Uncovering the potential of Twitch as a source for social media metrics
by Enrique Orduña-Malea and Carlos Lopezosa

Abstract
The social live streaming service Twitch was launched in 2008 as Justin.tv, rebranded as Twitch Interactive in 2011, and acquired by Amazon in 2014. Although launched originally as a portal to broadcast videogames, Twitch currently hosts a wide range of content, including science and technology channels. Yet, despite growing interest in this online video sharing platform, Twitch’s potential for the study of science videos has been underexploited to date. This paper seeks to go some way to remedying this by studying the potential of Twitch as a data source for social media academic metrics. To do so, a scientometrics-inspired framework (the OBA framework) is proposed to integrate the analysis of Twitch, science videos and research organizations under a common conceptual space. Then, a science-related Twitch channel — National Aeronautics and Space Administration (NASA) — is used as a case study. We analyse 197 videos published by NASA between March 2017 and December 2022, as well as 51,935 clips created from NASA videos. Data were collected from the official Twitch API, which is also analysed to identify the units and metrics available and the channel’s performance in retrospective quantitative studies (i.e., non-live broadcasts). The results show that Twitch allows in-depth metric analyses of science videos to be undertaken, facilitating identification of both the activity and output-level impact of a scientific organization such as NASA. However, the Twitch API presents a few constraints, due, in the main, to the limited availability of many metrics that are restricted in time range, quantity, accuracy, or access, and which as such limit comprehensive retrospective studies. Despite these technical limitations, it is estimated that Twitch offers considerable potential for the study of science-related activity. The OBA model proposed facilitates the analysis of the activity of specific scientific agents (not only organizations but journals or other aggregates) under a conceptual framework based on approaches applied in quantitative studies of science.

Contents

1. Introduction
2. Background
3. Methods
4. Results
5. Discussion
6. Conclusions
The “altmetrics movement” facilitated the measurement of a composite trace of the impact of scholarly publications captured from a diverse group of online research-related activities and products (Piwowar, 2013; Priem, et al., 2011; Adie, 2014), including scholarly blogs, policy citations, mass media, reference managers, and social networking sites (Moed, 2016). Since its formal inception (Priem and Hemminger, 2010), a large body of literature has been accumulated (Sugimoto, et al., 2017; Olmeda-Gómez and Perianes-Rodríguez, 2019).

The “social media” component of altmetrics works is mainly based on the analysis of those online research objects representing scholarly works (namely a DOI-based URL or other similar URL-based publication ID) that have been embedded a social networking platform publication (e.g. Facebook posts, Twitter microposts), obtaining a series of metrics and indicators, which have constituted the so-called social media metrics (Haustein, et al., 2015; Thelwall and Kousha, 2015; Díaz-Faes, et al., 2019; Wouters, et al., 2019).

However, the ecosystem of social media “academic” metrics goes beyond the metrics related to the URL of scholarly works (López-Pérez and Olvera-Lobo, 2016). On the one hand, the embedded URL can represent aggregates, such as journals (e.g., pnas.org), universities (e.g., harvard.edu), or other research organizations (e.g., csic.es). In these cases, the material published by these entities will cover not only scholarly works but other informative, legal, or even promotional content, thus remaining outside the scope of altmetrics. Cybermetrics is the field that has generally undertaken the study of these academic URLs (Thelwall, 2012). On the other hand, other research-related objects can be embedded such as images and videos, which can be of great value for studying the dissemination and consumption of information published by publishers or research organizations, situating these analyses on the frontiers between the quantitative studies of science, informetrics, communication, and institutional research.

This study places social media metrics in a broader conceptual and scientometrics-inspired framework, based on the establishment of three elements: objects, social media publication blocks, and social media accounts (the OBA framework). This model is compatible with the social acts leading to online events framework proposed by Haustein, et al. (2016), based on research objects, documents, and agents.

The OBA framework (Figure 1) considers the existence of a diverse range of online elements, including files and URLs. These elements are subsequently embedded in a publication block (e.g., a post on Facebook or a tweet on Twitter) by the responsible (i.e., an agent) of the account, which is created in one specific social networking site (i.e., the source).

Each publication can include several elements, being one of them the principal (e.g., YouTube’s main element is a video). Both the account, the publication and the element can attract metrics derived from social acts carried out by other agents, which in turn can be other social accounts or audience without a registered account.
Uncovering the potential of Twitch as a source for social media metrics

The understanding of social media metrics is derived from the social networking site operating mechanism, the analytical framework (metrics available, and functionalities to collect them), and the online behaviour of the agents linked to said social platform. In this way, the OBA framework can be used to measure the activity of academic agents (e.g., journals or research organizations) that operate as aggregates in social networks, using a conceptual scientometric-inspired framework.

To illustrate the OBA framework, the objective of this work focuses on testing the analytical framework of one specific social networking site (Twitch) to study the production and the output-level impact of specific objects (science videos) published by a research organization (NASA).

The rest of this study is structured as follows. Next, the Background section describes the importance of science videos in social media academic metrics studies, and details the role played by Twitch as a live streaming platform, describing its basic operation and data structure, as well as providing a review of related literature. The Methods section describes the procedures followed in the study (data collection and data analysis), while the Results section presents the main findings obtained. These are subsequently commented on in the Discussion section and the paper’s Conclusion considers the study’s main contributions and future lines of research.

2. Background

2.1. Science videos as social media objects
The use of online science videos (either pre-recorded or live streamed) has emerged as an especially engaging way to make research content more accessible, allowing direct, global access to science-related content, as well as facilitating a more interactive dialogue between science producers, mediators and science consumers (Smith, 2018; Boy, et al., 2020; Striewski, et al., 2022; Shaikh, et al., 2023).

