INTERNATIONAL SCIENTIFIC RESEARCH ON TWITCH.TV
Systematic Review of Academic and Scientific Output on the Livestreaming Platform

SIMEÓN DOMENECH-ESQUERDO [1], JESÚS SEGARRA-SAAVEDRA [1], TATIANA HIDALGO-MARÍ [1]
[1] University of Alicante, Spain

ABSTRACT

Twitch is a livestreaming platform that has gained popularity in recent years. This article presents a systematic review of the scientific literature found in WoS, Scopus, Proquest and Dialnet and published in journals. Both qualitative and quantitative analyses were carried out. Among the results, there is a notable increase in publications in 2022, a predominance of American authorship and article production, and a low interest in the communicative aspect (both commercial and non-commercial) of the platform.

1. Introduction and State of Play

Times change and with them society and everything that shapes it, such as people, Times change and with them society and everything that shapes it, such as people, education, values, etc. Among all these aspects, three could be highlighted: communication, technology and entertainment. All three are inextricably linked and interdependent. The most obvious example is the birth of the Internet, which changed the way we communicate and, later, the way we produce and consume entertainment. Similarly, advances in communication can stimulate the developments of technology, as has happened with the appearance of social networks and instant messaging applications.

Entertainment, on the other hand, is linked to both. As technology advances, new forms of entertainment have emerged, such as radio, cinema, television, video games, etc. All these media also have a very important communicative function, and the appearance of television and the Internet can be seen as relevant moments of change. Therefore, and without going through the history and chronology of the media, we can currently find the fusion of the aforementioned media. Television and the Internet are now almost inseparable, both in terms of content and media. An example of this is the television sets on the market that have or need an Internet connection to enjoy all the content that can be viewed through the electronic device. Much of this content is hosted on streaming platforms and applications, video on demand or live content that is available immediately and at any time.

One step beyond TVs are smartphones and computers. In addition to offering the same features as televisions, they give the user greater communicative power. Thanks to these electronic devices, it is possible to share and comment on content via messaging platforms or social networks. Most importantly, users have the power to create their own content, using their own technological means, and communicate it to whomever they choose. The content can range from simple comments in a forum to the production of live programmes as if they were made by professional television producers, with the added bonus that the audience can interact with the protagonists themselves. The traditional barrier between content creators and audiences is blurred until only physical distance remains as the only obstacle.

There is a proliferation of live streaming applications and platforms such as Facebook, Instagram, TikTok, Periscope and YouTube. However, due to its specificity in services and volume of users, Twitch is perhaps the most relevant platform.

Twitch.tv is a livestreaming platform that allows anyone with a good Internet connection, a camera and/or a medium-quality device (computer or mobile) to broadcast live video games and any other subject, including reactions, analyses, discussions and talks.

Authors such as Gutiérrez and Cuartero (2020) and Carrillo (2019) have already documented the beginnings of the platform. In 2007, Justin Kan, Emmet Shear, Michael Seibel and Kyle Vogt founded Justin.tv, a web platform that allowed users to broadcast and share content on any topic. Over time, users began broadcasting video game content and sharing sporting events, which led to an increase in the number of users. As a result of its growth thanks to this video game content, in 2011 the creators of Justin.tv decided to open another website called Twitch.tv, which is focused on this type of content related to this sector and e-sports (sports competitions borned from multiplayer video games). In 2014, Amazon decided to buy the platform.

In 2024, the Twitch platform will have officially been in existence for 10 years, during which time it has changed both in appearance and in the functions and services it offers. We will not go into all these changes here (this is a subject for another study), but it is interesting to highlight some aspects such as the creation and improvement of the chat, the appearance of advertisers, the possibility of financially supporting content creators, the storage of Videos on Demand (VODs) to be able to watch live broadcasts, among others. The platform has also improved its social aspect, facilitating live interaction between different stream channels and with the audience.

Since its creation in 2014, the site has been increasing the number of views, but it was in 2020 that the platform drastically increased the number of concurrent viewers, according to the website Twitchtracker.com, which specialises in the quantitative and statistical analysis of the platform, as can be seen in Image 1 (Twitch viewers statistics., n.d.). This increase in popularity temporarily coincides with the COVID-19 pandemic and its subsequent containment, which is one of the causes of its growth (as happened with other audiovisual content platforms) (Narassiguin and Garnès, 2020). However, this pandemic was not the only factor in the platform's growth.

Image 1. Evolution of Twitch.tv average and maximum monthly concurrent views.
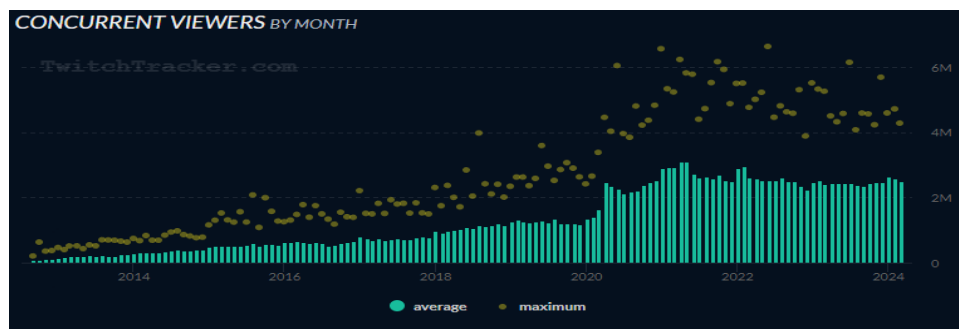
Source: twitchtracker.com

One possible factor is determined by the consumer's own consumption habits, since in recent years there has been a "disaffection with traditional television content and the consolidation of video games as a cultural product" (Padilla and Navarro, 2022, p. 68).

Another possible factor that may have motivated the increased consumption of the platform is psychological. Padilla and Navarro (2022, p. 69), paraphrasing Sheng and Kairam (2020), point to the existence of "affective gratifications and feelings of community or social integration", which may lead consumers to request similar content again when they feel identified in a group.

As can be seen from the above references, the scientific study of the platform and all that surrounds it has begun to proliferate in recent years. Some of this research focuses on the study of the platform's audience, such as that of Padilla and Navarro (2022), or the research on the media construction of the platform's most prominent streamers by Gutiérrez and Cuartero (2022).

Other recent studies have conducted research on the perception of brands on these and other platforms, such as that by García et al. (2022), while others have classified the types of content on the platform, such as the article by Olivares and Méndez (2022).

On the other hand, some of the aforementioned authors have carried out studies on the consumption habits of this platform for young people. However, these studies have focused on the impact of some commercial brands on this audience.

