

Literature Review on Self-harm Content on X (Twitter) Created and Shared by Young People and Teenagers

Revisión de estudios sobre contenido de autolesiones en X (Twitter) creado y compartido por jóvenes y adolescentes

Revisão de estudos sobre conteúdo de automutilação no x (Twitter) criado e compartilhado por jovens e adolescentes.

Esther Martínez-Pastor^{1*} 

M.ª Ascensión Miralles González-Conde^{2**} 

Pablo Sánchez-López^{3***} 

¹ Faculty of Communication Sciences, Universidad Rey Juan Carlos (URJC), Madrid, Spain

² Universidad Católica de San Antonio de Murcia, Spain

³ Audiovisual Communication and Advertising Department at Universidad Rey Juan Carlos, Madrid, Spain

* Professor of Advertising at the Faculty of Communication Sciences, Universidad Rey Juan Carlos (URJC), Madrid, Spain. Email: esther.martinez.pastor@urjc.es

** Lecturer at Universidad Católica de San Antonio de Murcia, Spain. Email: mmiralles@ucam.edu

*** PhD Assistant Lecturer at the Audiovisual Communication and Advertising Department at Universidad Rey Juan Carlos, Madrid, Spain. Email: pablo.sanchez@urjc.es



This work is licensed under a Creative Commons Attribution 4.0 International license

Received: 13/04/2024; **Revised:** 18/05/2024; **Accepted:** 08/10/2024;
Published: 24/07/2025

Translation to English: Hilary Jane Robinson.

To cite this article: Martínez-Pastor, Esther; Miralles González-Conde, María Ascensión; & Sánchez-López, Pablo (2025). Literature Review on Self-harm Content on X (Twitter) Created and Shared by Youth and Teenagers. *ICONO 14. Scientific Journal of Communication and Emerging Technologies*, 23(1): e2164. <https://doi.org/10.7195/ri14.v23i1.2164>

Abstract

The aim of this study is to gather and present the state of the art on non-suicidal self-harm content created and shared by young people and adolescents on X (Twitter). The methodology applied consists of a descriptive review of scientific articles published in journals (in English and Spanish) between 2013 and July 2023 in international multidisciplinary databases (WOS, SCOPUS and GOOGLE SCHOLAR). After searching, reviewing and screening the publications, 15 articles were selected and analysed. Based on a thorough analysis of the studies that inform these publications, the most relevant topics covered, and methodologies used have been identified. This review highlights the scarcity of studies on the subject specifically in the field of communication/sociology and serves as a basis for further studies of an interdisciplinary nature to address existing gaps.

Keywords

Non-suicidal self-harm; X/Twitter; Youngsters; Teenagers.

Resumen

El objetivo de la investigación es reunir y exponer el estado del arte sobre el contenido de autolesiones no suicidas creado y compartido por los jóvenes y ado-

lescentes en X (Twitter). La metodología aplicada en la investigación consiste en una revisión descriptiva de artículos científicos publicados en revistas (en inglés y español), entre 2013 y julio de 2023, en bases de datos multidisciplinares internacionales (WOS, SCOPUS y GOOGLE SCHOLAR). Mediante la estrategia de búsqueda, revisión y criba de las publicaciones se han seleccionado y analizado 15 artículos. El resultado del análisis exhaustivo de los estudios que informan estas publicaciones ha permitido identificar los temas tratados y las metodologías utilizadas más relevantes. Esta revisión pone de manifiesto la escasez de estudios al respecto desde el ámbito exclusivo de la comunicación/sociología y sirve de base para plantear estudios posteriores de carácter interdisciplinar que den respuesta a las lagunas existentes.

Palabras clave

Autolesiones No Suicidas; X/Twitter; Jóvenes; Adolescentes.