As a global, online video-sharing service (primarily oriented to pre-recorded content), the use of YouTube as a platform for the dissemination of science has increased notably (Amarasekara and Grant, 2019; Shaikh, et al., 2023; Yang, et al., 2022). Its international and inter-generational audiences (Thelwall, 2018), diverse content (Striewski, et al., 2022), social media features (Tsou, et al., 2014; Allgaier, 2020; Shaikh, et al., 2023), permanent access (Elberzhager, et al., 2022) and reach (Maynard, 2021) mean YouTube is a suitable tool for social media academic metrics, or “popularity” metrics (Yang, et al., 2022).

For this motive, YouTube has been extensively used as a data source, covering such topics as the specific features of science videos (Morcillo, et al., 2016; Welbourne and Grant, 2016), the output-level impact of science videos (Sugimoto and Thelwall, 2013; Yang, et al., 2022), the characteristics of science video presenters (Sugimoto, et al., 2013), analyses of science video comments (Tsou, et al., 2014; Striewski, et al., 2022), gender studies (Thelwall and Mas-Bleda, 2018; Amarasekara and Grant, 2019), categories of science videos (Kousha, et al., 2012), science video audiences (Boy, et al., 2020), best practices in making science videos (Beautemps and Bresges, 2021; Elberzhager, et al., 2022) or promoting research projects (Davies, 2019), and discussions centred on controversial issues, such as fracking (Jaspal, et al., 2014), genomics (Harris, et al., 2014), vaccination (Donzelli, et al., 2018), climate change (Shapiro and Park, 2015; Allgaier, 2019) and the COVID-19 pandemic (Orduña-Malea, et al., 2020). In addition, the use of YouTube by research organizations, especially universities (Martin-González and Llarena, 2017; Meseguer-Martínez, et al., 2019; Ros-Gálvez, et al., 2021).

Aside from YouTube, Twitch has emerged as another leading online video sharing platform, albeit in this case oriented primarily to live streaming videos. Although launched in 2011, Twitch did not gain any real momentum until 2020, coinciding with the COVID-19 lockdown (Leith and Gheen, 2022). This increase in user numbers attracted the interest of the scientific community (Törhönen, et al., 2021) and, subsequently, triggered a growth in scientific studies analysing the platform itself (see section 2.3). However, studies focusing specifically on analysing science-related videos, channels, or streamers remain few and far between, and no studies related to research organizations have been located.

### 2.2. Twitch as a source of social media metrics

#### Origins and basic operation

Twitch is a social live streaming service launched in 2008 as Justin.tv, a beta product specialising in electronic sports or esports that allowed any user to broadcast live videos, thus enabling synchronous communication between the broadcaster and the audience. Due to the rising popularity of the portal and some concerns over pirate streams, the Justin.tv matrix rebranded as Twitch Interactive in 2011 (Partin, 2019), which was subsequently acquired in its entirety by Amazon in 2014.

In 2013, following the closure of Own3d.tv, its direct competitor, Twitch became the most popular streaming portal for esports. The following year, it ranked as the fourth largest source of peak Internet traffic in the United States [1], its viewers climbing from 208,000 in 2013 to 351,000 in 2014 and reaching 539,000 in 2015 [2]. As of November 2022, Twitch was receiving 1.1 billion visits (i.e. the sum of all visits on desktop and mobile to the Twitch Web site) [3] with 7.6 million unique creators streaming each month and an average of 2.6 million concurrent viewers (i.e., number of users watching a live video at one time).

Although launched as a portal to broadcast videogames, Twitch currently hosts a wide variety of content, grouped into five basic communities: Games, IRL (In Real Life), Music, Esports and Creative. The Games community allows the user to consume live streams of different video games of all genres; the IRL community offers talks from content creators via streaming on a wide variety of topics, but mainly sport, fitness, travel, and science; the Music community allows the viewer to explore different streaming channels
Uncovering the potential of Twitch as a source for social media metrics

by music genre and to view live performances and special events and to learn about music production; the Esports community offers the possibility of watching live professional videogame tournaments; and, finally, the Creative community is designed for artists, crafters, and builders who use Twitch to broadcast their activities, including painting, cooking, and software development.

Users can adopt one of two roles: streamer (broadcaster) or viewer (user), and then interact with each other via chat. Broadcasters create channels (i.e., a profile) via which they broadcast their content (i.e., the streams). Twitch’s streamers are classified as normal, affiliate, or partner broadcasters.

To facilitate content findability, streams are labelled with specific tags and grouped into categories according to the type of content livestreamed. Each category can be found on tailored landing pages that include, among others, the following subsections: live channels (access to channels that are broadcasting live), videos (a list of streams which were recorded and recently uploaded to Twitch) and clips (short videos used by streamers to showcase uniquely Twitch moments from streams). Finally, each category is linked to one (or more) main Twitch community.

The specific channels also have tailored landing pages, which basically include the following sections: home (including recent broadcasts), about (basic descriptive information), a chat, a schedule of streams, and the videos available, which include on-demand videos (VODs), highlights (highlight reels of past streams), and uploaded videos (external videos that the broadcaster uploaded using the Twitch video producer).

Viewers can access a channel and watch the active stream, chat with the streamer and the other viewers, follow (i.e., free mode which allows them to receive alerts) or subscribe (i.e., paid mode which allows them to enjoy premium advantages) to specific channels or categories.

Twitch allows streamers to monetize their participation. Thus, users can purchase “bits” or coins that they can donate to streamers who have this possibility enabled on their channels. Donations are made through “cheers”, that is, chat messages that can include animated emoticons or cheermotes. Each cheer delivers a predetermined number of bits to the streamer. The number of bits that a user can donate with a particular cheermote is predefined by the tier level that each cheermote supports.