Little research has been carried out on how these new characters, such as streamers, affect the behaviour of adolescents. Gutiérrez and Cuartero (2022) approach this topic, but from a qualitative perspective, focusing almost exclusively on a study of the number of pieces that the mainstream media produce on some of the main streamers on the platform. On the other hand, there are academic articles that deal with the potential of Twitch.tv as an educational tool, such as the study by Pozo-Sánchez et al.

Finally, the study by Orduña-Malea and Lopezosa (2024) is also interesting, in which they carry out a general literature review on the platform, indicating only the number of publications and topics, but without focusing on the social sciences.

All the aforementioned research on Twitch can be found within our borders. Therefore, it is not risky to assume that there have also been studies on the platform outside Spain.

2. Hypotheses and Objectives

Following this brief introduction, the motivations for this research are outlined. The following research questions are proposed:

RQ1. What is the publishing ecosystem on the Twitch.tv platform like?
RQ2. Which countries have the highest number of publications indexed in Scopus and WoS?
RQ3. What percentage of scientific articles are open access?
RQ4. What are the subject areas of the scientific articles on the platform?

In order to try to answer the above research questions, the following hypotheses are formulated, which will be confirmed or refuted throughout the research process:

H1. The number of social science articles on the platform has increased progressively and steadily in recent years (during and after the COVID-19 pandemic).

H2. The vast majority of scientific articles are not open access.

H3. Studies on communication (both commercial and non-commercial) produced on the platform are in the minority, with studies mainly focusing on aspects related to COVID-19.

H4. Research methods on the platform are clearly qualitative.Therefore, the purpose of this work is to conduct various studies and analyses on the scientific literature written about Twitch.tv.

As the main objective (MO), the aim is to comprehensively understand the ecosystem of existing scientific publications about the platform. To achieve this, the object of study will be scientific articles published in academic and scientific journals, indexed in different databases, from the inception of the platform until the end of 2023.

The specific objectives include the following:

SO1. Classify the type of access to scientific articles about Twitch.tv.

SO2. Identify the most researched topics and the most commonly used keywords in studies about the platform.

SO3. Discover the methodologies employed for studying the platform.

3. Methodology

The methodology used in this study is a hybrid with both quantitative and qualitative aspects. The reason for this is to combine the most positive aspects of both working methods to achieve the greatest possible reliability. As Chaves Montero (2018, p. 165) points out:

Mixed research uses the pragmatic method and the system of philosophy, it is an inclusive and pluralistic method. The aim of mixed research is not to replace quantitative or qualitative research, but to use the strengths of both methods in combination and try to minimise their potential weaknesses.

Thus, the quantitative part of the research will focus on answering research questions and testing hypotheses through the collection of numerical data and statistical analysis of these data (Hernández et al., 2010).

On the other hand, the qualitative aspect of this study is related to the need to interpret and codify in some way the documents to be studied. It should be remembered that the aim is not only to carry out a bibliometric analysis of the data, but also to find out what these documents are about and to discover trends that may be of interest for research on the platform. Chaves Montero (2018, p. 172) argues that 'qualitative researchers use a wide range of interrelated methods, always hoping to get a better fix on the object of study at hand'. Thus, a qualitative approach is necessary to understand some aspects of the platform. These aspects are related to the research issues addressed in the articles. A purely qualitative approach, using the metadata provided by the databases and without reading and interpreting the documents, would not be as complete as it should be.

To achieve a systematic review of scientific articles, the Resiste-CHS framework (acronym for Systematised Reviews in Human and Social Sciences), developed by Lluís Codina in 2018, was used. This framework is, as the author argues, a variant of the SALSA (Search, Appraisal, Synthesis and Analysis) working method developed by Granth and Booth in 2009, whose aim was to establish a set of steps to be followed for the correct conduct of systematic reviews (Search, Appraisal, Analysis and Synthesis) (Codina, 2018b).

The reason for creating a specific framework for the human and social sciences is that systematic reviews themselves are related to the field of biomedicine and health, focusing mainly on quantitative outcome studies, with all their phases being highly standardised and with the synthesis of results using statistical techniques (Codina, 2018b).

Thus, the Resiste-CHS framework is carried out according to the idea of systematic reviews, which share the search and appraisal phases with systematic reviews but differ in the analysis and synthesis depending on the research objective, adding a final phase of presentation of results (Codina, 2018b).

Finally, it should be noted that the research carried out by Fernández-Gómez et al. (2023) was used as a guide for the systematic review. The following subsections present the methodology used in each stage of the framework.

3.1. Search Phase

The initial requirements for the search phase were to include all scientific documents categorised as scientific articles, books or book chapters. In addition, it was determined that the languages to be included in this study would be English (due to their importance in the academic field), Spanish and Catalan (as both are official languages of the autonomous community from which the study is being carried out).

For this purpose, different databases have been used that are relevant both to the research and to the context in which it is carried out:

Scopus and WOS: databases of great importance for the dissemination of scholarship, multidisciplinary and multi-publisher, more recognised by evaluation agencies and with growing collections in the social sciences (Codina, 2018a).

Proquest: reference database in the biological and health field, which also has multiple publications related to social sciences, especially doctoral theses and books in both electronic and physical format.

Dialnet and Índice CSIC: reference databases in the Spanish production of scientific documents in the social sciences (Codina, 2018a).

The decision to exclude other types of documents is based on the need to maintain the quality of the scientific information under investigation. To do this, it is necessary that the documents included in this study have undergone an evaluation prior to publication. Including non-evaluated documents such as interventions at conferences, gray literature or information present on the Internet would devalue the scientific quality of the information analyzed and would distance this study from its main objective.

The search terms used were different on each platform, depending on the search engine and the search options available. To optimize document management, the Zotero bibliographic reference manager has been used.

Before continuing, it is necessary to point out a problem that arose from the very meaning of the word "twitch", which means to flicker or move quickly. This fact, together with the fact that the authors themselves use Twitch as Twitch.tv (there are even those who use the word "twitchers" to refer to streamers) to talk about the platform, has motivated the use of search filters that go beyond Boolean operators.

This required a more complex search strategy in order to obtain as many results as possible without including results that are not very relevant to the area of study. For this purpose, the term twitch*, the Boolean operators AND and/or OR and terms related to the platform such as *stream*, *broadcast*, televi*, social AND network, entertein*, game, advertis*, business, influence, teen*, audience, monet*, mass AND media, content, communicat* and internet in Scopus, Web of Sciencie and Proquest were combined. As can be seen, truncations have been used in words containing derivatives. At the same time, fields of knowledge related to health sciences, biology, chemistry, pharmacy, etc., i.e. those not related to the platform, were also excluded. In the Proquest database, the typology of documents to be included in the search was added: articles in academic journals, books and book chapters, theses and dissertations. In addition, specific terms that classified the results, such as muscle contraction or proteins, were excluded from this database. Detailed search strategies for these three platforms are provided in Appendix 1.