Resumo

O objetivo da pesquisa é reunir e apresentar o estado da arte sobre conteúdos não suicidas de automutilação criados e compartilhados por jovens e adolescentes no X (Twitter). A metodologia aplicada na pesquisa consiste em uma revisão descriptiva de artigos científicos publicados em periódicos (em inglês e espanhol), entre 2013 e julho de 2023, em bases de dados multidisciplinares internacionais (WOS, SCOPUS e GOOGLE SCHOLAR). Através da estratégia de busca, revisão e triagem de publicações, foram selecionados e analisados 15 artigos. O resultado da análise exhaustiva dos estudos relatados nestas publicações permitiu identificar os temas mais relevantes discutidos e as metodologias utilizadas. Esta revisão destaca a escassez de estudos nesse sentido no campo exclusivo da comunicação/sociologia e serve de base para propor estudos interdisciplinares posteriores que respondam às lacunas existentes.

Palavras-chave

Automutilação Não Suicida; X/Twitter; Jovens; Adolescentes.

1. Introduction

Non-suicidal self-harm (NSSH) is a behaviour in which an individual repeatedly inflicts minor-to-moderate, often painful injuries to the surface of his or her body without suicidal intent (American Psychiatric Association, 2022). People who self-harm intentionally injure themselves through cutting, burning or hitting, to enable them to manage pain they are unable to cope with (Klonsky, 2011). It is important to differentiate this conduct from suicide, as the intention is different and although suicidal behaviour can be potentially self-injurious, the ultimate purpose is to die (American Psychiatric Association, 2022).

NSSH has increased among teenagers and young people, which has prompted a number of associations to sound the alarm, such as the American Psychiatric Association (2022), UNICEF (UNICEF, 2021a & 2021b; Leiva, 2023), Fundación ANAR (Informe anual. Teléfono chat ANAR, 2023), the Spanish Ministerio de Sanidad (2021), FAD-Fundación de Ayuda contra la Drogadicción (Kuric et al., 2023), Asociación Española de Pediatría (AEP, 2022) and Asociación Española de Pediatría de Atención Primaria (AEPAP) (López et al., 2023). There is no specific profile of the young people who engage in this behaviour, although some believe that it is more common among the female population, nor is it aligned with a single gender, origin, culture or social class. Furthermore, the data on self-harm among young people vary greatly, and only 0.21-6.5% of teenagers who self-harm have requested medical care (Faura-Garcia et al., 2022); this results in significant differences in the data published about teenagers: internationally, prevalence ranges between 1.5 and 54.8% while in Spain it is between 0.58% and 74.9% (Faura-Garcia, Orue, Calvete, 2021; Pérez et al., 2021). These works demonstrate growing concern about the increase in NSSH among academics, whose studies address and describe the profiles of the young people who self-harm (Cipriano, Cella & Cotrufo, 2017), the existence of communities of young people who discuss this subject on social media (Lerman et al., 2023) and how the act of self-injury is becoming normalised and romanticised (Khasawneh et al., 2021).

On social media, in the last few years the sharing of self-harm experiences and suicidal acts among teenagers has become increasingly common. A 15% in-

crease was observed in posts about stressful situations on social media between 2013 and 2017 (Abi-Jaoude et al., 2020) because they share their thoughts and opinions in order to be understood and not be judged (Wang et al., 2017). X/ Twitter is a social network with more than 556 million active users (Fernández, Rosa, 2024). The platform's users are able to find the content, which tends to be textual content, photographs and videos, by means of ambiguous hashtags such as *#blithe* (NSSH), *#Deb* (depression) or *#ouchtw* utilised to avoid being detected by the platform (Brown et al., 2018). X/Twitter has a policy relating to suicide and self-harm, with the goal of preventing this conduct; however, the images appear with a content warning text stating "sensitive content" (Twitter, 2023) and there are different opinions in the literature regarding the platform's content moderation stance. In this regard, acknowledging an emerging field in research has become of interest. This paper proposes to review how self-harm content created and shared by young people and teenagers on X/Twitter is being studied and identify gaps in the bibliography in order to inform future research in the field.

2. Methodology

The main aim of this study is to reveal the current situation regarding self-harm-related posts shared on social media by young people and teenagers. In this case, the study has focused on the social network known as X (Twitter until 2023). The secondary objectives set are as follows:

1. Identification of scientific research on self-harm content created and shared by young people and teenagers on X/Twitter.
2. Determination of the research topics covered in these studies and their specific domains.
3. Identification of the methods and tools used to locate the self-harm content on X/Twitter, as well as the young people and teenagers who share it.
4. Determination of the methodologies used to analyse the content shared about self-harm by young people and teenagers on X/Twitter.

