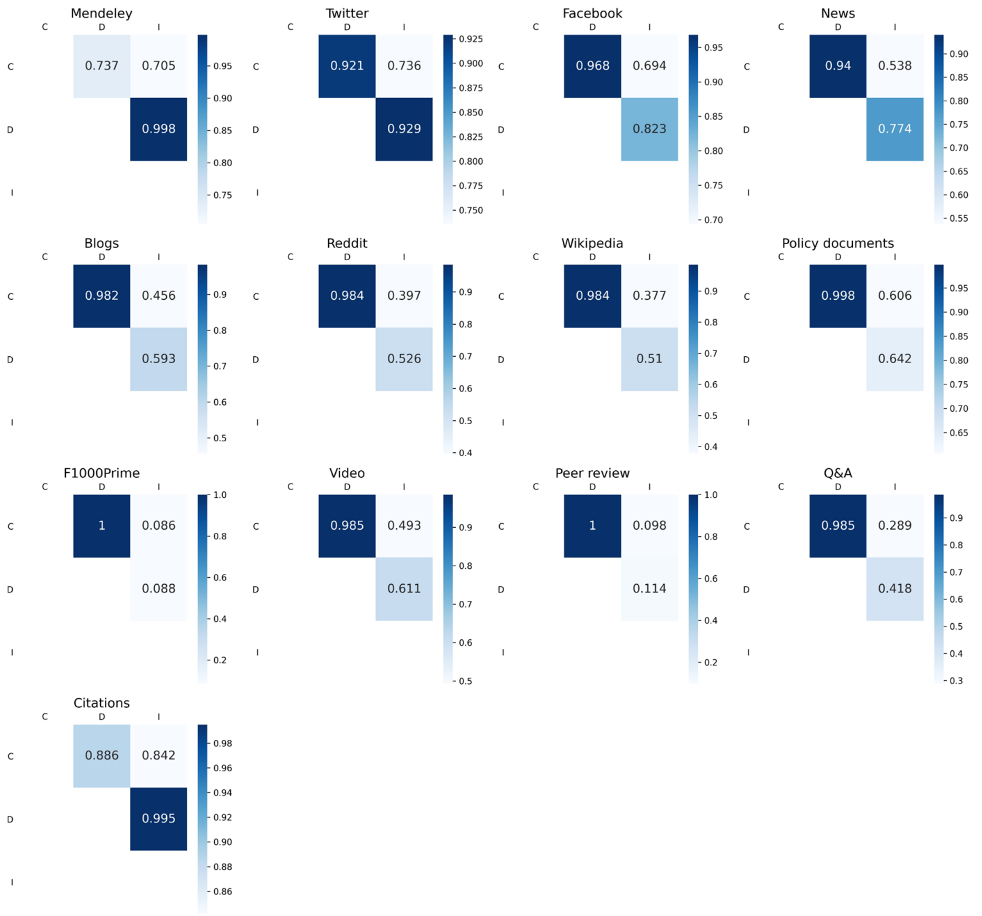

Figure 9. Spearman correlations among coverage (C), density (D), and intensity (I) at the micro-topic level

2

In consideration of the influence of zero values of some micro-topics on inflating the Spearman correlation coefficients, we did a complementary analysis by calculating the Spearman correlations for each pair of data after excluding those mutual micro-topics with zero values (Figure 10). Compared to the results shown in Figure 4, values in Figure 10 are clearly lower, especially for those pairs of altmetric data with relatively low data presence. However, the overall patterns are still consistent with what we observed in Figure 4.

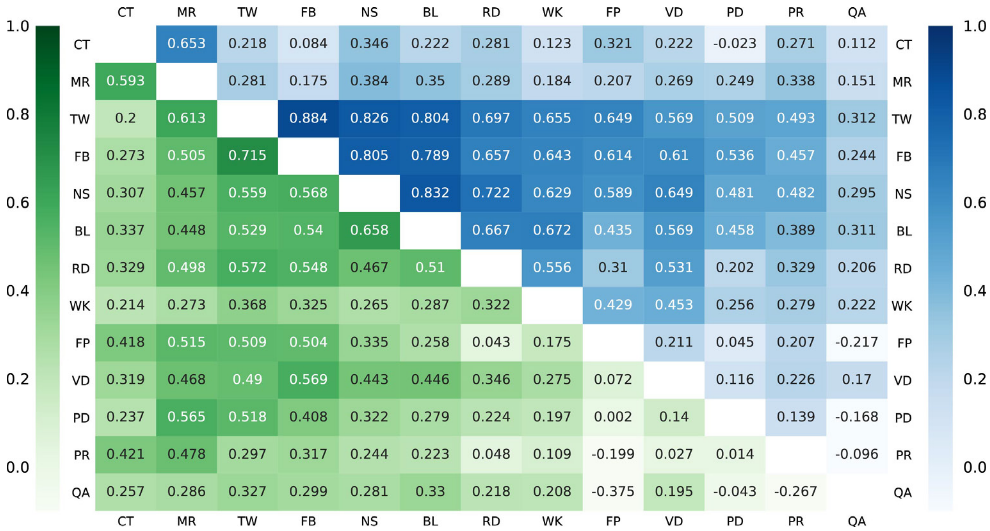

Figure 10. Spearman correlation analyses of coverage (upper-right triangle) and intensity (bottom-left triangle) among citations and 12 types of altmetric data at the micro-topic level (with mutual zero-value micro-topics excluded). WoS citations (CT), Mendeley readers (MR), Twitter mentions (TW), Facebook mentions (FB), news mentions (NS), blog citations (BL), Reddit mentions (RD), Wikipedia citations (WK), F1000Prime recommendations (FP), video mentions (VD), policy document citations (PD), peer review comments (PR), Q&A mentions (QA)

The 12,271,991 sample WoS papers were matched with their document types through the CWTS in-house database. Table 4 presents the number of papers and the coverage of altmetric data of each type. The types of Article, Review, and Letter, which are included in the CWTS classification system, account for about 87.5% in total. The altmetric data coverage varies across document types as observed by Zahedi et al. (2014). For most altmetric data, Review shows the highest altmetric data coverage, followed by Article, Editorial Material, and Letter.

Table 4. Coverage of 12 kinds of altmetric data of different document types

Indicator	Article	Review	Editorial Material	Meeting Abstract	Letter	Book Review	Other
Number of papers	9,851,747	616,514	595,577	527,049	273,819	227,369	179,916
Percentage	80.28%	5.02%	4.85%	4.29%	2.23%	1.85%	1.47%
Mendeley readers	94.27%	95.80%	77.02%	46.67%	75.02%	31.92%	54.99%
Twitter mentions	34.61%	55.24%	41.74%	2.21%	31.72%	10.49%	29.09%
Facebook mentions	8.30%	16.38%	14.97%	0.39%	7.79%	2.28%	9.03%
News mentions	4.04%	6.70%	5.58%	0.37%	3.10%	0.16%	4.44%
Blog citations	3.75%	6.18%	4.52%	0.10%	1.86%	0.62%	4.04%
Wikipedia citations	1.29%	4.38%	1.06%	0.03%	0.53%	0.46%	1.16%
Policy document citations	1.16%	2.56%	0.90%	0.06%	0.53%	0.03%	0.33%
Reddit mentions	0.56%	0.75%	0.81%	0.12%	0.38%	0.08%	1.38%
F1000Prime recommendations	0.63%	0.94%	0.15%	0.01%	0.17%	0.00%	0.05%
Video mentions	0.39%	1.20%	0.35%	0.01%	0.16%	0.01%	0.27%
Peer review comments	0.30%	0.20%	0.08%	0.00%	0.08%	0.00%	0.14%
Q&A mentions	0.06%	0.16%	0.04%	0.00%	0.02%	0.00%	0.05%

CHAPTER 3

Studying the accumulation velocity of altmetric data tracked by Altmetric.com[1]

[1] **This chapter is based on**:

Fang, Z., & Costas, R. (2020). Studying the accumulation velocity of altmetric data tracked by Altmetric.com. *Scientometrics*, *123*(2), 1077–1101. https://doi.org/10.1007/s11192-020-03405-9

Author contributions:

Fang, Z. (Conceptualization, Methodology, Formal analysis, Investigation, Visualization, Data Curation, Writing - Original Draft, Writing - Review & Editing)

Costas, R. (Conceptualization, Methodology, Investigation, Supervision, Writing - Review & Editing)

Abstract

This paper investigates the data accumulation velocity of 12 Altmetric.com data sources. DOI *created date* recorded by Crossref and altmetric event *posted date* tracked by Altmetric.com are combined to reflect the altmetric data accumulation patterns over time and to compare the data accumulation velocity of various data sources through three proposed indicators, i.e., *Velocity Index*, *altmetric half-life*, and *altmetric time delay*. Results show that altmetric data sources exhibit different data accumulation velocity. Some altmetric data sources have data accumulated very fast within the first few days after publication, such as Reddit, Twitter, news, Facebook, Google+, and blogs. On the opposite spectrum, research outputs are at relatively slow pace in accruing data on some data sources, like policy documents, peer review, Q&A, Wikipedia, video, and F1000Prime. Most altmetric data sources' velocity degree also changes by document types, subject fields, and research topics. The type Review is slower in receiving altmetric mentions than Article, while Editorial Material and Letter are typically faster. In general, most altmetric data sources show higher velocity values in the fields of *Physical Sciences and Engineering* and *Life and Earth Sciences*. Within each field, there also exist some research topics that attract social attention faster than others.

Keywords

Altmetrics, Crossref, data accumulation speed, Velocity Index, altmetric half-life, time delay

3.1 Introduction

“Speed” has been highlighted as one of the most important characteristics of altmetrics (Bornmann, 2014a; Wouters & Costas, 2012). Compared to citations, which has been often criticized for its time delay in providing reliable measurement for research impact (J. Wang, 2013), speed in the context of altmetrics is related to the idea that the impact of a given scholarly output can be measured and analyzed much earlier (Mohammadi & Thelwall, 2014; Priem et al., 2010). Publication delays are considered to substantially slow down the formal communication and dissemination of scientific knowledge (Amat, 2008; Björk & Solomon, 2013). In contrast, interactions around science on social media platforms are likely to happen within a very short time-frame. For instance, Twitter mentions of scientific papers may occur immediately within hours or even minutes after they were available online (Haustein, Bowman, et al., 2015; Shuai et al., 2012).

However, because of the strong heterogeneity of altmetrics (Haustein, 2016), which incorporate a wide range of metrics based on different types of data sources, it is difficult to establish a clear-cut and unified conceptual framework for the temporal analysis of all altmetrics. Each altmetric indicator, typically with unique functions and aimed at different audiences, may tell different stories about the reception of scientific information, and show distinguishing patterns in varying contexts. Lin and Fenner (2013a) concluded that altmetrics are very likely representing very different things. From this point of view, we argue that the interpretation of the characteristic properties of different altmetrics should be made for each metric separately, including among these properties also their “speed”.

3.1.1 Accumulation patterns and immediacy measurement of citations and usage metrics

In contrast to altmetric data, the accumulation patterns of citations have already been widely discussed in previous studies from several perspectives, such as their “obsolescence” (Line, 1993), “ageing” (Aversa, 1985; Glänzel & Schoepflin, 1995), “durability” (Costas et al., 2010), or “delayed recognition” (Garfield, 1980; Min et al., 2016). Citation histories, which relate to the analysis of the distribution of citations over time, were mainly studied from the synchronous or diachronous perspectives (Stinson & Lancaster, 1987). The former considers the distribution of the publication years of cited references, while the latter focuses on the distribution of received citations over time (Colavizza & Franceschet, 2016; Sun et al., 2016), which are also referred to as “retrospective citations” and “prospective citations”, respectively (Glänzel, 2004). These two approaches have been applied to study the accumulation patterns of usage metric data as well. With the development of digital publishing, usage metrics have been proposed and adopted by publishers during the last decades to supplement citations in reflecting how frequently scientific outputs are used and measuring their early impact to some extent (Schloegl & Gorraiz, 2011). From the synchronous perspective, Kurtz et al. (2005) concluded that most studies of obsolescence found that the use of literature declines exponentially with age. The diachronous accumulation patterns of usage metrics, like views, downloads, or reads, were investigated and often compared with citations. On the basis of page views data of *Nature* publications, X. Wang et al. (2014) explored the dynamic usage history over time and found that papers are used most frequently within a short period after publication, finding that in median it only takes 7 days for papers to reach half of their total page views. Schlögl et al. (2014) reported that citations take several years until they reach their peak, however most downloads of papers are quickly accrued in the same publication year. In a similar fashion, Moed (2005) already found that citations and downloads show different patterns of obsolescence, and about 40% of downloads accumulated within the first 6 months after publication. More recently X. Wang, Fang, & Sun (2016) used the article-level "usage counts" provided by Web of Science to investigate the usage patterns of indexed papers, and identified that newly published papers accumulated more Web of Science usage counts than older papers.

As to the measurement of the "speed" of citations and usage metrics, several indicators have been created and applied in practice. For example, based on the time elapsed between the publication date and the date of the first citation of a paper, Schubert and Glänzel (1986) developed the indicator *mean response time* (MRT) in order to measure the citation speed of journals, understood as the properly formed average number of years between the publication of articles in a journal and the time of their first citation. In order to measure how quickly articles in a journal are cited, the *Journal Citation Reports* (JCR) calculates the indicator named *Immediacy Index* for each journal in each year. This indicator is defined as the average number of times an article is cited in the same year it is published.[1] Besides, at the journal level, *Cited Half-Life* and *Citing Half-Life* are also calculated by JCR to measure how fast journals are accumulating half of their citations and how far back that citing relationship extends.[2] Analogous to the citation-based Immediacy Index and half-life, the "usage immediacy index" and "usage half-life" (Rowlands & Nicholas, 2007), "download immediacy index" (Wan et al., 2010) were proposed to describe the life cycle of usage metrics. By analyzing usage data in the field of Oncology collected from Science Direct, Schloegl and Gorraiz (2010) calculated the mean usage half-life and found that it is much shorter than the average cited half-life, observing also different obsolescence patterns between downloads and citations.

[1] See more information about *Immediacy Index* at: https://clarivate.com/webofsciencegroup/blog/know-your-metrics-immediacy-index/ (Accessed January 29, 2020).

[2] See more information about *Cited Half-Life* and *Citing Half-Life* at: https://clarivate.com/webofsciencegroup/blog/a-closer-look-at-cited-and-citing-half-lives/ (Accessed January 29, 2020).

3.1.2 Accumulation patterns and immediacy measurement of altmetric data

Since the emergence of altmetrics, most related studies have focused on the coverage of scientific papers across altmetric sources and their correlation with citation counts (Costas et al., 2015a; Haustein, Peters, Bar-Ilan, et al., 2014; Thelwall, Haustein, et al., 2013). Less attention was paid to the study of the accumulation velocity of altmetric data over time. Only a few altmetric data sources were investigated from the perspective of their immediacy. Maflahi and Thelwall (2018) conducted a longitudinal weekly study of the Mendeley readers of articles in six library and information science journals and found that they start to accrue early from when articles are first available online and continue to steadily build over time, being this the case even for journals with large publication delays. Thelwall (2017) also found that articles attracted between 0.1 and 0.8 Mendeley readers on average in the month they first appeared in Scopus, with some variability across subject fields. The results based on *PeerJ* social referrals data of X. Wang, Fang, & Guo (2016) suggested that the number of "visits" to papers from social media (Twitter and Facebook) accumulates very quickly after publication. By comparing the temporal patterns of Twitter mentions and downloads of arXiv papers, Shuai et al. (2012) found that Twitter mentions have shorter delays and narrower time spans than arXiv downloads. Ortega (2018b) made a comparison of temporal distribution at the month time interval among citations, views, downloads, Mendeley readership, tweets, and blog mentions recorded by PlumX, and concluded that tweets and blog mentions are the quickest available metrics. Yu et al. (2017) found that Twitter and Weibo are more immediate than citations, however they also suggested that not all altmetric data sources have the same degree of immediacy.

In contrast to citation histories, which are mainly analyzed at year or month levels, for altmetrics it is insufficient to use such large time aggregations, since the real-time update of social media metric data makes altmetric events around research outputs visible within smaller time scales (e.g. hours or days). Nevertheless, a large-scale quantitative analysis comparing the data accumulation patterns of different altmetric data sources at the micro-level time interval (i.e., day) is still missing in the literature in altmetrics, probably caused by the absence of a reliable and precise proxy for publication dates, a piece of information that is critical in order to study the accumulation patterns of altmetric data (Haustein, Bowman, et al., 2015). Crossref provides several publication dates for its recorded DOIs, such as DOI *created date* (date on which the DOI was first registered), *published-online date* (date on which the work was published online), and *published-print date* (date on which the work was published in print). The distribution and potential of these date information for altmetrics have been compared and analyzed in a previous study (Fang & Costas, 2018), as suggested by Haustein, Bowman, et al. (2015), the value of DOI created date as a fine-grained benchmark of publication date in the context of altmetrics was highlighted.

In this paper, on the basis of DOI *created date* recorded by Crossref, as well as the altmetric event *posted date*[3] recorded by Altmetric.com, we compare the accumulation velocity amongst different types of altmetric data from a diachronous perspective.

3.1.3 Objectives

The main objectives of this study are: (1) to measure the accumulation velocity of altmetric data of scientific papers on 12 Altmetric.com data sources, here velocity referring to the pace at which altmetric events accumulate over time, and (2) to compare altmetric data accumulation velocity of different altmetric data sources across document types, subject fields, and research topics. The specific research questions are as follows:

RQ1. What are the altmetric data accumulation patterns of various Altmetric.com data sources?

RQ2. On which data sources do newly published research outputs show higher velocity in accruing altmetric data (and which ones are relatively lower)?

RQ3. How do the data accumulation velocity of different Altmetric.com data sources vary across document types, subject fields, and research topics?

3.2 Data and methods

3.2.1 Altmetric.com data sources with altmetric event posted date

In this study altmetric event records of 12 Altmetric.com data sources with *posted date* are selected as research objects. The altmetric data for this study were provided by Altmetric.com in a dump file with their data until October 2017. Table 1 presents these 12 data sources with event posted date information tracked by Altmetric.com together with the date when they started their coverage.

[3] This is the date on which a given altmetric event (e.g., a tweet, a news mention, or a blog citation) was posted online or published (for policy documents).

Table 1. Altmetric.com data sources with altmetric event posted date[4]

Data source	Concept measured with regard to research outputs	Coverage began date
Twitter	Twitter mentions, including original tweets, reply tweets, quote tweets, and retweets.	Oct 2011
Facebook	Facebook mentions, including posts on a curated list of public pages only.	Oct 2011
News	News media mentions on a list of news sources tracked by Altmetric.com, which contains over 2,900 English and non-English global news outlets.	Oct 2011 & Dec 2015
Blogs	Blog citations on a list of blogs tracked by Altmetric.com, which contains over 14,000 academic and non-academic blogs.	Oct 2011
Google+	Google+ mentions.	Oct 2011
Wikipedia	Wikipedia citations on English Wikipedia pages only.	Jan 2015
Policy documents	Policy documents citations on a wide range of public policy documents tracked by Altmetric.com, including policy, guidance, or guidelines documents from a governmental or non-governmental organization.	Jan 2013
F1000Prime	F1000Prime recommendations.	May 2013
Reddit	Reddit mentions on all sub-reddits, including original posts only.	Oct 2011
Peer review	Post-publication peer review comments collected from two forums: PubPeer and Publons.	Mar 2013
Video	Video mentions on YouTube.	Apr 2013
Q&A	Q&A mentions on Stack Overflow.	Oct 2011

Note: As of 2017, Altmetric.com has stopped collecting data from CiteULike, Sina Weibo, LinkedIn, and Pinterest. Syllabus data only posted in 2015 were provided by Altmetric.com and almost all publications mentioned by Syllabus are not indexed by Web of Science. Mendeley and CiteULike, two online reference managers, lack proper posted date information. Therefore, these data sources have not been included in this study. Although Google+ has also been discontinued and thereby Altmetric.com has stopped tracking it since January 2019, it is still considered as one of the data sources to be studied in this paper due to the availability of data during our observation time window.

3.2.2 Dataset

Considering the posted dates of the different altmetric events, we could know the exact date on which an altmetric event was posted. In addition, in order to study the accumulation patterns of altmetric data at the day time interval, DOI created dates of research outputs recorded by Crossref are collected to serve as the proxy of publication dates. To obtain both

[4] See more information about different Altmetric.com data sources at: https://help.altmetric.com/support/solutions/articles/6000060968-what-outputs-and-sources-does-altmetric-track- (Accessed November 26, 2019); and their coverage dates at: https://help.altmetric.com/support/solutions/articles/6000136884-when-did-altmetric-start-tracking-attention-to-each-attention-source- (Accessed November 26, 2019).

altmetric event posted date and DOI created date for measuring accumulation velocity, Web of Science (WoS) papers with the following criteria were selected as research objects:

- Papers with DOI recorded by Crossref. In order to get the DOI created dates, selected papers must have DOIs recorded by Crossref.

- Papers with publication date ranging from 2012 to 2016 according to both WoS publication year and Crossref DOI created date. To filter out old papers with newly registered DOIs (Fang & Costas, 2018), WoS publication year is also used as a benchmark to restrict the publication year of samples.

- Papers with at least one altmetric event recorded from any altmetric data source listed in Table 1.

- Papers without arXiv preprint version tracked by Altmetric.com. The existence of preprint version makes research outputs available to social media before they are formally published (Darling et al., 2013), which may lead to the altmetric record posted dates to be earlier than the publication date. Therefore, papers with arXiv IDs tracked by Altmetric.com are not included in this study.

According to the above criteria, there are 2,597,339 papers extracted from the CWTS in-house WoS database. However, 204,387 of them (accounting for 7.9%) have at least one altmetric event posted date earlier than their DOI created dates. Except for the influence of preprint versions, in theory an altmetric event cannot mention a DOI before it exists. The possible reasons for the existence of these unreliable cases are the following:

- Crossref DOI created dates may contain errors and not always accurately reflecting the publication date.

- Papers' DOI created dates may be updated by publishers due to different reasons (e.g., publisher mergers).[5]

In order to ensure the highest precision in our analysis, papers with any altmetric event posted date before their DOI created date are excluded from our analysis, resulting in a total set of 2,392,952 papers that are finally analyzed in this study. Table 2 lists the number of papers mentioned by each data source and the total number of altmetric events they have accumulated in the dataset. Twitter contributes the most majority of altmetric data to the selected papers, followed by Facebook.

[5] Extracted from personal communication with Euan Adie from Altmetric.com.

Table 2. General presence of altmetric data for the dataset

Data source	Number of papers	Number of altmetric events	Coverage	Intensity
Twitter	2,157,556	14,853,823	90.2%	6.9
Facebook	545,370	1,375,880	22.8%	2.5
News	224,036	1,037,719	9.4%	4.6
Blogs	200,784	360,736	8.4%	1.8
Google+	84,754	216,787	3.5%	2.6
Wikipedia	75,693	106,917	3.2%	1.4
Policy documents	56,296	73,523	2.4%	1.3
F1000Prime	39,981	48,517	1.7%	1.2
Reddit	31,726	43,805	1.3%	1.4
Peer review	20,783	33,599	0.9%	1.6
Video	12,918	18,643	0.5%	1.4
Q&A	2,369	2,474	0.1%	1.0

Note: Coverage refers to the proportion of papers with at least one corresponding altmetric event of all papers in our dataset. Intensity refers to the mean number of altmetric events of papers with at least one corresponding altmetric event (Haustein, Costas, et al., 2015).

3.2.3 Indicators and analytical approaches

Considering the diverse nature, scale, and user types of different altmetric data sources, it is very likely that they exhibit also very different velocity degrees of accumulation in face of newly published research outputs. To reflect the velocity differences among altmetric data sources, we use three indicators to measure velocity from both flexible and fixed perspectives, including *Velocity Index*, *altmetric half-life*, and *altmetric time delay*.

For altmetric data accumulated on a specific data source, the *Velocity Index* (VI) refers to the proportion of altmetric events that happened in a specific time interval (e.g., 1 day, 1 month, or 1 year) after the publication of the papers. The calculation method is shown in the formula below.

$$Velocity\ Index = \frac{P_i}{TP_i}$$

Pi is the number of events accrued in a specific time interval after publication (e.g., 1 day, 1 month, or 1 year) for a set of papers, *TPi* indicates the total number of events during the observed time window. In general, the closer to 1 of the Velocity Index, the more immediate (faster) the altmetric data of new papers accumulated in the given observation period.

Conversely, the closer to 0, the lower the accumulation velocity (i.e., more events happened beyond the specified period of time).

Besides, in line with the *Twitter half-life* and *Twitter time delay* proposed by Haustein (2019), which refer to the number of days until 50% of all tweets have appeared and the number of days between the publication of a document and its first tweet, respectively, we generalize these indicators for all altmetric data sources. Consequently, the *altmetric half-life* of an altmetric data source is defined as the number of days until half of its events have appeared, and *altmetric time delay* of a research output on an altmetric data source is defined as the number of days between its publication and its first altmetric event on that data source.

Both Velocity Index and altmetric half-life are based on overall data distribution of all events received by a paper, while altmetric time delay focuses on a special altmetric event (the first one). Velocity Index provides a flexible perspective for the measurement of data accumulation velocity, since it allows for more nuanced time accumulation discussions considering different time intervals (i.e., days, months, or years). By comparison, altmetric half-life and altmetric time delay provide a fixed perspective at the day level. Therefore, these indicators work as relevant complements to each other in order to better characterize the tempo of altmetric data accumulation.

In addition, the Spearman correlation analysis is performed with IBM SPSS Statistics 25 to explore the relationships among Velocity Index, altmetric half-life, and altmetric time delay. Also, at the research topic level, in order to testify whether or not research topics with fewer papers and altmetric events are more likely to reach higher values of Velocity Index, the Spearman correlation analysis is applied to exhibit the relationships among number of papers, number of altmetric events, and the Velocity Index.

3.2.4 CWTS publication-level classification system

The CWTS classification is a publication-level subject field classification system developed by Waltman and Van Eck (2012). It has not only been applied in Leiden Ranking (https://www.leidenranking.com/), but has also been employed by many previous studies for subject field related analysis (Costas et al., 2015a; Didegah & Thelwall, 2018). In the 2019 version of the publication-level classification, only citable items (Article, Review, and Letter) indexed by Web of Science are clustered into 4535 micro-level fields. These micro-fields correspond to small research topics (*micro-topics*), and they are assigned to five main subject fields of science algorithmically obtained, including *Social Sciences and Humanities* (SSH), *Biomedical and Health Sciences* (BHS), *Physical Sciences and Engineering* (PES), *Life and*

Earth Sciences (LES), and *Mathematics and Computer Science* (MCS),[6] which are illustrated in Figure 1 with VOSviewer. The layout of Figure 1 is also used to exhibit the Velocity Index of each micro-topic in the Result section. For the selected papers in our dataset, 2,189,708 of them (accounting for 91.5%) have CWTS classification information. This set of papers is drawn as our final sample of papers for the comparison of altmetric data accumulation velocity across subject fields and research topics. Statistics on the general presence of different altmetric data across five main subject fields can be found in Table 4 in Appendix.

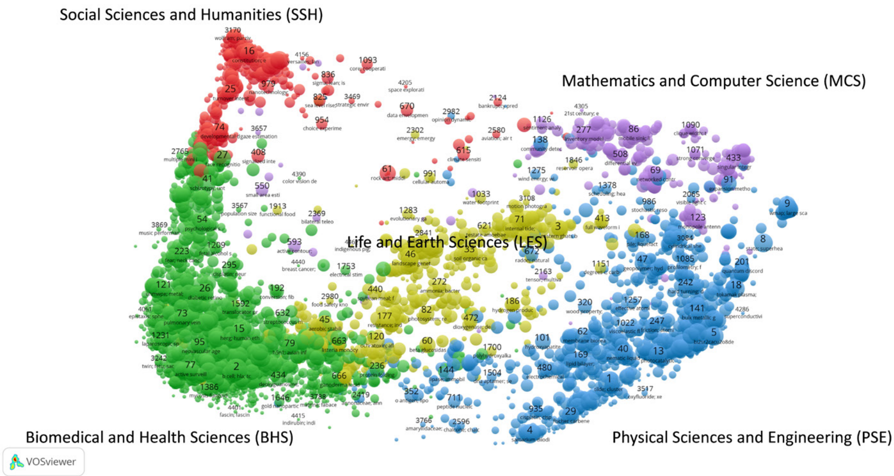

Figure 1. Five main subject fields of science of the CWTS publication-level classification system. Each circle represents a micro-level field clustered by papers with similar research topics (micro-topics)

3.3 Results

3.3.1 Altmetric data accumulation patterns

The intervals between publication dates and altmetric events posted dates are calculated for all altmetric events on each data source. Thus we can investigate the altmetric data accumulation patterns at the day time interval. Figure 2 shows the different data accumulation patterns of the 12 data sources within 1-year time interval (365 days) after publication. Data

[6] See more details about the CWTS classification system at: https://www.leidenranking.com/information/fields (Accessed January 29, 2020).

sources show different data accumulation patterns. Altmetric events to newly published research outputs on some data sources accumulated very fast, such as Reddit and Twitter, since half of their data accrued in the first 2 weeks (14 days) after the research outputs were published, and over 85% of their data happened within a year (365 days). Following Twitter and Reddit we have other pretty fast altmetric data sources including news, Google+, Facebook, and blogs. In contrast, policy documents, Wikipedia, Q&A, and peer review show much slower data accumulation patterns similar to that of traditional citations. Only 21.5% of policy document citations, 31.9% of peer review comments, 39.4% of Wikipedia citations, and 40.6% of Q&A mentions are accumulated within 1 year, which means that most of the events from these data sources happened more than a year after publication. Among these data sources, F1000Prime presents some uniqueness. In the first month after research outputs are published, the accumulation of F1000Prime recommendations is not very fast, but it speeds up over time, with more than 84% of data accrued within the first year.

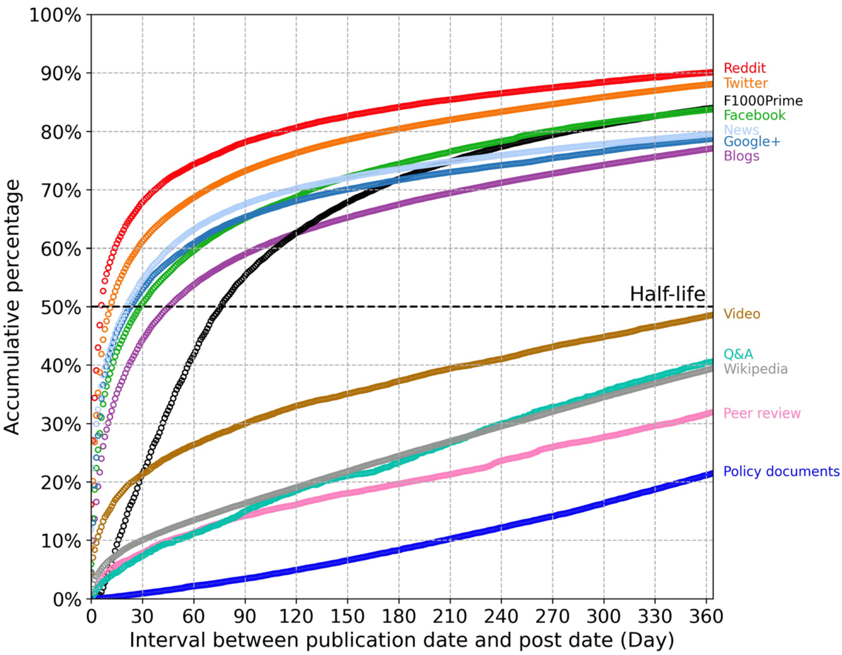

Figure 2. Altmetric data accumulation patterns of 12 Altmetric.com data sources within the first year (365 days) after publication

The dashed line at accumulative percentage of 50% in Figure 2 indicates the altmetric half-life, and Table 3 lists the altmetric half-lives of the 12 data sources analyzed. Reddit ranks first, with a half-life of 7 days, followed by Twitter (13 days), news (22 days), Google+ (25 days), and Facebook (30 days). Over half of altmetric events on these data sources happened within 1 month after the publication of research outputs. Other sources such as Wikipedia, peer review, and policy documents, need over 500 days to accumulate half of their event data. On the one hand, these data sources show lower reaction speed towards newly published papers. On the other hand, it suggests that they also pay more attention to papers with older publication time.

Table 3. Altmetric half-lives of 12 Altmetric.com data sources

Rank	Data source	Altmetric half-life (day)
1	Reddit	7
2	Twitter	13
3	News	22
4	Google+	25
5	Facebook	30
6	Blogs	47
7	F1000Prime	77
8	Video	394
9	Q&A	498
10	Wikipedia	515
11	Peer review	633
12	Policy documents	716

3.3.2 Generalizing the Velocity Index and altmetric time delay

The Velocity Indexes of each Altmetric.com data source at the day, month, and year time intervals are calculated respectively, and the rankings of sources by their Velocity Index are shown in Figure 3. The rankings vary at different time intervals. Reddit, Twitter, and news are the data sources showing the most immediate data accumulation patterns at the day, month, and year time intervals. Followed by Facebook, Google+, and blogs. While policy documents, peer review, Wikipedia, Q&A, and video perform more slowly in their Velocity Index values. F1000Prime, as mentioned above, although one of the slowest data sources at the day time interval, ranks the third at the year time interval. This means that the accumulation of F1000Prime recommendations of newly published papers is relatively slow in the short term, but it is faster at the year time interval (see also Figure 2). The case of F1000Prime highlights the importance of considering together the altmetric half-life of data sources and their Velocity Index, since both bring two different perspectives about the tempo of altmetric data.

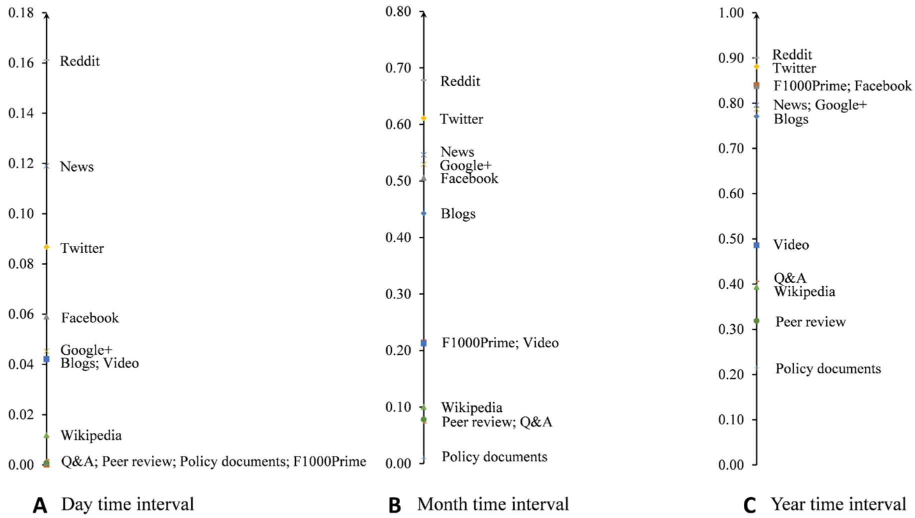

Figure 3. Velocity Index rankings at the **A** day, **B** month, and **C** year time intervals

Besides the Velocity Index and altmetric half-life which are based on overall altmetric data of each data source, we also consider the time delay of papers until they accrued their first altmetric event from different data sources, in which case only one specific altmetric event of papers is considered. The number of days between being published and being mentioned for the first time on a certain data source is calculated for each paper, and the distribution of altmetric time delays of the 12 Altmetric.com data sources is plotted in Figure 4. Each curve shows, for each specific data source, the proportion of papers that accrued the first altmetric event beyond certain number of days since being published. For instance, only about 37% of papers received their first Twitter mentions after the 10th day after being published (the vertical dashed line in Figure 4), while 94% of papers received their first Wikipedia citations after the 10th day after publication. In other words, around 63% of papers obtained their first Twitter mentions within 10 days after publication, and only 6% of papers got the first Wikipedia citations within the same time period. The more skewed the curve, the higher the proportion of papers accrued their first altmetric event after a long time. As a result, papers are faster to be visible on Twitter compared to other data sources, followed by Reddit, Google+, and Facebook. For various altmetric data sources, the patterns of accumulating the first altmetric event are quite similar with their Velocity Indexes at the month time interval and altmetric half-lives (in Appendix Table 5 provides the spearman correlations for the rankings based on these three indicators).

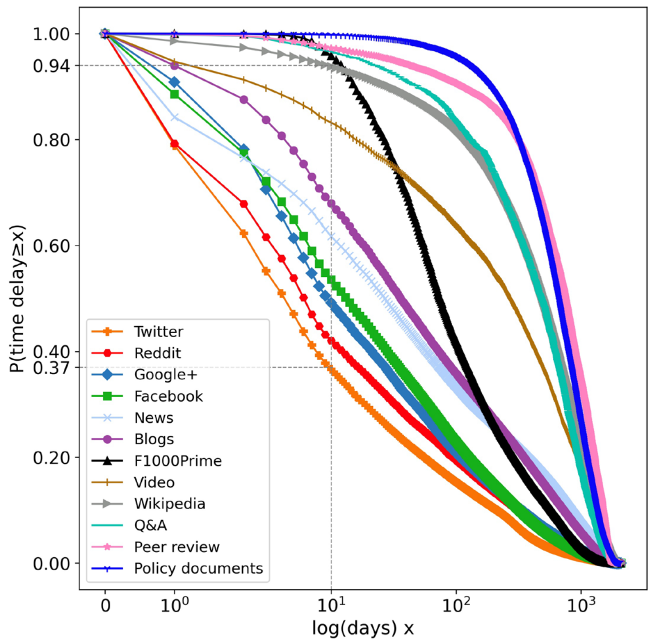

Figure 4. Distribution of altmetric time delay of 12 Altmetric.com data sources in log(days)

Overall, Twitter, Reddit, Google+, Facebook, news, and blogs can be categorized as *fast sources*, while in general, F1000Prime, video, Wikipedia, Q&A, peer review, and policy documents show lower velocity in mentioning scientific papers. These six data sources can be classified as *slow sources*.

3.3.3 Velocity Index variations across document types

For different document types, their altmetric data accumulation velocity might show some differences. So we utilize the Velocity Index at the month time interval to measure the altmetric data accumulation velocity for different document types across data sources. The differences in the Velocity Index across the four main document types with most number of papers: Article (N = 1,951,197, Coverage = 81.5%), Review (N = 196,722, Coverage = 8.2%), Editorial Material (N = 139,950, Coverage = 5.8%), and Letter (N = 52,038, Coverage = 2.2%), are illustrated in Figure 5. The presence of altmetric data across these four document types is listed in Table 6 in Appendix. The type of Article is the largest in number of papers, so its Velocity Index is very close to the overall Velocity Index of each data source. Review, Editorial Material, and Letter, in comparison, show differences with the overall Velocity Index, especially for data sources with relatively high Velocity Index values. Reviews are not as fast in accumulating altmetric data as compared to the other document types. Conversely, Editorial Material and Letter are document types more likely to be mentioned faster after publication. The Velocity Indexes of these two document types are higher than the overall Velocity Index for most data sources. In particular, Editorial Material and Letter hold relatively high Velocity Indexes on peer review platforms (Publons and PubPeer), which is among the group of “slower” data sources based on the overall Velocity Index (Figure 3) and its altmetric half-life (Table 3). The Review type also has a slightly higher Velocity Index than the overall and Article type on peer review events. Results show that peer review platforms seem to notice and comment on Editorial Materials, Letters and Reviews more quickly than regular Articles. Although the coverage of these three document types with peer review data is limited (0.20-0.27%), there are larger shares of peer review comments that happened soon after their publication compared to other altmetric events of slow sources.

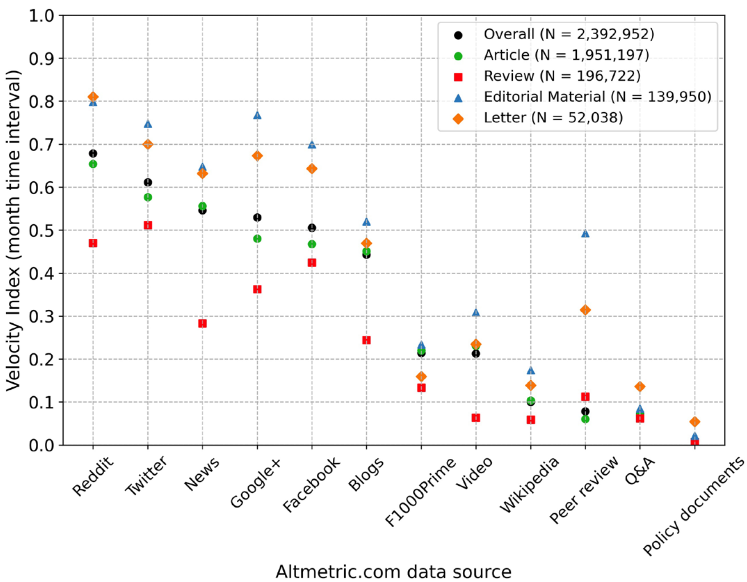

Figure 5. Velocity Index variations across four document types

3.3.4 Velocity Index variations across subject fields

The coverage of papers in Altmetric.com from different data sources differs by subject field (Costas et al., 2015b; Zahedi et al., 2014). In this study (Figure 6) we analyze the changes in the Velocity Index at the month time interval of different Altmetric.com data sources across five major subject fields of science (using the CWTS classification). Each row presents the

Velocity Indexes of different altmetric data sources ranked from high to low in each subject field. Each altmetric data source in Figure 6 is indicated with the same color, together with their specific Velocity Index. On the top of Figure 6, altmetric data sources are ranked by their overall Velocity Indexes at the month time interval. Colorful lines between two Velocity Indexes in the same color display the rank changes for the same data source across subject fields. According to these results, Twitter and Reddit are the most immediate data sources to newly published papers in all subject fields. By subject fields, the overall Velocity Indexes of all altmetric sources in *Physical Sciences and Engineering* (PSE) and *Life and Earth Sciences* (LES) are the highest. Facebook shows the higher immediacy degree in the fields of *Social Sciences and Humanities* (SSH) and *Mathematics and Computer Science* (MCS), although overall, the Velocity Index values of these subject fields are comparatively low. Conversely, news has relatively high Velocity Index in the fields of *Physical Sciences and Engineering*, *Life and Earth Sciences*, and *Biomedical and Health Sciences* (BHS), while it is slower in *Social Sciences and Humanities*. As to other data sources, they keep quite steady medium or low Velocity Indexes in all subject fields. For example, policy documents, peer review, and Q&A have the lowest Velocity Indexes across most subject fields, suggesting that these data sources are comparatively less focused on more recent papers as compared to the other sources regardless the subject fields of the papers.

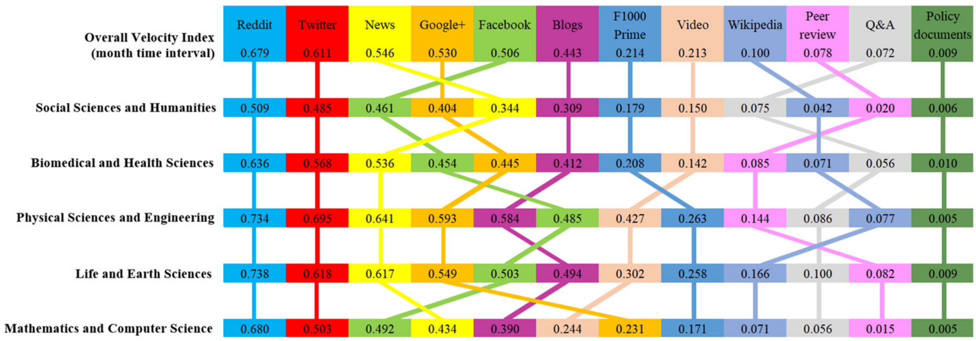

Figure 6. Velocity Index variations across the five subject fields

From the perspective of altmetric time delay, Figure 7 shows the distribution of altmetric time delay across the five subject fields for 12 Altmetric.com data sources respectively. For most data sources, although to different degrees, scientific papers in the fields of PSE and LES are faster to receive their first altmetric mention. In contrast, it took more days for papers in the fields of SSH and MCS to accumulate the first altmetric event record. Altmetric time delays of papers in BHS are in the middle on most data sources. Still, the accumulation velocity across subject fields in terms of altmetric time delay is similar with the results observed through the lens of Velocity Index.