Dialnet was searched using the word twitch due to the search limitations of the platform itself. Although not relevant and not included in the results, it is worth mentioning that the CSIC index was also searched using the word twitch. However, the few results obtained were already included in other databases.

All searches were performed on 17 July 2013 and repeated on 31 January 2024 in order to find documents indexed until the end of 2023.

Once the search results were obtained, we proceeded to eliminate duplicates, documents that had nothing to do with the object of research (such as those related to health sciences or biology, as well as those in other languages). Later, we proceeded to eliminate those documents that were not the aforementioned (scientific articles, books, book chapters and doctoral theses) and those that dealt with

the platform but from a more technical aspect, typical of computer and computational sciences. At another point in the research, we worked only with scientific articles, and later with those who only had access to the full document. Although the results are presented in advance, the whole process can be seen in Table 1.

Table 1. Search phases, inclusion and exclusion criteria and number of papers found.

Search phase	Description of the phase	Inclusion criteria	Exclusion criteria	Total number of documents
Phase 1	Initial search	Articles, theses, books and book chapters	None	2467
Phase 2	Elimination of duplicates		Duplicates	1954
Phase 3	Eliminate irrelevant results	Articles in Spanish, Catalan and English	Articles using "twitch" referring to biological or health sciences, authors' names or meanings unrelated to the platform.	712
Phase 4	Delete conference papers		Conference papers	467
Phase 5	2nd search (31 January)			549
Phase 6	Eliminate irrelevant results		As in phase 3, in addition to discarding written documents in 2024.	359
Phase 7	Deletion of articles		Documents dealing with technical and IT aspects of the platform.	319
Phase 8	Final sample	Scientific articles		247

Source: own elaboration

3.2. Assessment

Both the Zotero reference manager and Microsoft Excel were used to evaluate and record the search results. The Excel data collection template used to record the data was created using the resource "Bibliographic review: Data collection template" (Tur-Viñes, 2017) as a reference. The following data were included in this document: ID, type of document, category, subcategory (free), main research topic, citation, availability of full text, URL, methodology used, research tools, country of researchers, main findings, year of publication, title, journal, publisher, keywords, abstract, type of access and researcher comments. It should be noted that some of these items, such as "Main findings" and "Researcher comments" are subjective and optional.

To assess the type of access to scientific documents, different resources have been used. Firstly, different publication paths were taken into account, such as gold, green, hybrid and bronze (Biblioguías UCM: Open Access: Open Access Paths, (n.d.)). Two browser applications (Google Chrome) were also used to study access to scientific articles. These were Open Access Button and Unpaywall. The first of these scans the article page of databases such as Scopus and WOS to determine whether the article is legally open access when the application button is clicked. The second has similar functionality but differs in that it automatically scans the article page in the database and displays a padlock icon on the right-hand side of the screen which changes colour depending on the type of access the document has (green, gold or bronze).

3.3. Analysis, Synthesis and Presentation

Once the final sample was obtained, Google Spreadsheets were used to generate graphs of the most relevant numerical results. In order to complete all the items proposed in the working document of the previous phase, a content analysis of the sample texts was carried out to identify aspects such as category, subcategory (free), main research topic, main findings and researcher's comments. In addition, this analysis was also carried out to confirm the data for the other items listed above.

VOSviewer was also used to perform a co-occurrence analysis of the main terms used in the selected scientific papers.

The applications used were Zotero (to record the data) and the open-access programme VOSViewer (created by Van Eck and Waltman in 2010) to produce the relationship and keyword density maps. The data used come from the metadata extracted from the databases (Scopus, WoS and Proquest) and classified as "tags" in Zotero. These tags are made up of keywords and some terms that the databases themselves use to classify documents.

It should be noted that some articles, especially those obtained from Dialnet, do not include this metadata. In order to include them in the sample, the keywords of these articles were manually inserted in Zotero. The keywords that were in Spanish were also translated and the numerical elements that the platforms themselves use to classify and code their documents were eliminated.

Finally, it is necessary to mention the procedure used to create the VOS map. First of all, the article data were exported in RIS format to be included in the application. A new map was created from the bibliographic data and the following parameters were set:

Type of analysis: co-occurrence.

Counting method: Full count.

Minimum number of occurrences of a keyword: 5 (minimum number of keyword similarities to be counted). 73 out of 1448 (the keyword "muscle twitch" was removed as it is not related to social sciences.

Method of analysis: Association strength, where the strength of the relationships between the different keywords is calculated using the platform's own algorithm (Van Eck and Waltman, 2010).

As the last part of the study, an exhaustive and qualitative analysis of the scientific articles was carried out. For this purpose, a content analysis was carried out on those scientific articles to which access was available, either because the repository or database in which they are located is publicly accessible, or because they were consulted thanks to the institutional access provided by the University of Alicante.

Firstly, in order to categorise the articles in a way that is coherent with the research, it was decided to establish two categories that would help to give a small insight into the relevance of the platform in the scientific field. The categories used were the following:

Specialised document on Twitch.tv: those documents that talk only about the platform and study only phenomena, data and/or elements present on it.

Cross-sectional study alongside other platforms: articles that study the platform through a broader analysis, i.e. alongside other social networks or other platforms or media.

Secondly, the research topics have been grouped as much as possible according to the proximity of the aspects addressed and/or investigated in the articles:

Psychosocial and/or behavioural aspects: documents that focus on understanding and/or investigating the attitudes of platform users.

Commercial and non-commercial communication: studies aimed at analysing the communicative aspects of Twitch (chat communication, appearance of brands, type of language used, etc.).

Study of media and other platforms: articles that deal with the characteristics of the platform as a communication medium and its comparison and/or analysis with other platforms, both online and offline (traditional media).

Content and/or streamers: documents analysing the content shown on the platform and the streamers who produce it, through case studies or other related research tools.

Thirdly, the articles in the sample were read in order to identify both the research methodology and the research tools used. This reading focused mainly on the abstract and the methodological section to gain a correct understanding of the issues under study.

Finally, in terms of synthesis and presentation, it is important to point out that some results have been grouped together if their percentage or number is not too representative or interesting to study. On the

other hand, the presentation of the results has been made as visual as possible, using three-dimensional pie charts (whose data are visible in the graph itself and in which different colours are used), bar charts, tables and representative images of the analyses carried out with the VOSviewer programme.

4. Analysis and Results

After searching the different databases, a total of 2467 results were obtained. 889 in Scopus (36.04% of the sample), 491 (19.90%) in WoS, 757 (30.69%) in Proquest and 330 (13.38%) in Dialnet. After eliminating duplicates using the Zotero reference manager (and under human supervision), the final number of records was 1954.