The methodology followed is the methodology of a scientific literature review, in accordance with the stages and guidelines for systematic review of scientific articles (Sánchez-Meca, 2022). Research was gathered about communication about self-harm among young people of social media via X/Twitter. The tasks performed in the review (Emparanza & Urreta, 2005; Sánchez-Meca, 2022) were the following: defining the study selection criteria; searching for and locating the studies; selecting and screening the articles; extracting data from the studies; and summarising and interpreting the results: categorising the topics covered in the research studies.

2.1 Definition of the selection criteria

The chosen eligibility criteria consisted of content on non-suicidal self-harm (NSSH) created and shared by young people and teenagers on X/Twitter. This study also defined the following criteria for screening the articles:

- Target group: the study population must be teenagers and/or young people.
- Topic of the article: it must be related to self-harm.
- Distribution channel of the messages: the X/Twitter social network.
- Language: it was decided that the language of the publication should be English or Spanish.
- Source type: the article must be published in research journals.
- Time frame: the decade between 2013 and (July) 2023.

2.2 Search for sources and study localisation strategy

The literature search used words associated with self-harm content on X/Twitter and was carried out on the most well-known international multidisciplinary databases: Web of Science Core Collection, SCOPUS and Google Scholar. The descriptors applied were key words in English referring to the object of the study: "Twitter";

“adolescents”, “teen/s”, “young people”; and “self-harm”, “self-injury”, “cut”. With regard to how they were related to the Boolean operators, the term combinations used were:

1. Twitter AND self-harm AND cut AND teen.
2. Twitter AND self-harm AND cut AND teens.
3. Twitter AND self-harm AND cut AND adolescents.
4. Twitter AND self-harm AND cut AND young AND people.
5. Twitter AND self-injury AND cut AND teen.
6. Twitter AND self- injury AND cut AND teens.
7. Twitter AND self-injury AND cut AND adolescents.
8. Twitter AND self-injury AND cut AND young AND people.

The initial search retrieved a total of 63,304 articles (Table 1). Twitter’s name change, to X, began on 23 July 2023, while the review was being performed.

Table 1. *The phases of the search for and selection of articles on self-harm content shared by young people and teenagers on X (Twitter) (2013-2023)*

INITIAL SEARCH	Google Scholar	SCOPUS	WOS
– Twitter AND self-harm AND cut AND teen; – Twitter AND self-harm AND cut AND teens; – Twitter AND self-harm AND cut AND adolescents; – Twitter AND self-harm AND cut AND young AND people	42,990	39	3
– Twitter AND self-injury AND teen; – Twitter AND self- injury AND teens; – Twitter AND self- injury AND adolescents; – Twitter AND self- injury AND young AND people.	20,270	0	2
TOTAL number of publications found in the initial search	63,260	39	5
RESEARCH REVIEW EXCLUSION CRITERIA	74	36	0
REVIEW AND DISCUSSION BETWEEN RESEARCHERS	14	1	0
TOTAL ARTICLES FOR ANALYSIS	15		

Source: developed by the authors.

2.3 Selection and screening of articles

In the document selection phase, an initial assessment was performed of the total number of articles retrieved through access to the title and the abstract. Articles pertaining to messages about self-harm on Twitter by young people and teenagers were included in the Excel database for the review. For this initial screening, the following causes for exclusion were applied (Table 2):

- The article is not related to X (Twitter) or social media.
- The research uses X (Twitter) as a method to recruit the subjects who participate in the study.
- The article is not directly related to self-harm.
- The research does not focus on young people/teenagers.
- The article is written in other languages (neither English nor Spanish).
- The reference is not for a research journal publication, preprints or processing.
- The article is a duplicate.