Each channel has the option of making different types of cheermote available: global first-party (Twitch-defined cheermotes that are shown in the streamer’s bits card), global third-party (Twitch-defined cheermotes that are not shown in the streamer’s bits card), channel custom (broadcaster-defined cheermotes), sponsored (sponsor-defined cheermotes; when used, the sponsor adds additional bits to the amount that the user cheered), and “display only” cheermotes (i.e., for internal use only).

Analytic framework

Twitch provides a free application programming interface (API) for developers to facilitate the development of Twitch integrations. Thanks to this API, several external data providers have launched online portals offering stats, most notably TwitchTracker, TwitchStats, or Social Blade–Twitch.

The API offers around 60 different endpoints oriented to data collection (GET type). After filtering out the merely technical endpoints (e.g., “Get Stream Markers”, oriented to the marking of points in a live broadcast), we proceeded to review all the attributes that each endpoint provides the user when making a query, grouping these attributes into different units of measurement (e.g., broadcaster, channel, stream, category, etc.). This process yielded 26 units of measurement and 208 metrics. All the units and metrics identified, as well as the type of data provided and the name of the endpoint from which each metric can be obtained, are listed in the supplementary material. We then proceeded to connect all the measurement units in order to derive a “metric portrait” of Twitch (Figure 2).
Uncovering the potential of Twitch as a source for social media metrics

Literature review

Despite the growing interest in Twitch, no comprehensive bibliographic review has been conducted to date that might help identify the main features of this streaming video platform as examined by the scientific literature. To address this shortcoming, the term query “twitch AND streaming” was used to search for documents in the Web of Science Core Collection (title, abstract, keywords plus), Scopus (title, abstract and keywords), Dimensions (title, abstract), and Google Scholar (title), the four leading bibliographic databases. All queries were submitted on 4 December 2022, generating a total of 1,044 records (242 from Web of Science Core Collection, 204 from Dimensions, 304 from Scopus, and 204 from Google Scholar).

A data cleaning process was subsequently performed to filter out non-peer reviewed publications, including news posts (e.g., from Business Insider, CNN, and Washington Post), Bachelor’s dissertations and Master’s theses. Duplicates and false positives were also removed. This process left a total of 449 publications. The distribution of these documents by year of publication is shown in Figure 3, where the increased interest in Twitch as an object of study is especially evident from 2017 onwards.
Uncovering the potential of Twitch as a source for social media metrics

Figure 3: Number of publications about Twitch over time.
Note: error bars (±5%) are included to consider potential errors in the year of the publication metadata field as well as data processing human errors.

The full-text of each publication was obtained through institutional repositories and full-text bibliographic databases, access to which was provided by the authors’ institutions. This process yielded a total of 399 publications.

It was decided to limit the analysis to those languages in which the authors were fluent. Thus, only publications written in Spanish and English were included. This reduced the number of publications in the final dataset to 324 (that is, 79.2 percent of the 429 publications obtained after data cleaning).

Our initial results showed that the main research disciplines interested in Twitch can be divided into two distinct camps: those taking a social approach and those taking an eminently technical approach. In the case of the former, we find studies related to the influence of Twitch on the gaming industry and media landscape, examining, for example, business and marketing interests (Johnson and Woodcock, 2019; Lujia, et al., 2020; Tang and Huang, 2021; Mao, 2022), the labour market (Woodcock and Johnson, 2019), new social dynamics and user behaviour (Kaytoue, et al., 2012; Sjöblom and Hamari, 2017), gender (Ruvalcaba, et al., 2020; Zolides, 2021), race (Gray, 2017) and participatory culture (Konstantinova, et al., 2018). In the case of the latter, we find contributions related to the design and testing of recommendation systems, algorithms, video streaming, video quality enhancement and or multimedia systems, among others (e.g., Pires and Simon, 2014; Bilal, et al., 2017; Göring, et al., 2019; Ray, et al., 2019).

**Twitch and science videos**

Twitch can also be used as a source for undertaking research about science videos (Ather, 2019; Buitrago, et al., 2021). Indeed, a category designed for streaming science-related content (Science & Technology, part of the IRL main community) has been created [9], attracting 709,000 followers as of March 2023. Live online science videos allow agents (e.g., science communicators, research organizations) to establish a
natural, more direct connection with their audience, often transforming their broadcasts into live, interactive online classes, with potential learning benefits (Payne, et al., 2017; Steinbeck, et al., 2021; Gandolfi, et al., 2022).

However, despite the increasing body of literature centred on Twitch and the features of this platform, its use for science videos has received little attention from the scientific community. One exception is Buitrago and Ortiz (2022), who carried out a qualitative analysis of the streams created by leading Spanish scientific popularisers, focusing their attention on the discursive strategies they use for disseminating scientific content in online videos. Thus, the authors observed the use of rhetorical figures, humour, citations of primary sources, and expert interviews as specific mechanisms to stream scientific content covering everyday issues but with an emphasis on the empirical sciences and physics (Buitrago and Ortiz, 2022).

3. Methods

To carry out the case study, we opted to select a specific channel — the National Aeronautics and Space Administration (NASA) — whose videos are streamed in Twitch’s Science & Technology category.

NASA is the independent agency of the U.S. federal government responsible for the civil space program, aeronautics research, and space research. To date, it has co-authored around 231,000 publications indexed in Scopus. The official NASA channel has more than one million followers, and broadcasts videos that are not only of great scientific and technological interest, but also have the capacity to generate interest among the public, contributing in this way to popularising science. According to TwitchStats, in 2022, NASA was the Twitch channel with the highest peak viewer count while streaming in the Science & Technology category, highlighting its relevance in the science category and its suitability for the objectives of the current study.