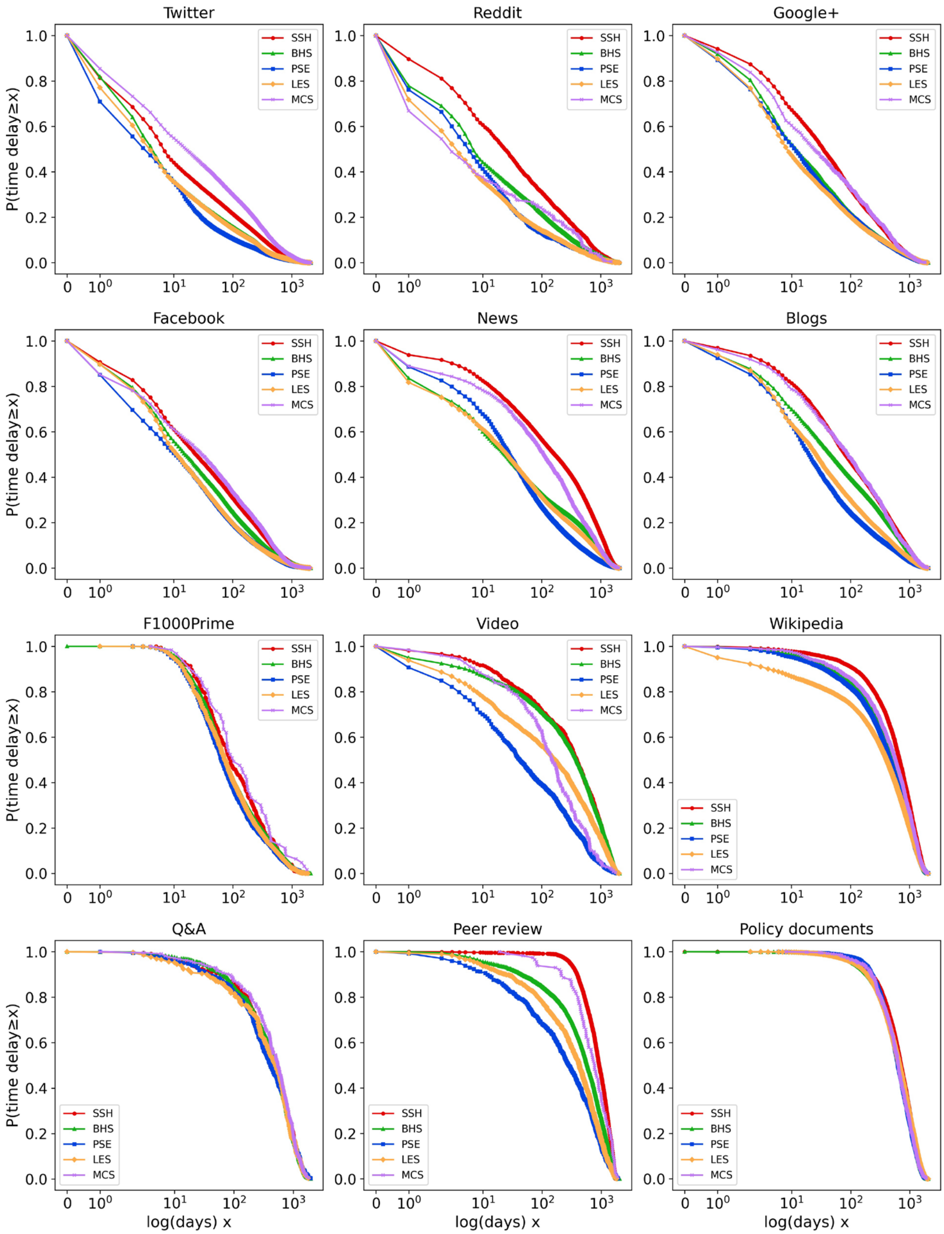

Figure 7. Distribution of altmetric time delays of 12 Altmetric.com data sources in log(days) across the five subject fields

3.3.5 Velocity Index variations across research topics

Considering the Velocity Index at the month time interval, we further investigate the variations across research topics to study which topics accumulated altmetric data faster than others. Twitter and Wikipedia are selected as two representatives for fast sources and slow sources, respectively, because they hold the largest data volume among their same types of data sources. Velocity Indexes are calculated for papers within each micro-level field sharing the similar research micro-topics based on Twitter mention data (Figure 8) and Wikipedia citation data (Figure 9). In both Figures 8 and 9, size of each circle is determined by the number of papers with Twitter mention/Wikipedia citation data in this micro-level field, while color is determined by the Velocity Index at the month time interval. Within micro-level fields, number of papers and number of altmetric events are very weakly correlated with the Velocity Index values based on Twitter data, and are moderately and positively correlated with those based on Wikipedia data (see Table 7 in Appendix), indicating that not all of micro-level fields with fewer papers are more likely to reach high Velocity Index, and vice versa. Some prominent research micro-topics with relatively high Velocity Index values in every main subject field are highlighted with annotation texts.

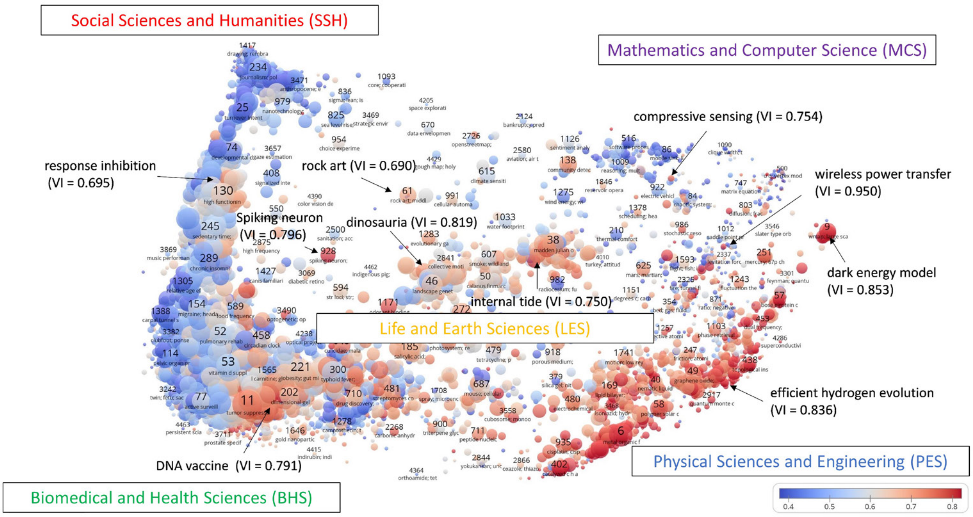

Figure 8. Velocity Index variations across research micro-topics (Twitter)

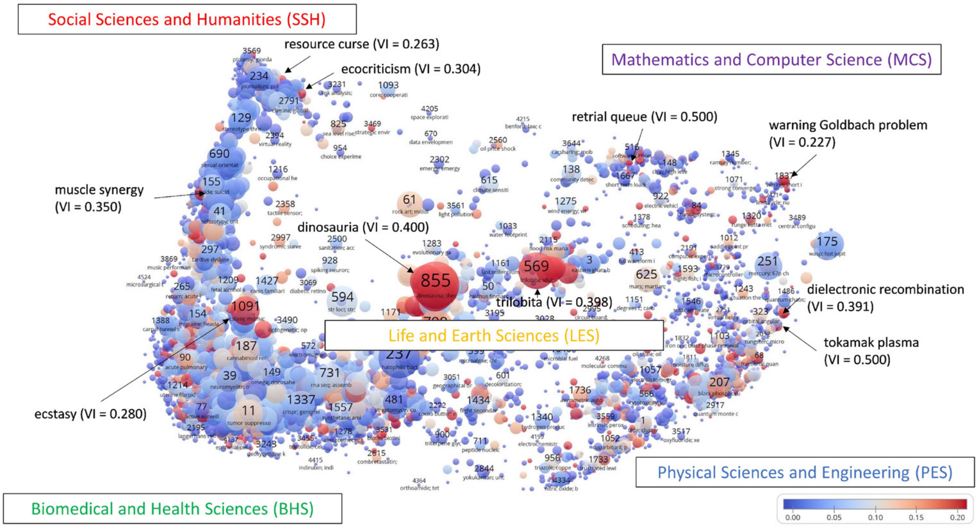

Figure 9. Velocity Index variations across research micro-topics (Wikipedia)

From the point of view of Twitter data, research micro-topics in the fields of PSE exhibit the highest Velocity Index values in contrast to the other fields, which is in correspondence with the above observations. Within the other subject fields, there are some research micro-topics that show quite high Twitter mention accumulation velocity as well. For example, “wireless power transfer” and “compressive sensing” in MCS accumulated the majority of their Twitter mentions in a short time, as well as “dinosauria” and “internal tide” in LES. In the fields of BHS and SSH, “DNA vaccine”, “spiking neuron”, “response inhibition”, and “rock art” drew attention on Twitter relatively fast too.

Compared to Twitter mentions, the overall accumulation velocity of Wikipedia citations is much lower, and the difference among main subject fields is not as obvious as Twitter. However, there also exist some research micro-topics showing higher data accumulation velocity. For instance, “dinosauria” and “trilobita” in LES are two micro-topics faster in Wikipedia. Papers about these two topics received more Wikipedia citations in a short time period compared to the others. Similarly, “ecstasy” (caused by drugs), “muscle synergy”, “warning Goldbach problem” and some other research micro-topics accumulate Wikipedia citations also relatively fast. In the field of SSH, although most research micro-topics were quite slow to be cited by Wikipedia, some environmental protection related micro-topics, such as “ecocriticism” and “resource curse”, show higher Velocity Index values.

3.4 Discussion

Speed has always been assumed as a characteristic property of altmetrics, however not much research has been done in characterizing the accumulation velocity of different altmetric data on a large scale. This study fills this gap by describing the immediacy of altmetric data accrued after the publication of scientific papers. Using the DOI created date and altmetric event posted date enables the possibility of studying the altmetric data accumulation patterns at the day level. The date when a DOI was assigned to a paper provided by Crossref has already been used to show the life cycle of some altmetric events at the month level by Ortega (2018b). This study investigates further on the accumulation velocity of various altmetric data at a more micro-level time interval and considering a larger scale of data samples.

As observed by Sun et al. (2016), citation histories typically show a pattern of just a few citations accrued within the first few years after publication, reaching a citation peak after 3-4 years, and then a decrease afterwards. Yet most kinds of altmetric data exhibit a different accumulation pattern compared with citations. We found that the accumulation velocity of different altmetric data vary substantially across data sources, document types, and subject fields.

3.4.1 Variations across altmetric data sources

It is demonstrated that various altmetric data sources vary in their data accumulation patterns, and the property of speed is not found to be owned by all of altmetric data sources. Some of the altmetric data sources accrue a considerable proportion of events very soon after the publication date of scientific papers. Among these outputs we have Reddit, Twitter, news, Facebook, Google+, and blogs. All these altmetric data sources exhibit short altmetric half-lives, short altmetric time delays, and relatively high Velocity Indexes. Therefore, it can be argued that their velocity aligns with the property of speed that altmetrics are expected to have, being possible to label these as fast sources. However, for policy documents, Q&A, peer review, Wikipedia, video, and F1000Prime events, only a very limited share of these altmetric events happened within a short time after publication, being these slow sources. The data accumulation velocity of some slow sources are similar to that of citations, with important delayed patterns after publication. For example, based on our dataset, half of policy document citations happened after 716 days since publication. Older papers, however, seem also to still be attractive for these slow data sources, so that their attention is not concentrated on just newly published scientific papers. As a whole, most social media platforms and mainstream media are more immediate in sharing, discussing, and reporting new research outputs.

Interestingly, different time windows may also show different sources as being fast or slow. For example, although F1000Prime is seen as a slow source in the short term (e.g., day or month level), it is one of the sources that accumulated the largest share of its events within 1 year. This reinforces the importance of combining different perspectives (e.g., different indicators, different time windows) to study the tempo of altmetrics to provide the most complete picture.

As a result, assumptions about the “speed” of types of events classified under the umbrella term “altmetrics” should be taken with particular caution. Not all of them are fast sources, and not all of them have the same accumulation pace. Thus, it is important to take the social media environment in which these events are produced into consideration (Alperin, 2015). Once again, caution about the merging of altmetric sources in compound metrics or global indicators must be observed, particularly considering that time affects differently to different sources. Keeping altmetric events separate seems to be an important recommendation, this given not only their fundamental differences (Haustein, Bowman, & Costas, 2016; Wouters et al., 2019) but also their time accumulation patterns as demonstrated in this study. Moreover, the pace and tempo of different altmetrics cannot be seen as equivalent and, similar to what happens with citations, these time differences need to be taken into account when considering different time windows in altmetric research.

3.4.2 Variations across document types

Zahedi et al. (2014) concluded that the coverage of several altmetric data sources varies across document types and subject fields. In this study, it is shown that the same type of variations applies also to the data accumulation velocity of different altmetric data sources. In terms of document types, Reviews (this document type mainly focuses on retrospectively reviewing existing findings) are overall the slowest in accumulating altmetric events. A possible reason for this slowest reception lies in the less innovative nature of Reviews. In other words, Review papers are less prone to provide new research discoveries and more to condense the state-of-the-art in a subject field or research topic, therefore lacking the novelty component of other document types. For example, the research topics presented in Editorial Materials and Letters may be more likely to evoke social buzz immediately, since they cover more novel topics, debates, scientific news, etc., without using a too complicated and technical language (Haustein, Costas, et al., 2015). The thematic property of these two document types might facilitate the users’ attention received more immediately, particularly on peer review platforms, a type of altmetric data source which is mainly used by researchers, who are faster to take notice of controversial topics emerging in the scientific community. This finding is quite similar with the ageing patterns of citations to different document types: Editorial Materials and Letters were found more likely to be the “early rise-rapid decline” papers with most citations accumulated in a relatively short time period, while Review was observed to be the delayed document type with a slower growth (Costas et al., 2010; J. Wang, 2013).

3.4.3 Variations across scientific fields and topics

In terms of scientific fields, research outputs from the fields of PSE and LES are more attractive to social media audiences shortly after publication, accruing altmetric events faster compared to other fields. Research outputs from the fields of both SSH and MCS are relatively slower to be disseminated on altmetric data sources, although papers in these two fields hold different altmetric data coverage, with the former much higher than the latter (Costas et al., 2015a; Fang, Costas, et al., 2020). Such field-related data accumulation dynamics was also observed in the context of citations, for instance, citation ageing in the social sciences and mathematics journals is similarly slower than in the medical and chemistry journals (Glänzel & Schoepflin, 1995), the physical, chemical, and earth sciences, fields in which the research fronts are fast-moving, have more papers showing rapidly declining citation pattern (Aksnes, 2003). From the perspective of first-citation speed, papers in the field of physics are faster in receiving the first citation, followed by biological, biomedical, and chemical research, while mathematics papers show lower first-citation speed (Abramo et al., 2011). Even though the overall accumulation patterns between citation data and most altmetric data are obviously different, they share very similar tempos across scientific fields.

Furthermore, the variations do not only exist at the main subject field level, but also the research topic level. Within each subject field, different research topics also show various velocity patterns in receiving altmetric attention, both on fast sources or slow sources. This signifies the thematic dependency of users in following up-to-date research outputs around some topics, just like some certain research topics drive more social attention over others (Robinson-Garcia et al., 2019). Thus, further research should focus on identifying the main distinctive patterns of papers and research topics to determine their faster/slower reception across altmetric sources, and how different observation time windows, and the selection of different data sources, may affect real-time assessment in altmetric practice.

3.4.4 Limitations

The main limitation of this study lies in the precision of Crossref's DOI created date as the proxy of actual publication date of scientific papers. There might still be a small distance between the date on which a DOI was created and the paper was actually made publicly available, which could result in some inaccuracies in our results. Besides, as we mentioned in the data part, DOI created dates might be updated due to the change of DOI status, thereby causing the unreliable time intervals. One of the effects of these inaccuracies is that some papers may have altmetric event posted date even earlier than DOI created dates. Therefore, papers with such unexpected time intervals have been excluded from this study to lower the negative influence made by questionable DOI created dates. Future research should focus on refining accurate methods of identifying the effective publication date of research outputs.

As shown in this study, they have important repercussion to determine accurate time windows for altmetric research.

3.5 Conclusions

Several conclusions can be derived from this study. First, we conclude that not all altmetrics are fast and that they do not accumulate at the same speed, existing a fundamental differentiation between fast sources (e.g., Reddit, Twitter, news, Facebook, Google+, and blogs) and slow sources (e.g., policy documents, Q&A, peer review, Wikipedia, video, and F1000Prime). Another important conclusion of this study is that the accumulation velocity of different kinds of altmetric data varies across document types, subject fields, and research topics. The velocity of most altmetric data of Review papers is lower than that of Articles, while Editorial Material and Letter are generally the fastest document types in terms of altmetric reception. From the perspective of scientific fields, the velocity ranking of different data sources changes across subject fields, and most altmetric data sources show higher velocity values in the fields of PSE and LES, and lower in SSH and MCS. Finally, with regards to individual research topics, substantial differences in the velocity of reception of altmetric events across topics have been identified, even among topics within the same broader field. Such topical difference in velocity suggests that it is worth studying the underlying reasons (e.g., hotness, controversies, scientific debates, media coverage) of why some topics within the same research area do receive social (media) attention much faster than others.

3.6 Appendix

Table 4. Descriptive statistics of altmetric events and papers mentioned by different altmetric data sources across the five main subject fields

Data source	SSH		BHS		PSE		LES		MCS	
	NP	NE	NP	NE	NP	NE	NP	NE	NP	NE
Reddit	3,349	5,258	18,137	22,554	1,814	2,405	4,462	5,906	244	316
Twitter	220,454	1,681,028	1,166,583	7,593,795	253,035	715,855	300,754	1,951,100	31,053	103,535
News	22,043	107,446	120,697	565,035	28,232	119,520	31,965	162,326	1,433	5,777
Google+	9,556	20,722	43,036	101,916	7,245	20,207	11,614	28,980	1,045	4,415
Facebook	54,884	107,947	305,928	774,662	44,676	79,992	75,564	198,963	5,143	7,105
Blogs	27,055	49,873	91,945	161,242	21,337	35,628	40,614	76,231	1,207	2,037
F1000Prime	361	424	34,422	41,976	1,018	1,214	3,674	4,335	63	70
Video	1,041	1,357	7,271	10,860	1,748	2,332	1,865	2,688	250	352
Wikipedia	7,379	9,588	35,517	47,549	9,003	13,783	17,806	28,290	1,491	1,882
Peer review	10,123	11,019	8,505	16,609	614	1,815	822	1,913	114	131
Q&A	225	239	1,109	1,149	380	397	293	299	214	232
Policy documents	10,001	13,368	32,519	42,244	1,802	2,136	8,862	11,816	549	650

Note: NP refers to number of papers with corresponding altmetric data; NE refers to total number of corresponding altmetric events.

Table 5. Spearman correlations for the rankings of altmetric data sources from the perspectives of Velocity Index at the month time interval, altmetric half-life, and altmetric time delay (proportion of altmetric events with altmetric time delays no more than 10 days as the benchmark)

Indicator	Velocity Index	Altmetric half-life	Altmetric time delay
Velocity Index	1.000	0.979	0.944
Altmetric half-life		1.000	0.937
Altmetric time delay			1.000

3

Table 6. Descriptive statistics of altmetric events and papers mentioned by different altmetric data sources across the four main document types

Data source	Article		Review		Editorial Material		Letter	
	NP	NE	NP	NE	NP	NE	NP	NE
Reddit	25,330	32,919	2,271	3,056	2,083	3,415	510	601
Twitter	1,752,296	10,359,094	179,941	1,463,407	130,161	1,866,897	48,355	272,291
News	183,647	884,239	17,353	59,932	14,283	58,626	3,740	17,269
Google+	62,463	153,770	8,826	20,556	8,991	21,786	1,435	2,433
Facebook	424,922	1,009,847	52,502	140,350	44,091	144,962	11,116	22,996
Blogs	162,979	293,559	17,684	28,813	12,724	23,720	2,349	4,066
F1000Prime	36,381	44,471	2,970	3,359	373	406	223	238
Video	10,619	14,873	1,448	2,553	550	733	133	200
Wikipedia	58,005	82,082	12,844	18,609	2,802	3,717	687	836
Peer review	19,706	31,171	535	847	276	1,030	137	340
Q&A	1,849	1,926	355	374	100	106	22	22
Policy documents	46,796	60,821	6,386	8,708	2,134	2,818	613	752

Note: NP refers to number of papers with corresponding altmetric data; NE refers to total number of corresponding altmetric events.

Table 7. Spearman correlations for micro-level fields (based on Twitter and Wikipedia data)

Twitter				Wikipedia			
	NP	NE	VI		NP	NE	VI
NP	1.000	0.949	0.031	NP	1.000	0.983	0.424
NE		1.000	-0.058	NE		1.000	0.432
VI			1.000	VI			1.000

Note: NP refers to number of papers with corresponding altmetric data in the micro-level field; NE refers to total number of corresponding altmetric events in the micro-level field; VI refers to the Velocity Index at the month time interval.

CHAPTER 4

User engagement with scholarly tweets of scientific papers: A large-scale and cross-disciplinary analysis[1]

[1] **This chapter is based on**:

Fang, Z., Costas, R., & Wouters, P. User engagement with scholarly tweets of scientific papers: A large-scale and cross-disciplinary analysis. (Under review).

Author contributions:

Fang, Z. (Conceptualization, Methodology, Formal analysis, Investigation, Visualization, Data Curation, Writing - Original Draft, Writing - Review & Editing)

Costas, R. (Conceptualization, Methodology, Investigation, Supervision, Writing - Review & Editing)

Wouters, P. (Conceptualization, Supervision, Writing - Review & Editing)

Abstract

This study investigates the extent to which scholarly tweets of scientific papers are engaged with by Twitter users through four types of user engagement behaviors, i.e., liking, retweeting, quoting, and replying. Based on a sample consisting of 7 million scholarly tweets of Web of Science papers, our results show that *likes* is the most prevalent engagement metric, covering 44% of scholarly tweets, followed by *retweets* (36%), whereas *quotes* and *replies* are only present for 9% and 7% of all scholarly tweets, respectively. From a disciplinary point of view, scholarly tweets in the field of *Social Sciences and Humanities* are more likely to trigger user engagement over other subject fields. The presence of user engagement is more associated with other Twitter-based factors (e.g., number of mentioned users in tweets and number of followers of users) than with science-based factors (e.g., citations and Mendeley readers of tweeted papers). Building on these findings, this study sheds light on the possibility to apply user engagement metrics in measuring deeper levels of Twitter *reception* of scientific information.

Keywords

Altmetrics, social media metrics, Twitter engagement, scholarly communication, retweet

4.1 Introduction

With the growing enthusiasm for sharing scientific information via Twitter, tweets mentioning scientific papers (hereinafter "scholarly tweets") are widespread. In the altmetrics realm, Twitter has arguably become one of the most crucial data sources, with more than one third of recent scientific papers being tweeted (Fang, Costas, et al., 2020). Scholarly tweets, therefore, have long been seen as measureable traces possibly capturing the impact of research outputs in a broader sense (Bornmann & Haunschild, 2016; Eysenbach, 2011).

Instead of merely serving as countable information carriers bringing scientific papers to the attention of Twitter users, scholarly tweets per se are also informative in terms of the content incorporated, the characteristics of users involved, as well as the possible user engagement triggered, collectively making Twitter a valuable source of social media metrics. In other words, the creation of scholarly tweets stands not only for an outcome of Twitter reception of science by users who posted them, but also a prologue of another narrative about how other users might interact with them in the Twitter universe, being relevant to quantitative elaboration of science-social media interactions (Costas et al., 2021).

4.1.1 Scholarly tweets as the objects of study

Díaz-Faes et al. (2019) proposed the umbrella term *secondary social media metrics* to conceptualize metrics taking "social media objects" (i.e., social media users and their online activities) as the objects of study, distinguishing them from *primary social media metrics* which focus on "research objects" (e.g., publications, datasets, journals, and individual scholars), in particular "the use and visibility of publications on social media". To date, in the direction of *secondary social media metrics*, many research efforts centering on scholarly tweets have been made to characterize the mechanisms of how Twitter users process, circulate, and engage with scientific information from different perspectives.

At the tweet level, content analyses provide straightforward insights into the tweeting behavior of users who are disseminating scientific information. For example, by scrutinizing the content of scholarly tweets received by the top ten most tweeted papers in the field of dentistry, Robinson-Garcia et al. (2017) exemplified the scarce existence of original thought but more mechanical nature of the bulk of tweet content. Similarly, with a case study containing 270 tweets, Thelwall, Tsou, et al. (2013) reported that the majority of the observed scholarly tweets only echoed a paper title or presented a brief summary. Regarding the sentiment of tweet texts, scholarly tweets were found to be generally neutral, with limited share showing positive or negative sentiment expressed by users (Friedrich et al., 2015; Thelwall, Tsou, et al., 2013). Besides, the use of some tweet features in scholarly tweets, such as hashtags (word or phrase prefixed with #) and user mentions (user's handle name prefixed with @), was also of interest by some altmetric research (Haustein, Bowman, Holmberg, Peters, et al., 2014; S. Xu et al., 2018), because it represents a particular form of user interactions enhancing the description and visibility of tweets and facilitating connections amongst users (Haustein, 2019; Holmberg et al., 2014).

At the user level, the presence of scholarly tweets makes it possible to recognize and characterize users discussing science on Twitter. Scholarly tweets, therefore, were drawn upon for identifying and classifying Twitter users participating in scholarly communication (Costas et al., 2020; Díaz-Faes et al., 2019; Vainio & Holmberg, 2017; Yu et al., 2019), and for further exploring how users by type performed differently while utilizing Twitter for scholarly communication (Didegah et al., 2018; Holmberg & Thelwall, 2014; Mohammadi et al., 2018; Yu, 2017). Moreover, the aforementioned objects derived from scholarly tweets, either at the tweet or user level, were not only studied separately, but sometimes networked in different ways to map the contexts in which Twitter interactions with science happened. The network methods include but not limited to co-occurrence of hashtags (Haunschild et al., 2019), co-occurrence of users and hashtags (Hellsten & Leydesdorff, 2020), user mentions network (Said et al., 2019), and follower/friend network of users (Alperin et al., 2019; Robinson-Garcia et al., 2018), which were collectively conceptualized as *heterogeneous couplings* by Costas et al. (2021).

4.1.2 User engagement behaviors around scholarly tweets

In addition to tweet content and user characteristics, user engagement behavior[1] around scholarly tweets is also a focal point of *secondary social media metrics*. Conceptually speaking, scholarly tweets offer the possibility for a wider range of users to participate in science-focused discussions through many engagement behaviors enabled by Twitter. In the current platform version (2021), Twitter provides several engagement functionalities for users to interact with tweets on their own initiative. As illustrated with a tweet example in Figure 1, there are four main types of engagement functionalities with corresponding metrics visible at the bottom of tweets and publicly retrievable through the Twitter API, including (1) *like*, (2) *retweet*, (3) *quote tweet*, and (4) *reply*. These engagement behaviors differ in both input and output. In terms of input, *liking* and *retweeting* are relatively basic and simple engagement behaviors because they are both devoid of extra original content added, whereas *quoting* and *replying* are comparatively more informative and conversational because they enable users to express original thought and content. As to output, except liking, the other

[1] We referred to the definition of “engagements” metrics interpreted by Twitter as “total number of times a user interacted with a tweet. Clicks anywhere on the tweet, including retweets, replies, follows, likes, links, cards, hashtags, embedded media, username, profile photo, or tweet expansion” (https://help.twitter.com/en/managing-your-account/using-the-tweet-activity-dashboard) (Accessed April 28, 2021). Therefore, in this study user engagement behavior refers to any interaction behavior performed by Twitter users on existing tweets.

three types of engagement behaviors can produce new tweets (i.e., retweets, quote tweets, and replies) which are accounted for users' total number of tweets posted.

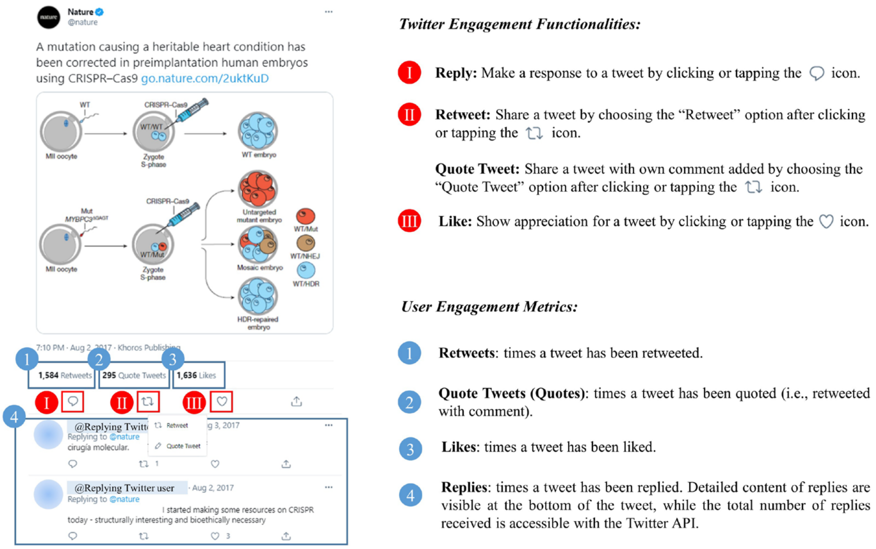

Figure 1. User engagement functionalities and metrics embedded in a tweet example

In existing altmetric literature, retweeting is the most analyzed user engagement behavior. In 2010, Priem & Costello (2010) found that retweets, as the outcomes of retweeting behavior, only made up 19% of a sample of scholarly tweets posted by 28 academic users. However, in more recent research, it was generally reported that retweets accounted for close to or over half of scholarly tweets of papers (Alperin et al., 2019; Didegah et al., 2018; Haustein, 2019), being a key component of the data base of studies related to scholarly Twitter metrics. For a sample of tweets posted by the Twitter accounts of 25 U.S. health agencies, Bhattacharya et al. (2014) found that about one third of them had zero retweet while the rest were retweeted at least once. As a form of information diffusion in nature, retweets were often analyzed to help capture topics of the public's interest in sharing (Bhattacharya et al., 2014; Kahle et al., 2016), or to construct Twitter dissemination networks of scientific knowledge across communities (Araujo, 2020; Hassan et al., 2019).

Besides retweeting, other types of user engagement behaviors, such as liking, replying, and clicking, were also studied to help yield insights into whether and how the public engages with scientific information on Twitter. For instance, considering a spectrum of user engagement metrics (e.g., retweets, likes, replies, clicks on tweeted URLs), Kahle et al. (2016)

studied the rates of user engagement with the tweets posted by the official Twitter accounts of the European Organization for Nuclear Research (CERN). Mohammadi et al. (2018) surveyed the motivations behind users' liking and retweeting behaviors in scientific contexts and reported that most survey respondents liked a tweet to "inform the authors that their tweets were interesting" and retweeted to disseminate the tweets. Based on the click metrics data provided by Bitly for its generated short links, Fang et al. (2021) analyzed the click rates on Twitter of short links referring to scientific papers and concluded that nearly half of the studied short links were not clicked by Twitter users at all.

4.1.3 Objectives

Although some attempts so far have been made to enhance the understanding of how people react to scholarly tweets, existing literature generally focused on either a specific user engagement behavior or a specific discipline. Little is known about the overall picture of the coverage of diverse types of user engagement with science on Twitter. Against this background, on the basis of a large-scale and cross-disciplinary dataset, the main research objective of this study is to systematically unravel the extent to which scholarly tweets are related to different categories of user engagement. Specifically, this study sets out to address the following research questions:

RQ1. To what extent are scholarly tweets engaged with by Twitter users through different types of engagement behaviors (i.e., liking, retweeting, quoting, and replying)?

RQ2. Which subject fields of science have scholarly tweets attracting higher levels of user engagement on Twitter?

RQ3. How does the presence of user engagement correlate with other science-based and Twitter-based factors of scholarly tweets (e.g., scholarly impact of tweeted papers, use of tweet features, user characteristics)?

4.2 Data and methods

4.2.1 Dataset

We retrieved a total of 6,229,001 Web of Science-indexed (WoS) papers published between 2016 and 2018 from the CWTS in-house database, and searched their scholarly tweets recorded by Altmetric.com until October 2019. For the matching with Altmetric.com data, WoS papers are restricted to those with DOI or PubMed ID assigned. On the whole, there are 1,999,199 WoS papers (accounting for 32.1%) with at least one scholarly tweet received, totally generating 7,037,233 unique original scholarly tweets.[2] Note that to explore user engagement behaviors, in this study the analyzed scholarly tweets are limited to original tweets which can be engaged with through the engagement functionalities provided by Twitter.

For the approximately 7 million scholarly tweets in our dataset, we retrieved their engagement metrics (i.e., number of likes, retweets, quotes, and replies received) with the Twitter API in February 2021.

4.2.2 The CWTS publication-level classification

To compare the user engagement situations of scholarly tweets across subject fields of science, we applied the CWTS publication-level classification system (Waltman & Van Eck, 2012) to assign scholarly tweets with subject field information based on their mentioned scientific papers. The CWTS classification clusters WoS papers into micro-level fields based on their citation relationships. These micro-level fields are then algorithmically assigned to five main subject fields, including *Social Sciences and Humanities* (SSH), *Biomedical and Health Sciences* (BHS), *Physical Sciences and Engineering* (PSE), *Life and Earth Sciences* (LES), and *Mathematics and Computer Science* (MCS).[3] For our dataset, there are a total of 5,932,279 scholarly tweets (accounting for 84.3%) referring to scientific papers with the subject field information assigned by the CWTS classification system. This set of scholarly tweets was drawn as a subsample for studying the subject field differences of user engagement. Table 1 presents the distribution of the analyzed scientific papers and scholarly tweets across the five subject fields of science.[4]

4

Table 1. Five subject fields of the CWTS publication-level classification system

Subject field	Abbreviation	Number of papers	Number of tweets
Social Sciences and Humanities	SSH	188,142	671,490
Biomedical and Health Sciences	BHS	968,605	3,544,755
Physical Sciences and Engineering	PSE	324,559	676,269
Life and Earth Sciences	LES	288,563	881,941
Mathematics and Computer Science	MCS	58,279	159,680

[2] We collected detailed Twitter information (e.g., tweet content and user demographics) in December 2019 for the tweet IDs provided by Altmetric.com (version: October 2019). Unavailable scholarly tweets caused by deletion of tweets, or suspension and protection of Twitter users' accounts (Fang, Dudek, et al., 2020) were not included in our dataset.

[3] See more introduction to the CWTS classification system (also known as the Leiden Ranking classification) at: https://www.leidenranking.com/information/fields (Accessed April 28, 2021).

[4] Full counting was applied for scholarly tweets that cite multiple papers belonging to different subject fields.

4.2.3 Science-based and Twitter-based factors of scholarly tweets

To explore how the presence of user engagement behaviors associates with scholarly, tweet, and user-related factors of the scholarly tweets, we extracted a total of ten factors from the following three dimensions: (1) *scholarly impact* of tweeted papers, (2) use of *tweet features*, and (3) *user characteristics* of those who posted scholarly tweets.

As listed in Table 2, in the dimension of scholarly impact of tweeted papers, we selected WoS citations and Mendeley readers to reflect the impact of tweeted papers in the science environment. In the dimension of tweet features, we focused on the use of hashtags and user mentions. Number of hashtags used, and number of users mentioned in tweets were analyzed to reflect how users edit their scholarly tweets with such interactive tweet features. Last but not least, in the dimension of user characteristics, we studied six factors related to users' demographics and behaviors on Twitter. Thus, number of followers and number of lists in which users are listed represent social media capital held by users, because these two factors largely affect how broad the audiences can be reached for posted tweets. Number of friends and number of likes given tell the story of how active users interact with other users by following others or liking their tweets. Number of tweets posted by users and their science focus depict users' overall tweeting behavior. The former indicates how frequent users post all kinds of tweets, while the latter implies how concentrated users are on tweeting scientific papers.

4

Table 2. Analyzed factors related to scholarly tweets

Dimension	Factor	Description
Scholarly impact of tweeted papers	Citations	Total number of WoS citations received by the papers mentioned in a tweet. Citation counts were retrieved from the CWTS in-house WoS database (version: March, 2020).
	Readers	Total number of Mendeley readers received by the papers mentioned in a tweet. Mendeley readership data were collected with the Mendeley API in July, 2020.
Tweet features	Hashtags	Number of hashtags used in a tweet.
	Mentioned users	Number of Twitter users mentioned in a tweet.
User characteristics	Followers	Number of Twitter users following a user.
	Lists listed	Number of lists in which a user is listed.
	Friends	Number of Twitter users followed by a user.
	Likes given	Number of likes given by a user since the account was created.
	Tweets posted	Number of tweets posted by a user since the account was created.
	Science focus	Proportion of scholarly tweets (recorded by Altmetric.com) among all tweets posted by a user. This indicator is equivalent to "ptws to papers" in Díaz-Faes et al. (2019). The higher the value of science focus of a user, the more concentrated the user is on tweeting scientific papers.

In the Results section, the correlations between the four analyzed user engagement metrics and the ten factors were studied based on not only the Spearman correlation analysis of counts (performed by R), but also the visualized change trend of the coverage of user engagement among scholarly tweets aggregated at the different levels of each studied factor (coverage of user engagement refers to proportion of scholarly tweets with at least one specific user engagement received).

4.2.4 Regression analysis: hurdle model

To further investigate how different factors can predict user engagement behaviors, we conducted regression analysis for each of the four user engagement metrics as a response variable. As presented later in the Results section, in consideration of that all of the four user engagement metrics are count data and the data are over-dispersed (data with the variance much greater than the mean value) and zero-inflated (data with excess zero values), we adopted hurdle models (Mullahy, 1986) as our regression models. Given that social media engagement data were generally found to be count data with the characteristics of zero-inflation and over-dispersion, hurdle models have been applied by many previous research to model user engagement on different social media platforms like Twitter (Bhattacharya et al., 2014), Facebook (Bhattacharya et al., 2017; Bohn et al., 2014), and Weibo (Fu & Chau, 2013).

Hurdle model is a two-part regression model that processes the count data as two separate components: one is the *zero portion* modeling whether an observation takes zero value or non-zero value (typically a binary logit model), while the other is the *count portion* fitting those non-zero values (a zero-truncated negative binomial model used in this study in consideration of the over-dispersion of the count data). In our case, the zero portion in the hurdle models determines whether a scholarly tweet gets at least one specific user engagement or not, while the count portion models how many times it is engaged with through certain behavior. Therefore, the hurdle models of user engagement metrics allow for the simultaneous modeling of both the likelihood for scholarly tweets of being engaged with, as well as the frequency of being engaged with by users. We employed the *pscl* package in R (Zeileis et al., 2008) to construct four hurdle models (mode 1: likes; model 2: retweets; model 3: quotes; model 4: replies).

4.3 Results

The Results section consists of four parts. The first part exhibits the overall presence of the four types of user engagement (i.e., likes, retweets, quotes, and replies) among the 7 million scholarly tweets in our dataset. The second part compares the presence of user engagement across scholarly tweets in different subject fields. The third part investigates how the presence of user engagement relates to different factors with respect to scholarly impact of tweeted papers, use of tweet features, and characteristics of users. The last part focuses on the hurdle regression of user engagement metrics.

4.3.1 Overall user engagement with scholarly tweets

Figure 2 illustrates the coverage of the four types of user engagement among the 7 million scholarly tweets. About 52% have been engaged with through at least one of the four analyzed engagement behaviors, namely, the overall coverage of user engagement among scholarly tweets is 52%. Around 20% of scholarly tweets were engaged with by users through only one type of engagement behavior, while as low as 2% got all the four types of engagement. More specifically, the coverage of likes is 44%, followed by retweets 36%. Liking and retweeting appear to be the most prevalent engagement behaviors around scientific information. In contrast, the coverage of both quotes and replies is relatively scarce. Only 9% of scholarly tweets got quoted by users, and as low as 7% received at least one reply.

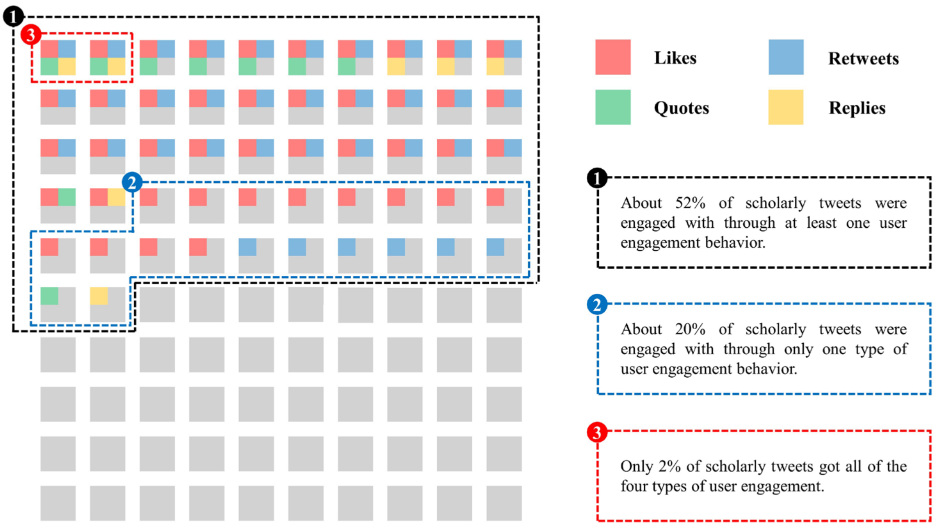

Figure 2. Coverage of the four types of user engagement. Each square represents 1% of scholarly tweets in our dataset. A square tinted with specific color(s) indicates that its represented 1% of scholarly tweets got corresponding type(s) of user engagement

Table 3 presents the descriptive statistics of the four engagement metrics to further reflect the extent to which scholarly tweets are engaged with. The coefficient of skewness and quartiles indicate that the distribution of all of the four types of engagement metrics is highly skewed. Only a few scholarly tweets got considerable user engagement, while the majority were never or rarely engaged with by Twitter users. Liking is the most widespread engagement behavior, contributing the most to user engagement metrics, followed by retweeting. On average, scholarly tweets in the dataset have been liked 2.95 times and retweeted 1.91 times. However, quoting and replying are more rare engagement behaviors, with only 1% of scholarly tweets being quoted for at least 3 times (99^{th} percentile of quotes is 3) or replied for at least 2 times (99^{th} percentile of replies is 2), suggesting that only a very limited share of scholarly tweets successfully aroused users' interest in sharing and communicating their thought within Twitter conversations.

Table 3. Descriptive statistics of the four user engagement metrics

Metrics	Sum	Mean	Min	Q1	Q2	Q3	90^{th} P	99^{th} P	Max	Skewness	SD
Likes	20,755,430	2.95	0	0	0	2	6	39	10,561	156.11	21.17
Retweets	13,429,713	1.91	0	0	0	1	4	26	9,983	218.44	16.89
Quotes	1,179,934	0.17	0	0	0	0	0	3	804	155.91	1.53
Replies	821,176	0.12	0	0	0	0	0	2	1,033	285.41	1.04

Note: Sum = total number of corresponding engagement metrics; Q1, Q2, Q3 = the first, second, and third quartile; 90th P = the 90th percentile; 99th P = the 99th percentile; Min, Max = the minimum and maximum value; Skewness = the coefficient of skewness; SD = standard deviation.

4.3.2 User engagement across subject fields

Figure 3 shows how the coverage of the four types of user engagement varies across the five subject fields of science: *Social Sciences and Humanities* (SSH), *Biomedical and Health Sciences* (BHS), *Physical Sciences and Engineering* (PSE), *Life and Earth Sciences* (LES), and *Mathematics and Computer Science* (MCS). Overall, scholarly tweets mentioning SSH papers are more likely to be engaged with through any type of engagement behavior studied. For the field of SSH, the proportion of scholarly tweets with at least one engagement record always ranks first, suggesting the higher probability for SSH-related scientific information to be engaged with by Twitter users over other subject fields. Besides, scholarly tweets from the fields of LES and BHS also present a relatively stronger potential in attracting different types of user engagement. In comparison, MCS and PSE are the two subject fields with sparser user engagement behaviors observed, showing the lowest coverage of all kinds of studied engagement metrics.