Table 2. *Selection criteria for the review of studies on self-harm content on X (Twitter) shared by young people and teenagers (2013-2023): inclusion and exclusion criteria*

SELECTION CRITERIA	INCLUSION CRITERIA	EXCLUSION CRITERIA
Target group	Young people and teenagers	Not focused on young people and teenagers
Topic of article	Self-harm	Does not allude to self-harm
Distribution channel	Twitter social network	Study is not related to Twitter or social media
Language	Spanish and English	Written in other languages
Type of source	Research article, preprints or processing	Not a research article, preprints or processing Is a duplicate
Date of publication	January 2013 to July 2023	Prior to January 2013 After July 2023

Source: *developed by the authors.*

As a result of the first screening, the number of articles was substantially narrowed down to 110 studies. Next, the filtered body of articles resulting from this screening was subjected to a more in-depth assessment by the researchers who, after an individual selection of the articles, discussed any discrepancies and reached a consensus on the most suitable studies to be included in the definitive group. Ultimately, the body of texts consisted of 15 articles which had been published between 2013 and 2023.

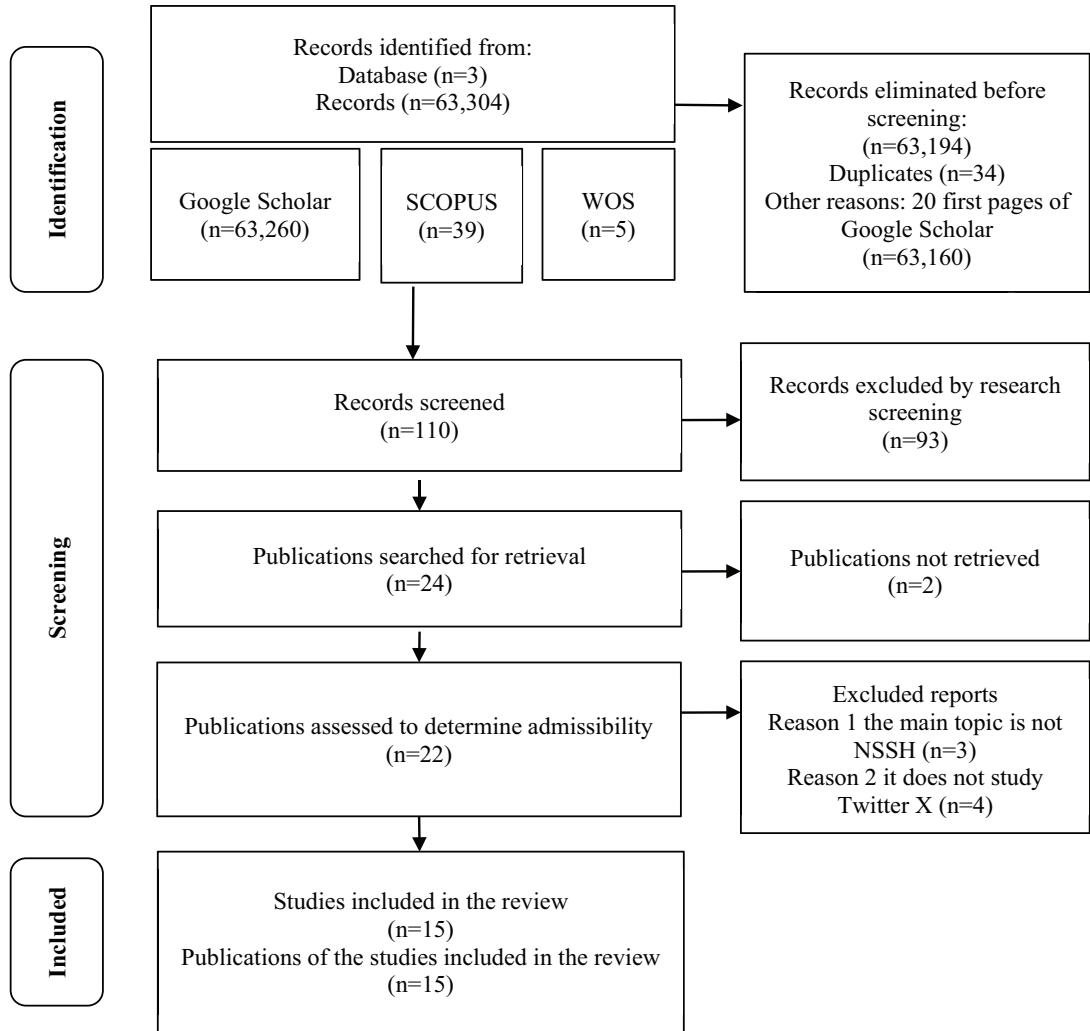
The presentation of the report for the systematic review follows the PRISMA statement 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) reporting guidelines, which can be applied to social or educational undertakings (Sánchez-Meca, 2022). The PRISMA 2020 flow diagram has been drafted following the indications of Yepes-Nuñez et al. (2021), which is the translation of the original document by Page et al. (2021). This tool sets out the flow of information through the different phases of a systematic review, indicating the number of records identified, excluded and ultimately included. Figure 1 specifies the process of identifying, screening and finally including the work to be analysed through this procedure and the flow diagram is used as the basis for new systematic reviews that only include searches in databases and records, as it is the most appropriate for the study being performed.