After the various stages of screening, 359 documents remained to be considered (stage 6), of which 287 were journal articles (79.94%), 44 book chapters (12.27%), 26 doctoral theses (7.24%) and 2 books (0.56%).

Eliminating documents not related to the social sciences (phase 7), which represent 11.08% of the sample (40 articles), the resulting sample consists of 319 documents that talk about the platform and are related to the social sciences. Of these, 247 are journal articles (77.43%), 44 book chapters (13.79%), 26 doctoral theses (8.15%) and 2 books (0.63%).

Finally, it was decided to work only with scientific articles (247) in order to carry out a relevant and fair documentary analysis. Several analyses were carried out on this sample. Firstly, we analysed the types of access to the scientific articles, and then we carried out a co-occurrence analysis of terms using the VOSviewer tool, but not before we analysed the years of publication of these documents, the authorship and the journals in which they were published.

4.1. Articles' Publication Date and Access Types

Figure 1 shows the evolution of the number of articles published since the creation of the platform until its peak in 2022, coinciding with the end of the health crisis caused by COVID-19. Although the number of articles published in 2023 decreased, the interest seems to have been maintained to some extent, as it is still higher than the number of articles published in 2021. It is important to highlight the fact that between 2020 and 2023, 176 of the 247 articles included in the sample were published (71.26% of the total). If we compare the data on scientific production with the data on visualisations obtained from the platform, shown in Image 1 (first point of this document), we can see that the evolution of the production of articles coincides in time with the increase in the use of the platform. It also coincides with the explosion of the fame of some streamers due to the most severe restriction of the pandemic (first half of 2020). From then on, we see that both views and posts increase (2021). From 2022, both graphs behave similarly, with both posts and views decreasing until 2023.

**Figure 1.** Scientific articles published by year.
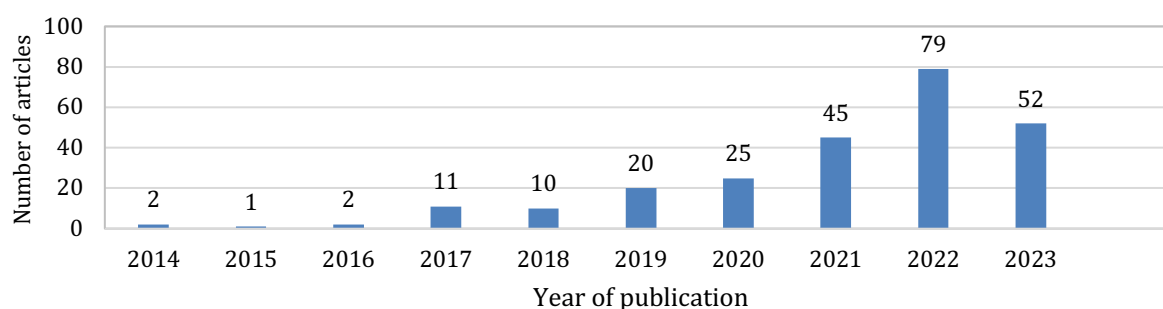
Source: own elaboration.

The classification of types of access was carried out using the information provided by some databases (such as Scopus) and the applications and resources mentioned in section 3.2 of this article. As can be seen in Figure 2, when all the types of access are added up, the number of documents is 207 (83.8% of the total number of articles). It is also important to highlight institutional access, which represents 36.4% of the sample, and Green Open Access, which represents 23.1%.

However, it is interesting to note that without institutional access through the University of Alicante, less than half of the sample (47.4%) could have been analysed.

**Figure 2.** Types of access to scientific articles
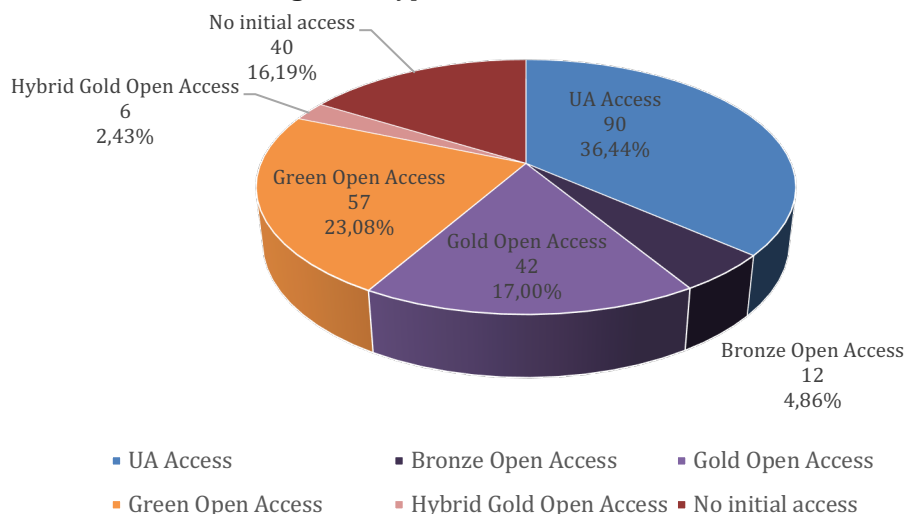
Source: own elaboration.

4.2. Authorship and Journals Analysis

The following two graphs show the study carried out on the authorship of articles.

The first (Figure 3) shows the countries of origin of the researchers as a percentage. When an article was written by more than one person, the country of the first researcher was chosen. This country was chosen based on the origin of the university to which the researcher belongs. Only articles from Scopus, WoS and Proquest were used for this graph, without considering those from Dialnet, as they could bias the sample towards Spain, since this repository mainly concentrates articles from this country. All countries with less than 10 publications (4.65% of the total) have been included in the "Others" sector.

**Figure 3.** Country of origin of investigations
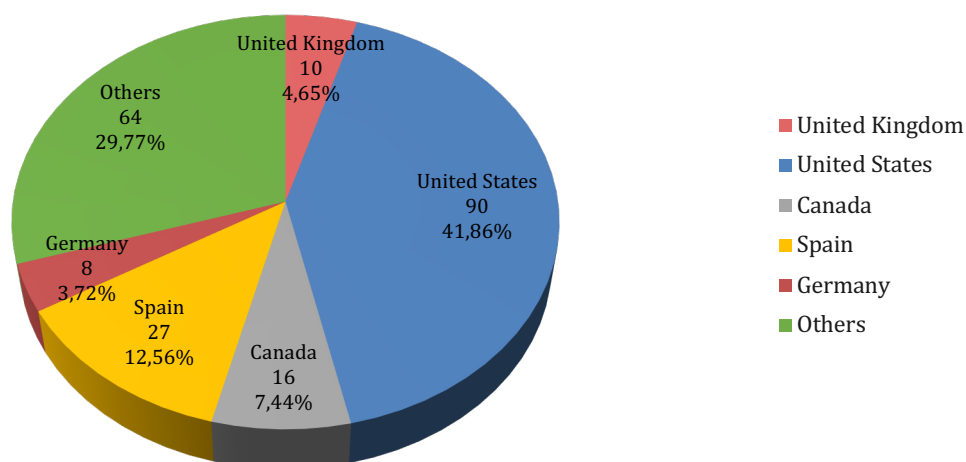
Source: own elaboration

Figure 4 takes into account the country of origin of the publisher, i.e. the company/institution that publishes the journal, without taking into account the results from Dialnet. As in the previous figure, countries with 10 or fewer results have been grouped in the "Other" sector.