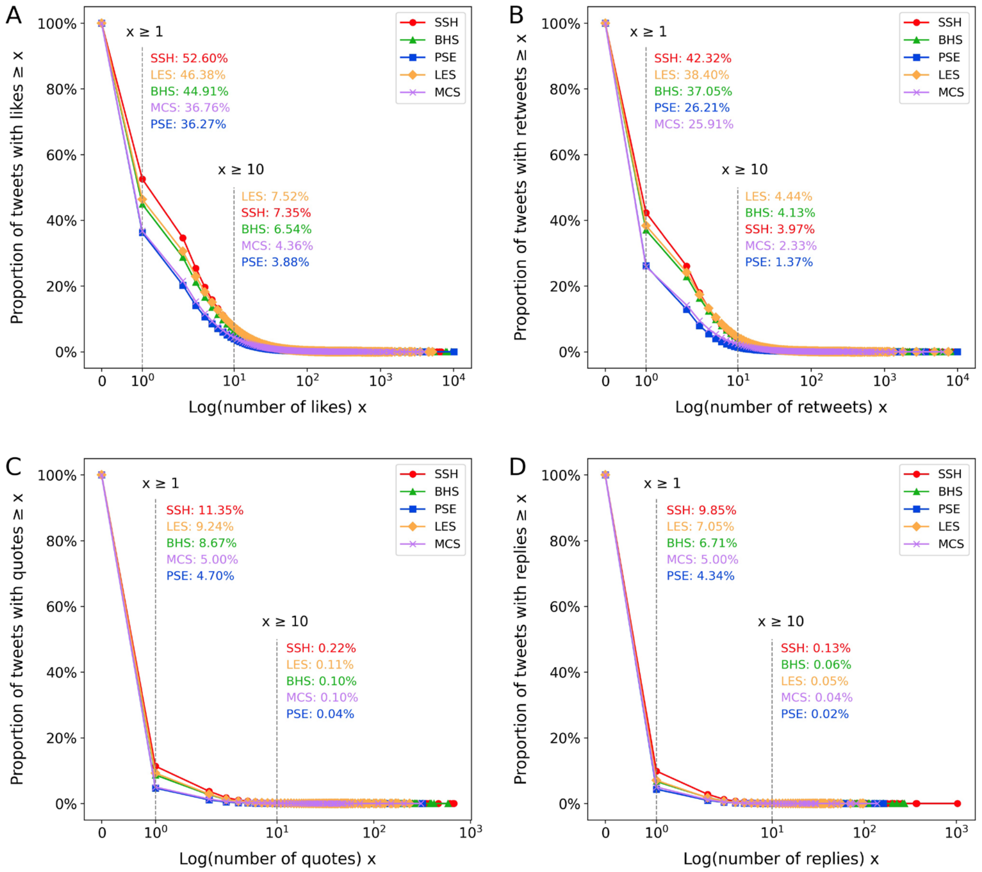

Figure 3. Proportion of scholarly tweets with different levels of **A** likes, **B** retweets, **C** quotes, and **D** replies across the five subject fields of science

The descriptive statistics of user engagement metrics across subject fields presented in Table 4 reinforces the disciplinary differences observed in Figure 3. Considering the greatest values of indicators highlighted in bold for each engagement metrics (if exists), SSH shows the most extensive distribution of all kinds of user engagement, thus acting as the most active subject field in giving rise to engagement with science on Twitter, followed by LES. BHS, as the subject field with the most scholarly tweets, contributes the most to the overall engagement metrics data due to the largest total number of corresponding engagement records. Besides, scholarly tweets of BHS papers also have a relatively higher presence of user engagement. However, user engagement is confirmed to be sparsely distributed among scholarly tweets in the fields of MCS and PSE.

Table 4. Descriptive statistics of the four engagement metrics across the five subject fields

Metrics	Field	Sum	Mean	Min	Q1	Q2	Q3	90th P	99th P	Max	SD
Likes	SSH	2,439,687	**3.63**	0	0	**1**	**3**	**7**	**43**	6,796	**29.03**
	BHS	**10,092,255**	2.85	0	0	0	2	6	38	9,336	17.66
	PSE	1,302,420	1.93	0	0	0	1	4	25	**10,105**	21.67
	LES	2,845,167	3.23	0	0	0	2	**7**	42	5,127	19.10
	MCS	370,053	2.32	0	0	0	1	4	32	3,904	22.73
Retweets	SSH	1,463,790	**2.18**	0	0	0	**2**	**5**	25	8,492	**24.06**
	BHS	**6,692,517**	1.89	0	0	0	1	4	25	8,317	15.59
	PSE	642,819	0.95	0	0	0	1	2	12	**9,983**	15.58
	LES	1,822,153	2.07	0	0	0	1	5	**26**	7,495	14.88
	MCS	209,015	1.31	0	0	0	1	2	19	6,255	20.70
Quotes	SSH	159,924	**0.24**	0	0	0	0	**1**	**4**	**673**	**2.49**
	BHS	**551,606**	0.16	0	0	0	0	0	3	584	1.22
	PSE	51,124	0.08	0	0	0	0	0	2	319	0.83
	LES	145,313	0.16	0	0	0	0	0	3	235	1.09
	MCS	15,316	0.10	0	0	0	0	0	2	308	1.31
Replies	SSH	121,918	**0.18**	0	0	0	0	0	**3**	**1,033**	**1.94**
	BHS	**388,313**	0.11	0	0	0	0	0	2	274	0.76
	PSE	43,566	0.06	0	0	0	0	0	1	161	0.58
	LES	95,471	0.11	0	0	0	0	0	2	98	0.63
	MCS	12,399	0.08	0	0	0	0	0	2	139	0.72

Note: Sum = total number of corresponding engagement metrics; Q1, Q2, Q3 = the first, second, and third quartile; 90th P = the 90th percentile; 99th P = the 99th percentile; Min, Max = the minimum and maximum value; SD = standard deviation.

4.3.3 Correlation analysis of user engagement metrics

In this part, to study how different types of user engagement metrics correlate with the factors listed in Table 2, through the lens of each factor, we graphically show the coverage of specific user engagement of scholarly tweets aggregated at the different levels of the corresponding factor, and further interpret the observed relationships by combining the results of the Spearman correlation analysis between studied factors and user engagement metrics at the tweet level.

From the perspective of scholarly impact of tweeted papers, Figure 4 plots the change trend of the coverage of user engagement with the increase of **A** citations and **B** Mendeley readers of tweeted papers. Overall, the coverage of all kinds of user engagement is slightly higher for scholarly tweets mentioning papers with higher levels of citation counts and Mendeley readers accrued, although the uptrends are not that strong, particularly for citations. According to the Spearman correlations (see Figure 7 in Appendix), the four types of user engagement metrics are all positively but negligibly correlated with citations and readers (the coefficient r_s ranges from 0.016 to 0.048 for citations, and ranges from 0.051 to 0.107 for readers).

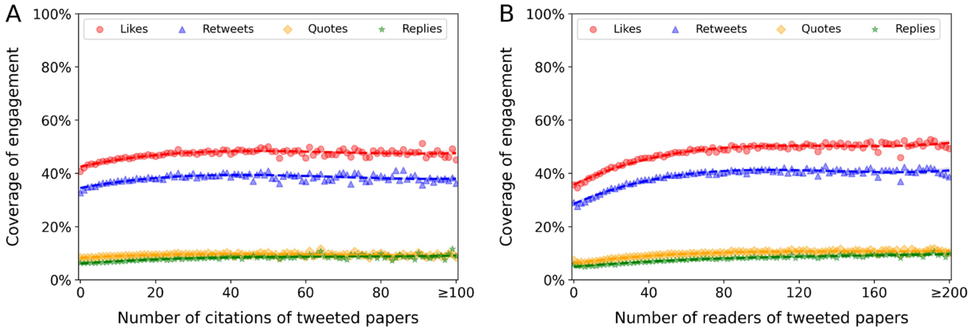

Figure 4. Coverage of the four types of user engagement among scholarly tweets with different levels of **A** WoS citations and **B** Mendeley readers received by tweeted papers

Regarding tweet features used in scholarly tweets, Figure 5 shows the coverage of user engagement when different **A** numbers of hashtags are used, and different **B** numbers of users are mentioned in tweets. These two tweet feature factors present different patterns in their relationships with user engagement. As the number of hashtags per tweet increases, a slight rise can be observed in the coverage of likes, retweets, and quotes, but not for replies. This is confirmed by the positive and negligible correlations found between number of hashtags and number of likes, retweets, and quotes received by tweets (r_s ranges between 0.042 and 0.113), whereas nearly no correlation found between number of hashtags and number of replies (r_s = -0.001). By comparison, the uptrend of the coverage of user engagement is stronger with the increasing number of mentioned users in tweets, especially for the coverage of likes and retweets. For scholarly tweets with more than two users mentioned, their likelihood of being liked exceeds 80% and the likelihood of being retweeted reaches 70%. Similarly, the coverage of both quotes and replies is relatively higher for scholarly tweets with more users mentioned than those without any mentioned users. Correspondingly, the Spearman correlations between user engagement metrics and number of mentioned users are comparatively stronger than other factors mentioned earlier, particularly for likes (r_s = 0.237) and retweets (r_s = 0.229).

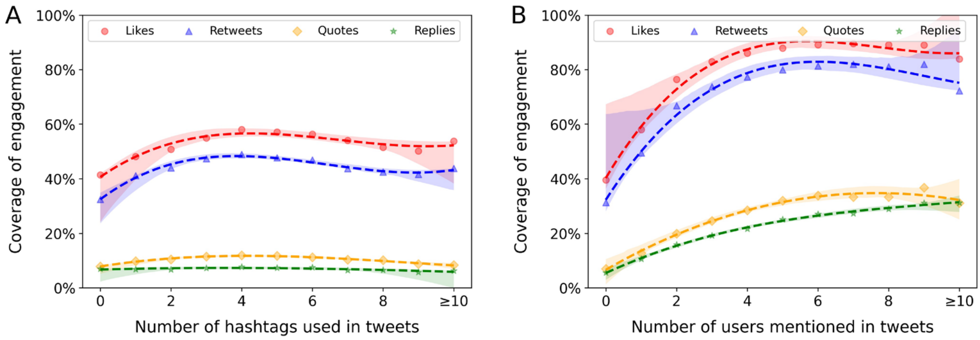

Figure 5. Coverage of the four types of user engagement among scholarly tweets with different numbers of **A** hashtags used and **B** users mentioned in tweets

In terms of user characteristics, Figure 6 shows the coverage of user engagement for scholarly tweets posted by users with different characteristics. Specifically, Figure 6**A** and 6**B** exhibit the change trend of the likelihood of being engaged with for scholarly tweets from users with different levels of followers and lists listed, respectively. These two factors, which to a large extent determine the audience size of posted tweets, are positively associated with the coverage of user engagement. The more followers that users accrue or the more lists that users are listed in positively correlate to the higher probability for their tweets of getting engagement. Based on the Spearman correlation analysis, number of followers of users is moderately correlated with both number of likes and retweets received (r_s = 0.349 and 0.368, respectively).

Figure 6**C** and 6**D** show the relations between user engagement and the two factors about users' interaction activity: number of friends and number of likes given. Similarly, these two factors also keep positive relationships with the coverage of user engagement. Overall, scholarly tweets posted by more active users (who interact more frequently with others by following other users and liking other users' tweets) tend to show a relatively higher probability to be engaged with. The Spearman correlation analysis proves that there exist weak to moderate correlations between user engagement metrics and the two factors about users' interaction activity (r_s ranges from 0.086 for the correlation between number of friends and number of quotes up to 0.313 for the correlation between number of likes given and number of likes received).

Different from the patterns observed for the above factors, as shown in Figure 6**E** and 6**F**, number of tweets posted and science focus of users, which depict the overall tweeting activity of users, show negative relationships with the coverage of user engagement among scholarly tweets. In general, the greater number of tweets posted, as well as the stronger science focus of users, the lower levels of coverage of user engagement. These negative relationships are reinforced by the negative coefficients reported in the Spearman correlation analysis between these factors and user engagement metrics (r_s ranges from -0.147 to -0.028).

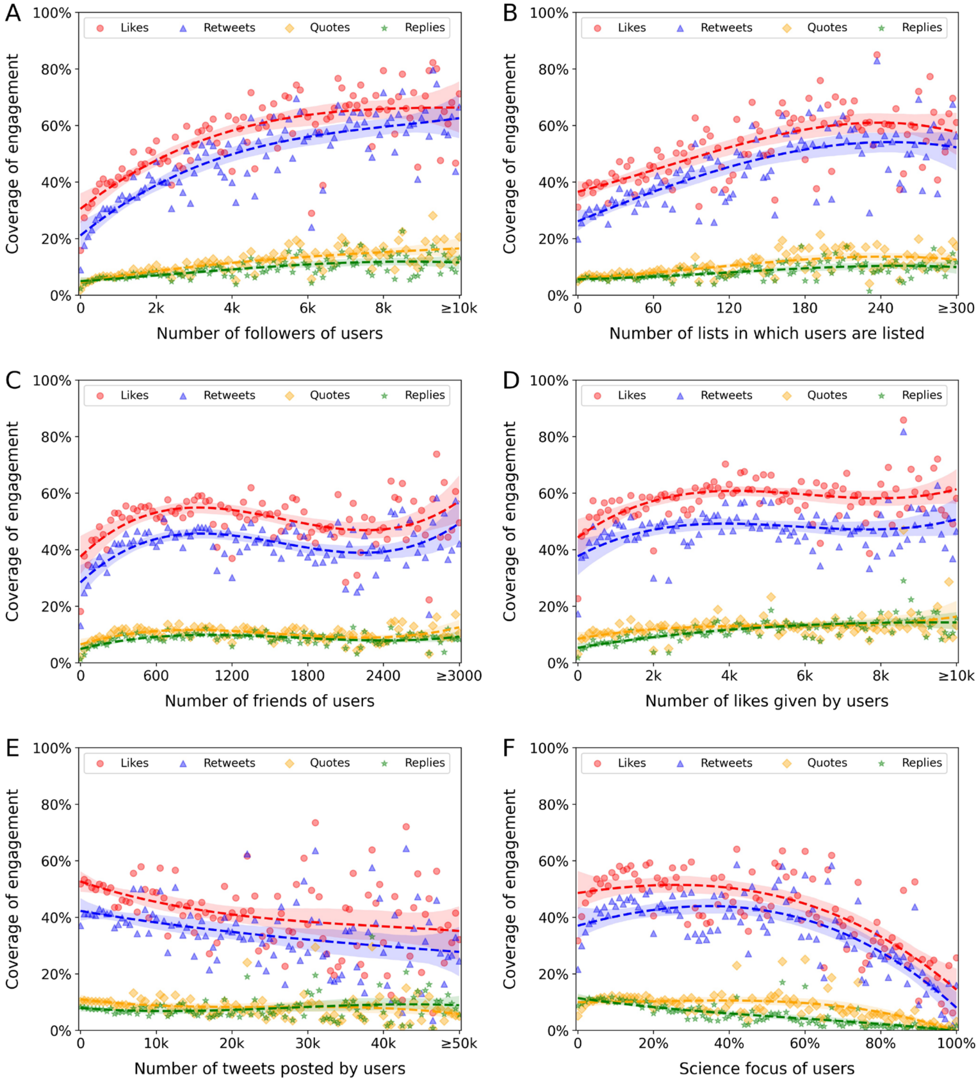

Figure 6. Coverage of the four types of user engagement among scholarly tweets posted by users with different levels of **A** followers, **B** lists listed, **C** friends, **D** likes given, **E** tweets posted, and **F** science focus

4.3.4 Regression analysis using hurdle models

To further compare how different science-based and Twitter-based factors serve as predictors of user engagement behaviors, we conducted regression analyses of the four types of user engagement using hurdle models. In order to avoid multicollinearity in the regression models, in each of the three dimensions discussed above, we selected several representative factors as the explanatory variables. For instance, in the dimension of scholarly impact of tweeted papers, since citations and Mendeley readers are strongly correlated with each other (r_s = 0.712), we selected citations as one of the explanatory variables. In the dimension of tweet features, both number of hashtags used and number of mentioned users were included because they are weakly correlated (r_s = 0.181). In the dimension of user characteristics, number of followers keeps moderate to strong correlations with other homogeneous user factors (r_s ranges from 0.433 to 0.859) except science focus of users (r_s = 0.015), so we adopted number of followers and science focus as representative variables in this dimension.

Table 5 reports the results of the zero portion of the four hurdle models of user engagement metrics (logit models): model 1 (likes), model 2 (retweets), model 3 (quotes), and model 4 (replies). Some variables were log-transformed for better model fitting. The zero portion of the models reflects how the selected explanatory variables relate to the likelihood for scholarly tweets of attracting at least one specific user engagement. The four models present a similar pattern in the zero portion, with citations, mentioned users and followers positively associated with the likelihood of getting at least one corresponding user engagement, whereas science focus presents a negative association. The exception is hashtags: in model 1 (likes), mode 2 (retweets) and model 3 (quotes), number of hashtags has a positive relationship with the likelihood that at least one like, retweet or quote occurs, however, in model 4 (replies), number of hashtags presents a negative association. The odds ratios (OR, exponent of regression coefficient in logit model) were calculated to help interpret the results. For example, in model 1 (likes), while all other variables remaining constant, a one-unit increase in the log-transformed number of mentioned users increases the odds of getting at least one like by 185.6% (OR = 2.856). However, while holding all other variables constant in model 1 (likes), a unit increase in science focus decreases the odds of getting at least one like by 42.7% (OR = 0.573).

Table 6 reports the results of the count portion of the hurdle models of user engagement metrics (zero-truncated negative binomial models). The count portion focuses on those scholarly tweets with at least one corresponding user engagement received and indicates how the explanatory variables associate with the increase of the frequency of user engagement. As is evident in all the four models, citations, mentioned users as well as followers are all positively associated with the frequency of user engagement, yet hashtags and science focus are negatively associated with the frequency of user engagement. Similarly, incidence rate ratios (IRR, exponent of regression coefficient in negative binomial model) were computed to help interpret the coefficient of a given variable while holding all other variables constant. For instance, in model 1 (likes), while all other variables remaining constant, a unit increase in the log-transformed number of followers increases the rate of receiving a like by a factor of 1.484, while a unit increase in the log-transformed number of hashtags decreases the rate of obtaining a like by a factor of 0.908.

Table 5. Results of the zero portion of the hurdle models of user engagement metrics

Variable	Model 1 (likes)		Model 2 (retweets)		Model 3 (quotes)		Model 4 (replies)	
	Estimate (SE)	OR	Estimate (SE)	OR	Estimate (SE)	OR	Estimate (SE)	OR
Citations (log-transformed)	0.058 (0.001)	1.059	0.060 (0.001)	1.062	0.030 (0.001)	1.030	0.058 (0.001)	1.060
Hashtags (log-transformed)	0.100 (0.002)	1.105	0.164 (0.002)	1.178	0.005 (0.003)	1.005	-0.245 (0.003)	0.782
Mentioned users (log-transformed)	1.049 (0.002)	2.856	1.035 (0.002)	2.815	0.866 (0.003)	2.378	0.791 (0.003)	2.205
Followers (log-transformed)	0.383 (0.000)	1.466	0.446 (0.001)	1.562	0.389 (0.001)	1.475	0.308 (0.001)	1.360
Science focus	-0.557 (0.003)	0.573	-0.378 (0.003)	0.685	-0.582 (0.005)	0.559	-2.228 (0.007)	0.108

Note: All estimates are significant at the 0.000 level. SE = Standard error. OR = Odds ratio (exponent of estimate in logit model).

Table 6. Results of the count portion of the hurdle models of user engagement metrics

Variable	Model 1 (likes)		Model 2 (retweets)		Model 3 (quotes)		Model 4 (replies)	
	Estimate (SE)	IRR	Estimate (SE)	IRR	Estimate (SE)	IRR	Estimate (SE)	IRR
Citations (log-transformed)	0.068 (0.001)	1.070	0.137 (0.001)	1.147	0.105 (0.002)	1.110	0.009 (0.002)	1.009
Hashtags (log-transformed)	-0.096 (0.002)	0.908	-0.067 (0.002)	0.936	-0.182 (0.005)	0.833	-0.117 (0.006)	0.890
Mentioned users (log-transformed)	0.661 (0.003)	1.936	0.475 (0.003)	1.609	0.238 (0.005)	1.268	0.189 (0.006)	1.208
Followers (log-transformed)	0.395 (0.001)	1.484	0.395 (0.001)	1.484	0.369 (0.001)	1.446	0.339 (0.002)	1.403
Science focus	-1.122 (0.005)	0.326	-0.922 (0.005)	0.398	-0.950 (0.011)	0.387	-1.595 (0.017)	0.203

Note: All estimates (coefficients) are significant at the 0.000 level. SE = Standard error. IRR = Incidence rate ratio (exponent of estimate in negative binomial model).

4.4 Discussion

As discussed by Brossard and Scheufele (2013), in the era of mass media, science stories as well as their accuracy, importance and popularity are no longer just "presented in isolation but instead are embedded in a host of cues that accompany nearly all online news stories", such as comments on blog posts and user engagement on social media. Such cues, according to Brossard and Scheufele (2013), "may add meaning beyond what the author of the original story intended to convey". In the context of scholarly Twitter metrics, this argument, on the one side, highlights the importance of the examination of user engagement in studying science-social media interactions, but on the other side, poses a question about how many scholarly tweets indeed triggered user engagement which are believed to contain extra meaning added to science stories.

Although user engagement with scholarly tweets have long been seen valuable for characterizing the interactions between scholarly objects and social media (Wouters et al., 2019), there is still an overall lack of evidence which can be drawn upon to mirror how effectively scholarly tweets attract the public's attention and further stimulate public engagement in Twitter conversations around science. Based on a large-scale and cross-disciplinary dataset, this study unravels the coverage of diverse types of user engagement among scholarly tweets, thus offering an answer to the question about the overall presence of public engagement with scientific information on Twitter.

4.4.1 Overall presence of user engagement with scholarly tweets

As conceptualized by Haustein, Bowman, & Costas (2016) in the context of *primary social media metrics* with scholarly objects as the objects of study, they classified acts referring to scholarly objects to three main categories, including *access*, *appraise*, and *apply*. *Access* refers to acts that involve "accessing and showing interest in the research objects", such as viewing and downloading a scientific paper. *Appraise* stands for acts of "mentioning the research objects on various platforms" like blogs and social media. *Apply* includes acts of "using significant parts of, adapting, or transforming the research objects", such as thoroughly discussing a scientific paper in a blog post or citing it in papers. Therefore, *apply* represents the highest level of engagement with research objects, followed by *appraise* and then *access*. Following this framework, we applied it in the context of *secondary social media metrics* (Díaz-Faes et al., 2019), in which social media users and their online activities become the objects of study, instead of the scholarly objects as in Haustein, Bowman, & Costas (2016). Correspondingly, in the specific case of Twitter engagement metrics, *access* would indicate acts of accessing and showing interest in the scholarly tweets and their constitutive elements, such as viewing a tweet and adding a tweet to bookmarks. *Appraise* would refer to acts of commending and further disseminating the scholarly tweets, such as liking a tweet or retweeting it. Those tweets liked or retweeted by a user would be displayed on the user's homepage and have notifications sent to other users involved in the tweets (e.g., authors of the tweets, and users mentioned in the tweets). Lastly, *apply* would include acts of participating in discussions and expressing views based on the scientific information tweeted, for example, retweeting with extra comments added (i.e., quoting) or making a response to a tweet (i.e., replying), which would contribute to the creation of another Twitter form of engagement that is a *conversation*[5]. The category of *apply* also has records publicly visible on actors' homepage and have notifications sent to users involved, and more importantly, such acts will generate more information (and possibly subsequent engagement) that the original scholarly tweets may not contain. From the standpoint of scholarly tweets, the level of engagement increases from *access* over *appraise* to *apply* as well.

It's not surprising to find that as the level of engagement grows, the coverage of user engagement behavior becomes lower. In this study, with likes, retweets, quotes and replies as the traces of user engagement, we found that likes and retweets, as the acts of *appraise* with the moderate level of engagement, were present for about 44% and 36% of the studied scholarly tweets, respectively. However, the coverage of quotes and replies, the two behaviors with the highest level of engagement (i.e., *apply*), is as low as 9% and 7%, respectively. The globally low presence of user engagement, particularly for the engagement behaviors with more informative outcomes generated, reveals the fact that the attention paid to scholarly tweets varied a lot on the one hand, and puts more emphasis on the significance of more in-depth measurement of Twitter reception of scientific papers on the other hand. For papers with exactly the same number of scholarly tweets accumulated, although the papers' Twitter reception appears to be equal only based on their absolute number of scholarly tweets, those with scholarly tweets being widely engaged with might be disseminated and perceived on Twitter in a more effective way. This is because engagement behaviors provide concrete evidence that they reached out to audiences who also showed further interest.

Moreover, the presence of user engagement differs by subject field. As the subject fields found to be more frequently mentioned in the Twittersphere (Costas et al., 2015a; Haustein, Costas, et al., 2015), SSH, LES, and BHS also have their scholarly tweets more actively engaged with by users through liking, retweeting, quoting, and replying, outperforming the fields of PSE and MCS. Behind the consistency of SSH, LES, and BHS shown in the vitality in the Twitter environment, there are multiple possible reasons such as the lay audiences' preference for topics related to social issues, environmental problems, and healthcare (Haustein, Costas, et al., 2015; Haustein, Peters, Sugimoto, et al., 2014), and the higher degree of Twitter uptake by scholars from these fields (Costas et al., 2020; Mohammadi et al., 2018). In addition, it has been reported that scholars from the field of social sciences and

[5] As defined by Twitter (https://help.twitter.com/en/using-twitter/twitter-conversations) (Accessed April 28, 2021), a conversation on Twitter is composed of an original tweet and its replies, as well as replies to those replies.

humanities more frequently communicate their research with the public as an important audience (Bentley & Kyvik, 2011) and more often devote to popularization activities than scholars from natural sciences and technology (Kreimer et al., 2011), which can be partly explained by the "less strict demarcation between internal scientific and public communication and between scientific and general knowledge" existed within social sciences and humanities than natural sciences (H. P. Peters, 2013). This might be another possible reason for the superiority of SSH in obtaining further engagement. To further interpret the subject field differences, future research is needed to scrutinize the contexts in which user engagement takes place (e.g., engaging users' identity and motivations) across subject fields.

4.4.2 Factors related to user engagement with scholarly tweets

On the basis of both correlation analysis and regression analysis, we investigated the relationships between user engagement and a spectrum of science-based and Twitter-based factors. Similar to previously reported weak or no correlations between citations and tweeting activities (Bardus et al., 2020; Zahedi et al., 2014), we found that user engagement with scholarly tweets was also negligibly correlated with scholarly impact factors of tweeted papers (i.e., citations and Mendeley readers), thus adding more empirical evidence to the idea that science and social media have different concerns about research outputs and conform to different spaces of engagement (Fang et al., 2021).

In contrast to science-based factors, Twitter-based factors generally tend to be more related to user engagement. Specifically, from the perspective of tweet features, although hashtags and user mentions are both tweet features increasing the visibility of tweets, the former is utilized to label and broadcast tweets to potential users interested in the same topics, while the latter is targeted to specific users with notifications delivered to them, showing a more conversational nature than the former. As a result, number of users mentioned in tweets is more related to user engagement and more effective in predicting user engagement. From the perspective of user characteristics, both users' social media capital (i.e., followers and lists listed) and interaction activity (i.e., friends and likes given) were positively correlated with user engagement around their tweets. Nevertheless, flooding the screen (i.e., too many tweets posted) and attaching to tweeting only scientific papers (i.e., too strong science focus) were found to be related to lower levels of user engagement. From a practical point of view, as suggested by Cheplygina et al. (2020) for scientists getting start on Twitter, building a community by interacting with others, as well as sharing something personal and non-academic can also be relevant to get support in science communication on Twitter.

This study took into account scholarly tweets written in all languages. With this we provide a relatively complete picture of user engagement regardless of the language of the tweets. However, more detailed tweet content analysis should be applied, considering different linguistic contexts, as well as more local topics and sentiment, which were not included in this study. Given that specific tweet content like those including awe-inspiring imagery and newsworthy items frequently attract high levels of engagement across social media platforms (Kahle et al., 2016), future research should also focus on how different tweet content might be related to subsequent user engagement in scholarly contexts.

4.4.3 Implications for social media studies of science

As an important part of the Twitter information ecosystem, user engagement behaviors leave digital traces of wider public interactions with science, thereby allowing both for investigation of online scholarly communication and civic participation in science-focused discussions, and for exploration of the deeper levels of Twitter reception of science from the standpoint of broader social media audiences. Correspondingly, the implications of studying user engagement for more advanced social media studies of science are two-fold.

On the one hand, in terms of science-social media interactions, user engagement provides additional information beyond what is delivered by scientific papers and original Twitter mentions, especially for those behaviors with higher levels of engagement such as quoting and replying. These engagement behaviors act as sources of information on how users communicate science in non-academic environments and how the public at large receives these messages related to science. For those scholarly tweets with informative commentaries or conversations, they may offer valuable evidence to develop a more comprehensive understanding of science-social media interactions.

On the other hand, in terms of impact measurement on Twitter, the presence of scholarly tweets has been regarded as an important indicator of social (media) attention paid to research outputs (Sugimoto, Work, et al., 2017). However, it only reflects the attention of users who brought scientific information to Twitter, but neglects the attention of those Twitter audiences who engaged with this scientific information through diverse engagement behaviors. User engagement metrics would then capture a more deep-seated reception of science in the Twitter universe, complementing the "science stories" in the social media environment (Brossard & Scheufele, 2013). Therefore, including user engagement metrics in the altmetric toolkit might open a novel window to characterize the popularity of research outputs. This argument, although based on the study of scholarly tweets, can be generalized to other altmetric data sources with user engagement metrics available (e.g., likes and shares on Facebook posts, views on YouTube videos) or potential (e.g., readers of blog posts, comments in news media platforms).

4.4.4 Limitations

There are several limitations in this study. First, there are more than four types of engagement behavior that users can take to interact with scholarly tweets, such as clicking on tweeted scholarly URLs, clicking on users' profile, and adding tweets to bookmarks. However, these engagement metrics are currently not publicly accessible on a large scale, and they were not included in this study. Should these engagement metrics be made publicly retrievable in the future, a more complete picture of user interaction behavior around science could be drawn. Second, although reply tweets and quote tweets, which are outcomes of replying and quoting behavior, can be further engaged with through Twitter functionalities as well (e.g., liking or retweeting replies and quote tweets), they were not included in the analyzed dataset to avoid double counting. Lastly, we explored how the characteristics of engaged users (i.e., users whose tweets are engaged with) relate to user engagement, while the characteristics of engaging users (i.e., users who engaged with tweets) are also of relevance to the interpretation of the occurrence of engagement behavior. However, due to the lack of detailed information of the engaging users, their demographics and behavioral patterns were not analyzed in this study.

4.5 Conclusions

This study contributes to the expansion of the idea of *secondary social media metrics* (Díaz-Faes et al., 2019) by presenting a large-scale and cross-disciplinary analysis of four types of user engagement (i.e., liking, retweeting, quoting, and replying) around scholarly tweets. Of the 7 million scholarly tweets in our dataset, 52% were engaged with by Twitter users through at least one engagement behavior. Likes and retweets are most widespread, with the highest coverage among scholarly tweets (44% and 36%, respectively). In contrast, the coverage of quotes and replies is much lower (9% and 7%, respectively), suggesting the overall low presence of user engagement amongst Twitter mentions of scientific papers, particularly for those behaviors with higher levels of engagement needed. Scholarly tweets from the fields of SSH, LES, and BHS tend to have more frequent user engagement distributed. Finally, the presence of user engagement is more related to other Twitter-based factors (mentioned users in tweets and number of followers of users in particular) than with science-based factors of papers (e.g., citations and Mendeley readers), implying both the intrinsically connected dynamics of Twitter elements and the distinguishing focuses between scientific and tweeting activities.

Our findings provide a first overview of the extent to which scholarly tweets are related to broader public engagement with science on Twitter, thereby paving the way towards the measurement of Twitter reception of science in a more interactive and comprehensive manner. Based on the exploratory results presented in this study, a series of research questions emerge, which will need to be examined in much greater detail, such as the motivations and behavioral patterns of engaging users, the differential aspects that increase the social media capital of Twitter users (e.g., by increasing their number of followers), and what topic-related factors (e.g., controversial topics) embodied in tweets can be related to be triggering more effective forms of engagement and Twitter communication. All in all, delving into user engagement behaviors may help delineate the role that Twitter plays in facilitating public understanding of science as well as the meaning that Twitter-based indicators may have in research evaluation and science communication.

4.6 Appendix

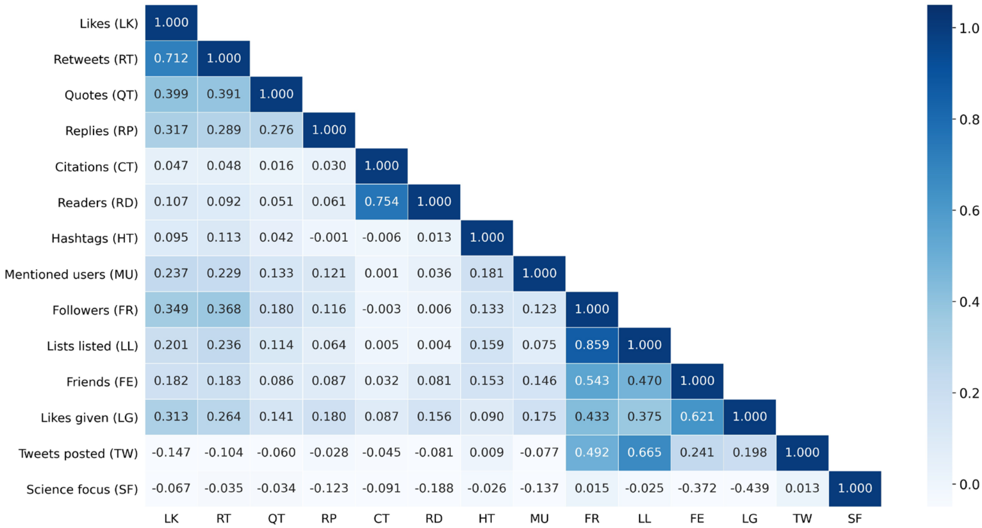

Figure 7. Spearman correlation analysis of the four user engagement metrics and studied factors

CHAPTER 5

The stability of Twitter metrics: A study on unavailable Twitter mentions of scientific papers[1]

[1] **This chapter is based on**:

Fang, Z., Dudek, J., & Costas, R. (2020). The stability of Twitter metrics: A study on unavailable Twitter mentions of scientific publications. *Journal of the Association for Information Science and Technology*, *71*(12), 1455–1469. https://doi.org/10.1002/asi.24344

Author contributions:

Fang, Z. (Conceptualization, Methodology, Formal analysis, Investigation, Visualization, Data Curation, Writing - Original Draft, Writing - Review & Editing)

Dudek, J. (Conceptualization, Methodology, Data Curation, Writing - Review & Editing)

Costas, R. (Conceptualization, Methodology, Investigation, Supervision, Writing - Review & Editing)

Abstract

This study investigated the stability of Twitter counts of scientific papers over time. For this, we conducted an analysis of the availability statuses of over 2.6 million Twitter mentions received by the 1,154 most tweeted scientific papers recorded by Altmetric.com up to October 2017. The results show that of the Twitter mentions for these highly tweeted papers, about 14.3% had become unavailable by April 2019. Deletion of tweets by users is the main reason for unavailability, followed by suspension and protection of Twitter user accounts. This study proposes two measures for describing the Twitter dissemination structures of papers: *Degree of Originality* (i.e., the proportion of original tweets received by an article) and *Degree of Concentration* (i.e., the degree to which retweets concentrate on a single original tweet). Twitter metrics of papers with relatively low Degree of Originality and relatively high Degree of Concentration were observed to be at greater risk of becoming unstable due to the potential disappearance of their Twitter mentions. In light of these results, we emphasize the importance of paying attention to the potential risk of unstable Twitter counts, and the significance of identifying the different Twitter dissemination structures when studying the Twitter metrics of scientific papers.

Keywords

Twitter metrics, altmetrics, data stability, Twitter unavailability rate, Twitter dissemination structures

5.1 Introduction

Twitter has become one of the most important dissemination tools of scientific information and scholarly communication, used not only by the scientific community, but also by the public in general (Kahle et al., 2016; Van Noorden, 2014). Twitter is also one of the most predominant altmetric data sources for scientific papers (Haustein, 2019; Robinson-Garcia et al., 2014). Several studies have discussed aspects of data coverage, density, and intensity (Haustein, Costas, et al., 2015; Thelwall, Haustein, et al., 2013), or the accumulation velocity of tweets to papers (Fang & Costas, 2020). It is assumed that Twitter mentions, as well as other types of social media metric data, are more likely to measure a broader impact of research that differs from the academic impact reflected by citations (Bornmann, 2015b; Robinson-Garcia et al., 2018). Therefore, Twitter metrics are usually calculated with the motivation of further application in research assessment and science policy (Haustein, 2019; Wilsdon et al., 2015). In this context, the stability of metrics can be seen as a key component of data quality, being of great significance for a reasonable and sustainable measurement of the reception and discussion of research outputs on Twitter.

5.1.1 Development of Twitter metrics

5

The characteristics of altmetric data, such as broadness, speed, openness, and transparency (Wouters & Costas, 2012), have raised expectations towards the development of alternative indicators that can measure research impact in an early stage following publication (Priem & Hemminger, 2010). As a result, numerous studies have analyzed the correlation between various altmetric indicators and citation-based indicators, testing whether the former might be applied for predicting highly cited articles – which otherwise is impaired owing to the citation delay (Costas et al., 2015a; Priem, Piwowar, et al., 2012; Waltman & Costas, 2014; Zahedi et al., 2014, 2017). As a source that contributed a considerable share of data about online activities associated with scholarly outputs only second to Mendeley (Sugimoto, Work, et al., 2017), Twitter has been widely discussed in previous research. There the impact of scientific papers on Twitter was usually measured by counting the total number of mentions they received or the total number of Twitter users who mentioned them in their tweets. These two counting methods of Twitter metrics are commonly employed by altmetric data aggregators.

In spite of this strong interest on the dissemination of scientific papers on Twitter, the calculation of Twitter metrics is not free of challenges and limitations (Haustein, 2016, 2019). *Heterogeneity*, which refers to the diversity of acts and online events (Haustein, 2016), is one of the biggest challenges for altmetrics. Heterogeneity is not only observable across altmetric data sources in general, but appears in the reception of scientific papers on Twitter in particular. For example, there are various actions users can take to interact with scholarly content on Twitter, such as originally tweeting, retweeting, replying, or liking tweets mentioning papers, among others (Haustein, Bowman, & Costas, 2016). There are multiple heterogeneous forms of co-occurrence that can happen in a single tweet, like hashtags, mentioned users, or URLs (Costas et al., 2021). Hence, when a Twitter mention is accrued, it is not just a simple number, but entails a multitude of information that refers to the different forms of interaction and exchange of information on Twitter. This lack of internal homogeneity (Wouters et al., 2019) of Twitter metrics represents both a challenge as well as an opportunity, as it makes possible the further exploration of underlying patterns and user motivations (Sud & Thelwall, 2014) in their Twitter interactions with scientific papers.

Therefore, researchers are increasingly paying attention to the content analysis of Twitter mentions and the behavioral analysis of Twitter users, going beyond the mere counting of tweets linking to scientific papers (Bornmann, 2014a; Haustein, 2019). Twitter users' identities, motivations, and related interactions or engagement behaviors have been analyzed in order to improve the understanding of Twitter metrics in a much more fine-grained manner (Díaz-Faes et al., 2019; Holmberg et al., 2014; Mohammadi et al., 2018). Nevertheless, rethinking the tweeting patterns and Twitter user behaviors in more detail comes with worries and problems that have aroused the concern of researchers. By scrutinizing the patterns of tweeting of the top-10 most tweeted scientific dental articles, Robinson-Garcia et al. (2017) observed the mechanical nature of the bulk of tweeting behavior. This indicated that Twitter metrics based on simple counting of tweets runs the risk of conflating multiple issues related to the tweeting activity, like obsessive single-user tweeting, duplicate tweeting, bots, and even human tweeting, but devoid of original thought or engagement of the user with the article in the tweet (Robinson-Garcia et al., 2017). Related concerns about Twitter data quality can be found in other studies as well (Haustein, Bowman, Holmberg, et al., 2016; Thelwall, Tsou, et al., 2013).

5.1.2 Consistency of altmetric data

Data consistency is of great concern in studies of altmetric data. As Wouters et al. (2019) pointed out, among the characteristics of altmetric data, transparency and consistency are particularly essential for new indicators to be used for research evaluation. The lack of consistency is seen as one of the most noteworthy data quality challenges that all altmetric indicators have to confront (Haustein, 2016). Related research questions have been discussed from both the conceptual and empirical perspectives, since article-level metrics emerged and were offered by several data providers with different data collection and integration principles (Chamberlain, 2013; Sutton, 2014; Zahedi et al., 2015).

Considering the strong dependency of altmetric data on commercial data providers, previous studies mainly focused on the consistency of various altmetric data among different data aggregators. For example, Ortega (2018a) analyzed the coverage differences amongst Altmetric.com, PlumX, and Crossref Event Data. These three altmetric data providers performed differently in each metric due to technical errors and extracting criteria; therefore, strategies of using specific services for particular metrics and combining different services for overall analysis were recommended. Meschede and Siebenlist (2018) also made a comparison between Altmetric.com and PlumX. They found that these two data aggregators achieved a moderate correlation overall but showed relatively weak consistency in some metrics, like Google+, Facebook, and news mentions. Zahedi and Costas (2018) presented an exhaustive study on the differences of data collection and reporting approaches among four major altmetric data providers, including Altmetric.com, PlumX, Lagotto, and Crossref Event Data. Similar results were found and further explored in their study. More specifically, values of each metric provided by the different data aggregators obviously differed from each other because of their specific choices for the data collection and aggregation approaches. In a case study on the altmetric performance of articles published in *Journal of the Association for Information Science and Technology* (JASIST) reported by Altmetric.com, PlumX, and Mendeley, the inconsistencies of metrics across data providers were observed by Bar-Ilan et al. (2019) in the same manner. Taken together, these results show that the data inconsistency at the data aggregator level is an important concern within the altmetric research community.

Moreover, as explained by Chamberlain (2013), altmetric data can be collected at different times, which potentially can also end up in obtaining different values of social media metrics, even when collected from the same source and for the same set of papers. This is one of the explanations for the differences in the data collected by different aggregators (Zahedi & Costas, 2018).

In this article we introduce a different form of altmetric data inconsistency, related to the ever-changing nature of social media data, in which data records and social media events can easily be deleted by their creators, or users may abandon the social media platforms, removing all their records from the platform. This form of inconsistency is therefore more related to the *stability* of altmetric data, and more specifically, of the Twitter metrics of publications. To the best of our knowledge, research on this type of inconsistency of Twitter metric data, as well as on their underlying causes, is still lacking in the social media metrics literature. In this article we intend to fill this gap through a large-scale study of Twitter counts of papers collected at different times, focusing also on conceptualizing the potential reasons and risks that the observed instability may pose for the consistent calculation of Twitter metrics.

5.1.3 Objectives

The main objectives of this study are: (1) to investigate the stability of Twitter metrics by identifying Twitter mentions that have become unavailable over time and (2) to explore the potential influence that these unavailable tweets may have on the overall Twitter metrics of papers. We addressed the following specific research questions:

RQ1. What is the number and share of Twitter mentions of highly tweeted scientific papers in Altmetric.com that have become unavailable over time?

RQ2. What are the most common reasons for tweets becoming unavailable?

RQ3. To what extent do unavailable Twitter mentions influence the temporal stability of Twitter metrics of scientific papers?

RQ4. Based on papers' unique Twitter dissemination structures consisting of original tweets, retweets, and retweeting links, is it possible to determine which scientific papers are at a higher risk of substantially decreased Twitter metrics when tweets become unavailable?

5.2 Data and Methods

5.2.1 Distribution of Twitter mention data recorded by Altmetric.com

The Twitter mention data of scientific papers used in this study were extracted from the historical data files provided by Altmetric.com in 2017. Until October 2017, Altmetric.com has tracked and recorded nearly 43 million Twitter mentions for around 5.4 million unique scientific papers (namely, Altmetric IDs). Altmetric.com provides two main indicators for measuring Twitter impact of scientific papers. One is the total number of tweets to the article (TWS), the other is the number of unique Twitter users who have mentioned the article (NUTU). Here we employ NUTU to present the distribution of Twitter mention data. Figure 1**A** and 1**B** shows the NUTU distribution of all scientific papers recorded by Altmetric.com under a log-log scale and its probability density function (PDF), respectively. Distributions of several kinds of bibliometric data, such as citations (Brzezinski, 2015) and usage counts (X. Wang, Fang, & Sun, 2016), have been found to follow typical power law distributions, which is also observed in Figure 1**B** for Twitter mention data. Figure 1**B** is visualized based on the Python *powerlaw* package (Alstott et al., 2014), the distribution of unique Twitter users fits a power law distribution with $\alpha = 2.87$. Only a few scientific papers have attracted a large number of unique Twitter users, while the Twitter counts of most scientific papers are relatively low. In order to examine the stability of Twitter metrics, 1,154 scientific papers with at least 1,000 unique Twitter users (NUTU $\geq$1,000) were selected as our research objects. Until October 2017, these were the most tweeted scientific papers from the perspective of unique Twitter users involved, showing the highest impact on Twitter.