The “Get User” endpoint was employed to obtain the NASA user-id from the channel’s display name (“nasa”). Then, different scripts written in python (available in supplementary material, at https://doi.org/10.4995/Dataset/10251/201304) were used to collect data from various endpoints: “Get Channel Information” (basic data about the NASA channel), “Get Videos” (all the videos created by NASA), “Get Clips” (all clips created from NASA videos), “Get Users Follows” (all NASA followers), and “Get Cheermotes” (all the cheermotes available on the NASA channel). Finally, the “Get User” endpoint was also used to collect data about NASA’s followers (e.g., the number of followers that each NASA follower has). All these endpoints constitute the main points of public access to a channel’s data, without their being the broadcaster.

All data were collected during the first two weeks of February 2023, covering all of the channel’s activity from its creation to 31 January 2022. Data were obtained in JSON format and subsequently transformed to CSV using OpenRefine, and then statistically analysed using XLStat.

4. Results

4.1. Analysis of videos

NASA’s channel on Twitch has the status of partner broadcaster. The channel published 197 videos between March 2017 and December 2022, a total of 486 broadcast hours attracting a total of 17,561,591 views. Quantitative statistics on the videos, views, and followers of NASA’s Twitch channel since the beginning of 2017 are shown in Figure 4.
Uncovering the potential of Twitch as a source for social media metrics
Publication frequency has been quite irregular over the years (Figure 3, top), with 2018 being the most productive (57 videos) and 2019 the least (8 videos). The years 2020 and 2021, however, also stand out in terms of their productivity (with 33 and 40 videos, respectively). Publications are predominantly on-demand videos (VODs) of past streams (186; 94.4 percent), while highlights (6) and uploaded videos (5) are uncommon.

The videos have an average duration of 2 h 28 min, and include 117 videos (59.4 percent) with a duration that exceeds one hour, the latter being precisely the videos that record, on average, the greatest output-level impact in terms of the number of views received (Table 1). The video with the longest duration (around 27 h) is “Making History: NASA and SpaceX Launch Astronauts to Space!”, broadcast on 30 May 2020 and which attracted 105,151 views (i.e., the number of times that users have watched the video).

**Table 1: Video duration and number of views received.**

<table>
<thead>
<tr>
<th>Duration (seconds)</th>
<th>Video count</th>
<th>Percentage</th>
<th>Seconds broadcast</th>
<th>View count</th>
<th>Percentage</th>
<th>Views per video</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 60</td>
<td>7</td>
<td>3.6</td>
<td>255</td>
<td>145,573</td>
<td>0.8</td>
<td>20,796</td>
</tr>
<tr>
<td>60 ≤ x &lt; 600</td>
<td>15</td>
<td>7.6</td>
<td>3,359</td>
<td>103,046</td>
<td>0.6</td>
<td>6,870</td>
</tr>
<tr>
<td>600 ≤ x &lt; 1,800</td>
<td>22</td>
<td>11.2</td>
<td>26,966</td>
<td>355,803</td>
<td>2.0</td>
<td>16,173</td>
</tr>
<tr>
<td>1,800 ≤ x &lt; 3,600</td>
<td>36</td>
<td>18.3</td>
<td>99,302</td>
<td>1,496,544</td>
<td>8.5</td>
<td>41,571</td>
</tr>
<tr>
<td>3,600 ≤ x &lt; 7,200</td>
<td>54</td>
<td>27.4</td>
<td>270,480</td>
<td>3,883,399</td>
<td>22.1</td>
<td>71,915</td>
</tr>
<tr>
<td>≥ 7,200</td>
<td>63</td>
<td>32.0</td>
<td>1,350,238</td>
<td>11,577,226</td>
<td>65.9</td>
<td>183,765</td>
</tr>
<tr>
<td><strong>All</strong></td>
<td><strong>197</strong></td>
<td><strong>100</strong></td>
<td><strong>1,750,600</strong></td>
<td><strong>17,561,591</strong></td>
<td><strong>100</strong></td>
<td><strong>89,145</strong></td>
</tr>
</tbody>
</table>

**Views (block-level metrics)**
While the first 15 videos published by NASA on Twitch did not attract much attention (489 views in total), the 16th video, published in August 2017 (24 h long and the second longest video in the dataset), marked the channel’s take-off at a communicative level, recording 79,456 views.

The audience is concentrated primarily between June 2020 and June 2021 (Figure 3, middle), with 72.1 percent of all views received by the channel being recorded in that 13-month period. This period coincides with the broadcasting of a series of videos showing the landing of Perseverance (a car-sized rover) on planet Mars in February 2021, as part of NASA’s Mars 2020 mission. The most viewed video broadcast by NASA, “Watch the Perseverance Mars Rover Landing” (1,579,273 views; 1 h 59 m 48 s), was published in that period (18 February 2021).

Overall, the channel has an average of 89,145 views per video (median = 5,138). The distribution of these views is highly dispersed (SD = 224,987) and slightly negatively skewed (Bowley’s coefficient of skewness = 0.86). Indeed, we find a few highly viewed videos (four videos attracting more than one million views each) and many poorly viewed videos (27 receiving fewer than 100 views each). Otherwise, the view counts exhibit an h-index of 160 (i.e., there are 160 videos with at least 160 views each).

**Followers (profile-level metrics)**

NASA has attracted 1,204,523 followers as of 31 December 2022. Of these, we were able to extract data (user type and creation date) from 1,197,598 users (99.4 percent), of which 91.3 percent are normal broadcasters, 8.4 percent are affiliates, and 0.4 percent are partners. A further 300 followers are Twitch staff users.