The United States emerges as the country with the highest scientific production, with 41.9% of indexed articles. However, it should be noted that Spain is in second place with 12.6% of the articles. The United States is also the main country of origin of publishers, i.e. the publishers of the journals in

which authors publish their articles, with 45.3%. The United Kingdom ranks second and the Netherlands third. Spain ranks fourth with almost 8%. Some of the US publishers are (Sage Journals, Association for Computer Machinery or IEEE Computational Intelligence Society). From the United Kingdom (Taylor & Francis), while from the Netherlands, the high result is due to the Dutch origin of Elsevier.

Figure 4. *Publisher's* country of origin
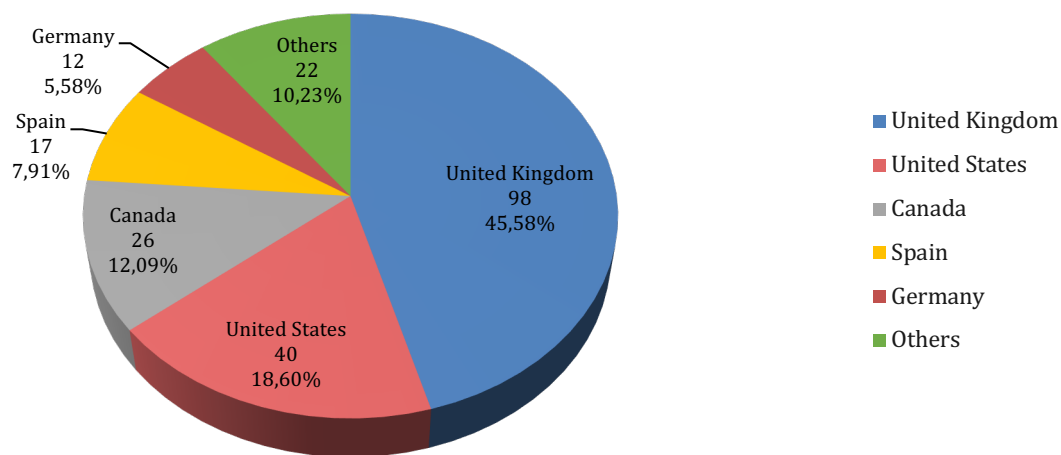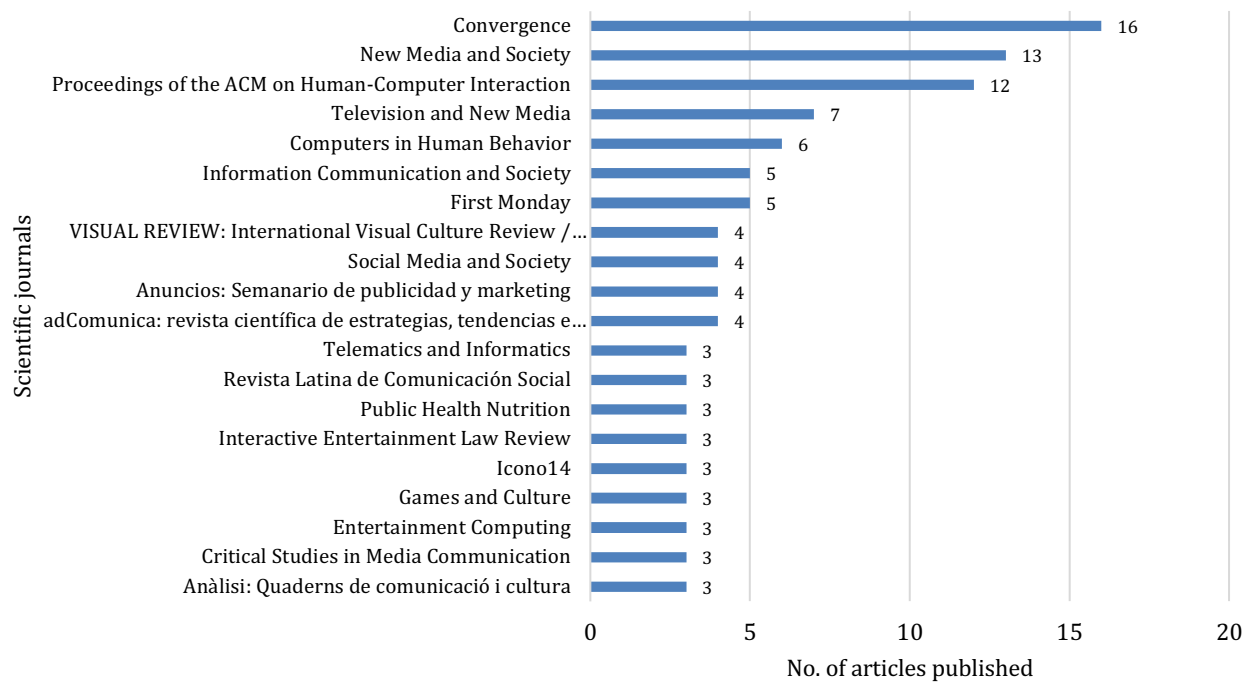

Source: own elaboration.

Table 2 shows the most prolific authors, i.e. those authors whose name appears in three or more documents. For this purpose, data were extracted from the entire sample of articles. It is interesting to highlight the scientific production of Buitrago, Torres Ortiz and Cabeza-Ramírez, as they are the Spanish authors with the highest number of papers.

Table 2. Author with three or more articles

Author	Number of articles
Johnson, M. R.	12
Gandolfi, E.	7
Wohn, D. Y.	5
Ferding, R. E. / Hamari, J. / Woodcock, J. / Ruberg, B.	4
Buitrago, A. / Edmond, J. A. / Masterson, T. D. / Pollack, C. C. / Sjöbolm, M. / Torres Ortiz, L. / Törhönen, M / Cullen, A. L. / Kairam, S. R. / Cabeza-Ramírez / Church, E.M. / Partin, W. C. / Partin, W.C.	3

Source: own elaboration

Figure 5 shows that three journals are the most prominent in publishing scientific articles on Twitch.tv. These are *Convergence* (16), *New Media and Society* (13) and *Proceedings of the ACM on Human-Computer Interaction* (12). Both *Convergence* and *New Media and Society* are published by the same publisher (SAGE). The Spanish magazines with the highest number of articles on the platform are *adComunica*; *Visual review* and Anuncios: *semanarios de publicidad y marketing*, with 4 articles each.