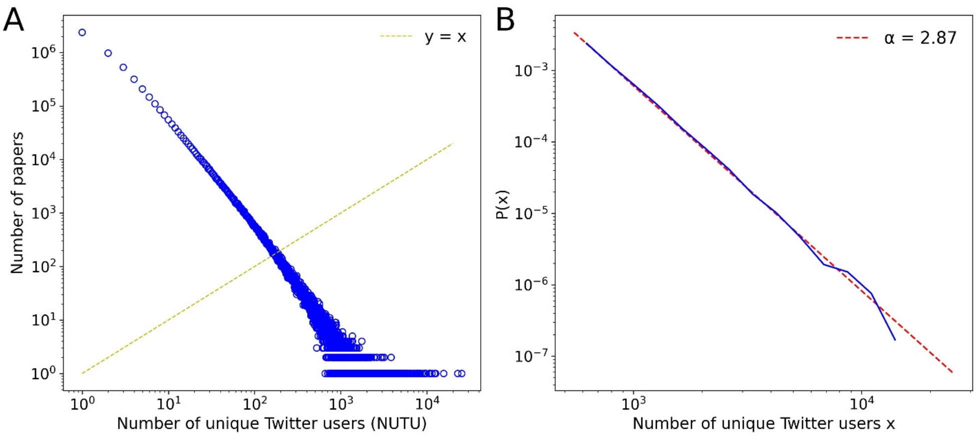

Figure 1. **A** distribution and **B** PDF of the number of unique Twitter users (NUTU)

5.2.2 Availability of Twitter mentions of the most tweeted scientific papers

Since 2016, Altmetric.com has made annual snapshots of its database available for researchers to study. These snapshots serve as an important reference point to study tweets that became unavailable at a later point in time. The snapshot data still provide evidence that an article was tweeted even in the case when the tweet has been removed from Twitter, although the content and details of the tweet are not available any longer. For the 1,154 scientific papers with Twitter mentions posted by at least 1,000 unique Twitter users, all the tweet IDs (unique identifier of tweets) were collected from the data files provided by Altmetric.com (version: October 2017). In total, there were 2,643,531 unique tweet IDs related to the selected papers.

On the basis of the tweet IDs previously identified by Altmetric.com, in April 2019 we rechecked all the tweets through the Twitter API in order to examine of which tweets the status changed. For all tweets that were still available, detailed metadata can be acquired, and for those tweets that are no longer available, the Twitter API responds with respective error codes and error messages. Both unavailable tweet IDs and their error codes were recorded for further analysis. For the 2,643,531 Twitter mentions recorded by Altmetric.com until October 2017, a total of 378,766 (14.3%) were unavailable by April 2019.

5.2.3 Indicators for describing Twitter dissemination structure

In order to provide some understanding of the influence that unavailable tweets can have for the calculation of Twitter metrics, we study the *Twitter dissemination structures* of scientific papers. Twitter dissemination structure refers to the dissemination form of research outputs on Twitter over time, which is composed of *original tweets*, *retweets*, and the *retweeting links*. Original tweets are defined as Twitter mentions of scientific papers originally posted by Twitter users; retweets refer to the re-dissemination of original tweets by Twitter users; finally, the term retweeting links refers to the relationship between a specific original tweet and its following retweets, which is established when an original tweet is retweeted. The Twitter dissemination structure reveals how many original tweets an article has accrued, how many retweets each original tweet has received, and how these original tweets and retweets connect over time.

As discussed before, a common Twitter metric for a scientific paper is the total count of tweets it has accumulated. However, the dissemination process of a scientific paper on Twitter is too intricate to be explained with a simple number. Studying the Twitter dissemination structures of scientific papers on Twitter can be seen as a more advanced approach to characterize the Twitter diffusion of scientific papers. *Originality* and *Concentration* are proposed as two dimensions for describing Twitter dissemination structures, which are based on the varieties that can be observed with scientific papers' original tweets, retweets, and their connections (i.e., retweeting links). Figure 2 illustrates four hypothetical examples of original tweet and retweet combinations in order to explain the two main dimensions for describing Twitter dissemination structures of papers. Blue nodes and yellow nodes represent original tweets and their related retweets, respectively. The four papers in the example (paper A, B, C, and D) do all have the same total number of Twitter mentions (TWS = 10). From the perspective of total tweet counts they show the same impact on Twitter, but they perform differently through the lens of Originality and Concentration.

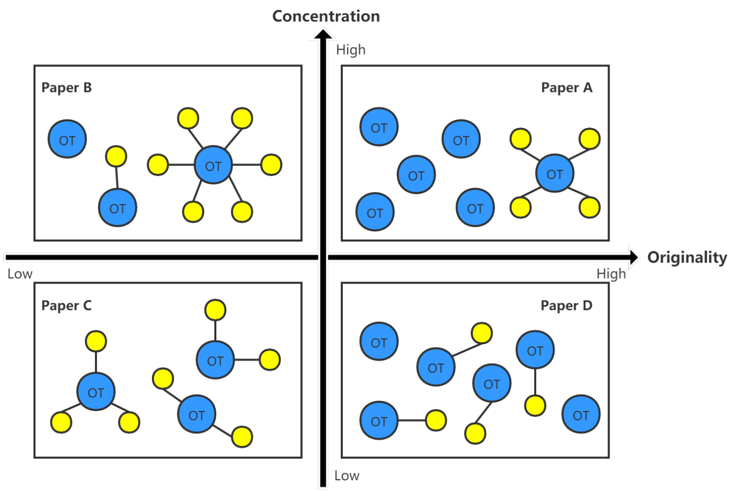

Figure 2. Two dimensions for describing Twitter dissemination structures of papers, resulting in four different diffusion scenarios

Originality is proposed to represent how many Twitter mentions of a specific scientific paper are posted originally by Twitter users rather than retweeting previous tweets. The more original tweets a paper has, the higher its degree of originality. The *Degree of Originality* (DO) of paper *x* is calculated as follows:

$$Degree\ of\ Originality_x = \frac{N(OT_x)}{TN(OT_x + RT_x)}$$

Where $N(OT_x)$ denotes the number of original tweets that paper *x* has received, while $TN(OT_x + RT_x)$ refers to the total number of Twitter mentions (including all original tweets and retweets) that paper *x* has accumulated. Essentially, DO reflects the proportion of original tweets a paper has received. In Figure 2, paper A (DO = 0.6) and paper D (DO = 0.6) fall into the category that has accumulated more original tweets, while paper B (DO = 0.3) and paper C (DO = 0.3) belong to the category that has received fewer original tweets.

5

Concentration is proposed to show the extent to which a paper's retweets are linked to its most retweeted original tweet. The more retweets concentrate on the most retweeted original tweet, the higher the paper's degree of concentration. The *Degree of Concentration* (DC) of paper *x* is given by:

$$Degree\ of\ Concentration_x = Max\left(\frac{N(RT_{OT_i})}{TN(RT_x)}\right)\ (i = 1,2,\dots,n)$$

Where $N(RT_{OTi})$ denotes the number of retweets that the original tweet *i* (*i* = 1, 2, …, *n*) for paper *x* has received, $TN(RT_x)$ denotes the total number of retweets that paper *x* has accumulated. DC reflects the maximum percentage of retweets linking to (at least) a single original tweet. The higher the maximum percentage, the higher proportion of retweets concentrate on a single original tweet, while a low maximum percentage reflects a more disperse distribution of retweets. For papers without any retweet, their DCs are zero by default. For each paper in Figure 2, the proportions of retweets that every original tweet received are calculated and the maximum one is the DC of that paper. Therefore, the DCs of paper A and paper B are 1.0 and 0.86, respectively, with most retweets of these two papers concentrating on a certain original tweet; while for paper C (DC = 0.43) and paper D (DC = 0.25), the retweets are distributed dispersedly.

All Twitter dissemination structures can be classified into the four categories in Figure 2 based on the two dimensions of Originality and Concentration. In order to study the dissemination structures of the highly tweeted papers selected for this study, their original tweets and retweets were distinguished at first. For Twitter mentions that are still available, the collected metadata indicate whether a tweet is an original tweet or a retweet, and in case it is a retweet, the tweet ID of its corresponding original tweet is returned as well, so that the retweeting links between original tweets and retweets can be identified. For Twitter mentions that are not available on Twitter any more, their status of original tweet and retweet, and their original tweet–retweet connections were established based on the data recorded by Altmetric.com, whenever this was possible. It should be noted that for some retweets, the corresponding original tweets are not always identified and recorded by Altmetric.com. Given that, in principle, the existence of a retweet relies on a corresponding original tweet, a possible explanation for the omission of original tweets is that during the data collection process by Altmetric.com, some retweets were identified and recorded first, and then the original tweets become unavailable before Altmetric.com could identify them, and therefore they were not included in the Altmetric.com data file. In those cases, we assumed that the original tweet must have existed at some point before the retweet. For the retweets without corresponding original tweets recorded, their original tweets are assumed for the sake of creating the retweeting links. Although these *assumed o*riginal tweets do not contribute to the total number of Twitter mentions of papers, they are considered to co-establish the Twitter dissemination structures of papers.

5.3 Results

The Results section consists of three main parts: The first one explains the major reasons for the unavailability of Twitter mentions and shows the distribution of unavailable Twitter mentions over the years. The second part presents the influence of unavailable Twitter mentions on Twitter metrics of scientific papers and explores the possible causes for the highly unstable Twitter metrics through a case study. The last part focuses on the potential risks for papers with different Twitter dissemination structures of being unstable in Twitter metrics.

5.3.1 Distribution of unavailable Twitter mentions

Table 1 presents the number of unavailable Twitter mentions arranged by the specific *error codes* directly provided by the Twitter API. There are four main error codes that signal the unavailability of Twitter mentions. The major reason for the unavailability is that the tweet has been deleted, with around 54.7% of unavailable Twitter mention records falling into this category. The second major reason is that the Twitter user accounts have been suspended because of a violation against Twitter rules,[1] leading to the unavailability of all their tweets. This accounts for 25.9% of all errors returned and is followed by the protection of tweets implemented by users.[2] Once a Twitter user has chosen this setting, unauthorized users cannot get access to their tweets (anymore), although the tweets themselves still exist. During our data collection, this error was found in the case of 16.7% of all unavailable Twitter mentions. Lastly, 2.7% of unavailable tweet IDs could not be found because the tweet IDs were directing to a page that does not exist anymore (e.g., users deactivated accounts). It should be noted that in those cases where the tweet IDs are no longer existent (error codes 144), the related Twitter mentions about scientific papers are unrecoverable. Concerning unavailable tweet IDs due to user suspension, deactivation, or tweet protection (error codes 63, 34, and 179), it is still possible that they become available to the public again once the suspended user accounts are unlocked, the deactivated accounts are reactivated, or the users cancel the protection of their tweets. Nevertheless, whether such reversion will take place is uncertain, thus the unavailability of these tweet IDs still has a negative effect on the stability of the Twitter metrics.

Table 1. Numbers of unavailable Twitter mentions and reasons for their unavailability

Error code	Twitter Error message	Description	N	P
144	No status found with that ID.	The requested Tweet ID is not found (if it existed, it was probably deleted).	207,147	54.7%
63	User has been suspended.	The user account has been suspended and information cannot be retrieved.	98,194	25.9%
179	Sorry, you are not authorized to see this status.	Thrown when a Tweet cannot be viewed by the authenticating user, usually due to the Tweet's author having protected their Tweets.	63,393	16.7%
34	Sorry, that page does not exist.	The specified resource was not found.	10,032	2.7%
Total			378,766	100.0%

Altmetric.com started tracking Twitter data from October 2011 onwards (Altmetric, 2020). Figure 3 shows the distribution of the Twitter mentions of the 1,154 most tweeted scientific papers over the years, as well as of the unavailable Twitter mentions. Each bar in Figure 3 presents the total number of Twitter mentions with posting date information every year, and the percentage of unavailable Twitter mentions is represented by the lined segments in the

[1] See more information about suspended Twitter accounts at: https://help.twitter.com/en/managing-your-account/suspended-twitter-accounts (Accessed December 17, 2019).

[2] See more information about public and protected tweets at: https://help.twitter.com/en/safety-and-security/public-and-protected-tweets (Accessed December 17, 2019).

bars, and numerically listed in brackets. Older Twitter mentions (e.g., from years 2011, 2012, or 2013) exhibit higher proportions of unavailable tweets, suggesting that the longer the time between the tweet and the data collection, the larger the chances of finding unavailable tweets.

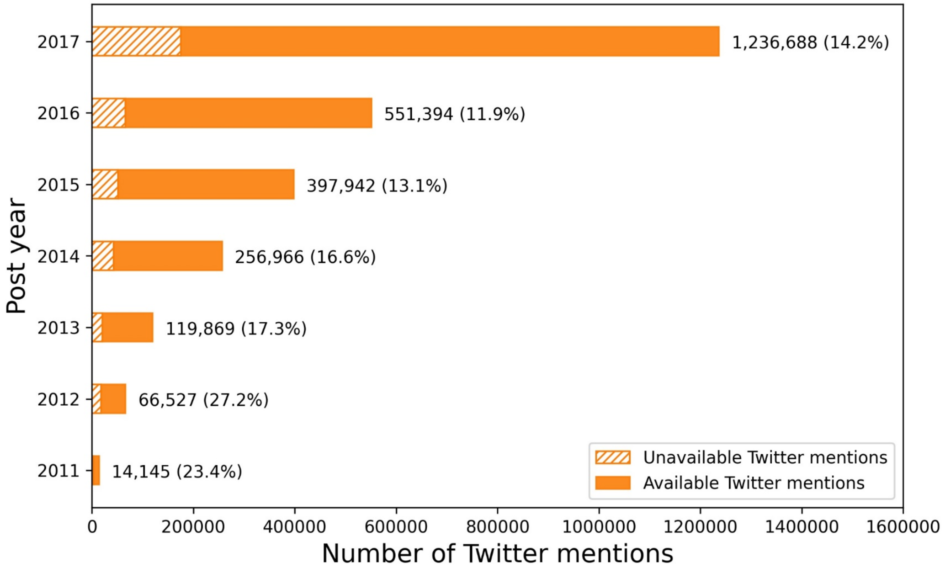

Figure 3. Distribution of Twitter mentions over the years. (Share of unavailable tweets per year listed in brackets)

5.3.2 Influence of unavailable Twitter mentions on the stability of Twitter metrics

Figure 4 shows the total number of Twitter mentions (blue line) and still available Twitter mentions (orange line) for 1,154 Altmetric IDs. The *Twitter unavailability rate*, namely, the percentage of unavailable Twitter mentions of each scientific paper, is presented as a yellow dashed line. For clearer visualization, the 1,154 papers are divided into three parts in the order of their total number of Twitter mentions and shown in Figure 4**A**-**C**, respectively. All highly tweeted papers have a certain number of unavailable tweets, and the amounts vary greatly across papers. Peaks of the yellow dashed line represent those papers with a large share of unavailable Twitter mentions. Due to these high unavailability rates, it can be argued that the Twitter metrics of the corresponding papers are unstable.

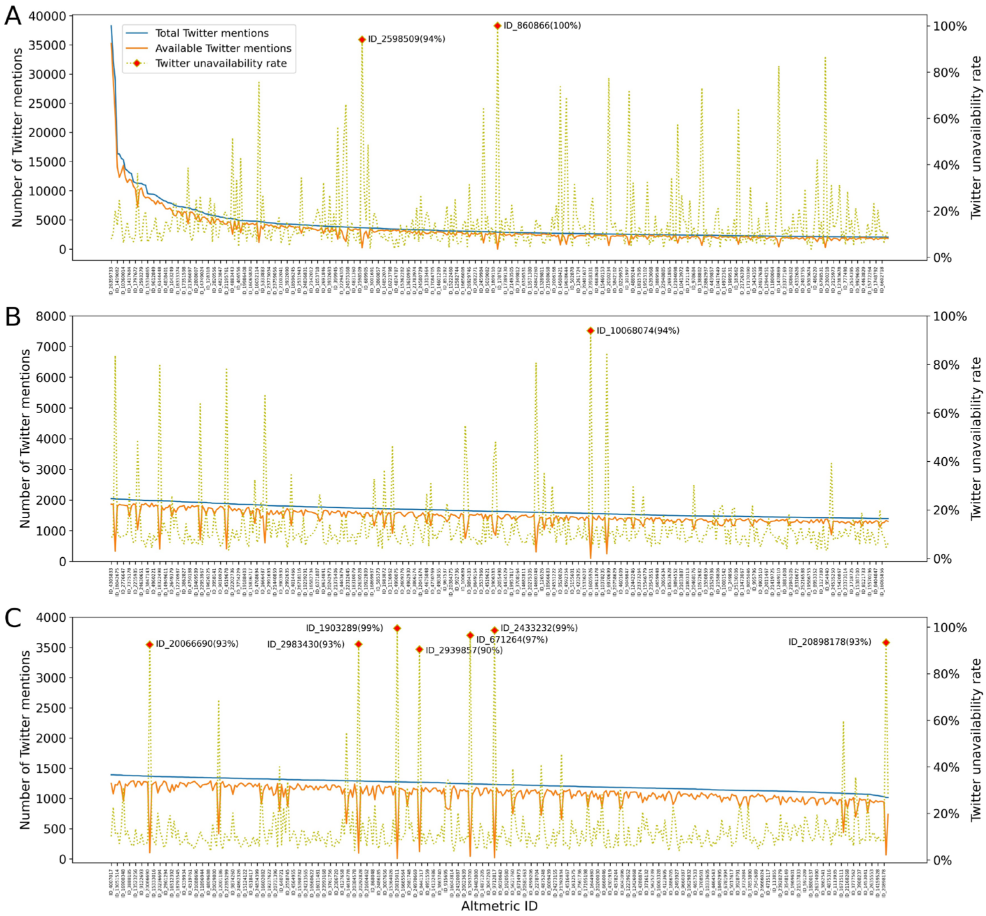

Figure 4. Twitter unavailability rates of the 1,154 most tweeted scientific papers

In order to investigate potential causes for the high Twitter unavailability rates of some papers, we selected the top-10 Altmetric IDs with the highest unavailability rate of Twitter mentions as a case study. In Figure 4, these top-10 Altmetric IDs are highlighted with red diamonds. The Twitter metrics of these scientific papers are most seriously affected by unavailable Twitter mentions, since their Twitter metrics sharply decreases, causing the *demotion* of these papers as highly tweeted papers. Table 2 presents details of their unavailable Twitter mentions from the aspects of original tweets and retweets in detail. The total number of tweets to the article, number of recorded original tweets, number of unavailable tweets, Twitter unavailability rate, number of unavailable original tweets, number of unavailable retweets, and maximum number of unavailable retweets related to an original tweet are calculated to reflect the composition of unavailable Twitter mentions.

Table 2. Top-10 Altmetric IDs with the highest unavailability rate of Twitter mentions

Altmetric ID	DOI	TWS	N_OT	N_UnT	TUnR	N_UnOT	N_UnRT	Max(N_UnRT)
860866	10.1088/1475-7516/2008/10/036	2,891	1	2,891	100.0%	1	2,890	2,890
1903289	10.2337/diacare.27.2007.s111	1,274	3	1,268	99.5%	0	1,268	1,268
2433232	10.1056/nejmoa1315231	1,230	11	1,213	98.6%	0	1,213	1,213
671264	10.1056/nejmoa1109017	1,241	23	1,198	96.5%	0	1,198	1,190
2598509	10.1080/17439884.2014.942666	3,659	122	3,440	94.0%	4	3,436	3,319
10068074	-	1,563	94	1,467	93.9%	17	1,450	1,426
20898178	10.1097/adm.0000000000000324	1,017	34	950	93.4%	0	950	950
2983430	10.2139/ssrn.2536258	1,290	76	1,195	92.6%	41	1,154	151
20066690	10.1038/nrmicro.2017.40	1,367	10	1,265	92.5%	1	1,264	1,253
2939857	-	1,266	86	1,145	90.4%	43	1,102	248

Note: TWS = total number of tweets to the paper; N_OT = number of recorded original tweets; N_UnT = number of unavailable Twitter mentions; TUnR = Twitter unavailability rate; N_UnOT = number of unavailable original tweets; N_UnRT = number of unavailable retweets; Max(N_UnRT) = maximum number of unavailable retweets related to a single original tweet. Altmetric IDs 10068074 and 2939857 are papers without DOI registered.

More than 90% of the Twitter mentions of these 10 scientific papers are unavailable, and almost all unavailable Twitter mentions are retweets. Therefore, for unavailable retweets of each paper, we explored the reasons for the high unavailability rate by calculating the maximum number of unavailable retweets related to a single original tweet for each paper. The results indicate that except for two Altmetric IDs (2983430 and 2939857), most unavailable retweets concentrate on a specific original tweet. For example, Altmetric ID 860866 has 2,891 Twitter mentions in total, consisting of only one original tweet and 2,890 retweets related to that original tweet. Therefore, when the original tweet became unavailable, according to the rules of Twitter,[3] all its related retweets that used Twitter's native "retweet" functionality turned unavailable as well, virtually decreasing the Twitter metrics of the paper to zero. The same happens to other Altmetric IDs, with most unavailable retweets concentrating around an original tweet that became unavailable. In Table 2, there are four Altmetric IDs where the number of unavailable original tweets equals zero. In fact, the unavailable retweets of these four papers direct to an unavailable original tweet as well according to our manual check. The zero values of N_UnOT are caused by the omission of

[3] See more information about rules of tweet deletion at: https://help.twitter.com/en/using-twitter/delete-tweets (Accessed December 17, 2019).

original tweets in Altmetric.com's data files, as we mentioned before. Based on these results we can state that the unavailability of an original tweet leads to the unavailability of a large number of retweets concentrating on it. This is the main reason for the high Twitter unavailability rates of papers listed in Table 2.

5.3.3 Twitter unavailability rates of papers with different Twitter dissemination structures

In order to further investigate the potential influence of different Twitter dissemination structures on the (in)stability of Twitter metrics, we calculated the DO and DC for the 1,154 sample scientific papers, with the distribution shown in Figure 5. Each dot represents a paper, and its color is determined by the Twitter unavailability rate shown in the color bar on the right. The dashed vertical and horizontal lines indicate the median DO (0.284) and median DC (0.203) of all papers, respectively. Moreover, the top-10 papers with the highest unavailability rate of Twitter mentions listed in Table 2 are marked by stars to highlight their location in the scatterplot. Most papers with high Twitter unavailability rates are located at the upper left part, especially for the eight starred papers with the highest unavailability rates. Their Twitter dissemination structures have very low DO and quite high DC, which means that once an original tweet with lots of retweets linking to it has been removed, most of that paper's Twitter mentions become unavailable. This results in the collapse of its Twitter metrics. Some papers at the left lower part, namely, those with both low DO and low DC, also show quite a high unavailability rate of Twitter mentions. This kind of papers has only a few original tweets but most of them received some retweets. Here, the distribution of retweets is more balanced, meaning that the risk of losing most of the retweets received once the original tweet becomes unavailable is not as high as for the papers at the upper left part. However, if the few original tweets received come from a specific Twitter user, and that user account is suspended, or that user decides to protect the tweets, the stability of Twitter metrics of those papers would be seriously affected as well. This is the case with the two starred papers at the left lower part. There are fewer papers with high Twitter unavailability rates in the right part. Papers in this part accumulated more original tweets, so they have fewer retweets that rely on the existence of original tweets. Throughout all four fields, the Twitter metrics of papers with high DO and low DC (right lower part) seem to be the most stable, since their dissemination structures consist of more independent original tweets and more decentralized retweets, which lowers the risk of losing a lot of Twitter records caused by the unavailability of several highly retweeted original tweets.

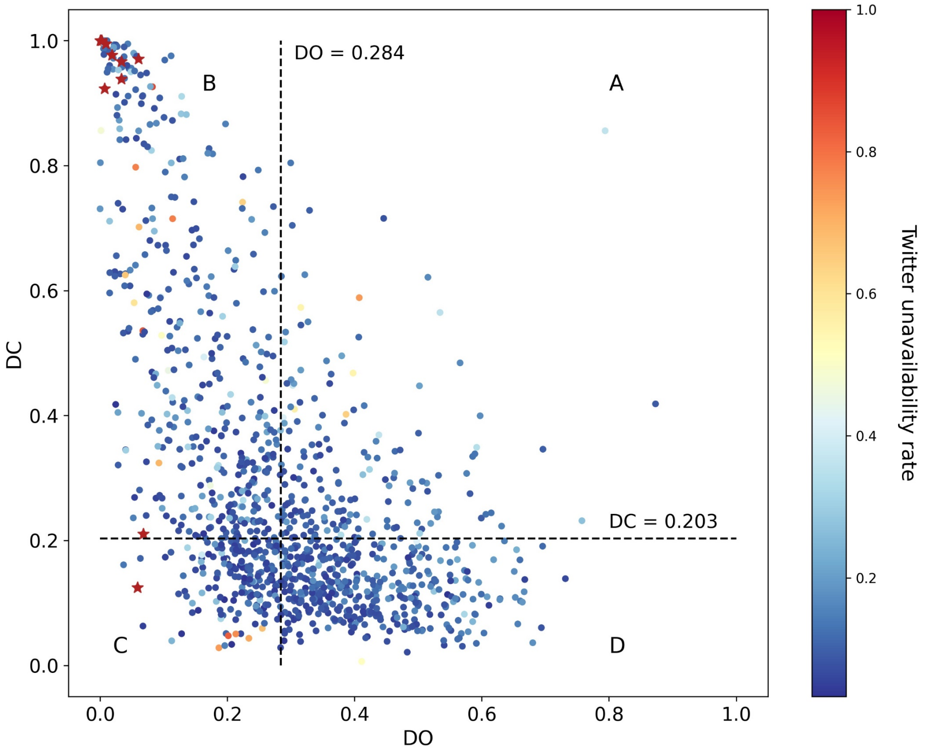

Figure 5. Distribution of the 1,154 scientific papers with different DO and DC

The dashed lines in Figure 5 represent the median value of DO and DC, respectively, and classify the papers included into four groups (A, B, C, and D). This is in correspondence with the four categories we introduced in Figure 2. The distribution of Twitter unavailability rates of these four groups of papers is shown in an associative plot (Figure 6, box plot and violin plot). With all four groups, most Twitter unavailability rates locate below 0.2, suggesting that most papers in these four groups have less than 20% of their Twitter mentions unavailable, their Twitter metrics are relatively stable regarding minor losses. However, the amount and distribution of outliers vary across groups. Group B and Group C have more outliers that hold extremely high Twitter unavailability rates, while those of Group A and Group D are fewer. Besides, most outliers of the latter are below 0.6; by contrast, Groups B and C have lots of outliers higher than 0.8, especially Group B. These results are in line with what we observed in Figure 5. Although most papers with different Twitter dissemination structures keep a relatively low Twitter unavailability rate, papers with extremely unstable Twitter metrics are more likely to occur when they have fewer original tweets and more concentrated retweets (Group B) or less original tweets and relatively deconcentrated retweets (Group C).

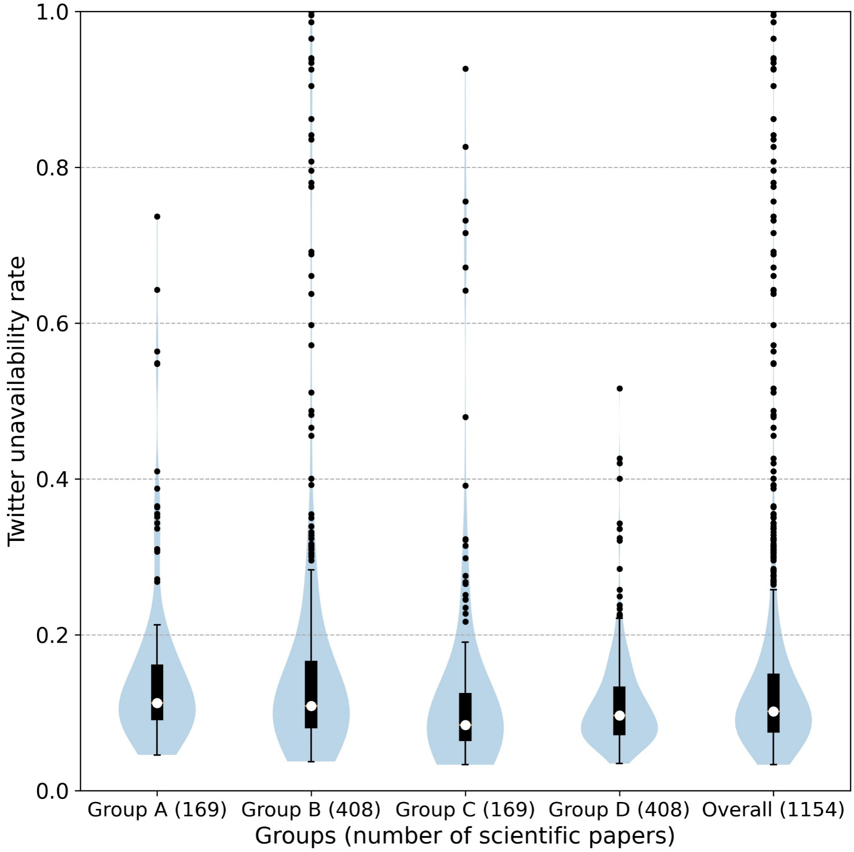

Figure 6. Distribution of Twitter unavailability rates of the four groups with different DO/DC characteristics

5.4 Discussion

5.4.1 The possible instability of Twitter metrics

Data consistency is essential for the measurement of impact in a sustainable and stable manner. In the context of altmetrics, data consistency is significantly affected by the dynamic nature of events (Haustein, 2016). Conceptually speaking, citations, once given, cannot disappear. Therefore, the decrease of citation counts of a specific paper is very rare, and is mostly caused by technical issues (e.g., changes in the coverage of the database, changes in the citation matching algorithms). For this reason, citation-based metrics of scholarly outputs are relatively stable over the course of time. On the other hand, there are no barriers for Twitter users to post a tweet or retweet, neither to delete a tweet or to cancel a retweet. A previously existing Twitter mention might become unavailable to the public for various reasons, and can no longer be identified or reused by following data aggregators and users, leading to the instability of Twitter counts of mentions to scientific papers. The same situation also happens to other altmetric indicators, for instance, Mendeley readership (Bar-Ilan, 2014). The number of Mendeley readership could decrease when users remove older references from their libraries (Zahedi et al., 2017), leading to the instability of readership counts as time goes by. Moreover, in the study of availability of blogs and news links, Ortega (2019b) observed that a considerable share of links in Altmetric.com and PlumX are broken due to the disappearance of some third parties that supply news and blog events, thereby making those news and blog records unavailable and which therefore cannot be audited.

In this study we checked the availability statuses of over 2.6 million Twitter mentions of the 1,154 most tweeted scientific papers recorded by Altmetric.com up to October 2017 to examine their Twitter unavailability rates, that is, the extent of Twitter mentions having become unavailable to the public. The status and reasons for unavailability were retrieved in April 2019. Our results indicate that for these most tweeted papers, around 14.3% of their Twitter mention records have become unavailable to the public. Twitter mentions that have been posted for a long time show a higher proportion of unavailability. Thus, the potential risk of Twitter mentions being unavailable for different reasons increases over time. Nevertheless, because Twitter users have become more active in sharing scientific information in recent years, the absolute number of unavailable Twitter mentions in 2017 is much higher than before. User deletion is the main reason for this high unavailability rate, accounting for 54.7% of unavailable Twitter records, followed by suspension and protection of Twitter user accounts (accounting for 25.9% and 16.7%, respectively).

Twitter unavailability rates vary markedly among scientific papers, hence influencing their Twitter metrics to different extents. In our study, all selected highly tweeted papers have a certain share of Twitter mentions unavailable at the time of data collection, and most of them have less than 20% of Twitter mentions that have become unavailable to the public. However, there are many papers that show extremely high unavailability rates. In our case study of the top-10 papers with the highest Twitter unavailability rates, over 90% of Twitter mentions directing to them have become unavailable. For these scientific papers, their Twitter metrics are among the highest when they were recorded by Altmetric.com, but if the unavailable Twitter mentions would be excluded from the counts, the overall Twitter counts of these papers would plummet dramatically. This is even more concerning given that Twitter data show a fast accumulation speed. In general, over 80% of Twitter data are accumulated within the first year after publication (Fang & Costas, 2020). This means that once the Twitter metrics of a relatively old paper has been affected by unavailable tweets, it is difficult for the paper to receive as many Twitter mentions as it had before to recover its Twitter metrics again. In this case, for papers that have been published for a long time, in general the loss of Twitter mentions is irreversible. What is more important is that those unavailable Twitter mentions cannot be detected and counted by other data aggregators that never recorded them before, which might exacerbate the inconsistency among Twitter data recorded by different data aggregators.

5.4.2 The influence of different forms of Twitter dissemination structures

In order to further explore the underlying reasons for high Twitter unavailability rates, we analyzed the Twitter dissemination structures of scientific papers based on the composition of their original tweets, retweets, and the connections between them. Originality and Concentration were introduced as two dimensions to classify these Twitter dissemination structures. Furthermore, DO and DC were proposed as two new measures to describe how many original tweets a paper has received (DO) and to what extent retweets concentrate around these original tweets (DC). On the basis of these two indicators, we found that scientific papers showing a relatively low DO and a relatively high DC are at a greater risk of losing larger numbers of Twitter mentions. This is because once a highly retweeted original tweet becomes unavailable, all its related retweets also become unavailable, generating a dramatic decrease in the overall Twitter metrics of the paper in question. In addition, some papers with extremely unstable Twitter metrics also show a relatively low DO and relatively low DC. In most cases, this is because the few original tweets were posted by the same user account, namely, those user accounts who tweeted the same article repeatedly, as observed by Robinson-Garcia et al. (2017). If the Twitter user sending original tweets repeated times is suspended, all of their original tweets become unavailable, and so do the related retweets. By comparison, among papers with a relatively high DO there are a few showing extremely unstable Twitter metrics, particularly when the DC is low. The high DO lowers to some extent the risk of losing the bulk of the Twitter records.

Given the diversity of users and complexity of engagement behaviors that happen on Twitter, the dissemination processes of scientific papers on Twitter are sophisticated, Twitter metrics can help to unveil such diversity and complexity (Haustein, 2019). The study of Twitter dissemination structures not only contributes to the identification of papers that may suffer from a stronger vulnerability of losing their Twitter counts, but also sheds light on the possibilities of measuring the performance of scientific papers on Twitter in a more fine-grained manner. The total count of Twitter mentions is one of the most common Twitter measures, but as we presented in Figure 2, papers with the same total Twitter mention counts might perform differently from the point of view of their Twitter dissemination structures. Didegah et al. (2018) studied the number of original tweets and retweets of papers and their qualities across different subject fields. But beyond this kind of statistic, it is relevant to organize these data to reveal the overall picture of Twitter dissemination structures of research outputs. Twitter impact is not only about how many times an article has been tweeted, but also about how it was tweeted. The reconstruction of the Twitter dissemination structure provides a partial answer to this question. Based on the Twitter dissemination structure, it is possible to unravel the underlying dissemination patterns and networks of papers that hide behind the total statistical numbers, with the latter compounding different types of Twitter mentions and their relationships in a simple way. As a result, the Twitter dissemination structure is supposed to contribute to a better understanding of the performance of papers on Twitter.

In future research, we will further optimize the indicators for describing Twitter dissemination structures. For example, in this article the DC was calculated based on the maximum percentage of retweets concentrating on a single original tweet. This method, derived from the case study of the top-10 papers with the highest Twitter unavailability rates, also has the advantage of simplicity. We will introduce multiple calculation methods for measuring the DC at both tweet and Twitter user levels in future studies. Particularly at the Twitter user level, in addition to taking the retweeting relationships among users into account, the status, the degree of activity, and diverse Twitter user profiles are expected to be considered to establish more fine-grained Twitter dissemination structures. Moreover, we will explore possible applications of Twitter dissemination structures in the measurement of Twitter reception of scientific information.

5.4.3 Overall situation of the stability of Twitter metrics

Besides rechecking the Twitter mentions of the 1,154 most tweeted papers presented above, in September 2019 we rechecked the statuses of all Twitter mentions recorded by Altmetric.com in the historical data files (version: October 2017) to reveal the overall situation of the stability of Twitter metrics for nearly 5.4 million papers. The results show that among the over 42.5 million unique recorded Twitter mentions, about 13.0% of them have become unavailable. Accordingly, the overall Twitter unavailability rate is slightly lower than that of the sample of highly tweeted papers (14.3%).

For understanding the overall influence of unavailable tweets on Twitter metrics at the paper level, Spearman correlation analyses between the total number of recorded Twitter mentions and the number of available Twitter mentions during data rechecking were conducted for both the sample of the 1,154 most tweeted papers and all recorded papers in Altmetric.com. For both datasets, these two numbers are highly correlated ($r_s = 0.91$ for the most tweeted papers, and $r_s = 0.93$ for all papers), which means that the majority of papers kept relatively stable Twitter metrics over time. This result is in line with the distribution of Twitter unavailability rates we observed for the most tweeted papers, with most papers having less than 20% of tweets unavailable and a limited share of papers showing extremely unstable Twitter metrics.

It should be noted that, although the value of Altmetric.com database snapshots is obvious for studying changes in altmetrics over time, due to the Twitter restrictions, Altmetric.com is no longer providing tweets that have been removed from Twitter, and researchers are now required to delete all unavailable tweets from their locally hosted snapshot files.[4] This implies that unavailable tweets cannot be studied in related future research. Moreover, except for tweet IDs and Twitter user IDs, Altmetric.com will no longer provide the content of Twitter mentions of papers in its snapshots, ensuring that the detailed information of potential unavailable tweets not be kept in the historical data files.

5.4.4 Limitations

There are some limitations that should be acknowledged in this study. First, as we mentioned in the Data and Methods section, there exist some retweets without corresponding original tweets recorded by Altmetric.com. Given that the existence of an original tweet is the basis of its following retweets, we assumed that there are some original tweets to complete the retweeting relationship, meaning that we had to work with "assumed" data instead of actual data. Second, for deleted original tweets, it would be interesting to analyze the motivations of users. However, this question is not further discussed in our article because of the lack of traceable evidence and the Twitter restrictions on deleted content. Lastly, Twitter dissemination structures were analyzed only from the perspective of the connections of different types of tweets (original tweets and retweets), whereas diversity of Twitter users in the Twitter dissemination process might be another factor that has an influence on the stability of Twitter metrics. In the case of some papers with relatively low DO and low DC, we could show that the reason why they have extremely high unavailability rates is that the few original tweets were posted by the same user account. Therefore, the composition of Twitter users involved and their identities should be further explored in the future, especially in the light of bot accounts playing a major role in the science communication landscape on Twitter (Didegah et al., 2018).

5.5 Conclusions

This study examined the stability of Twitter metrics of scientific papers by rechecking the statuses of their Twitter mentions. For over 2.6 million Twitter records of the 1,154 most tweeted papers recorded by Altmetric.com until October 2017, about 14.3% of them became unavailable by April 2019. The main reason for the high unavailability rate is deletion of tweets, followed by suspension and protection of Twitter user accounts. The stability of Twitter metrics varies among papers, most of them have Twitter unavailability rates of less than 20%, but there are some papers showing extremely high unavailability rates. The potential influence of Twitter dissemination structures on the stability of Twitter metrics was

[4] Extracted from personal communication with Stacy Konkiel from Altmetric.com.

investigated. DO and DC were proposed to describe Twitter dissemination structures based on original tweets, retweets, and original tweet-retweet connections. Twitter metrics of papers with a relatively low DO and relatively high DC are at greater risk of becoming highly unstable. Building on that, we discussed the stability and persistency of Twitter metrics of scientific papers and the potential risks they can be subject to. Thus, our study underlines the importance of distinguishing dissemination structures in the context of Twitter-based indicators.

CHAPTER 6

How is science clicked on Twitter? Click metrics for Bitly short links to scientific papers[1]

[1] **This chapter is based on**:

Fang, Z., Costas, R., Tian, W., Wang, X., & Wouters, P. (2021). How is science clicked on Twitter? Click metrics for Bitly short links to scientific publications. *Journal of the Association for Information Science and Technology, 72*(7), 918-932. https://doi.org/10.1002/asi.24458

Author contributions:

Fang, Z. (Conceptualization, Methodology, Formal analysis, Investigation, Visualization, Data Curation, Writing - Original Draft, Writing - Review & Editing)

Costas, R. (Conceptualization, Methodology, Investigation, Supervision, Writing - Review & Editing)

Tian, W. (Visualization, Data Curation, Writing - Review & Editing)

Wang, X. (Conceptualization, Methodology, Supervision, Writing - Review & Editing)

Wouters, P. (Conceptualization, Supervision, Writing - Review & Editing)

Abstract

To provide some context for the potential engagement behavior of Twitter users around science, this paper investigates how Bitly short links to scientific papers embedded in scholarly Twitter mentions are clicked on Twitter. Based on the click metrics of over 1.1 million Bitly short links referring to Web of Science papers, our results show that around 49.5% of them were not clicked by Twitter users. For those Bitly short links with clicks from Twitter, the majority of their Twitter clicks accumulated within a short period of time after they were first tweeted. Bitly short links to the papers in the field of *Social Sciences and Humanities* tend to attract more clicks from Twitter over other subject fields. This paper also assesses the extent to which Twitter clicks are correlated with some other impact indicators. Twitter clicks are weakly correlated with scholarly impact indicators (WoS citations and Mendeley readers), but moderately correlated to other Twitter engagement indicators (total retweets and total likes). In light of these results, we highlight the importance of paying more attention to the click metrics of URLs in scholarly Twitter mentions, to improve our understanding about the more effective dissemination and reception of scientific information on Twitter.

Keywords

Scholarly Twitter metrics, altmetrics, social media metrics, Twitter clicks

6.1 Introduction

Sharing research outputs and other relevant information on Twitter has arguably become a common way of scholarly communication, thereby making Twitter mentions one of the most important altmetric events for scientific papers (Haustein, 2019; Sugimoto, Work, et al., 2017). Such scholarly Twitter mentions imply that science is no longer restricted to the ivory tower, but expands beyond the borders of the scientific community and interests various types of people and institutions (Yu et al., 2019). The weak or negligible correlations confirmed between Twitter mentions and scholarly impact indicators, such as citation counts and journal citation scores (Bornmann, 2015a; Costas et al., 2015a; Zahedi et al., 2014), support the idea that scholarly Twitter mentions might reflect a wider and different type of influence of scientific papers beyond the science environment (Thelwall, Haustein, et al., 2013).

In order to better comprehend the impact that shared scientific papers made in the Twittersphere, instead of merely counting the absolute number of Twitter mentions accrued, it is necessary to explore users' online activities and specific interactions with research objects, which was referred to as "the second generation of Twitter metrics" by Díaz-Faes et al. (2019). This more interactive perspective will help to characterize the underlying mechanisms by which Twitter users interact with research outputs, and to further interpret the impact of research outputs generated through the processes of engagement among different stakeholders on Twitter (Robinson-Garcia et al., 2018).

6.1.1 Twitter interactions on the basis of scientific papers

A number of studies have investigated the interactions between Twitter users and scientific papers. For scientific papers, being shared on Twitter is usually coupled with or followed by a series of interaction behavior. On the one hand, Twitter users might organize their tweet content about the mentioned scientific papers in different ways, and on the other hand, those scholarly Twitter mentions might attract diverse types or levels of engagement from the audiences after being posted. The majority of prior research on Twitter users' interaction behavior has either focused on the tweet content or the engagement around scholarly Twitter mentions.

Tweet texts have usually been scrutinized to unravel the patterns of tweeting (Robinson-Garcia et al., 2017), the quality of interactions (Didegah et al., 2018), and the types of sentiment (Friedrich et al., 2015; Hassan, Saleem, et al., 2020) of scholarly Twitter mentions. Some functions used in tweets, like user mentions and hashtags, reflect specific interactions around scientific papers as well. For instance, user mentions establish the relationships among users who might be related to or interested in the mentioned research, based on which the communities of users sharing interest can be detected (Araujo, 2020; Pearce et al., 2014; Said et al., 2019; Van Schalkwyk et al., 2020). Hashtags added in scholarly Twitter mentions indicate particular concepts in relation to the mentioned papers (Haustein, Bowman, & Costas, 2016), therefore, the adoption of hashtags provides the opportunities of identifying not only the connections among tweets or users focusing on the same topics (Costas et al., 2021; Hellsten & Leydesdorff, 2020), but also the broader public concerns about some specific research topics (Haunschild et al., 2019; Lyu & Costas, 2020).