It is evident that the months with the highest increase in the number of followers coincide with the broadcast of videos with a high number of views (Figure 3, bottom). For example, in May 2020 the channel attracted 375,513 new followers, coinciding with the broadcasting of “Making History: NASA and SpaceX Launch Astronauts to Space!” (105,151 views). Likewise, in February 2021, the channel attracted 166,139 new followers, coinciding with the broadcasting of two highly viewed videos (1,579,273 and 1,235,509 views, respectively). Yet, another highly viewed video (458,678 views) broadcast in November 2022 is associated with just 6,724 new followers that month. This might indicate that the channel is reaching a saturation point in terms of followers and that new videos are no longer capable of generating new followers.

Interestingly, 57.5 percent of the channel’s followers created their Twitch profiles at least two years before they started following the NASA channel, that is, they were well-established Twitch users, while 1.5 percent (18,310) of followers began to follow NASA the same day they created their Twitch account (Table 2), possibly related to the broadcasting of a relevant stream. Indeed, further examination shows that 78.1 percent (935,575) of the channel’s followers began to follow the NASA channel the same day a new video was broadcast.

<table>
<thead>
<tr>
<th>Days from creation to follow</th>
<th>Scope</th>
<th>Users</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Same day</td>
<td>18,310</td>
<td>1.5</td>
</tr>
<tr>
<td>1 to 7</td>
<td>First week</td>
<td>14,548</td>
<td>1.2</td>
</tr>
<tr>
<td>8 to 30</td>
<td>First month</td>
<td>30,571</td>
<td>2.6</td>
</tr>
</tbody>
</table>
Uncovering the potential of Twitch as a source for social media metrics

<table>
<thead>
<tr>
<th></th>
<th>First semester</th>
<th>142,689</th>
<th>11.9</th>
</tr>
</thead>
<tbody>
<tr>
<td>31 to 180</td>
<td>First year</td>
<td>127,156</td>
<td>10.6</td>
</tr>
<tr>
<td>181 to 365</td>
<td>First 2 years</td>
<td>176,181</td>
<td>14.7</td>
</tr>
<tr>
<td>366 to 730</td>
<td>From 2 to 5 years</td>
<td>413,749</td>
<td>34.5</td>
</tr>
<tr>
<td>731 to 1,825</td>
<td>More than 5 years</td>
<td>274,394</td>
<td>22.9</td>
</tr>
</tbody>
</table>

To ascertain the relevance of NASA channel followers, the number of followers of each of the followers was extracted. Results show that 33.7 percent of all NASA followers have zero followers, suggesting they are passive Twitch users, without any activity as streamers. However, 9,119 NASA followers have obtained at least 1,000 followers each, highlighting their quite remarkable visibility, while 37 NASA followers have more than one million followers each ([Figure 5](#)), indicative of their remarkable output-level impact on the platform.

![Figure 5: Distribution of the number of followers obtained by NASA followers.](#)

*Note: Larger version of Figure 5 available [here](#).*

### 4.2. Analysis of clips

A total of 51,935 clips have been created from 166 videos (84.3 percent of the videos published by NASA), providing a total of 2,666,981 views. The average number of views per clip is, however, low (mean = 51; median = 1), the distribution of views being dispersed ($SD = 2,219.5$), peaked (kurtosis [Fisher] = 12,130.8) and skewed (Bowley’s coefficient of skewness = 1), with an h-index of 85 (i.e., there are 85 clips receiving at least 85 views each). This uneven distribution is exemplified by the fact that 29,681 clips (57.2
percent) received just one view each, while four clips have obtained more than 100,000 views each (max value = 333,353 views).

A total of 39,328 unique users have been identified as clip creators, of which 23,493 (59.7 percent) are also followers of the NASA channel. This means that 15,835 (40.3 percent) users created clips from NASA streams without being followers of the channel. Most clip creators 32,523 (82.7 percent) have only created one clip, while a small number of users have created many clips (e.g., 31 users have created at least 20 clips each). Yet, the most productive clip creators are not the ones that have attracted the largest audiences, as observed in Figure 6, where these productive creators obtain fewer total views than users who have created a few, or even one, clip.

![Figure 6: Scatterplot of the number of clips created by a user and the number of views received by these clips.](image)

Note: Larger version of Figure 6 available [here](#).

Table 3 ranks the videos according to the number of clips they have generated. As can be seen, “Making History: NASA and SpaceX Launch Astronauts to Space!” (ID 636407963) is the video with most clips (15,182). Of these, the most viewed clip obtained 110,190 views, while the total number of views received by all these clips amounts to 328,652, a number that surpasses that of the views received by the original video (105,151).

<table>
<thead>
<tr>
<th>Video ID</th>
<th>Number of clips</th>
<th>View count</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>636407963</td>
<td>15,182</td>
<td>328,652</td>
<td>110,190</td>
</tr>
<tr>
<td>168717817</td>
<td>3,814</td>
<td>115,009</td>
<td>60,975</td>
</tr>
</tbody>
</table>
Uncovering the potential of Twitch as a source for social media metrics

<table>
<thead>
<tr>
<th>Video ID</th>
<th>Clips (24h)</th>
<th>Views (24h)</th>
<th>Output Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>919215638</td>
<td>2,499</td>
<td>49,417</td>
<td>39,123</td>
</tr>
<tr>
<td>997654761</td>
<td>2,160</td>
<td>21,395</td>
<td>13,210</td>
</tr>
<tr>
<td>937725331</td>
<td>2,027</td>
<td>50,950</td>
<td>40,355</td>
</tr>
<tr>
<td>698106525</td>
<td>1,775</td>
<td>102,591</td>
<td>85,995</td>
</tr>
<tr>
<td>633533842</td>
<td>1,751</td>
<td>86,035</td>
<td>57,025</td>
</tr>
<tr>
<td>924423369</td>
<td>1,279</td>
<td>111,588</td>
<td>75,630</td>
</tr>
<tr>
<td>168479949</td>
<td>1,251</td>
<td>42,023</td>
<td>20,030</td>
</tr>
<tr>
<td>1609918806</td>
<td>1,022</td>
<td>50,157</td>
<td>31,978</td>
</tr>
<tr>
<td>998979775</td>
<td>1,003</td>
<td>57,624</td>
<td>41,140</td>
</tr>
</tbody>
</table>