Figure 5. Journals with more than three publications

Source: own elaboration

4. 3. Keywords and Relevant Terms

A VOS map was created to provide a visual and clear understanding of the results obtained on the topics covered in the articles analysed. VOS stands for *Visualisation of Similarities (Van Eck and Waltman, 2006). This type of map is used to graphically represent bibliometric data but focuses on the relationships between the articles studied* (Van Eck and Waltman, 2006). Image 2 shows the network generated by VOSviewer with the selected keywords. Five majority clusters can be observed (the program was set to join those clusters that contained less than 5 keywords, which is the default option of the program).

Blue cluster: it contains keywords that are more related to video games, motivation and psychology ("motivation", "video game", "psychology", "human", etc.).

Red cluster: contains a wide variety of words, but the words most closely related to other words in the same cluster are those related to marketing and advertising ("marketing", "advertising", "user generated content", "user behaviour").

Yellow cluster: could be identified as the cluster most related to gender studies and online communities, as words like "gender", "women", "moderators" and "online communities" appear.

Regarding the green and purple clusters, they are intermixed. "Twitch" and "social media" are the keywords with the most co-occurrences, together with "streaming". All three are green and/or purple and are located in the centre of the VOS map. The purple keywords could be categorised as those specific to media as platforms, such as 'communication', 'media', 'television', while the green ones are more diverse terms, such as 'internet', 'covid-19', 'streaming', 'social media', 'viewers', among others. This relationship between the clusters makes the link between these words strong.

Finally, it is important to highlight that in all clusters there is a term related to the world of streaming: "videostreaming", "twitch", "streamers", "streaming", "live streaming", etc. This is not insignificant, as it means that the platform has been studied from several fields of knowledge and not just from one perspective.

**Image 2.** Network visualization.
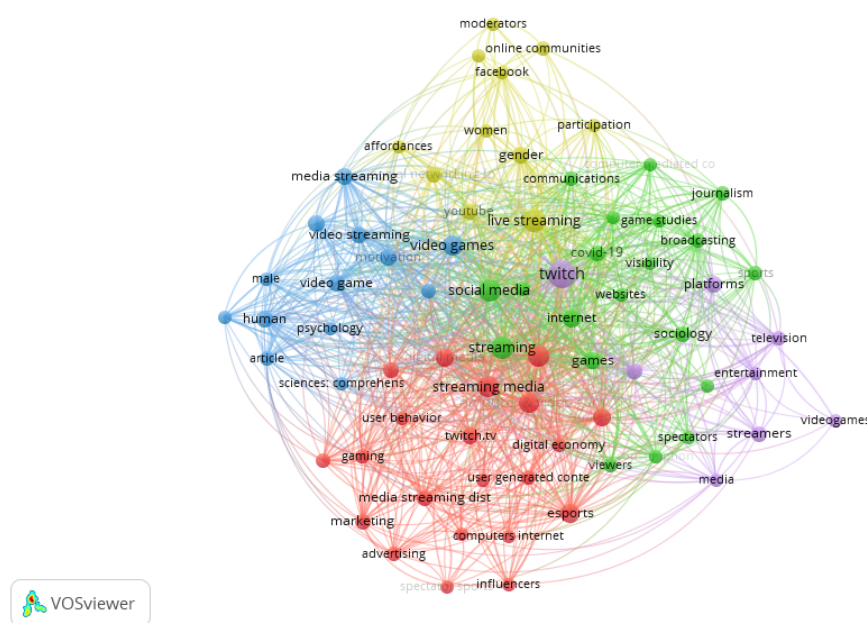
Source: own elaboration

Image 3 shows a keyword density map. "Twitch" is the word with the highest density, with 98 co-occurrences. "Live streaming" is second with 49 and "social media" is third with 45, tied with "streaming media". It is worth noting others with more than 20 co-occurrences, such as 'streaming' (40), 'social networks' (36), 'computer & video games' (29), 'video games' (27) and 'e-sports' (22).

It is important to highlight the use of keywords related to video games ("computer & video games", "video games" and "esports"). This shows the great weight that the video games sector has had and continues to have on the platform, as indicated above in the state of the question.

Image 3. Density visualization
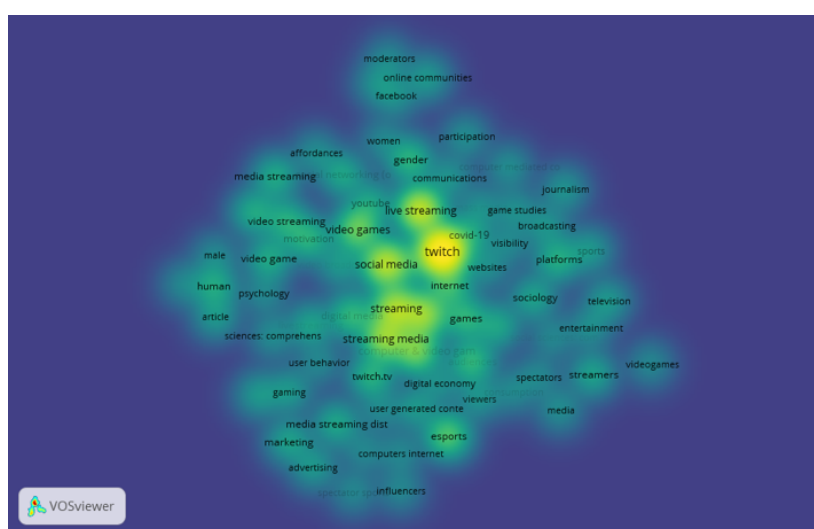
Source: own elaboration

Looking at more numerical data, Table 3 shows the average year of use for each keyword with more than 20 co-occurrences. This result is expressed with the year and a decimal part indicating the specific time of the year (e.g. the result for video games is 2021,5, i.e. its average time of use was in the middle of 2021). These keywords have been ordered from highest to lowest co-occurrence. The average year of publication of all the articles analysed is 2021,06.

Table 3. Most co-occurring *keywords* and their average year of use

| Keyword	Average year of use	Occurrence
Twitch	2021,37	98
Live streaming	2020,96	49
Social media	2021,87	45
Streaming	2020,95	40
Social networks	2021,39	36
Streaming media	2020,69	35
Computer & videogames	2020,59	29
Videogames	2021,5	27
Esports	2021,05	22

Source: own elaboration.

4.4. Analysis of Subject Areas and Methods used

70.5% of the sample of articles studies the platform without comparing or analysing it with other social networks, media or similar platforms.