Once duplicate articles had been excluded, the exclusion criteria applied and the articles reviewed by the researchers, out of the 63,304 references retrieved from the three databases chosen there were 15 articles that met the selection criteria (Table 3). Of these, 7% (1 article) were retrieved from SCOPUS and the remaining 93% (14 articles) from Google Scholar.

2.4 Data extraction

After selecting the definitive articles, the information pertaining to the focus of this study's analysis was extracted. Two relevant types of information can be discerned (Sánchez-Meca, 2022): information referring to the characteristics of the studies and the results from the question considered in the review. The data were included in a database, using a Microsoft Excel table.

Figure 1. Flow diagram for the review of articles on self-harm content on X (Twitter) shared by young people and teenagers (2013-2023)



Source: Compiled by the authors based on the PRISMA 2020 flow diagram for new systematic reviews that only included searches in databases and records from Page et al. (2021) and the translation to Spanish by Yepes-Nuñez et al. (2021).

The following characteristics were extracted from the research presented in the articles in order to perform the comprehensive review and text analysis: the identifying details for the articles (title, author, year and country), hypothesis and research questions, research methodology used and limitations, characteristics of the sample: selection criteria, subjects (age range) articles and main findings: result and conclusions (Table 3).

Table 3. *Main data of the selected articles*

No.	Title	Author(s)	Publication	Area of knowledge	Country
1	Analysis of non-suicidal self-injury posts on Twitter: A quantitative and qualitative research	Silva et al. (2021)	Research, Society and Development	Health sciences	Brazil
	Main findings Following a qualitative analysis of 663 tweets about self-harm, social networks are viewed as a promoter of self-harm practices and can also be used as a space for support. It highlights the importance of public health actions that include monitoring and managing online content on behaviour.				
2	Communication Patterns of Adolescent Self-Harm Suffering in Interpersonal Relationships	Gustina S. & Wirman, W. (2022).	Jurnal Kajian Komunikasi	Communication	Indonesia
	Main findings Surveys were performed on young people who self-harm and who upload content on Twitter. Twitter is a medium where young people who self-harm are looking for self-fulfilment and meeting up with other adolescents who understand them. It is suggested that physical and emotional closeness between parents and children should be built in order to reduce self-harming behaviour among teenagers.				
3	Exploring Suicidality on Social Media: Qualitative Analysis of Twitter	Perry, I. & Park, A. (2021)	IEEE Conference Publication	IT and health sciences	United States
	Main findings Through a qualitative and quantitative study and with the help of an application programming interface or API (retrieving the hashtags <i>#selfharm</i> and <i>#selfinjury</i>), it examines the suicidality content expressed on Twitter, in which self-harm appears incidentally. One of its most significant findings is identifying the jargon or jokes around these topics.				

(continued)

Table 3. Main data of the selected articles (continued)