In terms of engagement around scholarly Twitter mentions, there are some interactions, such as retweets, likes, and replies, that tell stories about the impact of scholarly Twitter mentions made in the Twitter environment. Since retweets account for a considerable share of scholarly Twitter mentions (Didegah et al., 2018), it is the most studied engagement behavior. As described by Haustein (2019), retweets represent a specific form of diffusing information, so they have been widely used to construct the retweeting networks and examine the diffusion patterns of scientific information (Alperin et al., 2019; Fang, Dudek, et al., 2020; Robinson-Garcia et al., 2017). Díaz-Faes et al. (2019) considered the number of likes given by Twitter users as one of the factors for measuring the social media activity of users around science. Overall, in contrast to the investigations of tweet content, there is less research of how Twitter audiences engage with scholarly Twitter mentions.

6.1.2 Click metrics for URLs embedded in tweets

In addition to the aforementioned Twitter functions, URLs have also been found to be frequently used by scholars on Twitter (Bowman, 2015). In a survey of 37 identified astrophysicists, Haustein, Bowman, Holmberg, Peters, et al. (2014) observed that more than one third of their tweets contained URLs. The ratio reported in a study by Weller and Puschmann (2011) was even higher: they found that more than 55% of tweets by scientists included at least one URL. Besides, scholars were found to be more inclined to frame their professional tweets with URLs, rather than their personal tweets (Bowman, 2015). Based on a coding sample of 2,322 tweets by scholars containing hyperlinks, Priem and Costello (2010) found that 6% of them referred to scientific papers. These embedding URLs offer audiences a portal to more abundant information than limited tweet texts would contain. They also serve as digital traces of the Twitter reception of scientific papers, leading to the scholarly Twitter metric data which are detected and tracked by many altmetric data aggregators (Zahedi & Costas, 2018).

Compared to other engagement behavior with scholarly Twitter mentions, clicking URLs cited in tweets to get access to the mentioned scientific papers has been less analyzed on a large scale due to the unavailability of traceable data on URL clicking. Relying on the referral sources[1] of visitors offered by *PeerJ* for its published papers, X. Wang, Fang, & Guo (2016) investigated how a selection of 110 *PeerJ* papers got visits from different web referral sources, and found that URLs shared on Twitter and Facebook attracted the majority of visits amongst social media platforms. Given that URLs embedded in tweets are usually shortened to comply with the maximum character length restriction of Twitter,[2] some publicly available *click metrics* for short links open a novel window for scrutinizing the visits to online resources driven by the shortened URLs posted on Twitter. For example, enabled by the click metrics provided by Bitly (https://bitly.com) for its generated short links, Gabielkov et al. (2016) studied the extent to which URLs from five leading news domains were clicked on Twitter. Similarly, both L. X. Wang, Ramachandran, & Chaintreau (2016) and Ramachandran et al. (2018) utilized Bitly click data to measure and model the click dynamics of the links to news tweeted by a group of BuzzFeed (https://buzzfeed.com) Twitter accounts. The same methodology was also employed by Holmström et al. (2019) to analyze the temporal accumulation dynamics of clicks of Bitly links associated with seven major news websites.

These previous studies experimentally showed that click metrics provide a practical method to gain a more in-depth understanding of the impact of Twitter mentions. Different from observable Twitter engagement behavior such as retweeting and liking, conceptually speaking, clicking is a type of digital behavior related to a deeper engagement with tweets by Twitter users, moving from merely viewing the tweets, to actually trying to access more detailed content by clicking the URLs included. Clicking behavior embodies a further Twitter reception of shared information by Twitter users, which could substantially increase the visits of the tweeted content. Therefore, click metrics capture a type of *potential* impact that Twitter mentions made in creating a greater awareness of shared information. Based on this idea we could argue that those Twitter mentions with URLs being more clicked (in contrast to those with URLs not or just scarcely clicked) are more effective in disseminating information.

6

6.1.3 Conceptualizing click metrics as social media metrics

In Figure 1 several forms of possible interactions *within* and *between* science and Twitter are represented. Conceptually speaking, there are two different interactive environments: one is the *science environment* where scientific work and papers are produced, and within the science environment there are interactions such as citing, reading, taking place. At the other side, there is the *Twitter environment*, where Twitter users interact by tweeting, liking, retweeting, and replying, etc. We argue that when a scientific paper is tweeted, the Twitter

[1] Referral sources are Internet addresses or hostnames that users used to visit the website where they are located now. The referral source information is one type of article-level metrics provided by *PeerJ* for tracking the web referrals through which visitors access *PeerJ* papers.

[2] Currently the text content of a tweet is allowed to contain up to 280 characters. See more details about the tweet length at: https://developer.twitter.com/en/docs/basics/counting-characters (Accessed July 26, 2020).

mention establishes a *bridge* connecting the two environments, through which an information flow moves from the science environment to the Twitter environment. The generated scholarly Twitter mentions offer Twitter users the chance to engage with scientific information within the Twittersphere. In addition, it is possible for Twitter users to click the URLs embedded in scholarly Twitter mentions to access the corresponding scientific papers. In those cases, we would argue that through the established bridge, the information flow would move back from the Twitter environment towards the science environment (in practical terms, any Twitter user clicking the URL of a scientific paper in a tweet would leave the Twitter platform, to move to another (scholarly) platform to view the paper or its metadata).

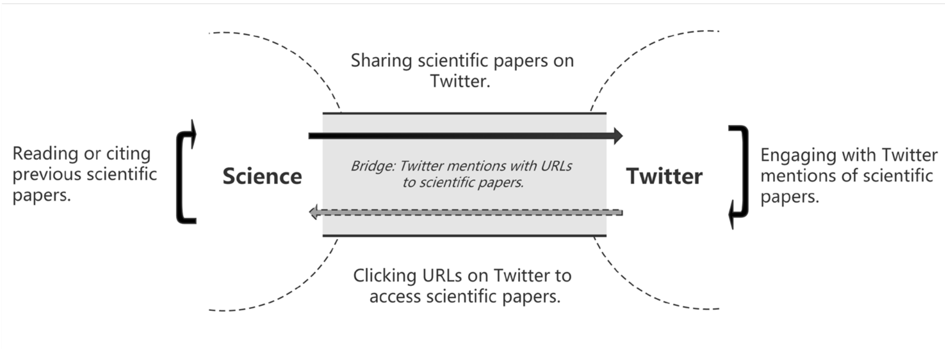

Figure 1. Four conceptual interactions within and between the science environment and the Twitter environment

In this model, the meaning of clicking behavior in the context of altmetric research is two-fold:

(1) From a social media metric point of view, for scholarly Twitter mentions, the fact that the embedding URLs are being clicked by different Twitter users implies that the audiences are motivated by the tweet content (or features) to seek for more details about the scholarly information. Clicks would then represent a type of impact related with the success of scholarly Twitter mentions in creating effective forms of scholarly dissemination and communication, thus offering a new perspective on “secondary social media metrics” (Díaz-Faes et al., 2019), which focus on the characterization of social media activities and interactions with science by social media users.

(2) From an impact measurement point of view, the act of clicking the URLs of papers embedded in tweets represents an expanded form of altmetric impact, capturing not only the interest raised by the tweeting users (who originally posted the tweets or retweeted them), but also by the clicking ones (i.e., the audiences further interested in the content posted by the tweets). Simply put, papers that get tweeted and clicked might exhibit a larger altmetric impact than those merely being tweeted (or retweeted).

Both points of view above support the idea that including click metrics in the altmetric toolset allows for a fundamental broadening of the analytical scope of altmetrics, moving beyond the notions of impact of scholarly outputs on social media (which is achieved by the information flow from science to Twitter), towards the notions of impact of social media on scholarly outputs (which is reflected by the information flow from Twitter back to science). Nevertheless, as mentioned above, due to the technical difficulties in obtaining click data on a large scale (since it is not evidence currently disclosed by Twitter), it remains unclear whether or not shared scientific papers' URLs directly drive traffic to the websites of these scientific papers. Against this background, click metrics for short links are expected to quantitatively depict such underlying Twitter reception of scientific information, and add a missing piece to the puzzle of interactions between the science and Twitter environments.

6.1.4 Objectives

By leveraging the click metrics tracked by Bitly for its generated short links, the main objective of this study is to disclose how Bitly short links to scientific papers are clicked on Twitter. Specifically, this study seeks to address the following explicit research questions:

RQ1. How frequently are Bitly short links to scientific papers clicked by Twitter users? And how do Twitter clicks accumulate over time?

RQ2. Do clicks of Bitly short links to scientific papers vary across subject fields? In which subject fields do Bitly short links have relatively more Twitter clicks?

RQ3. As a new type of indicator for measuring the potential impact of scholarly Twitter mentions, to what extent do clicks on Twitter correlate with other scholarly impact indicators (i.e., citation counts and Mendeley readers) and other Twitter engagement indicators (i.e., retweets and likes)?

6.2 Data and methods

6.2.1 Dataset

Bitly is a link management platform which was launched in 2008.[3] Its link shortening service has been employed by many Twitter users for the sake of, on the one hand, complying with the tweet length limit, and on the other hand, tracking the clicks for their generated short links. The latter offers the possibilities of monitoring how Bitly short links to scientific papers are clicked on different sources and on different dates.

This study focuses on the click metrics for Bitly short links to scientific papers that have been tweeted, to unravel to what extent scholarly content are accessed by being clicked on Twitter. Our dataset stems from the scholarly Twitter mention data recorded by Altmetric.com up to October 2017. Based on the tweet IDs provided by Altmetric.com, we further collected the detailed information of tweets through the Twitter API in September 2019. After excluding unavailable tweets caused by deletion, suspension or protection of Twitter accounts (Fang, Dudek, et al., 2020), we obtained 1,422,266 distinct original tweets[4] with 1,103,819 distinct short links using the "bit.ly" domain that refer to Web of Science (WoS) indexed papers. As Altmetric.com started to track Twitter data since October 2011, only WoS papers published from 2012 onwards were taken into account. Moreover, to ensure that selected Bitly short links refer to the webpages of scientific papers, Twitter mentions were restricted to those containing only one URL.[5]

6.2.2 Click data of Bitly short links

Bitly provides APIs for retrieving link-level analytics, making it possible to collect click metrics for the selected Bitly short links. In December 2019, for each Bitly short link in the above dataset, we collected their click metrics as follows:

---

[3] See more introduction to Bitly at: https://support.bitly.com/hc/en-us/articles/230895688-What-is-Bitly- (Accessed July 26, 2020).

[4] Here original tweets are tweets originally posted by Twitter users, including both original tweets and reply tweets. Retweets and quote tweets are not studied because they have the same embedding URLs as the corresponding original tweets.

[5] In the Altmetric.com database, there are a total of 7,446,310 unique tweeted URLs meeting the criteria we set. Table 4 in Appendix lists the usage rate of Bitly short links (i.e., the proportion of Bitly short links in all tweeted URLs) over the tweet post years, together with the usage rates of short links generated by three other frequently used link shortening services (i.e., Ow.ly, Goo.gl, and TinyURL) for comparison. Overall, as one of the most used link shortening services, the selection of 1,103,819 short links using the "bit.ly" domain accounts for about 14.8%.

(1) *Total number of clicks*. This is the overall number of accumulated clicks after the short link was generated, considering all referral sources together.[6]

(2) *Number of clicks on different sources.* This information details how many times the short link was clicked on each referral source, from which the number of clicks on Twitter can be extracted ("Twitter clicks" hereafter).

(3) *Number of clicks on different dates*. This information details how many times the short link was clicked on different dates after it was generated.

A total of 1,102,622 Bitly short links have valid and complete metric data extracted from the APIs, involving a total of 1,420,588 Twitter mentions of 783,433 distinct WoS papers. These Bitly short links were selected as the main dataset for this study.

When studying the temporal distribution of clicks for a given short link, an important limitation in Bitly is that the date information of the clicks (i.e., the dates when the clicks were performed) retrieved through the APIs is aggregated without distinguishing the sources from where the clicks were performed. This means that if a short link has been clicked from more than one source, it is not possible to isolate the clicks coming only from Twitter. As a solution, in order to explore the temporal distribution of clicks happened specifically only on Twitter, the set of 171,430 Bitly short links with all clicks only from the source Twitter was drawn as one of the subsamples.

6.2.3 CWTS publication-level classification system

For the comparison of click metrics amongst subject fields, the CWTS publication-level classification system (Waltman & Van Eck, 2012) was employed to assign papers with subject field information. This scheme clusters WoS papers into micro-level fields based upon their citation relationships and then algorithmically assigns them to five main subject fields of science, including *Social Sciences and Humanities* (SSH), *Biomedical and Health Sciences* (BHS), *Physical Sciences and Engineering* (PSE), *Life and Earth Sciences* (LES), and *Mathematics and Computer Science* (MCS).[7] For our dataset, a total of 697,644 distinct papers (accounting for around 89%) have the subject field information assigned by the CWTS classification, involving 944,686 distinct Bitly short links. This set of papers and short links

[6] Bitly records clicks from different types of referral sources, such as social media platforms (e.g., Twitter, Facebook, Reddit, LinkedIn, YouTube) and e-mail.

[7] More details about the CWTS classification system at: https://www.leidenranking.com/information/fields (Accessed July 26, 2020).

were used as a subsample to explore how Twitter clicks vary across the five subject fields. Table 1 lists the presence of these papers and short links in each subject field.[8]

Table 1. Number of papers (P) and Bitly short links (BL) in each subject field

Subject field	Abbr.	P	Percentage	BL	Percentage
Social Sciences and Humanities	SSH	77,031	11.0%	106,521	11.3%
Biomedical and Health Sciences	BHS	433,419	62.1%	612,628	64.9%
Physical Sciences and Engineering	PSE	83,465	12.0%	93,700	9.9%
Life and Earth Sciences	LES	88,442	12.7%	111,730	11.8%
Mathematics and Computer Science	MCS	15,287	2.2%	20,107	2.1%

6.2.4 Indicators and analytic approaches

To better understand the relationships between Twitter clicks and scholarly impact indicators and other Twitter engagement indicators, we calculated the following four indicators to measure their correlations with Twitter clicks for each Bitly short link:

Scholarly impact indicators:

6

(1) *WoS citations*: the number of WoS citations of the paper that the Bitly short link refers to.

(2) *Mendeley readers*: the number of Mendeley readers of the paper that the Bitly short link refers to.

Twitter engagement indicators:

(3) *Total retweets*: the total number of retweets received by all scholarly Twitter mentions containing the Bitly short link.

[8] As there are more tweeted papers from the field of BHS, this field contributed the most Bitly short links in our dataset. To reflect the variations of the usage rates of Bitly short links across subject fields, Table 5 in Appendix presents the proportion of Bitly short links in all URLs for each subject field. Among the total set of 7,446,310 unique tweeted URLs meeting the criteria we set, 6,507,861 of them (accounting for 87.4%) have the CWTS classification information. Although BHS is the field with quantitatively more Bitly short links, in relative terms the usage rates of Bitly short links are quite similar across the five main subject fields, ranging from 12.4% for LES up to 17.5% for PSE.

(4) *Total likes*: the total number of likes received by all scholarly Twitter mentions containing the Bitly short link.

Regarding the data collection date of these four indicators, WoS citations were retrieved from the CWTS in-house WoS database which contains WoS data up to March 2020, Mendeley readers were collected through the Mendeley API in July 2019, while total retweets and total likes were obtained in the process of Twitter data collection in September 2019 as well. Finally, Spearman correlation analysis was performed at the Bitly short link level to measure the extent to which Twitter clicks correlate with the above indicators.

6.3 Results

The Results section consists of three main parts: the first part presents the frequency and accumulation speed of Twitter clicks of Bitly short links to scientific papers. The second part shows the variations of Twitter clicks across papers from different subject fields. The last one focuses on the correlation analysis between Twitter clicks and two scholarly impact indicators (i.e., WoS citations and Mendeley readers) and other two Twitter engagement indicators (i.e., total retweets and total likes).

6.3.1 Overall distribution of Twitter clicks

On the whole, Bitly short links in our dataset received nearly 12 million clicks in total, 52% of which (about 6.2 million) are contributed by Twitter. Although Twitter plays a key role in directing traffic to scientific papers, as shown in Figure 2**A**, the overall distribution of Twitter clicks among Bitly short links is highly skewed. About 49.5% of Bitly short links were not clicked after being tweeted, and most Bitly short links only got a few clicks on Twitter - around 89.7% of short links were clicked by Twitter users no more than 10 times. For comparison, the distribution of total clicks with all referral sources considered is shown in Figure 2**B**. As the more visible in multiple platforms, the higher the possibility of being clicked, there are less Bitly short links without clicks (about 36.5%) and more being clicked at different levels.

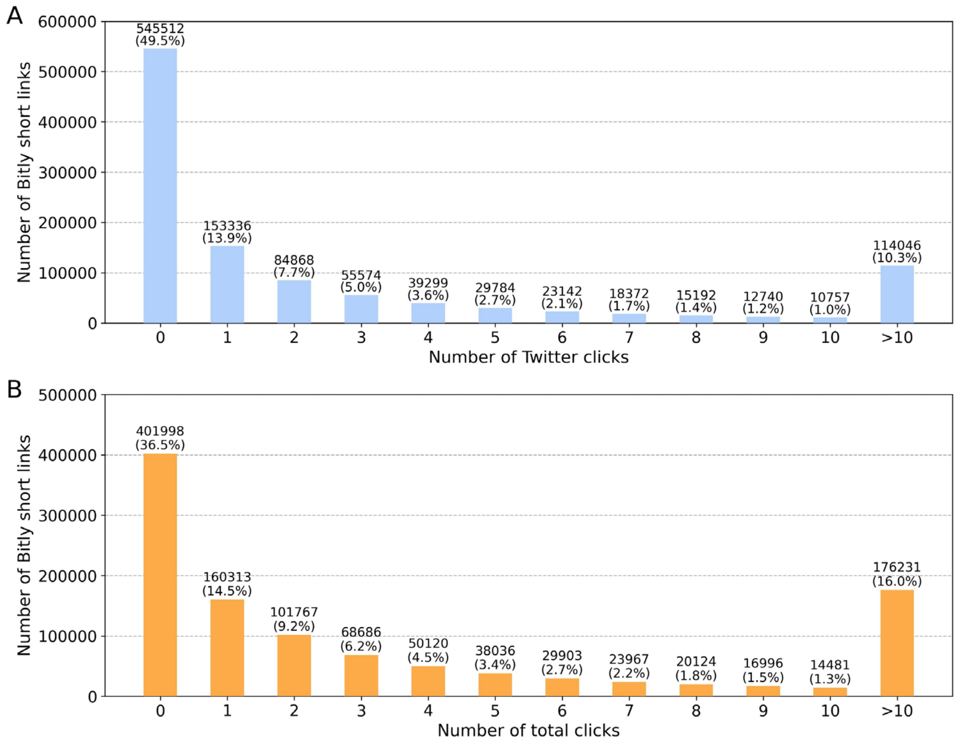

Figure 2. Distribution of Bitly short links with different numbers of **A** Twitter clicks and **B** total clicks

The majority of Twitter mentions of scientific papers accrued in a very short period of time after publication (Fang & Costas, 2020; Yu et al., 2017). As shown in Figure 3**A**, the same accumulation speed can be observed for Twitter clicks. For each Bitly short link, we selected the day when it was tweeted for the first time as the original point, and then calculated the time intervals between the original point and the date when it was clicked. Overall, around 26.3% of Twitter clicks happened in the day when the short links were first tweeted, and the number of Twitter clicks increases dramatically in the next few days. It exceeds 60% in the following day and reaches 80.5% in the first 10 days, and then flattens out. On the whole, there are about 86.2% of Twitter clicks accumulated in the first month after the Bitly short links appeared on Twitter. Similarly, Figure 3**B** exhibits the temporal accumulation pattern of total clicks with all referral sources counted. We used the same original point for calculating time intervals. Compared to Twitter clicks, there are slightly more clicks before the Bitly short links were tweeted because they might be posted on some other platforms earlier, yet the temporal accumulation pattern of total clicks is similar to that of Twitter clicks, with 24.6% of total clicks accrued by the day in which the first Twitter mentions came into view and 83.8% of total clicks accumulated in a month.

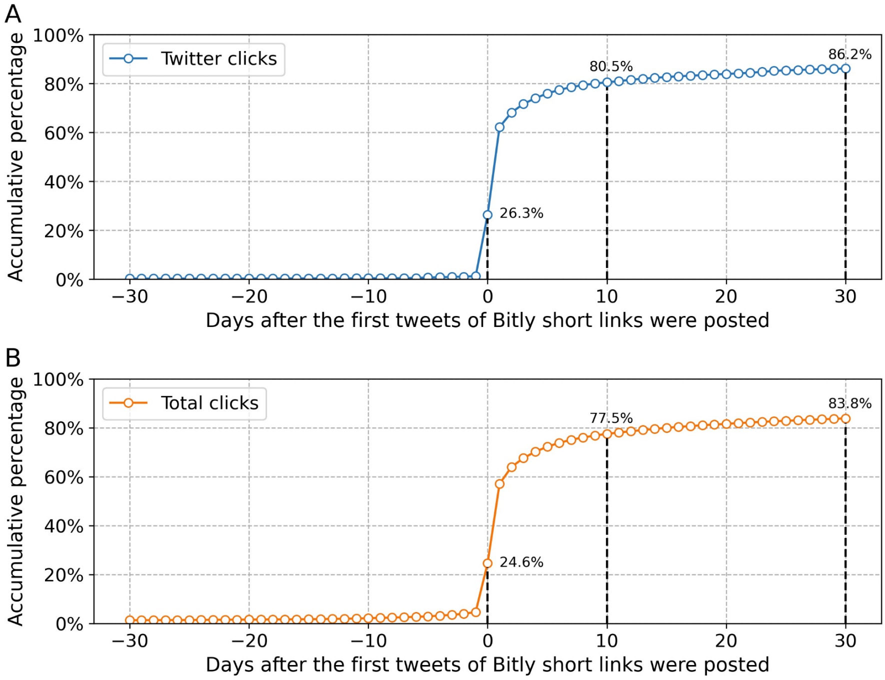

Figure 3. Temporal accumulation patterns of **A** Twitter clicks and **B** total clicks

6.3.2 Twitter clicks across subject fields

To make a comparison of the coverage of Twitter clicks amongst subject fields, Figure 4 plots the percentage of Bitly short links with different levels of Twitter clicks for the five main subject fields. Generally speaking, Bitly short links in the field of SSH show the highest coverage of Twitter clicks, followed by LES and BHS, while the coverage in the fields of MCS and PSE is relatively lower. As to the percentage of Bitly short links with at least one Twitter click received, obvious variations can be observed among subject fields ranging from about 42.8% for PSE up to 55.5% for SSH. Subject fields rank differently as the number of Twitter clicks increases, especially for BHS and LES. BHS, by contrast, tend to have higher percentage of Bitly short links with relatively larger quantity of Twitter clicks, but not for LES, which means that although LES has proportionally more Bitly short links with Twitter clicks received, it is proportionally less abundantly clicked than BHS does.

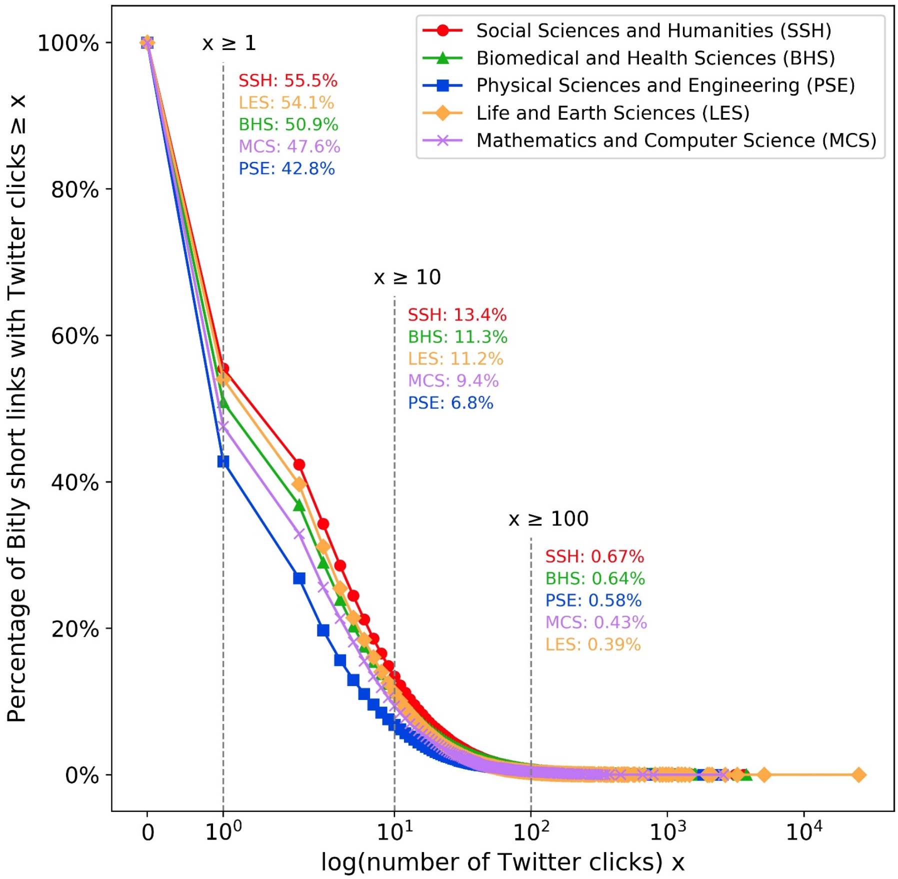

Figure 4. Percentage of Bitly short links with different levels of Twitter clicks across the five subject fields

Table 2 presents descriptive statistics of Twitter clicks of Bitly short links in the five subject fields. In general, Bitly short links in the field of SSH achieve not only the highest coverage but also the largest average number of Twitter clicks. As discussed above, even though the overall coverage of Twitter clicks for LES is higher than that for BHS, the latter has more Twitter clicks accrued on average. MCS and PSE always rank last in terms of both the coverage and the average Twitter clicks accumulated. The indicator of 90^{th} percentile tells the story in a consistent way. Bitly short links in the field of SSH have the top 10% received at least 13 Twitter clicks, both BHS and LES have their top 10% of Bitly short links received at least 11 Twitter clicks, while MCS and PSE come in last with at least 9 and 6 Twitter clicks for their top 10% of Bitly short links, respectively.

Table 2. Descriptive statistics of Twitter clicks of Bitly short links in the five subject fields

Subject field	TNL	TNC	PL ($C \geq 1$)	Mean (all)	Mean (clicked)	Min	Max	90th P	SD
SSH	106,521	644,164	55.5%	6.05	10.90	0	3,617	13	30.20
BHS	612,628	3,211,001	50.9%	5.24	10.30	0	3,777	11	24.70
PSE	93,700	355,963	42.8%	3.80	8.88	0	2,534	6	25.48
LES	111,730	553,588	54.1%	4.95	9.17	0	25,127	11	80.55
MCS	20,107	83,648	47.6%	4.16	8.74	0	2,557	9	24.34

Note: TNL = total number of Bitly short links; TNC = total number of Twitter clicks; PL($C \geq 1$) = percentage of Bitly short links with at least one Twitter click; Mean (all) = mean value of Twitter clicks of all Bitly short links; Mean (clicked) = mean value of Twitter clicks of Bitly short links with at least one Twitter click; Min = minimum value of Twitter clicks; Max = maximum value of Twitter clicks; 90th P = 90th percentile of Twitter clicks; SD = standard deviation.

6.3.3 Correlation analysis

In order to know more about the relationships between Twitter clicks and other indicators, we selected two scholarly impact indicators (i.e., WoS citations and Mendeley readers) and two Twitter engagement indicators (i.e., total retweets and total likes) to compare their correlations with Twitter clicks. Table 3 lists the results of the Spearman correlation analysis among Twitter clicks, WoS citations, Mendeley readers, total retweets, and total likes.

6

Table 3. Spearman correlation analysis of Twitter clicks, scholarly impact indicators, and Twitter engagement indicators

Indicator	Twitter clicks	WoS citations	Mendeley readers	Total retweets	Total likes
Twitter clicks	1.000	0.094	0.151	0.499	0.458
WoS citations		1.000	0.744	0.071	0.046
Mendeley readers			1.000	0.118	0.114
Total retweets				1.000	0.617
Total likes					1.000

Citation counts and Mendeley readers are moderately correlated (r_s = 0.744), which has also been confirmed by many previous studies (Thelwall & Wilson, 2016; Zahedi et al., 2014). The correlation between the two Twitter engagement indicators – total retweets and total likes received – is also relatively moderate (r_s = 0.617). According to the results, Twitter clicks of Bitly short links correlate positively to scholarly impact indicators and Twitter engagement indicators. In comparison, Twitter clicks are more correlated with the two indicators rooted in the Twitter environment than with the scholarly impact indicators. Both total retweets and total likes received by Bitly short links are moderately associated with Twitter clicks (r_s = 0.499 and r_s = 0.458, respectively), while Twitter clicks only show weak correlations with WoS citations and Mendeley readers of papers (r_s = 0.094 and r_s = 0.151, respectively). The results remain consistent even if only the Bitly short links with at least one Twitter click are considered (See Table 6 in Appendix).

To present the change trends of the analyzed indicators with the number of Twitter clicks increases, Figure 5 shows the relationships between Twitter clicks and **A** WoS citations, **B** Mendeley readers, **C** total retweets, and **D** total likes. For clear visualization, the change trends of the indicators were restricted to the Bitly short links with Twitter clicks ranging from 0 to 20, which account for 94.7% of all Bitly short links in our dataset. The indicators show an upward trend with the number of Twitter clicks. Since the number of Twitter clicks weakly and positively correlates with both WoS citations and Mendeley readers as presented in Table 3, the uptrend of these two indicators is relatively flat. The difference of accumulated citations and readers between relatively highly clicked Bitly short links and those with less Twitter clicks is not significant. By comparison, total retweets and total likes rise at a faster pace because they are moderately correlated with Twitter clicks. In general, Bitly short links with more retweets and likes also tend to receive more clicks on Twitter.

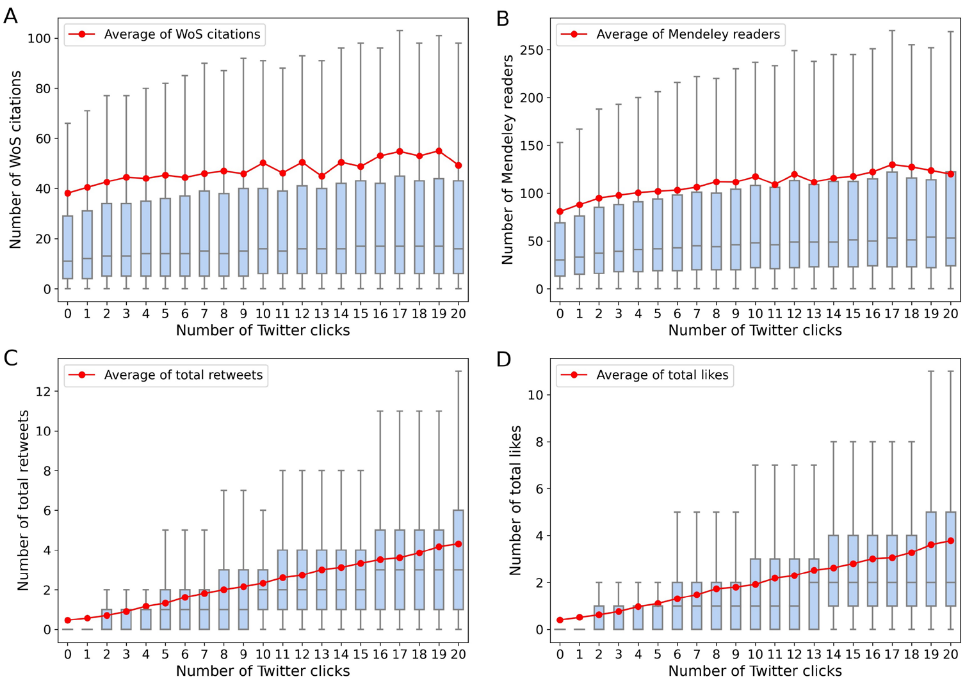

Figure 5. Relationships between the number of Twitter clicks and **A** WoS citations, **B** Mendeley readers, **C** total retweets, and **D** total likes

6.4 Discussion

Though scholarly Twitter mentions have been proven to be able to increase the online visibility and dissemination of scientific papers (Allen et al., 2013; Shu et al., 2018; X. Wang et al., 2017), as of yet, very little work has been done to uncover the mechanism of how scholarly Twitter mentions serve as bridges to direct users to view the tweeted scholarly content. That is to say, even though nowadays scholarly Twitter mentions are widespread and, therefore, are believed to represent a kind of social media attention that scientific papers received, it is still unclear whether they really result in any specific impact by attracting users to access the mentioned scientific papers.

A strong example of the importance of unveiling the mechanism behind the act of tweeting and its actual access to scientific content came from Twitter itself. On June 10, 2020, the official Twitter account of "@TwitterSupport" tweeted that, in order to promote informed Twitter discussion, they were testing a new prompt on Android devices to ask users if they would like to open the link of the article first before they retweet it.[9] This was interpreted as an attempt to nudge some users into rethinking their actions on the social network, and thus improving platform health (Hern, 2020), particularly given the common problem of users sharing links without reading them – it was estimated that 59% of the URLs mentioned on Twitter were not clicked at all (Gabielkov et al., 2016).

In the case of scholarly Twitter mentions, based on a large-scale analysis of over 1.1 million Bitly short links referring to WoS papers, we first found that there are about 49.5% of the Bitly short links that were never clicked on Twitter, while the remaining 50.5% successfully led at least one Twitter user to the detail pages of scientific papers. These empirical results indicate that, although the scholarly Twitter mentions improved the online visibility of papers, nearly half of the tweeted Bitly short links failed to stimulate Twitter users to open the links and increase access to the papers. Put this in the conceptual framework proposed in Figure 1, even if the bridges established by scholarly Twitter mentions widely brought scientific information into the Twitter environment, the assumed two-way information flows only happened on about half of the established bridges, while the remaining bridges did not guide users back to the science environment.

Just as conceptualized by Haustein, Bowman, & Costas (2016), there are various heterogeneous acts that relate to research objects. They summarized three act categories, including *access* (i.e., the acts of accessing and showing interest in the research objects), *appraise* (i.e., the act of mentioning the research objects on various platforms), and *apply* (i.e., the act of using significant parts of, adapting, or transforming the research objects), and

[9] See the detailed tweet texts at: https://twitter.com/TwitterSupport/status/1270783537667551233 (Accessed September 16, 2020).

assumed that the level of engagement of these acts increased from *accessing* over *appraising* to *applying*. Therefore, accessing can be seen as the starting point of a sequence of possible engagement around scientific papers. One may hypothesize that scholarly Twitter mentions could promote the access of papers, thus raising the possibility for them of being appraised or applied next. However, according to our findings, the capacity for increasing the access of scientific papers is limited for the majority of scholarly Twitter mentions, thus precluding the occurrence of subsequent academically related behavior such as viewing, downloading, reading or citing. This large-scale absence of the information flow bridging back from Twitter to science might partly contribute to explain the weak correlations between Twitter mentions and scholarly impact indicators usually found in the altmetric literature (Bardus et al., 2020; Haustein, Costas, et al., 2015; Sugimoto, Work, et al., 2017).

For those Bitly short links with Twitter clicks received, we investigated their temporal accumulation patterns and subject field variations. The distribution of Twitter clicks of Bitly short links shows very similar characteristics to Twitter mentions of scientific papers. For example, Twitter was found to be one of the fast sources as the majority of scholarly Twitter mentions emerged very soon after the papers were accessible (Fang & Costas, 2020; Shuai et al., 2012). Clicks by Twitter users also concentrate in a very short period of time after the Bitly short links were tweeted, with over 60% of Twitter clicks accrued by the following day after the Bitly short links were first tweeted, and about 80.5% accumulated within 10 days. In terms of subject field variations, many prior studies have confirmed that the fields of SSH, BHS, and LES had the highest share of papers with Twitter mention data, while the presence of Twitter mentions in the fields of PSE and MCS was much lower (Costas et al., 2015a; Fang, Costas, et al., 2020; Haustein, Costas, et al., 2015). Our results also show that Bitly short links to papers in the fields of SSH, BHS, and LES are more frequently clicked than those in the fields of MCS and PSE. Regarding the stronger attention paid to SSH and BHS by social media users, Haustein, Costas, et al. (2015) and Costas et al. (2015a) interpreted that it would be motivated by social media users having a preference for engaging with social and health-related topics over other more technical, mathematical or physical topics. Thus, social media users would more likely be intrigued and triggered by the research that are relatively easier to be understood or more closely bound up with social phenomena and healthcare. The results depicted in this paper support this interpretation, since a similar reason can be applied to the subject field variations in terms of Twitter clicks: those short links included in tweets related to SSH and BHS papers would be more frequently clicked due to the stronger interest of Twitter users in social and health-related topics.

In contrast to other engagement behavior with respect to science on Twitter, clicking represents a more deep-seated Twitter reception because Twitter users' attention is not only limited to the tweet content, but spreads to the original content of the tweeted papers. Clicking makes a potential impact on scholarly content by substantially increasing the visits. Therefore, click metrics mirror such potential impact. In this study we also assessed the correlations between this new type of impact indicator and other scholarly impact indicators and Twitter engagement indicators. WoS citations, Mendeley readers, total retweets, and total likes were aggregated at the Bitly short link level to conduct the Spearman correlation analysis with Twitter clicks. Given the confirmed weak or negligible correlations between Twitter mentions and scholarly impact indicators (Bardus et al., 2020; Haustein, Costas, et al., 2015; Sugimoto, Work, et al., 2017), it is not surprising that Twitter clicks showed relatively weak correlations with the two scholarly impact indicators, since the existence of Twitter mentions is the precondition for the generation of Twitter clicks. In comparison, Twitter clicks showed moderate correlations with the two other Twitter engagement indicators, indicating the possible intrinsic relationships among the engagement behavior in the Twitter universe. Put differently, it could be argued that clicks from Twitter to papers are more related to elements and dynamics coming from the Twitter environment (e.g., retweets and likes) than from the science environment (e.g., citations and readership), therefore reinforcing the interest in understanding the social media dynamics and factors that would enable a smoother interaction between the two environments.

This study provides an overview of the extent to which Bitly short links pointing to scientific papers are clicked on Twitter. It has particular implications for quantifying the traffic to scientific papers derived from Twitter and evaluating the Twitter reception of science information in depth. Another implication of this study is that we put clicks as a novel form of subsequent type of impact emanating from social media activities on the science environment. The study of click metrics can be seen as an enrichment of the scope of altmetrics, by incorporating a new quantitative and reproducible approach, able to expand the perspectives of how science is used and received in the Twittersphere. It must be noted that clicking the short links to scientific papers doesn't necessarily entail the reading of them. After clicking the tweeted URLs to papers, users might only briefly view the abstract or just look into the figures, but from a conceptual point of view, these two examples both represent deeper levels of engagement with the scholarly material than just the social media activities (e.g., retweeting or liking tweets) usually tracked by altmetric sources. As mentioned before, the opening of URLs before retweeting them is officially encouraged by Twitter for the sake of "promoting informed discussion". This suggests that the idea of bridging to and from social media regarding external information (e.g., scholarly papers, news media) is seen by Twitter as a relevant form of engagement, important to improve the validity and pertinence of the information circulating on Twitter.

In order to better understand what kind of Bitly short links to scientific papers are more likely to be clicked on Twitter (i.e., what are the features that improve bridging the information flow from the Twitter environment back to the science environment?), future research should explore the potential influencing factors on the clicking behavior of Twitter users, considering three main dimensions of relevant science-social media interaction features: (1) the bibliometric features of the tweeted papers (e.g., journal impact factor, authors' reputation, open access status); (2) the textual and interactive features of the tweets (e.g., originality of tweet texts, sentiment, mentions to other users, inclusion of hashtags, number of retweets or likes); and (3) the activities and profile features of the Twitter users (e.g., number of followers, degree of science focus, number of tweets posted, description as academic users). Moreover, given the different levels of Twitter clicks observed across subject fields, it would be necessary to explore the possible causes of these clicking disciplinary differences, particularly by comparatively studying the tweeting behavior of Twitter users across different disciplinary contexts.

There are some limitations that should be acknowledged in this study. First, the employed methodology is highly dependent on the click metrics provided by Bitly, so for short links generated through other services or those unshortened URLs there is no reliable way to harvest click metric data. Second, Twitter users come from very different social groups (e.g., academic users, professionals or the general public), and they may access Twitter from different devices (e.g., mobile phones, tablets or computers). This diversity of users and the devices used for accessing Twitter might have some influence on their clicking behavior. However, the click metrics provided by Bitly are aggregated numbers at the short link level. This means that it is not currently possible to conduct more in-depth research about who are the clicking users and from what kinds of devices they clicked the short links. Such lack of data, which is also bound by legal privacy constrains, hinders the possibility of studying the clicking behavior of Twitter users from a more fine-grained perspective, particularly in order to understand better whether the devices from where the users access social media platforms may also be related to their clicking behavior (e.g., it may be that users who access Twitter from their mobile phones are less prone to click the URLs of scientific papers, as this is a type of device less friendly for reading longer scholarly texts), or whether some types of academic users (e.g., researchers, students) are more prone to click the URLs of scholarly content, in contrast to other non-academic users (e.g., professionals, the general public). Third, only those Bitly short links using the "bit.ly" domain were considered, while customized domains of some companies were not included in this study (e.g., the "go.nature.com" domain customized by some *Nature* journals). Finally, there is still a remote possibility that some Twitter users would copy & paste or type the short links in their browsers instead of directly clicking them from the tweet. In such cases, although the paper would get actual visits, they would not be counted as valid Twitter clicks as in this study, and therefore the actual number of Twitter clicks would be underestimated. Ideally, it should be Twitter directly the one that could best report the number of clicks resulting from the tweets, thus making it possible the more thorough and systematic study of the clicking behavior of Twitter users around scientific papers. Such type of information, as announced by the aforementioned tweet of "@TwitterSupport", would also support the aim of Twitter to "help promote informed discussion" by better understanding how scientific content gets disseminated and actually accessed by its users.

6.5 Conclusions

The bridges established by scholarly Twitter mentions enable the analysis of bidirectional information flows between the science and Twitter environments. The information flow going from the science environment to the Twitter environment (e.g., by someone tweeting scientific papers) has been extensively studied, constituting an important part of the altmetric research literature. However, whether the tweeting of scientific papers was related to Twitter users also going back to the science environment (e.g., by users clicking the tweeted URLs of papers) is a form of information flow that has remained largely unexplored in the altmetric research area. This study represents the first attempt of studying how Twitter users try to access the scientific information embedded in tweets, by analyzing their clicks to Bitly short links to scientific papers. Based on the click metrics of over 1.1 million Bitly short links referring to WoS papers, we found that nearly half of them were not clicked on Twitter at all, and the majority of Bitly short links performed poorly in attracting Twitter users to access the original scholarly content, thereby revealing that most scholarly Twitter mentions played a relatively ineffective role in driving traffic from Twitter back to the science environment. There are still some Bitly short links that have received a substantial number of Twitter clicks. When this is the case, and there are substantial clicking activities around scientific papers, they showed similar characteristics as general scholarly Twitter mentions in terms of both accumulation speed and subject field variations. Timewise, Twitter users tended to click short links to scientific papers within a short period of time after they were tweeted. Twitter users also showed stronger preferences for clicking links to papers from the fields of SSH, BHS, and LES, arguably to access social and health-related research. Papers in the fields of PSE and MCS tended to be less frequently clicked, also in accordance with the lower tweeting activities previously reported for these fields. Finally, Twitter clicks were more correlated with other Twitter engagement indicators such as total retweets and total likes, rather than with scholarly impact indicators (WoS citations and Mendeley readers), suggesting that Twitter clicks are more a form of *Twitter engagement indicator*, rather than an academic-related impact indicator.

Building on the findings of this study, it is clear that there is a future research agenda regarding the understanding of the mechanisms of the Twitter reception of scientific information, particularly from an interactive point of view, in which many factors both from the Twitter environment (e.g., users, retweets, likes, conversations) as well as from the science environment (e.g., reputation of authors, journals, topics), interplay in order to attract broader audiences to scientific content, paving the way to the evaluation of the success of Twitter dissemination strategies of scientific knowledge.