On the other hand, in terms of research topics (Figure 6), 'content and streamers' is the most frequently addressed topic. The second is 'psycho-social and/or behavioural aspects', followed by 'studying media and other platforms'.

It is interesting to note that only 11.1% of articles on the platform deal with 'commercial/non-commercial communication'. While it is true that the vast majority of documents mention or talk about communicative aspects in one way or another, only a few (23 articles) deal comprehensively with this line of research.

**Figure 6.** Main research topic in percentages
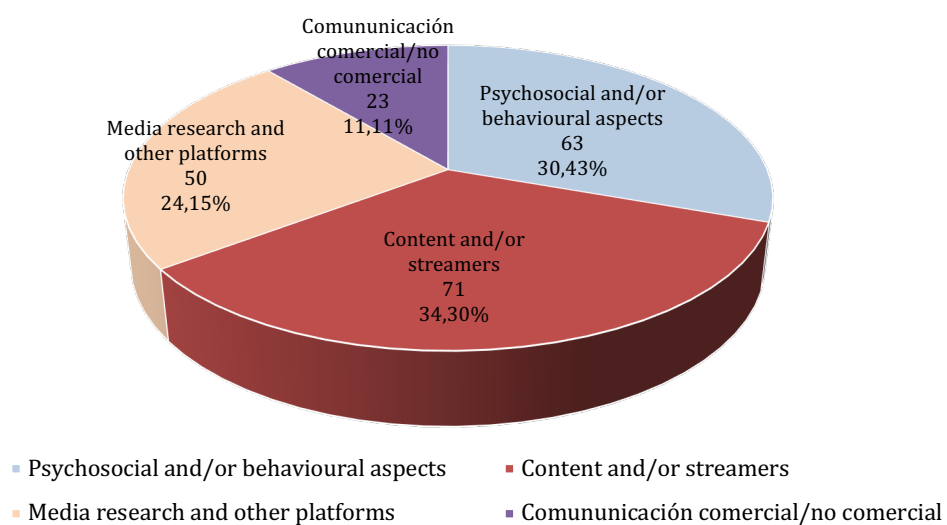
Source: own elaboration.

The methodology used is, in a slight majority, qualitative, in 51.7% of the articles reviewed. The other studies opt for a fully quantitative approach (30.9%) or use a mixed methodology (17.4%).

On the other hand, the research tools were as follows: content analysis (39.61%), data analysis (28.99%), survey (23.19%), interview (18.36%), case study (13.52%), observation (11.11%), research theory - understood as a specific existing process - (6.76%), literature review (6.28%), ethnography (4.83%), experiments (4.34%), focus group and documentation (2.42% each). It should be noted that some articles have more than one research tool, so all were counted.

Content analysis emerges as the most frequently used research tool. This result is somewhat logical given the perspective from which the Twitch.tv study is approached (social sciences). It is also relevant to note the research tools present in more than 20 articles (Figure 7): data analysis (60), survey (48), interview (38) and case study (28).

Figure 7. Most frequently used research tools
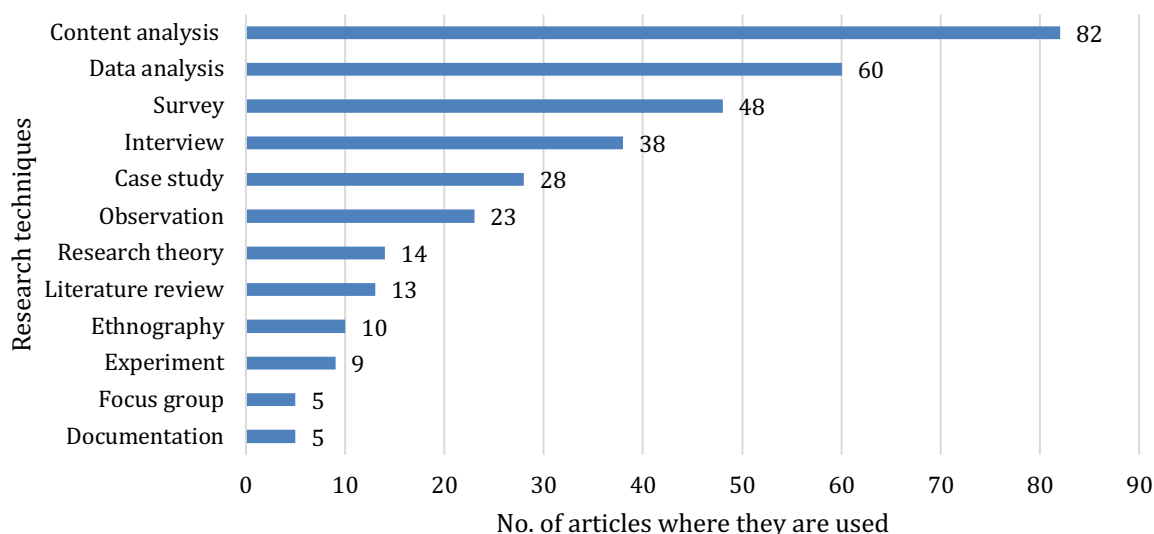
Source: own elaboration.

5. Discussion of Results, Conclusions and Future Lines of Research

As shown, the number of studies on the platform has increased in recent years, peaking in 2022 and decreasing in 2023. Nevertheless, their number is still higher in 2023 than in 2021. These results partially confirm H1, as the number of articles written only increases until 2022. H2 can be partially confirmed, because although the results show that 52.6% of articles are not OA, it cannot be said to be a "clear majority".

H3 is partially confirmed. It is true that the articles analysed that focus on communication (commercial or non-commercial) represent only 11% of the sample. However, the claim that these studies focus mainly on aspects related to COVID-19 is not correct. Although the term COVID-19 is one of the terms reflected in the VOS maps (especially in the density map), its co-occurrence is only 11, far below the more co-occurring terms such as "twitch", "live streaming" or "social media". Furthermore, the VOSviewer analysis did not reveal any clusters that could be identified with the pandemic.

H4 is also partially confirmed. Although 51.7% of the articles use a qualitative methodology, 17.4% use a mixed methodology and a significant 30.9% use a purely qualitative methodology.

After this review of the hypotheses formulated and their confirmation or refutation, it could be said that the main objective of this article has been largely achieved, which is to show what the publishing ecosystem on Twitch is like. In addition, the specific objectives have also been achieved: classifying the way in which scientific articles are accessed on Twitch.tv, identifying the main research topics and the most used keywords, and recording the methodologies and research tools used.

Regarding the limitations of this study, the most relevant has been the exclusion from the study of those documents published with restricted access. This involves the exclusion of information that could be relevant. It would be advisable to continue this study in a line of research with more time and resources which also includes those articles with restricted access.