No.	Title	Author(s)	Publication	Area of knowledge	Country
4	Identifying Depressive Symptoms from Tweets: Figurative Language Enabled Multitask Learning Framework	Yadav, S., et al. (2020)	Proceedings of the 28 th International Conference on Computational Linguistics	IT and health sciences	United States
Main findings					
The experimental results of the study demonstrated that the BERT-based multi-task learning framework (BERT: Bidirectional Encoder Representations from Transformers for pre-training natural language) was an effective method to identify depressive symptoms in tweets. They propose improving the model to detect symptoms of depression and other conduct through user communications.					
5	Investigating non-suicidal self-injury discussions on Twitter	Alhassan, M. A., & Pennington, D. (2021)	International Conference on Social Media and Data Mining-ICSMDDM.	IT	United Kingdom
Main findings					
The methodology involved retrieving tweets using <i>#selfharm</i> and <i>#selfinjury</i> hashtags through Twitter Archive Google Sheets (TAGS) and, by viewing the data, to know who had generated this content in the United Kingdom. The results included the finding that <i>#selfharm</i> and <i>#selfinjury</i> hashtags are generated by non-professional users (nearly half of the users, or 47%).					
6	Mentally Ill and Cute as Hell: Menhera Girls and Portrayals of Self-Injury in Japanese Popular Culture	Seko, Y., & Kikuchi, M. (2022).	Frontiers	Communication sciences and IT	Japan
Main findings					
An analysis of blogs, anime websites and games identified three self-harm-related types of girls: sad girls (vulnerable, abnormal, deranged, potentially erotic), mad women (obsessive, manipulative and unfeminine), and cuties (a subculture from Harajuku, known as the mecca of youth countercultures). The article proposes exploring the cultural environment in which people explain, perform and make sense of self-harm because this can help towards understanding their behaviour and that of their followers.					

(continued)

Table 3. Main data of the selected articles (continued)

No.	Title	Author(s)	Publication	Area of knowledge	Country
7	Peer Responses to Trans Youth Tweeting about Self-Harm and Suicidality	Simms, D. (2020)	Creative nursing	Health sciences	United Kingdom
Main findings					
Through interviews of young trans people aged 14-18 years who tweeted about self-harm, it was identified that they do not encourage self-injurious behaviour in their interactions or tweet dismissive replies to expressions of anguish in comparison with young people in general. The content included suicidal thoughts (45.5% - 48.1%) and thoughts of self-harming (11.1% - 14.7%); suicide attempts were relatively rare (2.8% - 9.3%).					
8	Self-harm: detection and support on Twitter	Alhassam Abubakar, M., et al. (2021)	arXiv	IT	United Kingdom and Nigeria
Main findings					
Through an analysis of self-harming content on Twitter (by means of an API), a number of ways are suggested to provide online help. One of the conclusions drawn is that social media could be a useful tool for support, advice and raising awareness. Hashtag strategies by associations can contribute to helping with these behaviours, such as @MindCharity with the hashtag #MentalHealthAwarenessWeek					
9	Semi-Supervised Approach to Monitoring Clinical Depressive Symptoms in Social Media	Yazdavar, A. H., et al. (2017)	IEEE	IT and Psychology	United States
Main findings					
A model was created to detect depressive symptoms in tweets. To do so, they used PHQ-9 3 on a scale of depression of nine items, which incorporated DSM-V. The PHQ-9 model detected the main topics and the most used words and terms, such as self-harm. Among their results, it was observed that the tool automatically detected clinical depression symptoms with an accuracy of 68% and precision of 72%.					
10	Suicide Risk Prediction by Tracking Self-Harm Aspects in Tweets: NUS-IDS at the CLPsych 2021 Shared Task	Gollapalli, S. et al. (2021)	Proceedings of the Seventh Workshop on Computational Linguistics and Clinical Psychology	IT and Psychology	Singapore

(continued)

Table 3. Main data of the selected articles (continued)