Taking video ID 636407963 as a case study, Figure 7 shows the chronology of all the clips generated from it. The first clip was published just 57 seconds after the publication of the original video, while a further 12,817 clips (84.2 percent) were created within the first 24 hours. After that, sporadic bursts are observed until the number of clips created and their output-level impact decrease over time. The last clips were created on 23 April 2021.

Figure 7: Evolution of the number of clips created from a video and their number of views.
Note 1: The figure covers only the first two days following the publication of the original video.
Note 2: The original video corresponds to video ID 636407963.
Note: Larger version of Figure 7 available [here](#).
The degree of influence of the creator of each clip, and the fact that each clip is likely to contain different fragments of the original video, is likely to explain the variation in output-level impact of the clips generated from the same video. Additionally, it is evident that the creators published the clips in different categories from that in which the original video was published (Science & Technology), thus disseminating the clips among different communities. Table 4 shows the number of clips generated from video ID 636407963 according to the category in which each clip was published. This highlights the fact that a high number of clips appeared in “Talk Shows & Podcasts” (38.6 percent), as well as the particularly high output-level impact of clips published in “Reality” (417.6 views per clip).

<table>
<thead>
<tr>
<th>Category name</th>
<th>Number of clips</th>
<th>Percentage</th>
<th>Views per clip</th>
</tr>
</thead>
<tbody>
<tr>
<td>Science &amp; Technology</td>
<td>24,050</td>
<td>46.3</td>
<td>41.6</td>
</tr>
<tr>
<td>Talk Shows &amp; Podcasts</td>
<td>20,042</td>
<td>38.6</td>
<td>33.4</td>
</tr>
<tr>
<td>Exploration</td>
<td>5,661</td>
<td>10.9</td>
<td>27.9</td>
</tr>
<tr>
<td>Reality</td>
<td>1,628</td>
<td>3.1</td>
<td>417.6</td>
</tr>
<tr>
<td>Just Chatting</td>
<td>245</td>
<td>0.5</td>
<td>8.5</td>
</tr>
<tr>
<td>Special Events</td>
<td>96</td>
<td>0.2</td>
<td>9.7</td>
</tr>
<tr>
<td>N/A</td>
<td>213</td>
<td>0.4</td>
<td>731.3</td>
</tr>
<tr>
<td>All</td>
<td>51,935</td>
<td>100</td>
<td>51.4</td>
</tr>
</tbody>
</table>

5. Discussion

The main contributions made by this study are, first, the proposal of a conceptual scientometric-inspired framework to broaden the study of social media metrics, second, the study of the Twitch’s analytical framework by the novel identification of the units of measurement and metrics offered by the Twitch API (section 2.2) and, third, the insights provided by the case study of the NASA channel on Twitch (section 4), as a research organization publishing science videos.

In the first place, the OBA framework is consistent with and integrates the “social acts” framework proposed by Haustein, et al. (2016), and follows other frameworks aimed at studying social media metrics following scientometrics concepts (Orduña-Malea and Costas, 2023). In addition, the potential relationship between objects (e.g., a science video and a URL-based DOI) aligns the OBA framework with the heterogeneous couplings concept (Costas et al., 2021), an aspect to be further developed in future studies.
Uncovering the potential of Twitch as a source for social media metrics

The consideration of videos and clips as cited objects, of views count as metrics at the block level, of followers count as metrics at the profile level, as well as of the different agents (both NASA as citing author and followers as triggers of social acts) allow for incorporating the analyzes carried out within a common framework inspired by scientometric measures, facilitating the study of aggregates, in this case, a research organization.

In the second place, our analysis of the Twitch’s analytical framework is restricted to the GET endpoints provided by the Twitch API, and limited to the data that any Twitch user can access from external channels without being the channel’s administrator. The 26 units of measurement and 208 metrics identified, as well as the connections between these units (Figure 2) should be of particular use to meta-researchers wishing to evaluate the possibilities of Twitch as a source of social media metrics, facilitating the understanding of the available metrics and their generation by the users (i.e., agents), given the operating mechanism of the platform.

Twitch’s emphasis on streaming video, coupled with a business model based on the allocation of bits managed via chat and the breadth of its API (with over 60 GET type endpoints), make Twitch a complex platform, but one which presents a series of limitations when seeking to extract data massively. For example, the number of channels making up any given category cannot be obtained directly, since the corresponding endpoint (search channels) returns only 1,000 results, limited, in this instance, to those channels that have broadcast in the last six months. Moreover, these channels cannot be filtered by category, as the system returns all the channels whose title or category begins with the search term. Similarly, the number of videos corresponding to a category or channel is limited to a maximum of 500 per day — a limit that virtually all the categories exceed — which prevents retrospective studies being conducted at the category level. Other metrics of interest, such as the number of subscribers, the number of polls (and votes received per poll) created by a channel or the number of users that connected to a channel’s chat room, are likewise restricted. Access to these metrics would, however, greatly facilitate the carrying out of multiple studies, both quantitative and qualitative.

In the third place, the case study undertaken here has allowed us to analyse the activity of a research organization (NASA) on Twitch and to learn more about the dissemination and output-level impact of the science videos broadcast via its channel.