6.6 Appendix

Table 4. Usage rates of short links generated by four link shortening services over the tweet post years

Post year	Total number of tweeted URLs	Bitly short links	Ow.ly short links	Goo.gl short links	TinyURL short links
≤ 2012	561,932	175,082 (31.2%)	51,681 (9.2%)	18,145 (3.2%)	8,575 (1.5%)
2013	833,308	172,168 (20.7%)	93,659 (11.2%)	28,050 (3.4%)	10,553 (1.3%)
2014	1,141,654	177,701 (15.6%)	113,018 (9.9%)	33,202 (2.9%)	11,377 (1.0%)
2015	1,655,277	226,516 (13.7%)	154,993 (9.4%)	40,780 (2.5%)	12,272 (0.7%)
2016	1,881,691	250,198 (13.3%)	161,301 (8.6%)	61,455 (3.3%)	11,691 (0.6%)
2017	1,533,613	113,547 (7.4%)	120,925 (7.9%)	66,630 (4.3%)	9,533 (0.6%)
Total	7,446,310	1,103,819 (14.8%)	693,386 (9.3%)	247,623 (3.3%)	63,334 (0.9%)

Note: For URLs which have been tweeted in multiple years, full counting was used to calculate the number of URLs in each year.

Table 5. Usage rates of Bitly short links across the five main subject fields

Subject field	Total number of tweeted URLs	Number of tweeted Bitly short links	Proportion
Social Sciences and Humanities	753,646	106,521	14.1%
Biomedical and Health Sciences	4,196,291	612,628	14.6%
Physical Sciences and Engineering	536,380	93,700	17.5%
Life and Earth Sciences	898,614	111,730	12.4%
Mathematics and Computer Science	122,930	20,107	16.4%

Table 6. Spearman correlation analysis of Twitter clicks, scholarly impact indicators, and Twitter engagement indicators considering only the Bitly short links with at least one Twitter click

Indicator	Twitter clicks	WoS citations	Mendeley readers	Total retweets	Total likes
Twitter clicks	1.000	0.083	0.141	0.562	0.519
WoS citations		1.000	0.752	0.045	0.013
Mendeley readers			1.000	0.082	0.078
Total retweets				1.000	0.615
Total likes					1.000

CHAPTER 7

Discussion and conclusions

7.1 Introduction

Social media have seen a remarkable rise in their scholarly use of communicating scientific information by both scholars and the public, leading to an abundance of traces left by diverse interaction behaviors around science in social media environments (Adie & Roe, 2013; Priem, Groth, et al., 2012). Given that these traces are usually generated in non-academic contexts, it has been discussed that they may hold particular promise for new lenses on plural societal considerations of science (Bornmann, 2014b; Noyons, 2019; Priem et al., 2010). Based on the traces, social media metrics came out as an endeavor to quantitatively measure and analyze interactions related to scholarly objects on social media platforms (Haustein, Bowman, & Costas, 2016; Wouters et al., 2019). The advent of social media metrics appeals to the increasing need in both the scientific community and science policy to value all kinds of research products and value their polymorphous benefits and impacts beyond academia (Piwowar, 2013; Wilsdon et al., 2015; Wouters & Costas, 2012).

As one of the most popular social media platforms, Twitter plays a significant role in shaping the realm of social media metrics because it contributes the second largest amount of evidence of online interactions regarding science, only behind Mendeley (Sugimoto, Work, et al., 2017). More importantly, enabled by the rich interactive features, Twitter interactions around science may come with extra information added by users and may be followed by a series of further interactions from broader audiences, potentially supplementing science stories with online evidence concerning the accuracy, importance, and popularity of scholarly outputs (Brossard & Scheufele, 2013). Therefore, *scholarly Twitter metrics* (Haustein, 2019), which specifically study Twitter interactions related to scholarly objects, open up a range of opportunities to depict social media attention and attitudes towards scientific developments, and in reverse, to add extra meaning to science stories through the lens of the broader public.

Based on the conceptual framework of studying science-social media interactions proposed in chapter 1, this PhD dissertation presents an overview of Twitter interactions related to scientific papers from three main perspectives: (1) *Twitter uptake of scientific papers* which results in scientific information flowing to Twitter, (2) *Twitter interactions with scholarly tweets* mentioning scientific papers which generate information flows within Twitter, and (3) *scholarly uptake of scholarly tweets* reflected in the behavior of clicking tweeted URLs to access scientific papers, which leads to information flowing from Twitter back to science. By exploring these interactions from both conceptual and empirical aspects, this dissertation seeks to unravel the diversity of Twitter interactions around science and severally describe their characteristics in a systematic manner. Furthermore, on the basis of a better understanding of the characteristics of diverse forms of Twitter interactions, this dissertation reflects on Twitter-based metrics and discusses the possibility of establishing a more fine-grained indicator system to assess Twitter reception of scientific information in greater depth.

As the discussion and conclusions part, this chapter, first, summarizes the main findings under each research question proposed in chapter 1. Then this chapter further discusses the implications of the main findings for scholarly Twitter metrics. Finally, this chapter presents several perspectives for future research.

7.2 Summary of main findings

For the five primary research questions proposed in chapter 1, this PhD dissertation presents answers in chapters 2 to 6, respectively. This section summarizes the main findings under each research question.

RQ1. *To what extent are scientific papers mentioned on Twitter? In particular, which subject fields and research topics are more likely to have related scientific papers mentioned on Twitter?*

To answer this research question, chapter 2 presents an extensive analysis of the overall Twitter uptake of scientific papers recorded by Altmetric.com, in comparison to other eleven types of altmetric data sources, including Mendeley, Facebook, news media, blogs, Wikipedia, policy documents, Reddit, Faculty Opinions (formerly F1000Prime), video platforms (i.e., YouTube), peer review platforms (i.e., PubPeer and Publons), and Q&A platforms (i.e., Stack Overflow).

For a total of nearly 12.3 million Web of Science-indexed (WoS) scientific papers published between 2012 and 2018, chapter 2 examines how many of them are mentioned by the selected altmetric data sources based on the altmetric events recorded by Altmetric.com until October 2019. Overall, Mendeley presents the highest level of uptake of scientific papers, with over 89% of scientific papers receiving at least one Mendeley reader, followed by Twitter which mentions 34% of scientific papers for at least once. In contrast, other altmetric data sources show much lower levels of uptake of scientific papers. The proportion of papers mentioned on other data sources ranges from as low as 0.06% for Q&A platforms up to 8.57% for Facebook. In a word, Mendeley, together with Twitter, are the most important altmetric data sources showing the highest levels of uptake of scientific papers, whereas the uptake of scientific papers on other altmetric data sources is substantially low. So except Mendeley and Twitter which are possibly useful to capture social media attention towards science on a large scale, as suggested by Thelwall, Haustein, et al. (2013), other altmetric data sources may only be applicable to "identify the occasional exceptional or above average article rather than as universal sources of evidence".

Compared to some earlier large-scale studies (Costas et al., 2015a; Haustein, Costas, et al., 2015; Robinson-Garcia et al., 2014), the level of Twitter uptake of scientific papers reported in chapter 2 is generally higher. As a state-of-the-art analysis, chapter 2 includes more recently published papers and they are found to be more advantaged in garnering Twitter mentions, confirming a “recent bias” (Costas et al., 2015a) of Twitter uptake of scientific papers. Such recent bias is also observed for some other sources like Facebook, news media, blogs, and Reddit. However, some altmetric data sources, such as Mendeley, Wikipedia, policy documents, and Q&A platforms, present a “past bias” in the light of their higher uptake of relatively old papers as traditional citations, hinting the different pace of altmetric data sources in the uptake of scientific papers which is further analyzed in chapter 3.

From a disciplinary point of view, Twitter uptake is relatively higher for papers from the fields of *Social Sciences and Humanities*, *Biomedical and Health Sciences*, and *Life and Earth Sciences*, while the uptake is lower for papers from the fields of *Physical Sciences and Engineering* as well as *Mathematics and Computer Science*, feeding into the narrative that more Twitter attention is paid to scientific information regarding society, healthcare, and environment than those more technical, mathematical, or physical/chemical (Costas et al., 2015a; Haustein, Peters, Sugimoto, et al., 2014). There are several possible causes behind these disciplinary biases, such as the lay audiences’ preference for topics related to social issues, environmental problems, and healthcare (Haustein, Costas, et al., 2015; Haustein, Peters, Sugimoto, et al., 2014), the higher degree of Twitter uptake by scholars from social sciences, health sciences, and life and earth sciences than those specializing in natural sciences and engineering (Costas et al., 2020; Mohammadi et al., 2018), and the more frequent science communication with the public conducted by scholars from social sciences and humanities than scholars from natural sciences and technology (Bentley & Kyvik, 2011; Kreimer et al., 2011; H. P. Peters, 2013). Furthermore, by digging deep into the research topic level, chapter 2 identifies research topics with most papers receiving intensive Twitter attention for each subject field studied, namely, the so-called “hot research topics” in the eyes of Twitter users. For instance, in the field of *Social Sciences and Humanities*, it is found that research topics with higher Twitter uptake are those related to gender/sex, education, climate, and psychological problems. In the field of *Biomedical and Health Sciences*, hot research topics on Twitter are those about daily health keeping, worldwide infectious diseases, lifestyle diseases, and emerging biomedical technologies. In the field of *Life and Earth Sciences*, research topics concerning animals and natural environment problems attract the most Twitter attention. As the two subject fields with the generally lowest Twitter uptake, *Physical Sciences and Engineering* and *Mathematics and Computer Science* also have some of their research topics relatively more frequently picked up by Twitter users, such as universe/astronomy and quantum within *Physical Sciences and Engineering*, and emerging information technologies and robotics within *Mathematics and Computer Science*.

Findings presented in chapter 2 confirm Twitter as a global source of evidence of social media uptake of scholarly outputs, on the basis of its second highest level of uptake of scientific papers as well as the increasing presence of Twitter mention data for scientific papers published over years. Chapter 2 also highlights the biases of Twitter uptake towards different publication years, subject fields, and research topics of scientific papers. Such biases, on the one hand, support the idea that scholarly Twitter metrics may have more added values for some subject fields where citations relatively sparsely distribute (e.g., *Social Sciences and Humanities*) (Costas et al., 2015a); on the other hand, Twitter biases towards specific research topics open the possibility to monitor research topics or trends of particular interest by broader audiences in social media environments. Besides, given the significant differences in the overall levels and disciplinary biases of the uptake of scientific papers across altmetric data sources, chapter 2 reinforces the necessity of keeping altmetrics separate in future analyses or research assessments, rather than compounding them as composite indicators at the expense of the robust interpretation of the metrics (Thelwall, 2020).

RQ2. *When are scientific papers mentioned on Twitter after publication? In other words, how fast do Twitter mentions of scientific papers accumulate after papers are published?*

To address this research question, chapter 3 studies the accumulation velocity of Twitter mentions to scientific papers recorded by Altmetric.com at the day level after papers were published. The velocity of the Twitter uptake of scientific papers is also compared with other eleven altmetric data sources, including Facebook, news media, blogs, Wikipedia, policy documents, Reddit, Google+ (now defunct), Faculty Opinions (formerly F1000Prime), video platforms (i.e., YouTube), peer review platforms (i.e., PubPeer and Publons), and Q&A platforms (i.e., Stack Overflow).

For a total of nearly 2.4 million WoS papers published between 2012 and 2016, chapter 3 investigates the uptake velocity of them on the selected altmetric data sources. In order to study the velocity on a smaller time scale, chapter 3 adopts the DOI *created date* recorded by Crossref as the more precise proxy of the publication date of scientific papers (Haustein, Bowman, et al., 2015). Combined with the *posted date* recorded by Altmetric.com for altmetric events, chapter 3 measures the uptake velocity of scientific papers on specific altmetric data sources by calculating the day time intervals between the publication date of scientific papers and the posted date of altmetric events. Through the lens of the three indicators used for measuring velocity from both flexible and fixed perspectives, i.e., *Velocity Index* (after the publication of scientific papers, the proportion of altmetric events that happened in a specific time interval, e.g., one day, one month, or one year), *altmetric half-life* (the number of days until half of specific altmetric events have appeared), and *altmetric time delay* (the number of days between the publication of a paper and its first altmetric event on a specific altmetric data source), chapter 3 reports that the uptake velocity of scientific papers varies substantially across altmetric data sources.

Although *speed* is regarded as one of the characteristic properties of altmetrics (Bornmann, 2014a; Wouters & Costas, 2012), in chapter 3 it is demonstrated that this property is not held by all sorts of altmetric data sources. Amongst the analyzed sources, Twitter shows the highest velocity in terms of the uptake of scientific papers after they were published. Overall, more than 60% of scholarly tweets mentioning scientific papers accumulated within the first month after papers had been published, and nearly 90% happened within the first year. In consideration of the relatively high velocity, Twitter is labeled as one of the *fast sources*, along with several other altmetric data sources showing a similar high velocity in the uptake of scientific papers, such as Reddit, news media, Facebook, Google+, and blogs. However, chapter 3 observes that there exist some altmetric data sources taking a relatively long time to accumulate substantial altmetric events regarding scientific papers, being possible to be labeled as *slow sources*. The slow sources include policy documents, Q&A platforms, Wikipedia, video platforms, and Faculty Opinions, etc.

Chapter 3 also presents the variations of the Twitter uptake velocity of scientific papers across document types and subject fields. In terms of document types, Twitter mentions to editorial materials and letters accumulate faster than document types of article and review. This pattern is similar to citations which accumulate in a relatively short time period to editorial materials and letters, but present a slower growth to reviews (Costas et al., 2010; J. Wang, 2013). One possible reason for the more immediate Twitter attention to editorial materials and letters is that, in contrast to the less innovative nature of reviews, they cover more novel topics, debates, and scientific news, etc., without using a too complicated and technical language (Haustein, Costas, et al., 2015). From the disciplinary perspective, scientific papers from the fields of *Physical Sciences and Engineering* and *Life and Earth Sciences* are found to accumulate Twitter mentions faster than other fields like *Social Sciences and Humanities*, *Mathematics and Computer Science*, and *Biomedical and Health Sciences*. Despite the obvious difference of the overall accumulation dynamics between Twitter mentions and citations, similar disciplinary variations were also observed in the context of citations, for example, citations to scientific papers in the fields of physical, chemical, and earth sciences were generally found to accrue faster and also decline faster than social sciences, mathematics (Abramo et al., 2011; Aksnes, 2003; Glänzel & Schoepflin, 1995).

Findings presented in chapter 3 empirically confirm that Twitter is one of the altmetric data sources holding the property of *speed*, with the majority of Twitter mentions accrued shortly after the publication of scientific papers. The high uptake velocity of scientific papers on Twitter enables more real-time assessments of the attention drawn by scholarly outputs in their early life cycle (Priem et al., 2010). Nevertheless, based on the comparison of the uptake velocity amongst altmetric data sources, it is demonstrated that the property of speed is not universally applicable to all kinds of altmetric events. The data accumulation pace takes on different patterns between fast sources (e.g., Twitter, Reddit, and news media) and slow sources (e.g., Wikipedia, Q&A platforms, and policy documents). Given that time affects differently to different altmetric data sources, it is important to take into account these time differences when selecting appropriate observation time windows in altmetric research. Besides, the fundamental differentiation amongst altmetric data sources in terms of both the broadness (observed in chapter 2) and velocity (observed in chapter 3) of the uptake of scientific papers, collectively supports the idea that keeping altmetric events separate, rather than merging them into compound metrics, seems to be an important recommendation (Haustein, Bowman, & Costas, 2016; Thelwall, 2020; Wouters et al., 2019).

RQ3. *To what extent do scholarly tweets get engaged with through different engagement behaviors (i.e., liking, retweeting, replying, and quoting)?*

This research question focuses on the Twitter interactions with scholarly tweets as the research objects, moving from primary Twitter metrics towards secondary Twitter metrics (Díaz-Faes et al., 2019). After making a distinction between different tweet types as well as different Twitter engagement behaviors, chapter 4 answers this research question based on a large-scale and cross-disciplinary analysis of how original scholarly tweets, which originally bring scientific information to the Twittersphere, are further engaged with by Twitter users through four types of engagement behaviors, i.e., liking, retweeting, replying, and quoting.

For a set of 7 million original scholarly tweets referring to WoS papers published between 2016 and 2018, chapter 4 collects their Twitter user engagement metrics, i.e., number of likes, number of retweets, number of replies, and number of quotes, with the Twitter API in February 2021. Overall, 52% of the studied original scholarly tweets have been engaged with through at least one of the four types of engagement behaviors. Specifically, the coverage of the four types of user engagement metrics varies a lot, with *likes* as the most popular engagement metric covering 44% of the scholarly tweets, followed by *retweets* which is present for 36% of the scholarly tweets. In contrast, the coverage of *quotes* and *replies* is as low as 9% and 7%, respectively. The behaviors of quoting and replying represent a higher level of engagement over retweeting and liking, because the former two have additional views expressed by users and may initiate public Twitter conversations while the latter two do not. Therefore, it is not surprising to find that as the level of engagement grows, the coverage of user engagement behaviors becomes lower.

A disciplinary point of view is also applied to compare the levels of engagement metrics across original scholarly tweets mentioning scientific papers from different subject fields. As the subject fields found to show higher Twitter uptake in chapter 2, *Social Sciences and Humanities*, *Life and Earth Sciences*, and *Biomedical and Health Sciences* also have their related scholarly tweets more actively engaged with by Twitter users via liking, retweeting, quoting, as well as replying, outperforming *Physical Sciences and Engineering* and *Mathematics and Computer Science*. So in the Twitter universe, Twitter attention is consistently biased towards scientific information concerning social issues, environmental problems, and healthcare, whether from the perspective of the overall Twitter uptake of scientific papers or the Twitter engagement triggered afterwards. As discussed earlier in chapter 2, such disciplinary biases might be driven by the particular interest of lay audiences in social and health-related topics (Haustein, Costas, et al., 2015; Haustein, Peters, Sugimoto, et al., 2014), or the stronger motivation of scholars from social sciences and humanities to leverage Twitter as the tool for scholarly communication or science communication (Bentley & Kyvik, 2011; Costas et al., 2020; Kreimer et al., 2011; Mohammadi et al., 2018).

To better understand the characteristics of user engagement metrics through their relationships with other factors, chapter 4 selects ten factors for analysis from three main dimensions: the first dimension is *scholarly impact* of tweeted scientific papers, including number of citations and Mendeley readers received by papers; the second dimension is use of *tweet features* in scholarly tweets, including number of hashtags and mentioned users in tweets; the third dimension is *user characteristics* of those who posted original scholarly tweets, including number of followers, lists listed, friends, likes given, tweets posted, and science focus (i.e., the proportion of scholarly tweets amongst all tweets posted by a user). Based on both correlation analysis and regression analysis, user engagement metrics of scholarly tweets are found to be negligibly correlated with scholarly impact factors of tweeted papers (i.e., science-based factors), consistent with the previous observations of the weak or no correlations between citations and the overall Twitter uptake of scientific papers (Bardus et al., 2020; Zahedi et al., 2014). The negligible correlations between science-based factors and user engagement metrics on Twitter add more empirical evidence to the idea that science and Twitter focus on different aspects of scholarly outputs, and conform to different spaces of interactions. In comparison, user engagement metrics generally tend to be more related to factors about tweet features and user characteristics, namely, Twitter-based factors, particularly for number of mentioned users in original scholarly tweets and number of followers owned by users posting original scholarly tweets. The stronger correlations observed between user engagement metrics and Twitter-based factors indicate the intrinsic relationships amongst Twitter elements and activities.

Chapter 4 provides a first overview of the Twitter interactions happened around original scholarly tweets. Findings in chapter 4 quantitatively prove the globally low presence of user engagement behaviors, especially for those with extra informative content added, such as quoting and replying. The fact that almost half of original scholarly tweets did not trigger any user engagement after being posted implies that Twitter users' attention paid to scholarly tweets varies significantly. Such variations indicate that original scholarly tweets, as the carriers of scientific information, perform differently in terms of the attention received from broader audiences on Twitter. In other words, different original scholarly tweets contribute to the visibility of scientific papers on Twitter to varying degrees. From a practical point of view, distinguishing between tweet types (e.g., original tweets, retweets, reply tweets, and quote tweets) and considering further user engagement triggered by scholarly tweets would be an important step towards more in-depth assessments of Twitter reception of scientific information.

RQ4. *To what extent and for what reasons do scholarly tweets become unavailable as time goes by? What is the potential effect that the unavailability of tweets may make on the stability of Twitter metrics?*

To answer these research questions, chapter 5 examines the stability of Twitter metrics with a case study consisting of the most tweeted scientific papers recorded by Altmetric.com. For a set of 1,154 scientific papers which have been tweeted by at least 1,000 unique Twitter users up to October 2017, chapter 5 rechecked their totally 2.6 million scholarly tweets in April 2019 with the Twitter API. For the unavailable tweets identified during the tweet recheck, the error codes responded by the Twitter API were collected to further explore the specific reasons why they became unavailable to the public.

Results presented in chapter 5 show that overall 14.3% of the scholarly tweets to the most tweeted scientific papers in the case study have become unavailable. It is not surprising to find that the potential risk of scholarly tweets being unavailable increases over time, as the scholarly tweets posted for a long time show a relatively higher proportion of unavailability. Based on the error codes responded by the Twitter API for the unavailable tweets, chapter 5 demonstrates that tweet unavailability is mainly caused by some specific interaction behaviors by users as well as some actions taken by Twitter itself. The most common reason for the tweet unavailability is tweet deletion by users, accounting for 54.7% of the unavailable tweets detected. The second most important unavailability reason comes from user account suspension implemented by Twitter for user accounts violating certain Twitter rules,[1] which results in 25.9% of unavailable tweets. In addition, the behavior of protecting accounts, which prevents other unauthorized users from accessing all the tweets posted by the accounts, contributes to 16.7% of the unavailable tweets.

The identified unavailable tweets influence the stability of Twitter metrics of scientific papers to different extents. Most highly tweeted papers studied in chapter 5 have less than 20% of Twitter mentions falling unavailable, however, there are also many scientific papers with their Twitter mentions showing extremely high unavailability rates. For example, there are ten papers with over 90% of their Twitter mentions identified unavailable during the tweet recheck, which means that if their unavailable tweets are removed, the overall count of Twitter mentions of these papers would plummet dramatically. As a result, the unavailability of scholarly tweets may exert serious effects on the stability of Twitter metrics.

[1] See more information about suspended Twitter accounts at: https://help.twitter.com/en/managing-your-account/suspended-twitter-accounts (Accessed August 31, 2021).

In order to figure out what kinds of scientific papers are more likely to have their Twitter metrics seriously affected by the unavailability of scholarly tweets, chapter 5 examines the *Twitter dissemination structures* of scientific papers. For each scientific paper, chapter 5 distinguishes between their original tweets and retweets and connects them according to the retweeting relationships, based on which two indicators *Degree of Originality* (the proportion of original tweets amongst all the scholarly tweets that a paper received) and *Degree of Concentration* (the degree to which retweets concentrate on a single original tweet that a paper received) are developed to depict the basic Twitter dissemination structure. On the basis of these two indicators, it is found that most scientific papers with relatively high tweet unavailability rates (i.e., the proportion of unavailable tweets) show a low *Degree of Originality* (i.e., fewer original tweets received) and a high *Degree of Concentration* (i.e., more retweets concentrating on a single original tweet). For papers with such Twitter dissemination structure, once a highly retweeted original tweet becomes unavailable, all its related retweets would become unavailable synchronously,[2] thus leading to a drastic decline in the overall Twitter metrics of the paper in question. Therefore, studying the Twitter dissemination structures of scientific papers has the potential to monitor scientific papers that are at a greater risk of suffering from extremely unstable Twitter metrics.

Findings reported in chapter 5 emphasize the importance of noticing the volatile nature of Twitter data in scholarly Twitter metrics. Different from the more traditional scientometric data (e.g., publications, citations) which are in theory stable and persistent, the state of Twitter data is much more vulnerable due to the diverse user interaction behaviors (e.g., deleting tweets, protecting accounts) as well as Twitter surveillance (i.e., suspending accounts) which can easily change the (un)availability of tweets. The unavailability of tweets would lead to the inconsistency of the observations even based on the same Twitter dataset but rechecked at different time points (Bastos, 2021; J. M. Xu et al., 2013; Zubiaga, 2018), thereby undermining the reliability of Twitter metrics and the robustness of interpretation. Related data quality issues were also observed by previous research for some other altmetric data sources, such as Mendeley (Bar-Ilan, 2014; Zahedi et al., 2017), blogs and news media (Ortega, 2019b), as well as Facebook (Yu et al., 2021). Moreover, in chapter 5 the attempt made to describe the Twitter dissemination structures of scientific papers demonstrates the relevance of investigating how papers are tweeted rather than just counting their Twitter mentions. Differentiating between tweet types and applying network approaches to unravel dissemination patterns are essential for a better understanding not only of the Twitter reception of scholarly outputs, but also the potential risk of unstable Twitter metrics.

[2] See more information about tweet deletion at: https://help.twitter.com/en/using-twitter/delete-tweets (Accessed August 31, 2021).

RQ5. *To what extent do scholarly URLs to scientific papers embedded in scholarly tweets get clicked?*

This research question focuses on clicking behavior around scholarly URLs to scientific papers, a special Twitter engagement which can generate information flowing back to science from Twitter by directing Twitter users to the webpages of the tweeted scholarly content. Chapter 6 provides answers to this research question relying on the click metrics recorded by Bitly (http://bitly.com), a link shortening service which can be used to not only generate short links but also track a spectrum of click metrics for the generated short links, such as total number of clicks received after a link was generated, number of clicks happened on different sources where the link was posted, and number of clicks happened on different dates.

For over 1.1 million unique Bitly short links referring to WoS papers (i.e., scholarly URLs) extracted from the scholarly tweets recorded by Altmetric.com until October 2017, chapter 6 collects their click metrics through the Bitly API in December 2019. From the click metrics collected, number of clicks happened on Twitter (i.e., Twitter clicks) are extracted to study how scholarly URLs are clicked by users after they were tweeted as the portals to scientific papers. Similar to the overall degree to which scholarly tweets are engaged with through the four types of engagement behaviors studied in chapter 4, nearly half of the analyzed Bitly short links were not clicked at all. Results show that of the 1.1 million Bitly short links, about 50.5% were clicked by Twitter users at least once and only 10.3% were clicked for more than ten times. Although the proportion of zero-click URLs in the scholarly context is slightly lower than the general estimation made by Gabielkov et al. (2016), which speculates that there may exist about 59% of the URLs posted on Twitter never clicked, it remains noteworthy that only a limited share of scholarly tweets successfully stimulate substantial numbers of Twitter users to access the original content of scholarly outputs.

For Bitly short links that received Twitter clicks, chapter 6 investigates the temporal accumulation pattern and subject field variations of Twitter clicks. It is observed that Twitter clicks of scholarly URLs show an immediate accumulation pattern, with more than 60% of Twitter clicks happened in two days after the Bitly short links were first tweeted, and more than 80% accrued within the first ten days. Such high accumulation velocity is similar to that of Twitter mentions to scientific papers as observed in chapter 3, suggesting a consistent feature of Twitter interactions in the temporal dimension. The consistency is also reflected in the disciplinary biases. As concluded in both chapter 2 and chapter 4, the overall Twitter uptake as well as Twitter user engagement is biased towards scientific papers from the fields of *Social Sciences and Humanities*, *Biomedical and Health Sciences*, and *Life and Earth Sciences*. Similarly, through the lens of Twitter clicks, Twitter users are found to click more on the scholarly URLs referring to scientific papers from the aforementioned subject fields which are more active in the Twitter context. In contrast, scholarly URLs related to *Physical Sciences and Engineering* and *Mathematics and Computer Science* are generally less clicked on Twitter. Moreover, the negligible correlations between scholarly impact indicators and Twitter-based metrics are again verified through the lens of Twitter clicks. Number of Twitter clicks also correlates weakly with scholarly impact indicators for scientific papers, like number of citations and Mendeley readers, whereas correlates moderately with other Twitter engagement metrics for scholarly tweets, such as number of retweets and likes.

In chapter 2, Twitter is confirmed to be one of the most important altmetric data sources making mention of about one third of the recent scientific papers, leading to substantial information flowing from science to Twitter through the bridges established by original scholarly tweets. Although in previous research Twitter was found to be the most important social media source that directs visitors to the webpages of scientific papers (X. Wang, Fang, & Guo, 2016), click metrics of scholarly URLs studied in chapter 6 suggest that two-way information flows between science and Twitter only happened on about half of the established bridges, as only half of the tweeted scholarly URLs succeed in guiding users back to science. These findings further reinforce the idea that scholarly tweets perform differently in attracting the attention of broader audiences and in promoting the visibility of scientific papers. Click metrics of scholarly URLs can serve as a novel mirror of the effectiveness of scholarly tweets in disseminating scientific information. In addition, unlike the four types of user engagement metrics examined in chapter 4, Twitter clicks represent a deeper level of Twitter reception that can materially increase the access to scientific papers, thus opening a window for monitoring the direct impact of Twitter dissemination on scientific knowledge consumption.

7.3 Implications of main findings

The research findings presented in chapters 2 through 6 are framed around the idea of revealing the diversity of Twitter interactions around science, and unveiling the characteristics unique to or shared by different Twitter interactions. Building on a better understanding of the diversity and characteristics of Twitter interactions, this section sets out to further discuss the implications of the main findings for approaching more advanced Twitter-based metrics.

7.3.1 Demonstrating Twitter as a global and immediate source of evidence of interactions around science

Based on the large-scale comparison of the presence of different sources of altmetric data for scientific papers, Twitter is confirmed as the most global source of evidence for social media interactions around science, only next to Mendeley. Given that Mendeley readership has been seen more as a proxy of scholarly impact due to the predominant user groups with academic backgrounds (Mohammadi et al., 2015; Zahedi & Van Eck, 2018) and its moderate to strong correlations with traditional scholarly impact indicator, i.e., citations (Li et al., 2012; Thelwall & Wilson, 2016; Zahedi et al., 2017), Twitter, instead, functions as the most important data source for capturing attention to science beyond academia. This argument is based on both the considerable number of non-academic Twitter users involved in the communication of scientific information (Mohammadi et al., 2018; Yu et al., 2019), and the negligible correlations found between Twitter mentions and citations received by scientific papers (Costas et al., 2015a; Haustein, Larivière, et al., 2014). Findings presented in this dissertation confirm not only the negligible correlation between the Twitter uptake of scientific papers and citations, but also the weak or negligible correlations between diverse Twitter interactions around scholarly tweets (e.g., likes, retweets, quotes, replies, or clicks) and scholarly impact indicators of the tweeted papers (e.g., citations or Mendeley readers). These negligible correlations between scholarly impact indicators and Twitter-based metrics reinforce the idea that Twitter interactions capture different aspects of scientific performance in contrast to scholarly activities, which was introduced as one of the important characteristics of altmetrics by Wouters & Costas (2012) and then termed as *broadness* by Bornmann (2014a).

Besides broadness, *speed* was also listed as an important characteristic property of altmetrics (Bornmann, 2014a; Wouters & Costas, 2012). This dissertation empirically demonstrates that speed is indeed a distinguishing property of Twitter interactions around science. Not only do scholarly tweets mentioning scientific papers emerge very fast after the publication of papers, but also the majority of Twitter clicks on tweeted scholarly URLs occur shortly after the exposure of scholarly tweets. The overall high accumulation velocity of Twitter interactions makes it possible to track reactions to scientific papers on a much shorter time scale, underpinning the value of Twitter as an immediate data source for offering real-time evidence of social media attention in the very early stage of the life cycle of scholarly outputs (Ortega, 2018b; Yu et al., 2017).

Although Twitter is demonstrated as a global and immediate source in general, Twitter interactions are found to be biased towards certain types of scholarly outputs. The biases of Twitter-based metrics are not only limited to the recent biases (i.e., biases towards recently published papers) and disciplinary biases (i.e., biases towards social sciences and humanities, biomedical and health sciences, and life and earth sciences) reported in this dissertation, but also include, for example, geographic biases towards scientific papers authored by the US and the UK (X. Wang, Fang, Li, et al., 2016; Zahedi & Costas, 2017). Therefore, it is necessary to be aware of these biases particularly when drawing general conclusions based on scholarly Twitter data, which may lead to underrepresentation in some subject fields (e.g., natural sciences, mathematics, and engineering), publication years (e.g., papers published for a long time), and countries or regions (e.g., China and Latin American countries) (Alperin, 2015; X. Wang, Fang, Li, et al., 2016).

7.3.2 Approaching a more fine-grained indicator system of Twitter-based metrics

The comparisons of the uptake broadness and velocity of scientific papers reveal fundamental differences across altmetric data sources, adding more evidence to the argument that "altmetrics are indeed representing very different things" (Lin & Fenner, 2013a). Moreover, as different altmetrics inherently describe different aspects of the broad spectrum of interactions around scholarly outputs, they cannot "be expressed in a single number" (Lin & Fenner, 2013b). Therefore, keeping altmetrics separate and avoiding misappropriating composite indicators has been widely encouraged in both altmetric research and practice (Gumpenberger et al., 2016; Thelwall, 2020; Wouters et al., 2019).

In the current altmetric indicator system which is largely shaped by the data and metrics processed and provided by major altmetric data aggregators like Altmetric.com and PlumX, not only do the different altmetrics as a whole face the challenges posed by the misuse of composite indicators,[3] but Twitter interactions specifically are also at risk of being merged as composite Twitter-based indicators. For example, although retweets are widely considered by main altmetric data aggregators in their incorporated Twitter-based metrics of scientific papers, retweets are not treated as the outcomes of Twitter engagement but are counted as Twitter mentions on an equal footing with their dependent original tweets. Given that retweets per se do not represent original contributions but are essentially re-dissemination of original tweets (Haustein, 2019), simply adding up original tweets and retweets as a composite indicator would mix up two categories of Twitter interactions and thus undermine the robustness of the interpretation. Besides retweets, there are traces left by a range of other Twitter engagement behaviors, such as likes, replies, quotes, and clicks studied in this dissertation. However, they are generally left out by the prevailing indicator system of Twitter-based metrics, in spite of their value in reflecting the broader Twitter attention.

By making a clear-cut distinction amongst different Twitter elements and interactions, this dissertation initiates an effort to develop a hierarchical framework for Twitter-based metrics to take into account different levels of Twitter interactions around scientific papers, as shown in Figure 1. The first layer concerns the measurement of the Twitter uptake of scientific papers. Within the current Twitter ecosystem, there are three tweet types – original tweets, quote tweets, and reply tweets – that are capable of including URLs to scientific papers (i.e., scholarly URLs) in their original tweet content. These *original scholarly tweets* can be considered as a form of Twitter uptake of scientific papers, because they are able to originally bring scientific information from science to Twitter, offering the possibility for audiences to

[3] For example, the Altmetric Attention Score developed by Altmetric.com, a composite indicator which compounds diverse sources of altmetric indicators, has been widely criticized for its lack of transparency, reliability, and reproducibility: https://scholarlykitchen.sspnet.org/2021/08/24/unpacking-the-altmetric-black-box/ (Accessed August 31, 2021).

further engage in. Next, the second layer concerns the measurement of the Twitter engagement with original scholarly tweets. As observed in the main findings of this dissertation, after being posted on Twitter, original scholarly tweets might be engaged with by other Twitter users through liking, retweeting, quoting, replying, or clicking, etc., thus leaving valuable traces to assess the role of original scholarly tweets in increasing the visibility of scientific papers in the Twitter universe, or even in directly increasing the visits to scientific papers (in the case of clicks). Based on this framework, it is possible to approach a more fine-grained indicator system of Twitter-based metrics for scientific papers, in which diverse forms of Twitter interactions are valued and kept separate at the paper level (i.e., metrics regarding Twitter uptake of scientific papers) and the tweet level (i.e., Twitter engagement metrics regarding original scholarly tweets), respectively.

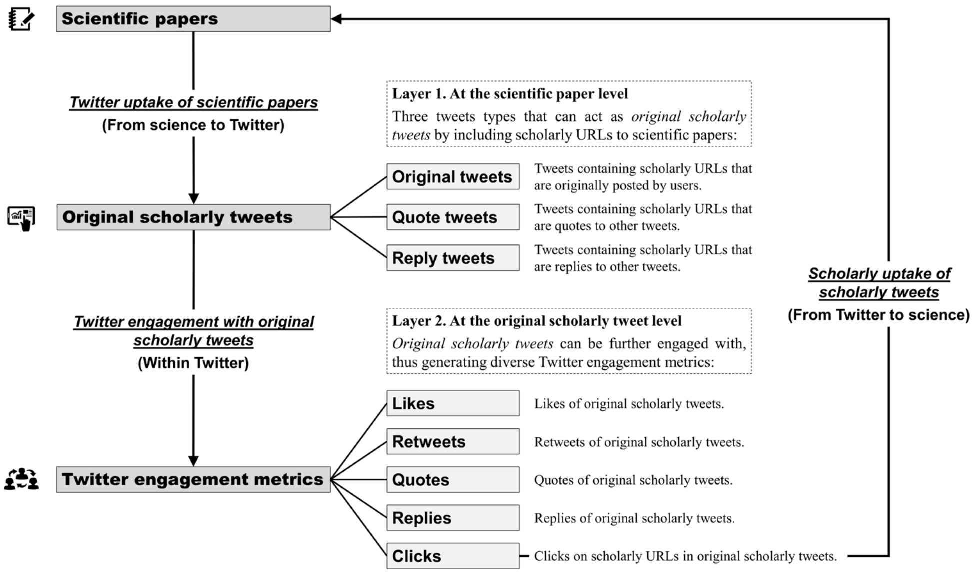

Figure 1. A hierarchical framework of Twitter elements and interactions involved in Twitter-based metrics

7.3.3 Measuring Twitter reception of scientific information in greater depth

On the basis of the above more fine-grained indicator system of Twitter-based metrics, it is also possible to measure Twitter reception of scientific information in greater depth. In the scope of primary Twitter metrics, number of Twitter mentions, as the indicator reflecting the frequency of the Twitter uptake of scientific papers, has long been used to measure the level of Twitter attention paid to scholarly outputs (Sugimoto, Work, et al., 2017). However, conceptually speaking, number of Twitter mentions only reflects the attention of Twitter users who brought scientific papers to Twitter, but neglects the attention of Twitter audiences who engaged with the tweeted scientific papers through diverse interaction behaviors on Twitter. As argued by Brossard (2013), in the online environment, scientific information is no longer consumed in isolated fashion but is now contextualized by, for instance, readers' comments, Facebook posts, and likes or short commentaries in tweets. On Twitter, the subsequent interactions around original scholarly tweets narrate the stories of how the tweeted scientific information is further discussed, disseminated (e.g., by liking, retweeting, quoting, or replying) in the Twitter environment, or even used as a portal to access detailed scholarly content (i.e., by clicking). Therefore, these Twitter interactions around original scholarly tweets represent a more deep-seated reception of scientific information by broader audiences.

From an empirical point of view, research findings presented in this dissertation reveal that original scholarly tweets trigger deeper levels of Twitter interactions to varying degrees, confirming the different effectiveness of original scholarly tweets in facilitating the dissemination and reception of scientific information. Thus, paying more attention to secondary Twitter metrics (Díaz-Faes et al., 2019), and specifically including Twitter engagement metrics around original scholarly tweets in the altmetric toolkit, would be an important step towards more in-depth measurements of Twitter reception of scientific information. Furthermore, as a special type of Twitter engagement behavior, clicking on the scholarly URLs embedded in original scholarly tweets leaves concrete evidence of the traffic drawn forth to the original webpages of scientific papers (X. Wang, Fang, & Guo, 2016). In this sense, Twitter engagement metrics also open up the possibility to evaluate the feedback effect of Twitter uptake on scientific papers.

7.3.4 Scrutinizing the effect of the volatility of Twitter data on the stability of Twitter-based metrics

Data quality issue has long been regarded as one of the grand challenges facing altmetrics (Haustein, 2016; Priem & Hemminger, 2010). Related research has so far mainly focused on the inconsistencies across altmetric data aggregators (Meschede & Siebenlist, 2018; Ortega, 2018a; Zahedi & Costas, 2018). According to the volatile nature of Twitter data observed in this dissertation, the challenge of data quality is rooted in the metadata as well. The unavailability of tweets, which is mainly caused by tweet deletion, user account protection or suspension, is a noticeable threat to the stability of Twitter-based metrics, particularly in consideration of the cascade of related Twitter engagement data becoming unavailable when some highly engaged original scholarly tweets fall into the unavailability state (Bastos, 2021).

Ephemerality is perhaps an expected characteristic of Twitter data, but it is not an expected design of stable and robust metrics. The volatility of Twitter data puts more emphasis on the necessity of adopting a dynamic perspective to deal with Twitter data, especially when working with snapshot files containing Twitter data collected at a certain time point. Tweets recorded in a snapshot file may have changed between the time of the creation of the snapshot and the time of the update or rehydration of the tweets, resulting in the potential inconsistency between the observations based on the tweets available at the two different time points.

In the currently prevalent indicator system where retweets are also counted as Twitter mentions to scientific papers equal to original scholarly tweets, the measurement of Twitter uptake of scientific papers risks more drastic fluctuations because retweets maintain the same (un)availability as their original scholarly tweets. Additionally, it has been found that almost half of the Twitter mentions recorded by Altmetric.com were actually retweets (Didegah et al., 2018; Haustein, 2019), potentially exacerbating the instability of the measurement of Twitter uptake particularly for those scientific papers with extensively retweeted original scholarly tweets received. Therefore, as to the measurement of Twitter uptake of scientific papers, the effect of the volatility of Twitter data would be mitigated if only original scholarly tweets are counted, as set forth earlier in the proposed hierarchical framework.

7.4 Future research prospects

Building on the main findings, it is clear that there is a promising research agenda regarding social media metrics. To further comprehend the nature and influencing factors of Twitter interactions around science, contextualizing Twitter interactions based on detailed tweet content and user information will be an important step. The generally negligible correlations found between scholarly impact indicators and Twitter-based metrics indicate that science and Twitter conform to fundamentally different spaces, and the agents active in the two spaces may follow different norms. As a result, it gains more importance to interpret Twitter interactions in the context of who interacts with science (i.e., Twitter users) and what is expressed in the interactions (i.e., tweet content) (Haustein, 2019). This section explicates five main future research prospects: the first one is about depicting Twitter interactions based on user information; the second one is about interpreting Twitter interactions in the context of tweet content; the third one is about understanding the nature of Twitter attention reflected by Twitter interactions; the fourth one is about unravelling the patterns and effectiveness of science communication on Twitter; and the last looks into the possibility of generalizing the conceptual framework of science-social media interactions to broader social media studies of science.

7.4.1 Depicting Twitter interactions based on user information

Painting a portrait of Twitter users who participate in Twitter interactions around science is relevant, because it offers insights into the type of Twitter communication observed, such as communication between scholars or between scholars and the general public. However, it is challenging to determine the types of Twitter users at scale as users present themselves in very different aspects and languages. For Twitter users who posted original scholarly tweets or related retweets, previous sampling research has found an almost fifty-fifty split between users with academic backgrounds (e.g., users possessing or pursing a PhD) and public users (Mohammadi et al., 2018; Tsou et al., 2015; Yu et al., 2019), based on which Thelwall (2020) argued that scholarly tweets might reflect “half academic, half nonacademic attention or impact”. Placing the identification of Twitter users in the hierarchical framework proposed in Figure 1, future research on the delineation of Twitter users can be conducted in two parts: first, who posted original scholarly tweets to achieve the Twitter uptake of scientific papers, and then, who engaged with these original scholarly tweets to facilitate the dissemination of scientific papers (e.g., replying users, retweeting users, or quoting users). As such, different types of Twitter users can be networked through their engagement relationships, helping to track and define Twitter communication of scientific information amongst Twitter users from a more interactive perspective (Robinson-Garcia et al., 2018).

7.4.2 Interpreting Twitter interactions in the context of tweet content

In addition to Twitter users, tweet content also plays an important role in interpreting the nature of Twitter interactions. Detailed tweet content serves as a mirror to reflect the motivations of users’ tweeting behaviors. For instance, based on the examination of tweet content, much Twitter uptake of scientific papers has been identified as the consequence of mechanical tweeting (Robinson-Garcia et al., 2017; Thelwall, Tsou, et al., 2013) or humor and hoax (Didegah et al., 2018; Sugimoto, 2015). In contrast, little is known about what motivates Twitter users to further engage with original scholarly tweets. As surveyed by Mohammadi et al. (2018), the majority of Twitter users like scholarly tweets in order to “inform the authors that their tweets were interesting”, while most retweet scholarly tweets to disseminate them. For other more informative Twitter engagement, such as quotes and replies, although they appear in only a limited share of original scholarly tweets, they are more meaningful due to the additional information added beyond what is delivered by scientific papers and original scholarly tweets. Content analysis of replies and quotes may provide valuable evidence to illuminate the mechanisms of how scientific papers appearing on Twitter are consumed in a more conversational and informative manner.