On the other hand, there are some aspects that have not been analysed that could be relevant. One of these is the field of knowledge of the journals in which the articles were published, so a future line of research could investigate this aspect. It would also be interesting to supplement the study with articles published in national databases and repositories in other countries (as was done in the case of Spain with Dialnet), to obtain a much more relevant and accurate sample of the scientific literature written on Twitch.

Another limitation is that only scientific articles were analysed, without examining in depth other documents such as doctoral theses, book chapters or books. An interesting line of research could also include these three documentary typologies, focus on only one or some of them, or look for the scientific production of the most prolific authorship in order to analyse these documents or interview the authors.

It is worth highlighting the relevance of this scientific article in the context of the study of the platform, since at the time of writing it represents a first rigorous approximation of the scientific literature written on the Twitch.tv platform, which will also serve as a starting point for the subsequent writing of scientific articles on it, which will constitute a doctoral thesis based on a compendium of publications.

In conclusion, having this compilation and classification of scientific articles on the platform, and knowing the topics they cover, is a great help for further studies on Twitch.

References

Biblioteca Universidad Peruana de Ciencias Aplicadas (n.d.). *Data analysis and visualisation tools: VOSviewer*. UPC Library. https://biblioteca.upc.edu.pe/analisisdedatos/vosviewer

Complutense University of Madrid Library (2023). *Open Access: Open Access Pathways*. Complutense Library. https://biblioguias.ucm.es/acceso-abierto/vias-acceso-abierto

Carrillo J. A. (2019). *New scenarios of video game consumption: Twitch.tv and esports* [Doctoral dissertation, Universidad de Murcia]. Digitum. http://hdl.handle.net/10201/85701

Chaves Montero, A. (2018). The use of a mixed methodology in social research. In S. G. Kenneth Delgado, W. F. Gadea, S. Vera-Quiñonez (Coords.), *Breaking barriers in research* (pp. 164-184). UTMACH.

Codina, L. (2018a). *Academic databases for research in Social Communication: proposal and characterization of the optimal group*. Master's Degree in Social Communication Research (MUCS). Department of Communication UPF. https://bit.ly/3TMjd9t

Codina, L. (2018b). Systematized bibliographic reviews: General procedures and Framework for Human and Social Sciences. Master's Degree in Social Communication Research (MUCS). Department of Communication UPF. https://bit.ly/3PUwlrN

Fernández-Gómez, E., Segarra-Saavedra, J., and Feijoo, B. (2023). Advertising literacy and minors. Literature review from the Web of Science (WOS) and Scopus (2010-2022). *Revista Latina de Comunicación Social*, 81, 1-23. https://www.doi.org/10.4185/RLCS-2023-1892

Gálvez, C. (2018). Co-word analysis applied to highly cited articles in Library and Information Science (2007-2017). *Transinformacao*, *30*(3), 277-286. https://doi.org/10.1590/2318-08892018000300001.

García A., Citlali E., and Bonales G. (2022). TikTok and Twitch: new media and formulas to impact Generation Z. *Icon 14, 20*(1). https://doi.org/10.7195/ri14.v20i1.1770

Gutiérrez, J. F., and Cuartero, A. (2020). The rise of Twitch: New audiovisual offers and changes in television consumption among young audiences. *Ámbitos. International Journal of Communication*, (50), 159-175. https://doi.org/10.12795/Ambitos.2020.i50.11

Gutiérrez, J. F., and Cuartero, A. (2022). The media construction of Twitch youth idols: Content creators and their growing presence in the Spanish media. *adComunica. Scientific Journal of Strategies, Trends and Innovation in Communication (*23), 251-274. http://dx.doi.org/10.6035/adcomunica.5936.

Hernández, R., Fernández, C., & Baptista, P. (2010). Research methodology. New York: McGraw-Hill Education.

Narassiguin, A., & Garnès, V. (2020). *The influence of COVID-19 on Twitch audience: How lockdown measures affect live streaming usage?* Influence of COVID-19 on Twitch audience (pp. 1-9). https://upfluence-common.s3.amazonaws.com/Covid19_Twitch.pdf

Olivares, F. J., and Méndez, I. (2022). Journalists and communicators on Twitch: media beyond social networks. *Anàlisi: Quaderns de Comunicació i Cultura, 66*, 45-61. https://doi.org/10.5565/rev/analisi.3473

Orduña-Malea, E., & Lopezosa, C. (2024). Uncovering the potential of Twitch as a source for social media metrics. *First Monday, 29*(1). https://doi.org/10.5210/fm.v29i1.13214

Padilla A., and Navarro C. (2022). Audiences and streamers on Twitch: consumption and production patterns in the Spanish-speaking world. *Quaderns del CAC, 25*(48), 67-77. https://bit.ly/3TIoCON

Pozo-Sánchez, S., López-Belmonte, J., Fuentes-Cabrera, A., and López-Nuñez, J. A. (2021). Twitch as a technopedagogical resource to complement the flipped learning methodology in a time of academic uncertainty. *Sustainability, 13*(9), 4901. https://doi.org/10.3390/su13094901

VOSViewer (n.d.). *Products*. VOSviewer. Visualizing scientific landscapes. https://www.vosviewer.com/products

Ruíz-Orjuela, E. T., Gatica, G., and Adarme-Jaimes, W. (2023). Literature review with bibliometric analysis of the hospital supply chain. *Engineering*, *28*(2), e18987. https://doi.org/10.14483/23448393.18987

TwitchTracker (n.d.). *Twitch viewers statistics*. Twitch Tracker. https://twitchtracker.com/statistics/viewers

Sheng, J. T., & Kairam, S. R. (2020). From Virtual Strangers to IRL Friends: Relationship Development in Livestreaming Communities on Twitch. Proceedings of the ACM on Human-Computer Interaction, 4(CSCW2), 1-34. https://doi.org/10.1145/3415165

Tur-Viñes, V. (2017). *Literature review: data collection template*. Repository of the University of Alicante. http://hdl.handle.net/10045/70867

Van Eck, N. J., & Waltman, L. (2007). VOS: A New Method for Visualizing Similarities Between Objects. In *Studies in classification, data analysis, and knowledge organization* (pp. 299-306). https://doi.org/10.1007/978-3-540-70981-7_34

Van Eck, N. J., & Waltman L. (2009). VOSviewer: A Computer Program for Bibliometric Mapping. Van Eck, N. J., & Waltman, L. (2009). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523-538. https://doi.org/10.1007/s11192-009-0146-3

Van Eck, N. J., Waltman, L., Dekker, R., & Van Den Berg, J. (2010). A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS. *Journal Of The Association For Information Science And Technology*, *61*(12), 2405-2416. https://doi.org/10.1002/asi.21421