While NASA’s activity (197 videos published) and output-level impact on Twitch is notable, it remains significantly less than that achieved by NASA on YouTube, where its main channel has received 921,932,379 views (compared to 17,561,591 on Twitch) and 11,000,000 subscribers (compared to 1,204,523 followers on Twitch). Moreover, NASA operates a further 13 channels on YouTube in addition to its main channel (e.g., @NASAgovVideo, @nasa_es, @NASAMarshall).

Although the YouTube account was created at a much earlier date (June 2008) than its Twitch account (March 2017), the greater outreach effort and output-level impact achieved on YouTube is more than evident. If we compare the number of views and followers obtained during 2021 on the two platforms (Table 5), the data clearly show that Twitch is a less relevant broadcast channel for NASA. In fact, if we compare NASA’s most viewed stream (i.e., “Watch the Perseverance Mars Rover Landing”), it has received 1,579,281 views as of 24 March 2023, whereas the same broadcast streamed on YouTube has attracted around 25 million views [11].

<table>
<thead>
<tr>
<th>Month</th>
<th>YouTube</th>
<th>Twitch</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5: Number of views and followers attained by NASA on YouTube and Twitch. Source: Social blade. Note: Number of subscribers (or followers) is the sum of new followers and lost followers.
Although NASA has used Twitch to broadcast full-length videos and to stream historical events (such as the landing of the Perseverance robot on Mars), the channel shows evident signs of the absence of strategies of optimization. For example, only eight of the 197 videos posted include a ‘Description’ text, which might facilitate the findability of the videos deposited.

Participation through cheermotes, moreover, is somewhat limited. NASA makes available just 35 cheermotes (including three “display only”, six “global third-party”, and 26 “global first-party”), and has created no “channel custom” or “sponsored” cheermotes. As a government research agency, NASA does not seek to profit from the channel, but this could be perceived as an indication of less engagement with users, who might provide bits symbolically or even participate in citizen science projects, taking advantage of the audience of their streams via young people, who might be interested in science.

All the videos are event streams; that is, the channel has not been used to broadcast other types of scientific video (informative explanations for the lay public, interviews, games, or challenges for young people, etc.), where some of these activities (though not all) are available on NASA’s YouTube channel.

On Twitch, NASA does not follow other users and, moreover, all its videos are published in English, despite a large community of followers in other languages. Here, Spanish speakers stand out, as evidenced by the creation of the NASA channel on YouTube in Spanish (NASA in Spanish; @nasa_es) which attracts 230,000 followers. Equally striking is the fact that the YouTube channel does not include the Twitch channel among NASA’s links on social media, when links to other social platforms do appear (i.e., Facebook, Instagram, Twitter, and Flickr).

NASA’s strategy in this regard seems to be indicative of the fact that science appears to be less popular on Twitch than it is on YouTube. Thus, “Science and Technology” is one of the most prominent video categories on YouTube (Yang, et al., 2022), containing many different types of science-related videos (Thelwall and Mas-Bleda, 2018), and, according to Channel Crawler [12], up to 386,360 channels as of March 2023; in contrast, the “Science & Technology” category on Twitch is a minority interest. According to Twitchstats, the number of Twitch channels streaming in this category stood at just 3,216 in 2022 (4,150 in 2020; 4,344 in 2021). To understand the reasons underpinning these differences, further studies are needed on the motivations and rewards of using Twitch (Gros, et al., 2017), especially in the field of science.

<table>
<thead>
<tr>
<th></th>
<th>Views</th>
<th>Subscribers</th>
<th>Views</th>
<th>Subscribers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan-21</td>
<td>15,702,632</td>
<td>140,000</td>
<td>52,502</td>
<td>670</td>
</tr>
<tr>
<td>Feb-21</td>
<td>70,014,600</td>
<td>790,000</td>
<td>3,201,674</td>
<td>178,749</td>
</tr>
<tr>
<td>Mar-21</td>
<td>17,260,774</td>
<td>140,000</td>
<td>895,445</td>
<td>140,867</td>
</tr>
<tr>
<td>Apr-21</td>
<td>22,395,260</td>
<td>160,000</td>
<td>607,959</td>
<td>67,739</td>
</tr>
<tr>
<td>May-21</td>
<td>15,417,300</td>
<td>100,000</td>
<td>141,540</td>
<td>16,392</td>
</tr>
<tr>
<td>Jun-21</td>
<td>8,272,691</td>
<td>60,000</td>
<td>1,917,984</td>
<td>114,017</td>
</tr>
<tr>
<td>Jul-21</td>
<td>6,743,460</td>
<td>40,000</td>
<td>7,074</td>
<td>-1,966</td>
</tr>
<tr>
<td>Aug-21</td>
<td>6,564,299</td>
<td>30,000</td>
<td>77,925</td>
<td>3,006</td>
</tr>
<tr>
<td>Sep-21</td>
<td>6,590,457</td>
<td>40,000</td>
<td>176,816</td>
<td>-1,236</td>
</tr>
<tr>
<td>Oct-21</td>
<td>6,223,308</td>
<td>30,000</td>
<td>31,643</td>
<td>-1,962</td>
</tr>
<tr>
<td>Nov-21</td>
<td>8,163,190</td>
<td>60,000</td>
<td>335,784</td>
<td>11,206</td>
</tr>
<tr>
<td>Dec-21</td>
<td>18,101,742</td>
<td>220,000</td>
<td>270,720</td>
<td>21,365</td>
</tr>
</tbody>
</table>
Uncovering the potential of Twitch as a source for social media metrics

Twitch’s metric features have been found to be wide ranging, with a varied number of endpoints where data can be freely collected. In this regard, clear parallels can be drawn between YouTube and Twitch in terms of the availability of metrics at both the user (numbers of followers obtained and videos published, and user features) and video levels (number of views and video features). However, major differences were also detected.