7.4.3 Understanding the nature of Twitter attention reflected by Twitter interactions

On the basis of the investigation of Twitter users and tweet content, it is also possible to gain a deeper understanding of the nature of Twitter attention as reflected by different Twitter interactions. Although Twitter interactions have been considered to be one of the potential proxies for the societal impact generated by scholarly objects (Bornmann, 2014b; Noyons, 2019), concrete evidence on how Twitter interactions around science fits into the framework of research impact evaluation is still lacking, particularly in consideration of the complexity of the notion of impact (Holmberg et al., 2019). For example, in Research Excellence Framework (REF) 2021, impact is broadly defined as the "effect on, change or benefit to the economy, society, culture, public policy or services, health, the environment or quality of life beyond academia".[4] The Economic and Social Research Council (ESRC) of the UK defines economic and societal impact in particular as the "demonstrable contribution that excellent social and economic research has on society and the economy, and its benefits to individuals, organizations or nations".[5] Under these definitions, it is still unclear whether Twitter interactions capture some aspects of societal impact, because it remains unknown if scientific information disseminated on Twitter can substantially change human cognition and behavior, or benefits society at large. Therefore, many researchers argued that attention is a more compelling claim than impact when describing the objects that Twitter metrics mirror (Bornmann et al., 2019; Sugimoto, 2015). Due to the heterogeneity of Twitter users and the motivations behind different Twitter interactions, the nature of Twitter attention attracted by scientific information is far less understood. For future research, ascertaining who is tweeting about science and what is expressed along with the tweeted scientific information will be helpful for better understanding the nature of Twitter attention.

7.4.4 Unravelling the patterns and effectiveness of science communication on Twitter

On the basis of the information regarding Twitter users and tweet content, it is possible to identify science communication that occurs between the scientific community and members of the general public active on Twitter. The scholarly use of Twitter has accelerated to some extent the transition of science communication from the one-way mode with scientists or science communicators transporting scientific information to the public (Davies, 2008; Durant et al., 1989; Ziman, 1991), to the two-way mode in which the scientific community both provides and receives information from the public achieved by open and bidirectional dialogues (Leshner, 2003; Reincke et al., 2020; Schäfer, 2009). By analyzing the interaction behaviors around science of both interacting users and interacted users, future research can delve into the patterns of science communication on Twitter, unravelling how scientific

[4] https://www.ref.ac.uk/media/1426/guide-for-testimonies.pdf (Accessed November 15, 2021).

[5] https://www.ukri.org/councils/esrc/impact-toolkit-for-economic-and-social-sciences/defining-impact/ (Accessed November 15, 2021).

information is communicated via a two-way mode between the scientific community and the public. Furthermore, "communicating science effectively" has been called for as an important agenda for science communication (NASEM, 2017). For science communication taking place in the Twitter environment, Twitter-based metrics can provide evidence to quantitatively evaluate the effectiveness of science communication by measuring how many audiences are reached and are motivated to engage in scientific discussions or access scientific knowledge.

7.4.5 Generalizing the conceptual framework to other social media studies of science

The conceptual framework of science-social media interactions, although specifically applied in the Twitter context in this dissertation, can be generalized to a wider range of social media studies of science in future research. Despite the differences in the specific interactive features available, social media platforms share the same patterns in terms of their interaction relationships with the science environment. For example, Facebook posts mentioning scientific papers represent the Facebook uptake of scientific papers (i.e., scientific information flowing from science to Facebook). Around Facebook posts, there are also several engagement metrics available, such as likes, shares, and comments (i.e., information flowing within Facebook). Similarly, scholarly URLs embedded in Facebook posts provide Facebook users with a portal to access the original webpages of scientific papers (i.e., information flowing from Facebook back to science). Such homogeneity amongst social media sources allows for the generalization of the conceptual framework and research methodologies introduced in this dissertation to broader social media studies of science.

References

Abramo, G., Cicero, T., & D'Angelo, C. A. (2011). Assessing the varying level of impact measurement accuracy as a function of the citation window length. *Journal of Informetrics*, *5*(4), 659–667. https://doi.org/10.1016/j.joi.2011.06.004

Adie, E., & Roe, W. (2013). Altmetric: Enriching scholarly content with article-level discussion and metrics. *Learned Publishing*, *26*(1), 11–17. https://doi.org/10.1087/20130103

Aichner, T., & Jacob, F. (2015). Measuring the degree of corporate social media use. *International Journal of Market Research*, *57*(2), 257–276. https://doi.org/10.2501/IJMR-2015-018

Aksnes, D. W. (2003). Characteristics of highly cited papers. *Research Evaluation*, *12*(3), 159–170. https://doi.org/10.3152/147154403781776645

Allen, H. G., Stanton, T. R., Di Pietro, F., & Moseley, G. L. (2013). Social media release increases dissemination of original articles in the clinical pain sciences. *PLoS ONE*, *8*(7), e68914. https://doi.org/10.1371/journal.pone.0068914

Alperin, J. P. (2015). Geographic variation in social media metrics: An analysis of Latin American journal articles. *Aslib Journal of Information Management*, *67*(3), 289–304. https://doi.org/10.1108/AJIM-12-2014-0176

Alperin, J. P., Gomez, C. J., & Haustein, S. (2019). Identifying diffusion patterns of research articles on Twitter: A case study of online engagement with open access articles. *Public Understanding of Science*, *28*(1), 2–18. https://doi.org/10.1177/0963662518761733

Alstott, J., Bullmore, E., & Plenz, D. (2014). powerlaw: A python package for analysis of heavy-tailed distributions. *PLoS ONE*, *9*(1), e85777. https://doi.org/10.1371/JOURNAL.PONE.0085777

Altmetric. (2020, September 17). *Attention sources coverage dates*. https://help.altmetric.com/support/solutions/articles/6000240455-attention-sources-coverage-dates

Amat, C. B. (2008). Editorial and publication delay of papers submitted to 14 selected Food Research journals. Influence of online posting. *Scientometrics*, *74*(3), 379–389. https://doi.org/10.1007/S11192-007-1823-8

Aral, S., & Walker, D. (2012). Identifying influential and susceptible members of social networks. *Science*, *337*(6092), 337–341. https://doi.org/10.1126/science.1215842

Araujo, R. F. (2020). Communities of attention networks: Introducing qualitative and conversational perspectives for altmetrics. *Scientometrics*, *124*(3), 1793–1809. https://doi.org/10.1007/s11192-020-03566-7

Aversa, E. S. (1985). Citation patterns of highly cited papers and their relationship to literature aging: A study of the working literature. *Scientometrics*, *7*(3), 383–389. https://doi.org/10.1007/BF02017156

Bar-Ilan, J. (2014). JASIST@Mendeley Revisited. *Altmetrics14: Expanding Impacts and Metrics Workshop at the Web Science Conference*. https://doi.org/10.6084/M9.FIGSHARE.1031681.V1

Bar-Ilan, J., Halevi, G., & Milojević, S. (2019). Differences between altmetric data sources – A case study. *Journal of Altmetrics*, *2*(1), 1. https://doi.org/10.29024/JOA.4

Bardus, M., El Rassi, R., Chahrour, M., Akl, E. W., Raslan, A. S., Meho, L. I., & Akl, E. A. (2020). The use of social media to increase the impact of health research: Systematic review. *Journal of Medical Internet Research*, *22*(7), e15607. https://doi.org/10.2196/15607

Bastos, M. (2021). This account doesn't exist: Tweet decay and the politics of deletion in the Brexit debate. *American Behavioral Scientist*, *65*(5), 757–773. https://doi.org/10.1177/0002764221989772

Beaver, D. D., & Rosen, R. (1978). Studies in scientific collaboration. *Scientometrics*, *1*, 65–84. https://doi.org/10.1007/BF02016840

Bentley, P., & Kyvik, S. (2011). Academic staff and public communication: A survey of popular science publishing across 13 countries. *Public Understanding of Science*, *20*(1), 48–63. https://doi.org/10.1177/0963662510384461

Bhattacharya, S., Srinivasan, P., & Polgreen, P. (2014). Engagement with Health Agencies on Twitter. *PLoS ONE*, *9*(11), e112235. https://doi.org/10.1371/journal.pone.0112235

Bhattacharya, S., Srinivasan, P., & Polgreen, P. (2017). Social media engagement analysis of U.S. Federal health agencies on Facebook. *BMC Medical Informatics and Decision Making*, *17*(1), 49. https://doi.org/10.1186/s12911-017-0447-z

Bik, H. M., & Goldstein, M. C. (2013). An introduction to social media for scientists. *PLoS Biology*, *11*(4), e1001535. https://doi.org/10.1371/journal.pbio.1001535

Björk, B. C., & Solomon, D. (2013). The publishing delay in scholarly peer-reviewed journals. *Journal of Informetrics*, *7*(4), 914–923. https://doi.org/10.1016/J.JOI.2013.09.001

Bohn, A., Buchta, C., Hornik, K., & Mair, P. (2014). Making friends and communicating on Facebook: Implications for the access to social capital. *Social Networks*, *37*(1), 29–41. https://doi.org/10.1016/j.socnet.2013.11.003

Bollen, J., & Van De Sompel, H. (2006). Mapping the structure of science through usage. *Scientometrics*, *69*(2), 227–258. https://doi.org/10.1007/s11192-006-0151-8

Bornmann, L. (2014a). Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *Journal of Informetrics*, *8*(4), 895–903. https://doi.org/10.1016/j.joi.2014.09.005

Bornmann, L. (2014b). Validity of altmetrics data for measuring societal impact: A study using data from Altmetric and F1000Prime. *Journal of Informetrics*, *8*(4), 935–950. https://doi.org/10.1016/j.joi.2014.09.007

Bornmann, L. (2015a). Alternative metrics in scientometrics: A meta-analysis of research into three altmetrics. *Scientometrics*, *103*(3), 1123–1144. https://doi.org/10.1007/s11192-015-1565-y

Bornmann, L. (2015b). Usefulness of altmetrics for measuring the broader impact of research: A case study using data from PLOS and F1000Prime. *Aslib Journal of Information Management*, *67*(3), 305–319. https://doi.org/10.1108/AJIM-09-2014-0115

Bornmann, L. (2016). What do altmetrics counts mean? A plea for content analyses. *Journal of the Association for Information Science and Technology*, *67*(4), 1016–1017. https://doi.org/10.1002/asi23633

Bornmann, L., & Haunschild, R. (2016). How to normalize Twitter counts? A first attempt based on journals in the Twitter Index. *Scientometrics*, *107*(3), 1405–1422. https://doi.org/10.1007/s11192-016-1893-6

Bornmann, L., & Haunschild, R. (2017). Measuring field-normalized impact of papers on specific societal groups: An altmetrics study based on Mendeley data. *Research Evaluation*, *26*(3), 230–241. https://doi.org/10.1093/reseval/rvx005

Bornmann, L., Haunschild, R., & Adams, J. (2019). Do altmetrics assess societal impact in a comparable way to case studies? An empirical test of the convergent validity of altmetrics based on data from the UK research excellence framework (REF). *Journal of Informetrics*, *13*(1), 325–340. https://doi.org/10.1016/j.joi.2019.01.008

Bowman, T. D. (2015). Differences in personal and professional tweets of scholars. *Aslib Journal of Information Management*, *67*(3), 356–371. https://doi.org/10.1108/AJIM-12-2014-0180

Brossard, D. (2013). New media landscapes and the science information consumer. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(SUPPL. 3), 14096–14101. https://doi.org/10.1073/pnas.1212744110

Brossard, D., & Scheufele, D. A. (2013). Science, new media, and the public. *Science*, *339*(6115), 40–41. https://doi.org/10.1126/science.1232329

Brzezinski, M. (2015). Power laws in citation distributions: Evidence from Scopus. *Scientometrics*, *103*(1), 213–228. https://doi.org/10.1007/S11192-014-1524-Z

Carr, C. T., & Hayes, R. A. (2015). Social media: Defining, developing, and divining. *Atlantic Journal of Communication*, *23*(1), 46–65. https://doi.org/10.1080/15456870.2015.972282

Chamberlain, S. (2013). Consuming article-level metrics: Observations and lessons. *Information Standards Quarterly*, *25*(2), 4–13. https://doi.org/10.3789/isqv25no2.2013.02

Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature.

R

Journal of the American Society for Information Science and Technology, *57*(3), 359–377. https://doi.org/10.1002/ASI.20317

Cheplygina, V., Hermans, F., Albers, C., Bielczyk, N., & Smeets, I. (2020). Ten simple rules for getting started on Twitter as a scientist. *PLoS Computational Biology*, *16*(2), e1007513. https://doi.org/10.1371/journal.pcbi.1007513

Chretien, K. C., Azar, J., & Kind, T. (2011). Physicians on Twitter. *JAMA*, *305*(6), 566–568. https://doi.org/10.1001/jama.2011.68

Colavizza, G. (2020). COVID-19 research in Wikipedia. *Quantitative Science Studies*, *1*(4), 1349–1380. https://doi.org/10.1162/QSS_A_00080

Colavizza, G., Costas, R., Traag, V. A., Eck, N. J. van, Leeuwen, T. van, & Waltman, L. (2021). A scientometric overview of CORD-19. *PLoS ONE*, *16*(1), e0244839. https://doi.org/10.1371/JOURNAL.PONE.0244839

Colavizza, G., & Franceschet, M. (2016). Clustering citation histories in the Physical Review. *Journal of Informetrics*, *10*(4), 1037–1051. https://doi.org/10.1016/J.JOI.2016.07.009

Cole, S., & Phelan, T. J. (1999). The scientific productivity of nations. *Minerva*, *37*(1), 1–23. https://about.jstor.org/terms

Collins, K., Shiffman, D., & Rock, J. (2016). How are scientists using social media in the workplace? *PLoS ONE*, *11*(10), e0162680. https://doi.org/10.1371/journal.pone.0162680

Costas, R. (2017). Towards the social media studies of science: Social media metrics, present and future. *Bibliotecas. Anales de Investigación*, *13*(1), 1–5. https://doi.org/20.500.11959/brapci/57683

Costas, R., Mongeon, P., Ferreira, M. R., Van Honk, J., & Franssen, T. (2020). Large-scale identification and characterization of scholars on Twitter. *Quantitative Science Studies*, *1*(2), 771–791. https://doi.org/10.1162/qss_a_00047

Costas, R., Rijcke, S., & Marres, N. (2021). "Heterogeneous couplings": Operationalizing network perspectives to study science-society interactions through social media metrics. *Journal of the Association for Information Science and Technology*, *72*(5), 595–610. https://doi.org/10.1002/asi.24427

Costas, R., Van Leeuwen, T. N., & Van Raan, A. F. J. (2010). Is scientific literature subject to a 'sell-by-date'? A general methodology to analyze the 'durability' of scientific documents. *Journal of the American Society for Information Science and Technology*, *61*(2), 329–339. https://doi.org/10.1002/asi.21244

Costas, R., Zahedi, Z., & Wouters, P. (2015a). Do "altmetrics" correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective. *Journal of the Association for Information Science and Technology*, *66*(10), 2003–2019. https://doi.org/10.1002/asi.23309

Costas, R., Zahedi, Z., & Wouters, P. (2015b). The thematic orientation of publications mentioned on social media: Large-scale disciplinary comparison of social media. *Aslib Journal of Information Management*, *67*(3), 260–288.

Crotty, D. (2014). Altmetrics: Finding meaningful needles in the data haystack. *Serials Review*, *40*(3), 141–146. https://doi.org/10.1080/00987913.2014.947839

D'Angelo, C. A., & Di Russo, S. (2019). Testing for universality of Mendeley readership distributions. *Journal of Informetrics*, *13*(2), 726–737. https://doi.org/10.1016/j.joi.2019.03.011

Darling, E., Shiffman, D., Côté, I., & Drew, J. (2013). The role of Twitter in the life cycle of a scientific publication. *Ideas in Ecology and Evolution*, *6*, 32–43. https://doi.org/10.4033/iee.2013.6.6.f

Davies, S. R. (2008). Constructing communication: Talking to scientists about talking to the public. *Science Communication*, *29*(4), 413–434. https://doi.org/10.1177/1075547008316222

de Winter, J. C. F. (2015). The relationship between tweets, citations, and article views for PLOS ONE articles. *Scientometrics*, *102*(2), 1773–1779. https://doi.org/10.1007/s11192-014-1445-x

Díaz-Faes, A. A., Bowman, T. D., & Costas, R. (2019). Towards a second generation of 'social media metrics': Characterizing Twitter communities of attention around science. *PLoS ONE*, *14*(5), e0216408. https://doi.org/10.1371/journal.pone.0216408

Didegah, F., Mejlgaard, N., & Sørensen, M. P. (2018). Investigating the quality of interactions and public

R

engagement around scientific papers on Twitter. *Journal of Informetrics*, *12*(3), 960–971. https://doi.org/10.1016/j.joi.2018.08.002

Didegah, F., & Thelwall, M. (2018). Co-saved, co-tweeted, and co-cited networks. *Journal of the Association for Information Science and Technology*, *69*(8), 959–973. https://doi.org/10.1002/asi.24028

Ding, W., & Chen, C. (2014). Dynamic topic detection and tracking: A comparison of HDP, C-word, and cocitation methods. *Journal of the Association for Information Science and Technology*, *65*(10), 2084–2097. https://doi.org/10.1002/asi.23134

Durant, J. R., Evans, G. A., & Thomas, G. P. (1989). The public understanding of science. *Nature*, *340*(6228), 11–14. https://doi.org/10.1038/340011a0

Erdt, M., Nagarajan, A., Sin, S. C. J., & Theng, Y. L. (2016). Altmetrics: An analysis of the state-of-the-art in measuring research impact on social media. *Scientometrics*, *109*(2), 1117–1166. https://doi.org/10.1007/s11192-016-2077-0

Eysenbach, G. (2011). Can tweets predict citations? Metrics of social impact based on Twitter and correlation with traditional metrics of scientific impact. *Journal of Medical Internet Research*, *13*(4), e2012. https://doi.org/10.2196/jmir.2012

Fang, Z., & Costas, R. (2018). Studying the posts accumulation patterns of Altmetric.com data sources. *The 2018 Altmetrics Workshop (Altmetrics18), London, UK*. http://altmetrics.org/wp-content/uploads/2018/04/altmetrics18_paper_5_Fang.pdf

Fang, Z., & Costas, R. (2020). Studying the accumulation velocity of altmetric data tracked by Altmetric.com. *Scientometrics*, *123*(2), 1077–1101. https://doi.org/10.1007/s11192-020-03405-9

Fang, Z., Costas, R., Tian, W., Wang, X., & Wouters, P. (2020). An extensive analysis of the presence of altmetric data for Web of Science publications across subject fields and research topics. *Scientometrics*, *124*(3), 2519–2549. https://doi.org/10.1007/s11192-020-03564-9

Fang, Z., Costas, R., Tian, W., Wang, X., & Wouters, P. (2021). How is science clicked on Twitter? Click metrics for Bitly short links to scientific publications. *Journal of the Association for Information Science and Technology*, *72*(7), 918–932. https://doi.org/10.1002/asi.24458

Fang, Z., Dudek, J., & Costas, R. (2020). The stability of Twitter metrics: A study on unavailable Twitter mentions of scientific publications. *Journal of the Association for Information Science and Technology*, *71*(12), 1455–1469. https://doi.org/10.1002/asi.24344

Fenner, M. (2013). What can article-level metrics do for you? *PLoS Biology*, *11*(10), e1001687. https://doi.org/10.1371/journal.pbio.1001687

Friedrich, N., Bowman, T. D., Stock, W. G., & Haustein, S. (2015). Adapting sentiment analysis for tweets linking to scientific papers. *Proceedings of ISSI 2015 Istanbul: 15th International Society of Scientometrics and Informetrics Conference*, 107–108. https://www.issi-society.org/proceedings/issi_2015/0107.pdf

Fu, K., & Chau, M. (2013). Reality check for the Chinese microblog space: A random sampling approach. *PLoS ONE*, *8*(3), e58356. https://doi.org/10.1371/journal.pone.0058356

Gabielkov, M., Ramachandran, A., Chaintreau, A., & Legout, A. (2016). Social clicks: What and who gets read on twitter? *Proceedings of the 2016 ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Science*, 179–192. https://doi.org/10.1145/2896377.2901462

Gan, C., & Wang, W. (2015). Research characteristics and status on social media in China: A bibliometric and co-word analysis. *Scientometrics*, *105*(2), 1167–1182. https://doi.org/10.1007/S11192-015-1723-2

Garfield, E. (1972). Citation analysis as a tool in journal evaluation. *Science*, *178*(4060), 471–479. https://doi.org/10.1126/science.178.4060.471

Garfield, E. (1980). Premature discovery or delayed recognition-Why? *Current Contents*, *21*, 5–10. http://garfield.library.upenn.edu/essays/v4p488y1979-80.pdf

Glänzel, W. (2004). Towards a model for diachronous and synchronous citation analyses. *Scientometrics*, *60*(3), 511–522. https://doi.org/10.1023/B:SCIE.0000034391.06240.2A

Glänzel, W., & Czerwon, H. J. (1996). A new methodological approach to bibliographic coupling and its application to the national, regional and institutional level. *Scientometrics*, *37*(2), 195–221.

R

https://doi.org/10.1007/BF02093621

Glänzel, W., & Schoepflin, U. (1995). A bibliometric study on ageing and reception processes of scientific literature. *Journal of Information Science*, *21*(1), 37–53. https://doi.org/10.1177/016555159502100104

Glänzel, W., & Thijs, B. (2012). Using 'core documents' for detecting and labelling new emerging topics. *Scientometrics*, *91*(2), 399–416. https://doi.org/10.1007/S11192-011-0591-7

Gumpenberger, C., Glänzel, W., & Gorraiz, J. (2016). The ecstasy and the agony of the altmetric score. *Scientometrics*, *108*(2), 977–982. https://doi.org/10.1007/S11192-016-1991-5

Gunn, A., & Mintrom, M. (2016). Higher education policy change in Europe: Academic research funding and the impact agenda. *European Education*, *48*(4), 241–257. https://doi.org/10.1080/10564934.2016.1237703

Hammarfelt, B. (2014). Using altmetrics for assessing research impact in the humanities. *Scientometrics*, *101*(2), 1419–1430. https://doi.org/10.1007/s11192-014-1261-3

Hassan, S. U., Aljohani, N. R., Idrees, N., Sarwar, R., Nawaz, R., Martínez-Cámara, E., Ventura, S., & Herrera, F. (2020). Predicting literature's early impact with sentiment analysis in Twitter. *Knowledge-Based Systems*, *192*, 105383. https://doi.org/10.1016/J.KNOSYS.2019.105383

Hassan, S. U., Aljohani, N. R., Shabbir, M., Ali, U., Iqbal, S., Sarwar, R., Martínez-Cámara, E., Ventura, S., & Herrera, F. (2020). Tweet Coupling: A social media methodology for clustering scientific publications. *Scientometrics*, *124*(2), 973–991. https://doi.org/10.1007/s11192-020-03499-1

Hassan, S. U., Bowman, T. D., Shabbir, M., Akhtar, A., Imran, M., & Aljohani, N. R. (2019). Influential tweeters in relation to highly cited articles in altmetric big data. *Scientometrics*, *119*(1), 481–493. https://doi.org/10.1007/s11192-019-03044-9

Hassan, S. U., Saleem, A., Soroya, S. H., Safder, I., Iqbal, S., Jamil, S., Bukhari, F., Aljohani, N. R., & Nawaz, R. (2020). Sentiment analysis of tweets through Altmetrics: A machine learning approach. *Journal of Information Science*. https://doi.org/10.1177/0165551520930917

Haunschild, R., & Bornmann, L. (2017). How many scientific papers are mentioned in policy-related documents? An empirical investigation using Web of Science and Altmetric data. *Scientometrics*, *110*(3), 1209–1216. https://doi.org/10.1007/s11192-016-2237-2

Haunschild, R., & Bornmann, L. (2021). Can tweets be used to detect problems early with scientific papers? A case study of three retracted COVID-19/SARS-CoV-2 papers. *Scientometrics*, *126*(6), 5181–5199. https://doi.org/10.1007/S11192-021-03962-7

Haunschild, R., Leydesdorff, L., Bornmann, L., Hellsten, I., & Marx, W. (2019). Does the public discuss other topics on climate change than researchers? A comparison of explorative networks based on author keywords and hashtags. *Journal of Informetrics*, *13*(2), 695–707. https://doi.org/10.1016/j.joi.2019.03.008

Haustein, S. (2016). Grand challenges in altmetrics: Heterogeneity, data quality and dependencies. *Scientometrics*, *108*(1), 413–423. https://doi.org/10.1007/s11192-016-1910-9

Haustein, S. (2019). Scholarly Twitter metrics. In W. Glänzel, F. H. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer Handbook of Science and Technology Indicators* (pp. 729–760). Heidelberg: Springer. https://doi.org/10.1007/978-3-030-02511-3_28

Haustein, S., Bowman, T. D., & Costas, R. (2016). Interpreting "altmetrics": Viewing acts on social media through the lens of citation and social theories. In C. R. Sugimoto (Ed.), *Theories of Informetrics: A Festschrift in Honor of Blaise Cronin* (pp. 372–405). Berlin : De Gruyter. https://arxiv.org/abs/1502.05701

Haustein, S., Bowman, T. D., & Costas, R. (2015). When is an article actually published? An analysis of online availability, publication, and indexation dates. *Proceedings of the 15th International Society of Scientometrics and Informetrics Conference*, 1170–1179. https://arxiv.org/abs/1505.00796v1

Haustein, S., Bowman, T. D., Holmberg, K., Peters, I., & Larivière, V. (2014). Astrophysicists on Twitter: An in-depth analysis of tweeting and scientific publication behavior. *Aslib Journal of Information Management*, *66*(3), 279–296. https://doi.org/10.1108/AJIM-09-2013-0081

Haustein, S., Bowman, T. D., Holmberg, K., Tsou, A., Sugimoto, C. R., & Larivière, V. (2016). Tweets as impact indicators: Examining the implications of automated "bot" accounts on Twitter. *Journal of the Association for Information Science and Technology*, *67*(1), 232–238. https://doi.org/10.1002/asi.23456

R

Haustein, S., Costas, R., & Larivière, V. (2015). Characterizing social media metrics of scholarly papers: The effect of document properties and collaboration patterns. *PLoS ONE*, *10*(3), e0120495. https://doi.org/10.1371/journal.pone.0120495

Haustein, S., Larivière, V., Thelwall, M., Amyot, D., & Peters, I. (2014). Tweets vs. Mendeley readers: How do these two social media metrics differ? *It - Information Technology*, *56*(5), 207–215. https://doi.org/10.1515/ITIT-2014-1048

Haustein, S., Peters, I., Bar-Ilan, J., Priem, J., Shema, H., & Terliesner, J. (2014). Coverage and adoption of altmetrics sources in the bibliometric community. *Scientometrics*, *101*(2), 1145–1163. https://doi.org/10.1007/s11192-013-1221-3

Haustein, S., Peters, I., Sugimoto, C. R., Thelwall, M., & Larivière, V. (2014). Tweeting biomedicine: An analysis of tweets and citations in the biomedical literature. *Journal of the Association for Information Science and Technology*, *65*(4), 656–669. https://doi.org/10.1002/asi.23101

Hellsten, I., & Leydesdorff, L. (2020). Automated analysis of actor–topic networks on Twitter: New approaches to the analysis of socio-semantic networks. *Journal of the Association for Information Science and Technology*, *71*(1), 3–15. https://doi.org/10.1002/asi.24207

Hern, A. (2020). *Twitter aims to limit people sharing articles they have not read.* https://www.theguardian.com/technology/2020/jun/11/twitter-aims-to-limit-people-sharing-articles-they-have-not-read

Hicks, D., & Wang, J. (2013). The New York Times as a resource for mode 2. *Science, Technology, & Human Values*, *38*(6), 851–877. https://doi.org/10.1177/0162243913497806

Hoang, J. K., McCall, J., Dixon, A. F., Fitzgerald, R. T., & Gaillard, F. (2015). Using social media to share your radiology research: How effective is a blog post? *Journal of the American College of Radiology*, *12*(7), 760–765. https://doi.org/10.1016/j.jacr.2015.03.048

Holmberg, K., Bowman, S., Bowman T., Didegah, F., & Kortelainen, T. (2019). What is societal impact and where do altmetrics fit into the equation? *Journal of Altmetrics*, *2*(1): 6. DOI: https://doi.org/10.29024/joa.21

Holmberg, K., Bowman, T. D., Haustein, S., & Peters, I. (2014). Astrophysicists' conversational connections on Twitter. *PLoS ONE*, *9*(8), e106086. https://doi.org/10.1371/journal.pone.0106086

Holmberg, K., & Thelwall, M. (2014). Disciplinary differences in Twitter scholarly communication. *Scientometrics*, *101*(2), 1027–1042. https://doi.org/10.1007/s11192-014-1229-3

Holmström, J., Jonsson, D., Polbratt, F., Nilsson, O., Lundström, L., Ragnarsson, S., Forsberg, A., Andersson, K., & Carlsson, N. (2019). Do we read what we share? Analyzing the click dynamic of news articles shared on twitter. *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2019*, 420–425. https://doi.org/10.1145/3341161.3342933

Howell, E. L., Nepper, J., Brossard, D., Xenos, M. A., & Scheufele, D. A. (2019). Engagement present and future: Graduate student and faculty perceptions of social media and the role of the public in science engagement. *PLoS ONE*, *14*(5), e0216274. https://doi.org/10.1371/journal.pone.0216274

Huang, W., Wang, P., & Wu, Q. (2018). A correlation comparison between Altmetric Attention Scores and citations for six PLOS journals. *PLoS ONE*, *13*(4), e0194962. https://doi.org/10.1371/journal.pone.0194962

Huber, B., Barnidge, M., Gil de Zúñiga, H., & Liu, J. (2019). Fostering public trust in science: The role of social media. *Public Understanding of Science*, *28*(7), 759–777. https://doi.org/10.1177/0963662519869097

Jabaley, C. S., Groff, R. F., Stentz, M. J., Moll, V., Lynde, G. C., Blum, J. M., & O'Reilly-Shah, V. N. (2018). Highly visible sepsis publications from 2012 to 2017: Analysis and comparison of altmetrics and bibliometrics. *Journal of Critical Care*, *48*, 357–371. https://doi.org/10.1016/j.jcrc.2018.09.033

Jung, H., Lee, K., & Song, M. (2016). Examining characteristics of traditional and Twitter citation. *Frontiers in Research Metrics and Analytics*, *1*, 6. https://doi.org/10.3389/frma.2016.00006

Kahle, K., Sharon, A. J., & Baram-Tsabari, A. (2016). Footprints of fascination: Digital traces of public engagement with particle physics on CERN's social media platforms. *PLoS ONE*, *11*(5), e0156409. https://doi.org/10.1371/journal.pone.0156409

Kalia, V., Ortiz, D. A., Patel, A. K., Moriarity, A. K., Canon, C. L., & Duszak, R. (2018). Leveraging Twitter to maximize the radiology meeting experience. *Journal of the American College of Radiology*, *15*(1), 177–183. https://doi.org/10.1016/J.JACR.2017.10.022

Kamel Boulos, M. N., & Anderson, P. F. (2014). Preliminary survey of leading general medicine journals' use of Facebook and Twitter. *Journal of the Canadian Health Libraries Association*, *33*(2), 38–47. https://doi.org/10.5596/c2012-010

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of Social Media. *Business Horizons*, *53*(1), 59–68. https://doi.org/10.1016/j.bushor.2009.09.003

Ke, Q., Ahn, Y. Y., & Sugimoto, C. R. (2017). A systematic identification and analysis of scientists on Twitter. *PLoS ONE*, *12*(4). https://doi.org/10.1371/journal.pone.0175368

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons*, *54*(3), 241–251. https://doi.org/10.1016/j.bushor.2011.01.005

Kolahi, J., & Khazaei, S. (2018). Altmetric analysis of contemporary dental literature. *British Dental Journal*, *225*(1), 68–72. https://doi.org/10.1038/sj.bdj.2018.521

Kouper, I. (2010). Science blogs and public engagement with science: Practices, challenges, and opportunities. *Journal of Science Communication*, *9*(1), 1–10. https://doi.org/10.22323/2.09010202

Kreimer, P., Levin, L., & Jensen, P. (2011). Popularization by Argentine researchers: The activities and motivations of CONICET scientists. *Public Understanding of Science*, *20*(1), 37–47. https://doi.org/10.1177/0963662510383924

Kurtz, M. J., Eichhorn, G., Accomazzi, A., Grant, C., Demleitner, M., Murray, S. S., Martimbeau, N., & Elwell, B. (2005). The bibliometric properties of article readership information. *Journal of the American Society for Information Science and Technology*, *56*(2), 111–128. https://doi.org/10.1002/ASI.20096

Larsen, P. O., & Von Ins, M. (2010). The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index. *Scientometrics*, *84*(3), 575–603. https://doi.org/10.1007/S11192-010-0202-Z

Late, E., Tenopir, C., Talja, S., & Christian, L. (2019). Reading practices in scholarly work: From articles and books to blogs. *Journal of Documentation*, *75*(3), 478–499. https://doi.org/10.1108/JD-11-2018-0178

Lee, S., & Bozeman, B. (2005). The impact of research collaboration on scientific productivity. *Social Studies of Science*, *35*(5), 673–702. https://doi.org/10.1177/0306312705052359

Lee, W. H. (2008). How to identify emerging research fields using scientometrics: An example in the field of Information Security. *Scientometrics*, *76*(3), 503–525. https://doi.org/10.1007/S11192-007-1898-2

Leshner, A. I. (2003). Public engagement with science. *Science*, *299*(5609), 977. https://doi.org/10.1126/science.299.5609.977

Li, X., Thelwall, M., & Giustini, D. (2012). Validating online reference managers for scholarly impact measurement. *Scientometrics*, *91*(2), 461–471. https://doi.org/10.1007/S11192-011-0580-X

Lin, J., & Fenner, M. (2013a). Altmetrics in evolution: Defining and redefining the ontology of article-level metrics. *Information Standards Quarterly*, *25*(2), 20–26. https://doi.org/10.3789/isqv25no2.2013.04

Lin, J., & Fenner, M. (2013b). The many faces of article-level metrics. *Bulletin of the American Society for Information Science and Technology*, *39*(4), 27–30. https://doi.org/10.1002/BULT.2013.1720390409

Line, M. B. (1993). Changes in the use of literature with time - Obsolescence revisited. *Library Trends*, *41*(4), 665–683. https://www.ideals.illinois.edu/bitstream/handle/2142/7869/librarytrendsv41i4i_opt.pdf?sequence=1

Lotka, A. J. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, *16*(12), 317–323. https://www.jstor.org/stable/pdf/24529203.pdf

Lulic, I., & Kovic, I. (2013). Analysis of emergency physicians' Twitter accounts. *Emergency Medicine Journal*, *30*(5), 371–376. https://doi.org/10.1136/emermed-2012-201132

Lyu, X., & Costas, R. (2020). How do academic topics shift across altmetric sources? A case study of the research area of Big Data. *Scientometrics*, *123*(2), 909–943. https://doi.org/10.1007/s11192-020-03415-7

Maflahi, N., & Thelwall, M. (2018). How quickly do publications get read? The evolution of mendeley reader counts for new articles. *Journal of the Association for Information Science and Technology*, *69*(1), 158–167. https://doi.org/10.1002/asi.23909

Maggio, L. A., Meyer, H. S., & Artino, A. R. (2017). Beyond citation rates: A real-time impact analysis of health professions education research using altmetrics. *Academic Medicine*, *92*(10), 1449–1455. https://doi.org/10.1097/ACM.0000000000001897

Marx, W., & Bornmann, L. (2015). On the causes of subject-specific citation rates in Web of Science. *Scientometrics*, *102*(2), 1823–1827. https://doi.org/10.1007/s11192-014-1499-9

Meschede, C., & Siebenlist, T. (2018). Cross-metric compatability and inconsistencies of altmetrics. *Scientometrics*, *115*(1), 283–297. https://doi.org/10.1007/s11192-018-2674-1

Min, C., Sun, J., Pei, L., & Ding, Y. (2016). Measuring delayed recognition for papers: Uneven weighted summation and total citations. *Journal of Informetrics*, *10*(4), 1153–1165. https://doi.org/10.1016/J.JOI.2016.10.001

Moed, H. F. (2005). Statistical relationships between downloads and citations at the level of individual documents within a single journal. *Journal of the American Society for Information Science and Technology*, *56*(10), 1088–1097. https://doi.org/10.1002/ASI.20200

Moed, H. F., Aisati, M., & Plume, A. (2013). Studying scientific migration in Scopus. *Scientometrics*, *94*(3), 929–942. https://doi.org/10.1007/s11192-012-0783-9

Mohammadi, E., & Thelwall, M. (2014). Mendeley readership altmetrics for the Social Sciences and Humanities: Research evaluation and knowledge flows. *Journal of the Association for Information Science and Technology*, *65*(8), 1627–1638. https://doi.org/10.1002/asi.23071

Mohammadi, E., Thelwall, M., Haustein, S., & Larivière, V. (2015). Who reads research articles? An altmetrics analysis of Mendeley user categories. *Journal of the Association for Information Science and Technology*, *66*(9), 1832–1846. https://doi.org/10.1002/asi.23286

Mohammadi, E., Thelwall, M., Kwasny, M., & Holmes, K. L. (2018). Academic information on Twitter: A user survey. *PLoS ONE*, *13*(5), e0197265. https://doi.org/10.1371/journal.pone.0197265

Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics*, *33*(3), 341–365. https://doi.org/10.1016/0304-4076(86)90002-3

Mund, C., & Neuhäusler, P. (2015). Towards an early-stage identification of emerging topics in science—The usability of bibliometric characteristics. *Journal of Informetrics*, *9*(4), 1018–1033. https://doi.org/10.1016/J.JOI.2015.09.004

NASEM. (2017). *Communicating science effectively: A research agenda*. National Academies Press. https://doi.org/10.17226/23674

National Science Board. (2020). *Science and technology: Public attitudes, knowledge, and Interest*. https://ncses.nsf.gov/pubs/nsb20207/interest-information-sources-and-involvement#involvement-in-s-t-activities

Noruzi, A. (2009). Wikipedia popularity from a citation analysis point of view. *Webology*, *6*(2), 1–3. www.webology.org/2009/v6n2/editorial20.html

Noyons, E. (2019). Measuring societal impact is as complex as ABC. *Journal of Data and Information Science*, *4*(3), 6–21. https://doi.org/10.2478/jdis-2019-0012

Nugroho, R., Paris, C., Nepal, S., Yang, J., & Zhao, W. (2020). A survey of recent methods on deriving topics from Twitter: Algorithm to evaluation. *Knowledge and Information Systems*, *62*(7), 2485–2519. https://doi.org/10.1007/S10115-019-01429-Z

Obar, J. A., & Wildman, S. (2015). Social media definition and the governance challenge: An introduction to the special issue. *Telecommunications Policy*, *39*(9), 745–750. https://doi.org/10.1016/j.telpol.2015.07.014

Ortega, J. L. (2018a). Reliability and accuracy of altmetric providers: A comparison among Altmetric.com, PlumX and Crossref Event Data. *Scientometrics*, *116*(3), 2123–2138. https://doi.org/10.1007/s11192-018-2838-z

Ortega, J. L. (2018b). The life cycle of altmetric impact: A longitudinal study of six metrics from PlumX. *Journal of Informetrics*, *12*(3), 579–589. https://doi.org/10.1016/J.JOI.2018.06.001

Ortega, J. L. (2019a). Exploratory analysis of Publons metrics and their relationship with bibliometric and altmetric impact. *Aslib Journal of Information Management*, *71*(1), 124–136. https://doi.org/10.1108/AJIM-06-2018-0153

Ortega, J. L. (2019b). Availability and audit of links in altmetric data providers: Link checking of blogs and news in Altmetric.com, Crossref Event Data and PlumX. *Journal of Altmetrics*, *2*(1), 4. https://doi.org/10.29024/joa.14

Ortega, J. L. (2020). Altmetrics data providers: A meta-analysis review of the coverage of metrics and publication. *Profesional de La Información*, *29*(1), e290107. https://doi.org/10.3145/EPI.2020.ENE.07

Osterrieder, A. (2013). The value and use of social media as communication tool in the plant sciences. *Plant Methods*, *9*(1), 1–6. https://doi.org/10.1186/1746-4811-9-26

Park, H., Reber, B. H., & Chon, M.-G. (2016). Tweeting as health communication: Health organizations' use of Twitter for health promotion and public engagement. *Journal of Health Communication*, *21*(2), 188–198. https://doi.org/10.1080/10810730.2015.1058435

Park, T. K. (2011). The visibility of Wikipedia in scholarly publications. *First Monday*, *16*(8). https://doi.org/10.5210/FM.V16I8.3492

Pearce, W., Holmberg, K., Hellsten, I., & Nerlich, B. (2014). Climate change on twitter: Topics, communities and conversations about the 2013 IPCC Working Group 1 report. *PLoS ONE*, *9*(4), e94785. https://doi.org/10.1371/journal.pone.0094785

Peoples, B. K., Midway, S. R., Sackett, D., Lynch, A., & Cooney, P. B. (2016). Twitter predicts citation rates of ecological research. *PLOS ONE*, *11*(11), e0166570. https://doi.org/10.1371/JOURNAL.PONE.0166570

Peters, H. P. (2013). Gap between science and media revisited: Scientists as public communicators. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(SUPPL. 3), 14102–14109. https://doi.org/10.1073/pnas.1212745110

Peters, I., Kraker, P., Lex, E., Gumpenberger, C., & Gorraiz, J. (2016). Research data explored: An extended analysis of citations and altmetrics. *Scientometrics*, *107*(2), 723–744. https://doi.org/10.1007/s11192-016-1887-4

Pew Research Center. (2017). *How Americans get science news and information today*. https://www.journalism.org/2017/09/20/science-news-and-information-today/

Pew Research Center. (2021). *Social media fact sheet*. https://www.pewresearch.org/internet/fact-sheet/social-media/

Piwowar, H. (2013). Altmetrics: Value all research products. *Nature*, *493*(7431), 159. https://doi.org/10.1038/493159a

Priem, J., Costello, K., & Dzuba, T. (2011). Prevalence and use of Twitter among scholars. *Metrics 2011 Symposium on Informetric and Scien- Tometric Research*. http://jasonpriem.org/self-archived/5uni-poster.png

Priem, J., & Costello, K. L. (2010). How and why scholars cite on Twitter. *Proceedings of the American Society for Information Science and Technology*, *47*(1), 1–4. https://doi.org/10.1002/meet.14504701201

Priem, J., Groth, P., & Taraborelli, D. (2012). The altmetrics collection. *PLoS ONE*, *7*(11), e48753. https://doi.org/10.1371/journal.pone.0048753

Priem, J., & Hemminger, B. M. (2010). Scientometrics 2.0: Toward new metrics of scholarly impact on the social Web. *First Monday*, *15*(7). https://doi.org/10.5210/fm.v15i7.2874

Priem, J., Piwowar, H. A., & Hemminger, B. M. (2012). Altmetrics in the wild: Using social media to explore scholarly impact. *ArXiv: 1203.4745*. https://arxiv.org/abs/1203.4745v1

Priem, J., Taraborelli, D., Groth, P., & Neylon, C. (2010). *Altmetrics: A manifesto*. http://altmetrics.org/manifesto/

Pulido, C. M., Redondo-Sama, G., Sordé-Martí, T., & Flecha, R. (2018). Social impact in social media: A new method to evaluate the social impact of research. *PLoS ONE*, *13*(8), e0203117. https://doi.org/10.1371/journal.pone.0203117

Ramachandran, A., Wang, L., & Chaintreau, A. (2018). Dynamics and prediction of clicks on news from Twitter. *HT 2018 - Proceedings of the 29th ACM Conference on Hypertext and Social Media*, 210–214. https://doi.org/10.1145/3209542.3209568

Regenberg, A. C. (2010). Tweeting science and ethics: Social media as a tool for constructive public engagement. *The American Journal of Bioethics*, *10*(5), 30–31. https://doi.org/10.1080/15265161003743497

Reincke, C. M., Bredenoord, A. L., & van Mil, M. H. (2020). From deficit to dialogue in science communication: The dialogue communication model requires additional roles from scientists. *EMBO Reports*, *21*(9), e51278. https://doi.org/10.15252/embr.202051278