No.	Title	Author(s)	Publication	Area of knowledge	Country
	Main findings				
	The learning model demonstrated its effectiveness in identifying a user's risk of suicide and self-harming. Through a content analysis model, it is possible to anticipate the risk of suicide and self-harm. Keywords were identified through the SHTM (Self-Harm Topic Model), including "death", "bruises", "emotional", "cut", "panic", "sadness" and "guilt", among others.				
11	The Instagram/Facebook ban on graphic self-harm imagery: A sentiment analysis and topic modeling approach	Smith, H. & Cipolli, W. (2022)	Policy & Internet	Law and mathematics	United States
	Main findings				
	An analysis was conducted through a quantitative study and an API of users' feelings about the banning of images of self-harm by Instagram. The keywords used were "self-injurious behaviour", "non-suicidal self-injury" and "self-harm". Following the banning of certain words on the social network, the adolescents' reaction was to feel their body had been censured by not being able to communicate, with an increase in anger and sadness.				
12	Tracking Suicide Risk Factors Through Twitter in the US	Jashinsky, J., et al. (2014)	Crisis	Health sciences, IT and artificial intelligence	United States
	Main findings				
	Through an analysis of 1,659,274 tweets, a significant correlation was observed between state Twitter-derived data and actual suicide and self-harm data. It is possible that Twitter may be a viable tool for large-scale, real-time monitoring of suicide risk factors. A directed response mechanisms via tweets could be created.				
13	Unveiling self-harm behaviour: what can social media site Twitter tell us about self-harm? A qualitative exploration	Hilton, C. (2017)	Journal of clinical nursing	Psychology and communication	United Kingdom
	Main findings				
	After an analysis of 362 tweets related to self-harm, the reaction from the general public towards self-harm content demonstrated that the content was misunderstood and even ridiculed. It was concluded that Twitter could serve as support for people who self-harm and could at the same time normalise self-harming behaviour and perpetuate it.				

(continued)

Table 3. *Main data of the selected articles (continued)*

No.	Title	Author(s)	Publication	Area of knowledge	Country
14	Visualización e interpretación de las interacciones en los mensajes de autolesiones no suicidas (ANS) en Twitter	Martínez-Pastor, E., et al.(2023)	Revista Redes	Communication sciences and IT	Spain
Main findings					
Through an analysis of the hashtags <i>#shtwt</i> and <i>#ouchietwt</i> (via an API) two types of conversations were detected: one between associations or other organisations that consume/produce NSSH-related content led by experts and focusing on the <i>#selfharm</i> hashtag; and another between young people and teenagers in which they shared their distress, among peers, focused around hashtags such as <i>#shtwt</i> and <i>#ouchietwt</i> .					
15	What twitter can tell us about user experiences of crisis text lines: A qualitative study	Coady, A., et al. (2022)	Internet Interventions	Psychology	Canada
Main findings					
The study analysed the experience of users who had resorted to the <i>@CrisisTextLine</i> service, which offers free help 24 hours a day, seven days a week, via text-based messages on Twitter about anxiety, depression, suicide and self-harm, among others. It was concluded that this was a good mechanism for providing solutions for the young and adolescent population.					

3. Results

The findings obtained from the review, with regard to the objectives set out in this paper, are the following:

3.1 Identification of scientific research on self-harm content created and shared by young people and teenagers on X/Twitter

A total of 15 articles were located and reviewed, the identifying details of which are provided in Table 3. In terms of the geographic origin of the studies by country,

five were from the United States of America (3, 4, 9, 11, 12), four from the United Kingdom (5, 7, 8, 13), one from Spain (14), one from Canada (15), one from Japan (6), one from Singapore (10) and one from Brazil (1).

3.2 Determination of the topics covered in these research studies and their specific domain.

A qualitative analysis was conducted (Gámez-Guadix et al., 2020) of the 15 articles selected to identify the main topics covered in the research. The results revealed four main topics:

1. Identification of the profiles, users and messages about self-harm (n=6).
2. Systems to detect and prevent self-harm (n=4).
3. Content analysis of the messages (n=2).
4. Analysis of the reception of the messages in the audience (n=1).

As regards the areas of knowledge of the studies, nine have been identified as multidisciplinary studies – Communication Sciences+IT (6, 14), Health sciences+IT (3, 4, 12), Psychology+IT (9, 10), Communication Sciences+Psychology (13), Law+Mathematics (11)– while six are attached to specific areas of knowledge – Health sciences (1, 7), IT (5, 8), Communication Sciences (2), Psychology (15).

3.3 Identification of the methods and tools they used to locate the self-harm content on X/Twitter, as well as the young people and teenagers who share it.