While YouTube places the emphasis on comments received asynchronously for each video, Twitch focuses on comments received synchronously (i.e., the chat). However, due to data sharing concerns, the Twitch API limits access to chat data to just the streamer and moderator. This means external researchers cannot extract historical chat data, and other data collecting techniques have to be used, such as Web scraping or external ad hoc applications. YouTube, which via its YouTube Live service competes with Twitch in the streaming sector, does provide a command (LiveChatMessages) in its Life Streaming API [13] for extracting comments from chats, while an event is being broadcast.

A further difference with respect to YouTube is the existence of metrics related to clips (i.e., videos made from a previously published video). Here, with the original video used as a baseline, data of great interest can be obtained about the dissemination and re-dissemination of a publication.

Finally, while the social media metrics analysed can be understood as “transparent markers of popularity and engagement” (Baym, 2013), their interpretation should be taken cautiously, as they face a series of shortcomings, such as their lack of representativeness (e.g., dependence on factors such as location, age, or education) or the effect of the platform’s algorithms or deceptive practices on users’ activities (Baym, 2013).

6. Conclusions

The productivity of the NASA channel and its output-level impact on Twitch is notable — 2.8 videos per month, 89,145 views per video on average, and 1,205,553 followers; yet, this output is not as great as that achieved by NASA on YouTube. This can probably be attributed to two factors: the lower degree of relevance attached to science channels on Twitch and the smaller amount of effort expended by NASA in the optimization of its Twitch channel. Moreover, we report that the channel’s output-level impact (i.e., new followers and views) is closely linked to its activity (i.e., the broadcasting of videos); thus, when no videos are streamed, the numbers of new followers and new views falls markedly. This might indicate that there is no consumption of videos on this channel once the stream has been broadcast, unlike on YouTube, which would account for the differences in video consumption between the two platforms.

Our analysis of clips verifies their importance as a means for disseminating videos once they have been broadcast. These clips, created by other users (some of whom are more influential than the original creator), exploiting different content (given that a creator selects the content to be extracted from the original stream) and rebroadcast on other channels (including those in categories other than Science & Technology), allow them to reach a larger audience. The case study undertaken here shows that the number of views of the various clips made from the same original video can be much higher than the number of views obtained by the original video itself. This maintains interest in the publication for a longer period, a situation that can be exploited both by NASA and by other agents.

Because of Twitch’s operating mechanism, it appears that the platform may be optimal for individual science disseminators; however, for organizations, groups, research centres, or even universities that do not seek to monetize their channel, it might mean they tend to engage less with users and the social dynamics of the service. This may well account for the low presence of research institutions on Twitch.

The Twitch API presents a few constraints, mainly attributable to the limited availability of certain metrics.
Uncovering the potential of Twitch as a source for social media metrics that are restricted in terms of their time range (e.g., the last six months only), quantity (e.g., no more than 500 results per day) or accuracy (e.g., unable to filter channels by category), which hinders comprehensive retrospective studies. Likewise, access to many metrics is limited to those responsible for the channels. However, it should be borne in mind that this mode of operation is no different from that of many other platforms, including Twitter Analytics and YouTube Analytics. But, although the API performs well and offers high-value data for free, the above limitations prevent scholars from exploiting the full potential of the platform. The availability of an API for academic use (free or otherwise) would be of great interest to the scientific community.

Finally, Twitch makes it possible to undertake an in-depth metric analysis of science videos, facilitating identification of the activity and output-level impact of a scientific organization like NASA through its streams on the platform. Future work needs to focus on establishing adequate methods to analyse the interactivity of users with the streams via the chat room during the broadcast, as this would allow us to discover additional aspects about how other agents (the public) interact with science videos and streamers. Additionally, a comprehensive analysis should be undertaken to identify and characterize all channels streaming in the “Science & Technology” category, to determine the scope and impact of scientific content on Twitch, the presence of research institutions and researchers, the type of videos generated, and their audience.

All of these future studies can be carried out following the OBA framework proposed in this paper, thus facilitating the analysis of the activity of specific scientific agents (not only organizations but journals or other aggregates) under a conceptual framework based on approaches applied in quantitative studies of science.

About the authors

Enrique Orduña-Malea is an associate professor in the Department of Audiovisual Communication, Documentation and History of Art at Universitat Politècnica de València in Valencia (Spain). 
E-mail: enorma [at] upv [dot] es

Carlos Lopezosa is a visiting researcher in the Departament d’Informació i Mitjans Audiovisuals at Universitat de Barcelona in Barcelona (Spain). 
E-mail: lopezosa [at] ub [dot] edu

Acknowledgements

This project has been supported by the project UNIVERSEO (Ref. GV/2021/141), funded by the regional government (Generalitat) of Valencia (Spain).

Notes


3. Data obtained from the professional analytics tool Similarweb (https://www.similarweb.com/website/twitch.tv).

4. Twitch is moving from Twitch-defined tags to channel-defined tags. As part of this move, Twitch is
Uncovering the potential of Twitch as a source for social media metrics

deprecating this endpoint and will remove it in 2023.

11. https://www.youtube.com/watch?v=gm0b_iJaYMQ.

References


Uncovering the potential of Twitch as a source for social media metrics


Uncovering the potential of Twitch as a source for social media metrics


doi: https://doi.org/10.24908/ss.v17i1/2.13021, accessed 30 December 2023.


doi: https://doi.org/10.1038/493159a, accessed 30 December 2023.


Uncovering the potential of Twitch as a source for social media metrics


Uncovering the potential of Twitch as a source for social media metrics

e62403.