Robinson-Garcia, N., Arroyo-Machado, W., & Torres-Salinas, D. (2019). Mapping social media attention in Microbiology: Identifying main topics and actors. *FEMS Microbiology Letters*, *366*(7), fnz075. https://doi.org/10.1093/femsle/fnz075

Robinson-Garcia, N., Costas, R., Isett, K., Melkers, J., & Hicks, D. (2017). The unbearable emptiness of tweeting—About journal articles. *PLoS ONE*, *12*(8), e0183551. https://doi.org/10.1371/journal.pone.0183551

Robinson-Garcia, N., Torres-Salinas, D., Zahedi, Z., & Costas, R. (2014). New data, new possibilities: Exploring the insides of altmetric.com. *El Profesional de La Información*, *23*(4), 359–366. https://doi.org/10.3145/epi.2014.jul.03

Robinson-Garcia, N., van Leeuwen, T. N., & Ràfols, I. (2018). Using altmetrics for contextualised mapping of societal impact: From hits to networks. *Science and Public Policy*, *45*(6), 815–826. https://doi.org/10.1093/SCIPOL/SCY024

Rosenkrantz, A. B., Ayoola, A., Singh, K., & Duszak, R. (2017). Alternative metrics (“Altmetrics”) for assessing article impact in popular general radiology journals. *Academic Radiology*, *24*(7), 891–897. https://doi.org/10.1016/j.acra.2016.11.019

Rowlands, I., & Nicholas, D. (2007). The missing link: Journal usage metrics. *Aslib Proceedings*, *59*(3), 222–228. https://doi.org/10.1108/00012530710752025

Rowlands, I., Nicholas, D., Russell, B., Canty, N., & Watkinson, A. (2011). Social media use in the research workflow. *Learned Publishing*, *24*(3), 183–195. https://doi.org/10.1087/20110306

Said, A., Bowman, T. D., Abbasi, R. A., Aljohani, N. R., Hassan, S. U., & Nawaz, R. (2019). Mining network-level properties of Twitter altmetrics data. *Scientometrics*, *120*(1), 217–235. https://doi.org/10.1007/s11192-019-03112-0

Schäfer, M. S. (2009). From public understanding to public engagement: An Empirical Assessment of Changes in Science Coverage. *Science Communication*, *30*(4), 475–505. https://doi.org/10.1177/1075547008326943

Schloegl, C., & Gorraiz, J. (2010). Comparison of citation and usage indicators: The case of oncology journals. *Scientometrics*, *82*(3), 567–580. https://doi.org/10.1007/S11192-010-0172-1

Schloegl, C., & Gorraiz, J. (2011). Global usage versus global citation metrics: The case of pharmacology journals. *Journal of the American Society for Information Science and Technology*, *62*(1), 161–170. https://doi.org/10.1002/ASI.21420

Schlögl, C., Gorraiz, J., Gumpenberger, C., Jack, K., & Kraker, P. (2014). Comparison of downloads, citations and readership data for two information systems journals. *Scientometrics*, *101*(2), 1113–1128. https://doi.org/10.1007/S11192-014-1365-9

Schubert, A., & Glänzel, W. (1986). Mean response time — A new indicator of journal citation speed with application to physics journals. *Czechoslovak Journal of Physics B*, *36*(1), 121–125. https://doi.org/10.1007/BF01599743

Seglen, P. O. (1992). The skewness of science. *Journal of the American Society for Information Science*, *43*(9), 628–638. https://doi.org/10.1002/(SICI)1097-4571(199210)43:9<628::AID-ASI5>3.0.CO;2-0

Shema, H., Bar-Ilan, J., & Thelwall, M. (2014). Do blog citations correlate with a higher number of future citations? Research blogs as a potential source for alternative metrics. *Journal of the Association for Information Science and Technology*, *65*(5), 1018–1027. https://doi.org/10.1002/asi.23037

Shibata, N., Kajikawa, Y., Takeda, Y., & Matsushima, K. (2008). Detecting emerging research fronts based on topological measures in citation networks of scientific publications. *Technovation*, *28*(11), 758–775. https://doi.org/10.1016/J.TECHNOVATION.2008.03.009

Shu, F., Lou, W., & Haustein, S. (2018). Can Twitter increase the visibility of Chinese publications? *Scientometrics*, *116*(1), 505–519. https://doi.org/10.1007/s11192-018-2732-8

Shuai, X., Pepe, A., & Bollen, J. (2012). How the scientific community reacts to newly submitted preprints: Article downloads, Twitter mentions, and citations. *PLoS ONE*, *7*(11), e47523. https://doi.org/10.1371/journal.pone.0047523

Small, H. (2006). Tracking and predicting growth areas in science. *Scientometrics*, *68*(3), 595–610. https://doi.org/10.1007/S11192-006-0132-Y

Small, H., Boyack, K. W., & Klavans, R. (2014). Identifying emerging topics in science and technology. *Research Policy*, *43*(8), 1450–1467. https://doi.org/10.1016/J.RESPOL.2014.02.005

Sonnenwald, D. H. (2007). Scientific collaboration. *Annual Review of Information Science and Technology*, *41*(1), 643–681. https://doi.org/10.1002/aris.2007.1440410121

Stinson, E. R., & Lancaster, F. W. (1987). Synchronous versus diachronous methods in the measurement of obsolescence by citation studies: *Journal of Information Science*, *13*(2), 65–74. https://doi.org/10.1177/016555158701300201

Sud, P., & Thelwall, M. (2014). Evaluating altmetrics. *Scientometrics*, *98*(2), 1131–1143. https://doi.org/10.1007/s11192-013-1117-2

Sugimoto, C. R. (2015, June 24). *"Attention is not impact" and other challenges for altmetrics*. https://www.wiley.com/network/researchers/promoting-your-article/attention-is-not-impact-and-other-challenges-for-altmetrics

Sugimoto, C. R., Robinson-Garcia, N., Murray, D. S., Yegros-Yegros, A., Costas, R., & Larivière, V. (2017). Scientists have most impact when they're free to move. *Nature*, *550*(7674), 29–31. https://doi.org/10.1038/550029a

Sugimoto, C. R., Thelwall, M., Larivière, V., Tsou, A., Mongeon, P., & Macaluso, B. (2013). Scientists popularizing science: Characteristics and impact of TED talk presenters. *PLoS ONE*, *8*(4), e62403. https://doi.org/10.1371/journal.pone.0062403

Sugimoto, C. R., Work, S., Larivière, V., & Haustein, S. (2017). Scholarly use of social media and altmetrics: A review of the literature. *Journal of the Association for Information Science and Technology*, *68*(9), 2037–2062. https://doi.org/10.1002/asi.23833

Sun, J., Min, C., & Li, J. (2016). A vector for measuring obsolescence of scientific articles. *Scientometrics*, *107*(2), 745–757. https://doi.org/10.1007/S11192-016-1884-7

Sutton, S. W. (2014). Altmetrics: What good are they to academic libraries? *Kansas Library Association College and University Libraries Section Proceedings*, *4*(2), 1. https://doi.org/10.4148/2160-942X.1041

Tenopir, C., Volentine, R., & King, D. W. (2013). Social media and scholarly reading. *Online Information Review*, *37*(2), 193–216. https://doi.org/10.1108/OIR-04-2012-0062

Terry, M. (2009). Twittering healthcare: Social media and medicine. *Telemedicine and E-Heath*, *15*(6), 507–510. https://doi.org/10.1089/tmj.2009.9955

Thelwall, M. (2017). Are Mendeley reader counts high enough for research evaluations when articles are published? *Aslib Journal of Information Management*, *69*(2), 215–229. https://doi.org/10.1108/AJIM-01-2017-0028

Thelwall, M. (2020). Measuring societal impacts of research with altmetrics? Common problems and mistakes. *Journal of Economic Surveys*, joes.12381. https://doi.org/10.1111/joes.12381

Thelwall, M., Haustein, S., Larivière, V., & Sugimoto, C. R. (2013). Do altmetrics work? Twitter and ten other social web services. *PLoS ONE*, *8*(5), e64841. https://doi.org/10.1371/journal.pone.0064841

Thelwall, M., & Kousha, K. (2015). ResearchGate: Disseminating, communicating, and measuring Scholarship? *Journal of the Association for Information Science and Technology*, *66*(5), 876–889. https://doi.org/10.1002/asi.23236

Thelwall, M., & Sud, P. (2016). Mendeley readership counts: An investigation of temporal and disciplinary differences. *Journal of the Association for Information Science and Technology*, *67*(12), 3036–3050. https://doi.org/10.1002/asi.23559

Thelwall, M., Tsou, A., Weingart, S., Holmberg, K., & Haustein, S. (2013). Tweeting links to academic articles. *International Journal of Scientometrics, Informetrics and Bibliometrics*, *17*(1), 1–8. http://cybermetrics.cindoc.csic.es/articles/v17i1p1.html

Thelwall, M., & Wilson, P. (2016). Mendeley readership altmetrics for medical articles: An analysis of 45 fields. *Journal of the Association for Information Science and Technology*, *67*(8), 1962–1972. https://doi.org/10.1002/asi.23501

Tseng, Y.-H., Lin, Y.-I., Lee, Y.-Y., Hung, W.-C., & Lee, C.-H. (2009). A comparison of methods for detecting hot topics. *Scientometrics*, *81*(1), 73–90. https://doi.org/10.1007/S11192-009-1885-X

Tsou, A., Bowman, T., Ghazinejad, A., & Sugimoto, C. (2015). Who tweets about science? *Proceedings of the 2015 International Society for Scientometrics and Informetrics*, 95–100. https://www.issi-society.org/proceedings/issi_2015/0095.pdf

Tsou, A., Thelwall, M., Mongeon, P., & Sugimoto, C. R. (2014). A community of curious souls: An analysis of commenting behavior on TED Talks videos. *PLoS ONE*, *9*(4), e93609. https://doi.org/10.1371/journal.pone.0093609

Vainio, J., & Holmberg, K. (2017). Highly tweeted science articles: Who tweets them? An analysis of Twitter user profile descriptions. *Scientometrics*, *112*(1), 345–366. https://doi.org/10.1007/s11192-017-2368-0

Van Eperen, L., & Marincola, F. M. (2011). How scientists use social media to communicate their research. *Journal of Translational Medicine*, *9*(1), 199. https://doi.org/10.1186/1479-5876-9-199

Van Noorden, R. (2014). Online collaboration: Scientists and the social network. *Nature*, *512*(7513), 126–129. https://doi.org/10.1038/512126a

Van Schalkwyk, F., Dudek, J., & Costas, R. (2020). Communities of shared interests and cognitive bridges: The case of the anti-vaccination movement on Twitter. *Scientometrics*, *125*(2), 1499–1516. https://doi.org/10.1007/s11192-020-03551-0

Veletsianos, G., Kimmons, R., Shaw, A., Pasquini, L., & Woodward, S. (2017). Selective openness, branding, broadcasting, and promotion: Twitter use in Canada's public universities. *Educational Media International*, *54*(1), 1–19. https://doi.org/10.1080/09523987.2017.1324363

Waltman, L., & Costas, R. (2014). F1000 recommendations as a potential new data source for research evaluation: A comparison with citations. *Journal of the Association for Information Science and Technology*, *65*(3), 433–445. https://doi.org/10.1002/asi.23040

Waltman, L., & Van Eck, N. J. (2012). A new methodology for constructing a publication-level classification system of science. *Journal of the American Society for Information Science and Technology*, *63*(12), 2378–2392. https://doi.org/10.1002/asi.22748

Wan, J., Hua, P., Rousseau, R., & Sun, X. (2010). The journal download immediacy index (DII): experiences using a Chinese full-text database. *Scientometrics*, *82*(3), 555–566. https://doi.org/10.1007/S11192-010-0171-2

Wang, J. (2013). Citation time window choice for research impact evaluation. *Scientometrics*, *94*(3), 851–872. https://doi.org/10.1007/s11192-012-0775-9

Wang, L. X., Ramachandran, A., & Chaintreau, A. (2016). Measuring click and share dynamics on social media: A reproducible and validated approach. *The Workshops of the Tenth International AAAI Conference on Web and Social Media*, 108–113. https://www.aaai.org/ocs/index.php/ICWSM/ICWSM16/paper/view/13213

Wang, Q. (2018). A bibliometric model for identifying emerging research topics. *Journal of the Association for Information Science and Technology*, *69*(2), 290–304. https://doi.org/10.1002/ASI.23930

Wang, X., Cui, Y., Li, Q., & Guo, X. (2017). Social media attention increases article visits: An investigation on article-level referral data of PeerJ. *Frontiers in Research Metrics and Analytics*, *2*, 11. https://doi.org/10.3389/frma.2017.00011

Wang, X., & Fang, Z. (2016). Detecting and tracking the real-time hot topics: A study on computational neuroscience. *ArXiv 1608.05517*. https://arxiv.org/abs/1608.05517v1

Wang, X., Fang, Z., & Guo, X. (2016). Tracking the digital footprints to scholarly articles from social media. *Scientometrics*, *109*(2), 1365–1376. https://doi.org/10.1007/s11192-016-2086-z

Wang, X., Fang, Z., Li, Q., & Guo, X. (2016). The poor altmetric performance of publications authored by researchers in mainland China. *Frontiers in Research Metrics and Analytics*, *1*, 8. https://doi.org/10.3389/frma.2016.00008

Wang, X., Fang, Z., & Sun, X. (2016). Usage patterns of scholarly articles on Web of Science: A study on Web of

Science usage count. *Scientometrics*, *109*(2), 917–926. https://doi.org/10.1007/s11192-016-2093-0

Wang, X., Mao, W., Xu, S., & Zhang, C. (2014). Usage history of scientific literature: Nature metrics and metrics of Nature publications. *Scientometrics*, *98*(3), 1923–1933. https://doi.org/10.1007/S11192-013-1167-5

Wang, X., Wang, Z., & Xu, S. (2013). Tracing scientist's research trends realtimely. *Scientometrics*, *95*(2), 717–729. https://doi.org/10.1007/S11192-012-0884-5

Weingart, P., & Guenther, L. (2016). Science communication and the issue of trust. *Journal of Science Communication*, *15*(5), 1–11. https://doi.org/10.22323/2.15050301

Weller, K., & Puschmann, C. (2011, June). Twitter for scientific communication: How can citations/references be identified and measured? *Proceedings of the 3rd ACM International Conference on Web Science (WebSci '11)*. https://websci11.org/fileadmin/websci/Posters/153_paper.pdf

Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., Jones, R., Kain, R., Kerridge, S., Thelwall, M., Tinkler, J., Viney, I., Wouters, P., Hill, J., & Johnson, B. (2015). *The metric tide: Report of the independent review of the role of metrics in research assessment and management*. http://eprints.whiterose.ac.uk/117033/1/2015_metric_tide.pdf

Wouters, P. (1999). *The citation culture* [University of Amsterdam]. https://dare.uva.nl/search?identifier=b101b769-100f-43e5-b8d2-cac6c11e5bbf

Wouters, P., & Costas, R. (2012). Users, narcissism and control-tracking the impact of scholarly publications in the 21st century. In *Utrecht: SURFfoundation*. http://research-acumen.eu/wp-content/uploads/Users-narcissism-and-control.pdf

Wouters, P., Zahedi, Z., & Costas, R. (2019). Social media metrics for new research evaluation. In W. Glänzel, F. H. Moed, U. Schmoch, & M. Thelwall (Eds.), *Springer Handbook of Science and Technology Indicators* (pp. 687–713). https://doi.org/10.1007/978-3-030-02511-3_26

Xia, F., Su, X., Wang, W., Zhang, C., Ning, Z., & Lee, I. (2016). Bibliographic analysis of Nature based on Twitter and Facebook altmetrics data. *PLOS ONE*, *11*(12), e0165997. https://doi.org/10.1371/JOURNAL.PONE.0165997

Xu, J. M., Burchfiel, B., Zhu, X., & Bellmore, A. (2013). An examination of regret in bullying tweets. *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 697–702. https://www.aclweb.org/anthology/N13-1082/

Xu, S., Yu, H., Hemminger, B. M., & Dong, X. (2018). Who, what, why? An exploration of JoVE scientific video publications in tweets. *Scientometrics*, *117*(2), 845–856. https://doi.org/10.1007/s11192-018-2880-x

Yu, H. (2017). Context of altmetrics data matters: An investigation of count type and user category. *Scientometrics*, *111*(1), 267–283. https://doi.org/10.1007/s11192-017-2251-z

Yu, H., Murat, B., Li, L., & Xiao, T. (2021). How accurate are Twitter and Facebook altmetrics data? A comparative content analysis. *Scientometrics*, *126*(5), 4437–4463. https://doi.org/10.1007/s11192-021-03954-7

Yu, H., Xiao, T., Xu, S., & Wang, Y. (2019). Who posts scientific tweets? An investigation into the productivity, locations, and identities of scientific tweeters. *Journal of Informetrics*, *13*(3), 841–855. https://doi.org/10.1016/j.joi.2019.08.001

Yu, H., Xu, S., Xiao, T., Hemminger, B. M., & Yang, S. (2017). Global science discussed in local altmetrics: Weibo and its comparison with Twitter. *Journal of Informetrics*, *11*(2), 466–482. https://doi.org/10.1016/j.joi.2017.02.011

Zahedi, Z., & Costas, R. (2017, October). How visible are the research of different countries on WoS and Twitter? An analysis of global vs. local reach of WoS publications on Twitter. *Proceedings of the 16th International Society of Scientometrics and Informetrics Conference*. https://doi.org/10.6084/m9.figshare.5481283.v1

Zahedi, Z., & Costas, R. (2018). General discussion of data quality challenges in social media metrics: Extensive comparison of four major altmetric data aggregators. *PLoS ONE*, *13*(5), e0197326. https://doi.org/10.1371/journal.pone.0197326

Zahedi, Z., Costas, R., & Wouters, P. (2014). How well developed are altmetrics? A cross-disciplinary analysis of the presence of 'alternative metrics' in scientific publications. *Scientometrics*, *101*(2), 1491–1513. https://doi.org/10.1007/s11192-014-1264-0

R

Zahedi, Z., Costas, R., & Wouters, P. (2017). Mendeley readership as a filtering tool to identify highly cited publications. *Journal of the Association for Information Science and Technology*, *68*(10), 2511–2521. https://doi.org/10.1002/asi.23883

Zahedi, Z., Fenner, M., & Costas, R. (2015, October 9). Consistency among altmetrics data provider/aggregators: What are the challenges? *Proceedings of the 2015 Altmetrics Workshop*. http://altmetrics.org/wp-content/uploads/2015/09/altmetrics15_paper_14.pdf

Zahedi, Z., & Haustein, S. (2018). On the relationships between bibliographic characteristics of scientific documents and citation and Mendeley readership counts : A large-scale analysis of Web of Science publications. *Journal of Informetrics*, *12*(1), 191–202. https://doi.org/10.1016/j.joi.2017.12.005

Zahedi, Z., & Van Eck, N. J. (2018). Exploring topics of interest of Mendeley users. *Journal of Altmetrics*, *1*(1), 5. https://doi.org/10.29024/joa.7

Zedda, M., & Barbaro, A. (2015). Adoption of Web 2.0 tools among STM publishers. How social are scientific journals? *Journal of the European Association for Health Information and Libraries*, *11*(1), 9–12. http://ojs.eahil.eu/ojs/index.php/JEAHIL/article/view/21

Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression models for count data in R. *Journal of Statistical Software*, *27*(8), 1–25. https://doi.org/10.18637/jss.v027.i08

Zheng, H., Aung, H. H., Erdt, M., Peng, T. Q., Sesagiri Raamkumar, A., & Theng, Y. L. (2019). Social media presence of scholarly journals. *Journal of the Association for Information Science and Technology*, *70*(3), 256–270. https://doi.org/10.1002/asi.24124

Ziman, J. (1991). Public understanding of science. *Science, Technology & Human Values*, *16*(1), 99–105. https://doi.org/10.1177/016224399101600106

Zubiaga, A. (2018). A longitudinal assessment of the persistence of twitter datasets. *Journal of the Association for Information Science and Technology*, *69*(8), 974–984. https://doi.org/10.1002/asi.24026

Summary

With the increasing popularity of the scholarly use of social media, numerous digital traces of social media interactions around science are left in the online environments every day. The analysis of these traces is what gave rise to the emergence of *social media metrics of science*. Twitter is arguably the most popular social media platform for communicating scientific information. The analysis of Twitter engagement around science opens a range of possibilities to capture and characterize the attention towards scientific developments beyond academia. There exist diverse forms of Twitter interactions, enabled by the multiple interactive features provided by the platform, which capture the stories of how scientific information is shared, disseminated, discussed, and used by Twitter users.

Building on a proposed conceptual framework of science-social media interactions, the main objective of this PhD dissertation is to characterize the various forms of Twitter interactions around science, and approach more advanced Twitter-based metrics by systematically considering the diversity and characteristics of Twitter interactions. This dissertation consists of seven chapters:

Chapter 1 presents a general introduction to science-social media interactions. This chapter starts with the definitions of social media, scholarly use of social media, and social media metrics of science. The chapter proposes a conceptual framework of science-social media interactions, which conceptually outlines and categorizes the diverse forms of interactions happening within and between the science and social media environments, as well as the information flows aroused by the corresponding interactions. Based on the proposed conceptual framework, this chapter applies it to systematically review the relevant literature regarding the interactions between science and Twitter. Finally, considering the opportunities and challenges facing scholarly Twitter metrics, this chapter puts forward the main objective and research questions to be addressed.

Chapter 2 presents a state-of-the-art analysis of the presence of Twitter mention data amongst scientific papers, in comparison with other eleven types of altmetric data sources tracked by Altmetric.com. Different altmetric data sources show significant differences in the uptake of scientific papers, confirming the heterogeneity of altmetrics and the importance of keeping them separate in both research and practice. Overall, Twitter mentions cover more than one third of the recent scientific papers, being the most global source of evidence of social media interactions around science, only second to Mendeley. This chapter also highlights the recent biases of Twitter uptake towards scientific papers published in recent years, and the disciplinary biases towards papers in the fields of *Social Sciences and Humanities*, *Biomedical and Health Sciences*, and *Life and Earth Sciences*. Finally, this

chapter compares the Twitter uptake of scientific papers at the research topic level, shedding light on a new way to identify hot research topics in the eyes of Twitter users.

Chapter 3 studies the Twitter uptake of scientific papers from the aspect of velocity, namely, how fast scientific papers are mentioned on Twitter after they were published. Chapter 3 also presents a comparison analysis of the data accumulation velocity amongst twelve altmetric data sources tracked by Altmetric.com. Different sources show significant discrepancy in the uptake velocity of scientific papers. Overall, there exist *fast sources* which show a relatively higher velocity in making mention of scientific papers after their publication, such as Twitter, Reddit, Facebook, and news media, and *slow sources* which exhibit a relatively lower velocity, such as policy documents, Wikipedia, and peer review platforms. This discrepancy reinforces the idea that keeping altmetrics separate is an important recommendation, and highlights the necessity of selecting appropriate time windows for different sources of altmetric data in altmetric research. Chapter 3 also observes that Twitter, as one of the fastest sources, tend to accumulate faster for document types of editorial materials and letters, and scientific papers from the fields of *Physical Sciences and Engineering* and *Life and Earth Sciences*.

Chapter 4 focuses on the user engagement behaviors around original scholarly tweets mentioning scientific papers, to explore how original scholarly tweets are further engaged with by Twitter users through *liking*, *retweeting*, *quoting*, and *replying*. It is found that only half of the original scholarly tweets triggered at least one of the four types of user engagement behaviors, implying that original scholarly tweets perform differently in drawing broader attention in the Twittersphere. Original scholarly tweets regarding *Social Sciences and Humanities*, *Biomedical and Health Sciences*, and *Life and Earth Sciences* are more likely to attract further user engagement on Twitter. Based on correlation and regression analyses, the results show that user engagement metrics correlate more with other Twitter-based factors (e.g., followers of Twitter users or mentioned users in scholarly tweets) than with science-based factors (e.g., citations or Mendeley readers of scientific papers), suggesting the intrinsic relationships amongst Twitter elements and activities, as well as the differential propensities of interactions taking place in science and Twitter for scholarly outputs.

Chapter 5 explores the possible unavailability of scholarly tweets and its influence on the stability of Twitter-based metrics. By rechecking the (un)availability of a set of scholarly tweets of highly tweeted scientific papers recorded by Altmetric.com up to October 2017, this chapter reports that in April 2019 there were 14.3% of the tweets had become unavailable to the public mainly due to tweet deletion, user account suspension, and user account protection. The unavailability of scholarly tweets may seriously influence the stability of Twitter-based metrics. By distinguishing between original tweets and retweets, and then networking between them based on the retweeting relationships, chapter 5 shows that scientific papers with fewer original tweets and high levels of retweets are more vulnerable in their Twitter dissemination structure, thus being at a greater risk of creating unstable Twitter-based metrics. These findings reflect the potential instability of Twitter-based metrics due to the volatile nature of Twitter data, and emphasize the necessity of not only differentiating between original tweets and retweets, but also analyzing the Twitter dissemination structure of papers with a network approach.

Chapter 6 investigates the clicking behavior around scholarly URLs referring to scientific papers embedded in scholarly tweets, which leads Twitter users to access the original webpages of the tweeted scientific papers. Relying on the click metrics provided by Bitly for its generated short links, chapter 6 observes that only about half of the scholarly URLs successfully received clicks on Twitter and thus directed Twitter users to visit the original scholarly content. The majority of Twitter clicks on the tweeted scholarly URLs concentrate in the first few days after the scholarly URLs were tweeted. Similar to other Twitter interactions, Twitter clicks appear to be more frequent on scholarly URLs regarding *Social Sciences and Humanities*, *Biomedical and Health Sciences*, and *Life and Earth Sciences*. The discrepancy of Twitter clicks across scholarly URLs also indicates the different performance of scholarly tweets in triggering wider attention, offering the possibilities to measure the effectiveness of scholarly tweets in disseminating scientific information and the feedback of Twitter dissemination on science itself (e.g., by increasing the visits to scientific papers).

Chapter 7 summarizes the main findings presented in chapters 2 through 6 and further discusses the implications and some future prospects based on the main findings. Research on the Twitter uptake of scientific papers (chapters 2 and 3) demonstrates that Twitter is a global and immediate source of evidence of social media interactions around science. Research on the diverse types of Twitter engagement metrics for scholarly tweets (chapters 4 and 6) reveals the possibility of establishing a more fine-grained indicator system of Twitter-based metrics, which also enables the measurements of more deep-seated Twitter reception of scientific information. Research on the possible unavailability of scholarly tweets (chapter 5) reflects the influence of the volatile nature of tweets may have on the stability of Twitter-based metrics, which is recommended to be treated with caution in Twitter-related research and evaluation. To better understand the nature of the diverse Twitter interactions, chapter 7 also proposes some directions for future research, particularly the contextualization of Twitter interactions by taking into account the information of the involved Twitter users and the detailed tweet content. Besides, it would be an important future step to generalize the conceptual framework of science-social media interactions (in chapter 1) and the design of a more fine-grained indicator system of social media metrics (in chapter 7) to a broader range of social media sources.

Nederlandse Samenvatting

Met de toenemende populariteit van het wetenschappelijke gebruik van sociale media, worden er dagelijks talloze digitale sporen van sociale media-interacties rond wetenschap achtergelaten in de online omgevingen. De analyse van deze sporen heeft geleid tot het ontstaan van de *sociale mediametriek van de wetenschap*. Twitter is aantoonbaar het populairste sociale media platform voor de communicatie van wetenschappelijke informatie. De analyse van Twitter-engagement rond wetenschap opent een waaier van mogelijkheden om de aandacht voor wetenschappelijke ontwikkelingen buiten de academische wereld te vatten en te karakteriseren. Er bestaan verschillende vormen van Twitter-interacties, mogelijk gemaakt door de vele interactieve functies die het platform biedt, die de verhalen vastleggen over hoe wetenschappelijke informatie wordt gedeeld, verspreid, besproken en gebruikt door Twitter-gebruikers.

Voortbouwend op een voorgesteld conceptueel kader van wetenschap-sociale media interacties, is het hoofddoel van deze doctoraatsverhandeling het karakteriseren van de verschillende vormen van Twitter interacties rond wetenschap, en het benaderen van meer geavanceerde Twitter-gebaseerde metrieken door systematisch rekening te houden met de diversiteit en karakteristieken van Twitter interacties. Dit proefschrift bestaat uit zeven hoofdstukken:

Hoofdstuk 1 geeft een algemene inleiding tot de interacties tussen wetenschap en sociale media. Dit hoofdstuk begint met de definities van sociale media, wetenschappelijk gebruik van sociale media, en sociale mediametriek van de wetenschap. Het hoofdstuk stelt een conceptueel kader voor van interacties tussen wetenschap en sociale media, dat de diverse vormen van interacties binnen en tussen de wetenschappelijke en sociale media-omgevingen conceptueel schetst en categoriseert, evenals de informatiestromen die door de overeenkomstige interacties worden opgewekt. Op basis van het voorgestelde conceptuele kader wordt in dit hoofdstuk een systematisch overzicht gegeven van de relevante literatuur over de wisselwerking tussen wetenschap en Twitter. Ten slotte worden in dit hoofdstuk, rekening houdend met de mogelijkheden en uitdagingen voor wetenschappelijke Twitter-metrieken, de hoofddoelstelling en de te behandelen onderzoeksvragen naar voren gebracht.

Hoofdstuk 2 presenteert een state-of-the-art analyse van de aanwezigheid van Twitter-vermeldingsgegevens bij wetenschappelijke artikelen, in vergelijking met andere elf soorten altmetrische gegevensbronnen die door Altmetric.com worden gevolgd. Verschillende altmetrische gegevensbronnen vertonen aanzienlijke verschillen in de opname van wetenschappelijke papers, wat de heterogeniteit van altmetrische gegevens bevestigt en het belang om ze zowel in onderzoek als in de praktijk gescheiden te houden. Over het geheel genomen bestrijken Twitter-vermeldingen meer dan een derde van de recente wetenschappelijke papers, en vormen ze de meest globale bron van bewijsmateriaal voor sociale media-interacties rond wetenschap, slechts tweede na Mendeley. In dit hoofdstuk wordt ook gewezen op de recente vertekening van Twitter ten aanzien van wetenschappelijke artikelen die in de afgelopen jaren zijn gepubliceerd, en op de vertekening in de verschillende disciplines ten aanzien van artikelen op het gebied van *sociale en geesteswetenschappen*, *biomedische en gezondheidswetenschappen*, en *biowetenschappen en aardwetenschappen*. Ten slotte vergelijkt dit hoofdstuk de Twitter-acceptatie van wetenschappelijke papers op het niveau van het onderzoeksthema, wat licht werpt op een nieuwe manier om populaire onderzoeksonderwerpen in de ogen van Twitter-gebruikers te identificeren.

Hoofdstuk 3 bestudeert de Twitter-opname van wetenschappelijke papers vanuit het oogpunt van snelheid, namelijk hoe snel wetenschappelijke papers op Twitter worden vermeld nadat ze zijn gepubliceerd. Hoofdstuk 3 presenteert ook een vergelijkende analyse van de gegevens accumulatie snelheid tussen twaalf altmetrische gegevensbronnen die worden bijgehouden door Altmetric.com. Verschillende bronnen laten significante discrepantie zien in de opnamesnelheid van wetenschappelijke artikelen. In het algemeen bestaan er *snelle bronnen* die een relatief hogere snelheid vertonen bij het vermelden van wetenschappelijke papers na hun publicatie, zoals Twitter, Reddit, Facebook en nieuwsmedia, en *trage bronnen* die een relatief lagere snelheid vertonen, zoals beleidsdocumenten, Wikipedia en platforms voor collegiale toetsing. Deze discrepantie versterkt het idee dat het gescheiden houden van altmetrische gegevens een belangrijke aanbeveling is, en benadrukt de noodzaak van het selecteren van geschikte tijdvensters voor verschillende bronnen van altmetrische gegevens in altmetrisch onderzoek. In hoofdstuk 3 wordt ook opgemerkt dat Twitter, als een van de snelste bronnen, de neiging heeft zich sneller te accumuleren voor documenttypes als redactioneel materiaal en brieven, en wetenschappelijke papers op het gebied van *natuur- en ingenieurswetenschappen* en *biowetenschappen en aardwetenschappen*.

Hoofdstuk 4 richt zich op het gedrag van gebruikers rond originele wetenschappelijke tweets waarin wetenschappelijke artikelen worden genoemd, om te onderzoeken hoe originele wetenschappelijke tweets door Twittergebruikers verder worden gebruikt door ze te *liken*, te *retweeten*, te *citeren* en te *beantwoorden*. Het blijkt dat slechts de helft van de originele wetenschappelijke tweets minstens één van de vier types van gebruikersbetrokkenheid uitlokten, wat impliceert dat originele wetenschappelijke tweets anders presteren in het trekken van bredere aandacht in de Twittersphere. Originele wetenschappelijke tweets over *sociale en geesteswetenschappen*, *biomedische en gezondheidswetenschappen*, en *biowetenschappen en aardwetenschappen* zullen waarschijnlijk meer gebruikersbetrokkenheid op Twitter aantrekken. Gebaseerd op correlatie- en regressieanalyses tonen de resultaten aan dat metriek van gebruikersbetrokkenheid meer correleert met andere Twitter-gebaseerde factoren (bv. volgers van Twitter-gebruikers of vermelde gebruikers in wetenschappelijke tweets) dan met wetenschapsgebaseerde factoren (bv. citaties of Mendeley-lezers van wetenschappelijke papers), wat de intrinsieke relaties tussen Twitter-elementen en -activiteiten suggereert, alsook de differentiële neigingen van interacties die plaatsvinden in de wetenschap en op Twitter voor wetenschappelijke outputs.

Hoofdstuk 5 onderzoekt de mogelijke onbeschikbaarheid van wetenschappelijke tweets en de invloed daarvan op de stabiliteit van op Twitter gebaseerde metrieken. Door het opnieuw controleren van de (on)beschikbaarheid van een set wetenschappelijke tweets van hooggetwitterde wetenschappelijke papers die tot oktober 2017 door Altmetric.com zijn geregistreerd, meldt dit hoofdstuk dat er in april 2019 14,3% van de tweets onbeschikbaar was geworden voor het publiek, voornamelijk als gevolg van het verwijderen van tweets, schorsing van gebruikersaccounts en bescherming van gebruikersaccounts. Het niet beschikbaar zijn van wetenschappelijke tweets kan de stabiliteit van op Twitter gebaseerde metrieken ernstig beïnvloeden. Door een onderscheid te maken tussen originele tweets en retweets, en daartussen vervolgens te netwerken op basis van de retweeting-relaties, toont hoofdstuk 5 dat wetenschappelijke artikelen met minder originele tweets en veel retweets kwetsbaarder zijn in hun verspreidingsstructuur op Twitter, en dus een groter risico lopen om onstabiele, op Twitter gebaseerde statistieken te creëren. Deze bevindingen weerspiegelen de potentiële instabiliteit van op Twitter gebaseerde metingen als gevolg van de veranderlijke aard van Twitter-gegevens, en benadrukken de noodzaak om niet alleen een onderscheid te maken tussen originele tweets en retweets, maar ook om de Twitter-verspreidingsstructuur van papers te analyseren met een netwerkbenadering.

Hoofdstuk 6 onderzoekt het klikgedrag rond wetenschappelijke URL's die verwijzen naar wetenschappelijke artikelen die zijn ingebed in wetenschappelijke tweets, waardoor Twitter-gebruikers toegang krijgen tot de originele webpagina's van de getweete wetenschappelijke artikelen. Gebaseerd op de klikstatistieken van Bitly voor zijn gegenereerde korte links, stelt hoofdstuk 6 vast dat slechts ongeveer de helft van de wetenschappelijke URL's met succes op Twitter werden aangeklikt en Twitter-gebruikers dus naar de oorspronkelijke wetenschappelijke inhoud leidden. De meeste Twitter-kliks op de getweete wetenschappelijke URL's concentreren zich in de eerste paar dagen nadat de wetenschappelijke URL's waren getweet. Net als bij andere Twitter-interacties lijken Twitter-kliks vaker voor te komen bij wetenschappelijke URL's met betrekking tot *sociale en geesteswetenschappen*, *biomedische en gezondheidswetenschappen*, en *biowetenschappen en aardwetenschappen*. Het verschil in het aantal Twitter-kliks tussen wetenschappelijke URL's wijst ook op de verschillende prestaties van wetenschappelijke tweets bij het op gang brengen van een bredere aandacht, wat de mogelijkheid biedt om de doeltreffendheid te meten van wetenschappelijke tweets bij de verspreiding van wetenschappelijke informatie en de feedback van Twitter-verspreiding op de wetenschap zelf (bv. door meer bezoeken aan wetenschappelijke papers).

Hoofdstuk 7 vat de belangrijkste bevindingen van de hoofdstukken 2 tot en met 6 samen en bespreekt de implicaties en enkele toekomstperspectieven op basis van de belangrijkste bevindingen. Onderzoek naar de opname van wetenschappelijke papers via Twitter (hoofdstukken 2 en 3) toont aan dat Twitter een wereldwijde en onmiddellijke bron van bewijs is van sociale media-interacties rond wetenschap. Onderzoek naar de verschillende soorten Twitter-engagement metrieken voor wetenschappelijke tweets (hoofdstukken 4 en 6) onthult de mogelijkheid om een meer fijnmazig indicatorsysteem van Twitter-gebaseerde metrieken op te zetten, dat ook de metingen van meer diepgewortelde Twitter-ontvangst van wetenschappelijke informatie mogelijk maakt. Het onderzoek naar de mogelijke onbeschikbaarheid van wetenschappelijke tweets (hoofdstuk 5) weerspiegelt de invloed die de vluchtige aard van tweets kan hebben op de stabiliteit van op Twitter gebaseerde metriek, die wordt aanbevolen om in Twitter-gerelateerd onderzoek en evaluatie met de nodige voorzichtigheid te behandelen. Om de aard van de diverse Twitter-interacties beter te begrijpen, worden in hoofdstuk 7 ook enkele richtingen voor toekomstig onderzoek voorgesteld, met name de contextualisering van Twitter-interacties door rekening te houden met de informatie van de betrokken Twitter-gebruikers en de gedetailleerde tweet-inhoud. Bovendien zou het een belangrijke toekomstige stap zijn om het conceptuele raamwerk van wetenschap-sociale media-interacties (in hoofdstuk 1) en het ontwerp van een fijnmaziger indicatorsysteem van sociale media-metrieken (in hoofdstuk 7) te veralgemenen naar een breder scala van sociale media bronnen.

Curriculum Vitae

Zhichao Fang was born on February 25, 1992 in Yueyang City, Hunan Province (China). Zhichao graduated from Yueyang First Senior High School in 2010 and was admitted to Dalian University of Technology in China. He obtained his Bachelor's degree in Public Administration in 2014. After that, he was given exemption from the examination to be admitted to a Master's programme at WISE Lab of Dalian University of Technology. In 2017, Zhichao obtained his Master's degree in Science of Science and Science & Technology Management. After receiving the financial support from the China Scholarship Council (the CSC scholarship), Zhichao moved to the Netherlands in September 2017 and started his PhD research at the Centre for Science and Technology Studies (CWTS) of Leiden University, under the co-supervision of Prof. dr. Paul Wouters and Dr. Rodrigo Costas.

Zhichao's PhD project mainly focuses on diverse Twitter interactions around scientific publications, to characterize the patterns and effectiveness of Twitter reception and dissemination of scientific knowledge. The fundamental objective of his PhD dissertation is to approach more advanced social media metrics of science by taking into consideration the diverse science-social media interactions. In 2019, this PhD dissertation was awarded an honorable mention of the Eugene Garfield Doctoral Dissertation Award by the International Society for Scientometrics and Informetrics (ISSI). During his master and PhD programmes, Zhichao has published a number of peer-reviewed journal papers and conference papers. In addition, Zhichao served as reviewers for a range of scholarly journals, including *Journal of the Association for Information Science and Technology*, *Scientometrics*, *Journal of Information Science*, *Quantitative Science Studies*, *Journal of Medical Internet Research*, and *Journal of Data and Information Science*.

List of publications

International peer-reviewed journal papers:

Fang, Z., Costas, R., Tian, W., Wang, X., & Wouters, P. (2021). How is science clicked on Twitter? Click metrics for Bitly short links to scientific publications. *Journal of the Association for Information Science and Technology*, 72(7), 918-932.

Fang, Z., Costas, R., Tian, W., Wang, X., & Wouters, P. (2020). An extensive analysis of the presence of altmetric data for Web of Science publications across subject fields and research topics. *Scientometrics*, 124(3), 2519–2549.

Fang, Z., Dudek, J., & Costas, R. (2020). The stability of Twitter metrics: A study on unavailable Twitter mentions of scientific publications. *Journal of the Association for Information Science and Technology*, 71(12), 1455-1469.

Fang, Z., & Costas, R. (2020). Studying the accumulation velocity of altmetric data tracked by Altmetric.com. *Scientometrics*, 123(2), 1077-1101.

International peer-reviewed conference papers:

Fang, Z., Costas, R., & Wouters, P. (2021). User engagement with scholarly Twitter mentions: A large-scale and cross-disciplinary analysis. *The 18th International Conference of the International Society for Scientometrics and Informetrics (ISSI 2021)*, Jul 12-15, Leuven, Belgium.

Fang, Z., Dudek, J., Noyons, E., & Costas, R. (2020). Science cited in policy documents: Evidence from the Overton database. *The 2020 Altmetrics Workshop*, Nov 17, online conference.

Fang, Z., Costas, R., Tian, W., Wang, X., & Wouters, P. (2020). How is science clicked on Twitter? Click metrics for Bitly short links to scientific publications. *The 7:AM Altmetrics Conference*, Oct 21 & 28, online conference.

Li, K., Chen, P., & **Fang, Z.** (2019, poster). Disciplinarity of software papers: A preliminary analysis. *Annual Conference of the American Society for Information Science and Technology (ASIS&T 2019)*, Oct 19-23, Melbourne, Australia.

Fang, Z., Dudek, J., & Costas, R. (2019). The stability of Twitter metrics: A study on unavailable Twitter mentions of scientific publications. *The 2019 Altmetrics Workshop*, Oct 11, Stirling, the UK.

Fang, Z., Lamers, W., & Costas, R. (2019). Studying the scientific mobility and international collaboration funded by the China Scholarship Council. *The 17th International Conference of the International Society for Scientometrics and Informetrics (ISSI 2019)*, Sep 2-5, Rome, Italy.

Fang, Z., & Costas, R. (2018). Studying the posts accumulation patterns of various Altmetric.com data sources. *The 2018 Altmetrics Workshop*, Sep 25, London, the UK.

Fang, Z., & Costas, R. (2018). Studying the Velocity Index for various Altmetric.com data sources. *The 23rd International Conference on Science and Technology Indicators (STI 2018)*, Sep 12-14, Leiden, the Netherlands.

Acknowledgments

It has been a long journey for me to get here, temporally, spatially, and epistemically. What has supported me through this bittersweet journey has been the encouragement and help of many people. I would like to express my gratitude to them.

I cannot be more grateful for the supervision of Prof. dr. Paul Wouters and Dr. Rodrigo Costas. I feel extremely fortunate to have them as my PhD supervisors. It was an unforgettable day when Paul and Rodrigo offered me the opportunity to start my PhD research at CWTS four years ago. They are always there to provide insightful feedback and valuable advice that inspired me on how to conduct scientific research and, more importantly, how to deal with unexpected challenges in my scientific career. This dissertation could not have come out without their expertise and patience in guiding me to become an independent researcher. I am also very grateful for the encouragement and support from my Master's supervisor, Prof. dr. Xianwen Wang. He opened the doors of scientometrics for me and led me into the research field of altmetrics.

I treasure the past four years I spent in the Netherlands because I was always surrounded by a lovely group of colleagues and friends. Special thanks to my best friends Wanbin and Juan Pablo for accompanying me to the defense as my paranymphs. Particular thanks to Zohreh, Jonathan, Felipe, Petra, Teresa, and Ali for their very kind help and companionship when I just arrived in the Netherlands. Thanks to my friendly neighbor Koen for helping me check the translation of Dutch summary. Many thanks to Sarah, Ludo, Ed, Nees Jan, Thed, Vincent, Clara, Alfredo, Martijn, Karin, Inge, Lidia, Henri, Patrick, Bram, Wout, Tung-Tung, Qianqian, Kai, Haiyun, Andrea, Judit, Jochem, Carole, Ingeborg, Sonia, André, Soohong, Eleonora, Huizhen, Xiaozan for their kind assistance and comfort when I was facing challenges.

Last but not least, I would like to thank my parents for their generous love and unconditional support. Due to the ongoing pandemic, I was only able to get together with them twice during the four-year PhD project. Eventually:

“陌上花开，可缓缓归矣。”