When selecting the elements from X/Twitter to be analysed, the research studies adopted different search and categorisation criteria, divided between those which focus on users (2, 7, 9, 10, 4), on hashtags (5, 14, 8), on keywords (1, 3, 11, 12, 13, 14) or on corporate accounts (15). There is a significant difference in the research studies in the first block, as the searches for NSSH-related words, expressions or hashtags are performed after selecting specific users, thereby limiting the

study population. Specific criteria are followed for the selection, such as having published multimedia NSSH content (2), trans users aged between 14 and 18 years old (7), users who have been diagnosed with depressive symptoms (9), self-reported depressed users (4), or registered users who attempted or committed suicide (10). As regards the hashtags or keywords selected, in general the most repeated NSSH-related terms are “self-harm” and “self-injury”, with and without a hyphen, although specific categorisation systems also feature, using tags (3), hashtags in the ambiguous language of the community of young people who self-harm such as *#shtwt* or *#ouchietwt* (14) or searches that include phrases or expressions containing several words for example, expressing a desire to hurt themselves (1).

With regard to the universe under analysis, the research studies focus on selecting tweets without specifying a target in the majority (75%) of the cases (1, 3, 4, 5, 6, 8, 9, 11, 12, 14, 15), and analyse the demographic or social characteristics after a screening process conducted through keywords or hashtags. Segmentation by target population, when performed, is carried out in accordance with the specific objectives of each study. In the review, we observed different target populations, such as young people aged 20 years old (2), trans youth aged between 14 and 18 years old (7), men and women aged between 17 and 54 years old (13) and X/Twitter users in Singapore who attempted or committed suicide (10).

3.4 Determination of the methodologies used to analyse the self-harm-related content shared by young people and teenagers on X/Twitter

As regards the first topic (identification of the profiles, users and messages about self-harm), all the articles reviewed, except one (6), use quantitative methodologies, focusing on analysing story-based fictional characters who demonstrate NSSH behaviours. The quantitative methodology strategies centre on searching for content or users related to NSSH, and include the use of APIs (3, 14), tools such as Twitter Archive Google Sheets (TAGS) (5), as well as content analysis of the tweets (7) and use of algorithm-aided topic modelling (5, 9). The qualitative methodologies applied are phenomenological approaches (2) and thematic analysis (3, 13) of the publications. Meanwhile, for the second topic (systems to detect and prevent

self-harm), quantitative methodology predominates (15, 10, 4, 12), with data mining techniques applied to text and the creation of predictive models (4, 10), whereas the qualitative methodology centres on qualitative content analysis of the tweets (15, 12). For topic 3 (content analysis of the messages) data mining and probabilistic data modelling techniques are again used (11), as well as qualitative data analysis (1). Lastly, for the fourth topic (analysis of the reception of the messages in the audience), a quantitative methodological approach has been found, based on a search for hashtags, in addition to content analysis (8).

4. Discussion

Numerous studies have shown how social media platforms are spaces used mostly by young people (Newman et al., 2023), and X/Twitter is among the platforms widely used by teenagers (Anderson, Faverio & Gottfried, 2023). There has been an emergence of **communities of young people** who share their thoughts about self-harm (Lerman et al., 2023). The characteristic of these communities is that they create their own language through hashtags, keywords and expressions, images or memes, excluding those who do not share their interests (Martínez & Gaete-Salgado, 2023). This is a **language** that a user needs to be familiar with in order to be able to draw nearer to the community, know what is being talked about and even be able to help them (Martínez & Gaete-Salgado, 2023; Guccini & McKinley, 2022). One of the pillars in building these communities is the development of their own language, which is evident in the hashtag. This reality is the starting point for nine of the studies when screening their analysis samples (1, 4, 7, 9, 10, 11, 12, 13, 15) mainly in the selection of keywords. Although there are studies where the keyword is a single word or a limited number of them – “self-harm” (13), “Crisis Text Line” (15), “self-mutilation” (1), “self-harm/non suicidal self-injury” (11) – in the majority of this group the keywords range in number from 10 to 20 (7, 9, 10, 12) and even more than 100 (4). Among this selection, the most repeated keywords are: “self-harm”, “anorexia”, “cut”, “bulimia”, “scar”, “razor”, “blade”, “suicide”, “anxiety”, “bruise”, “sadness”, “rip-skin” and “scratch”. The articles include research studies where screening is based on hashtags (3, 8, 14), with *#selfharm* and *#selfinjury* featuring repeatedly. In one case (14), up to 43

hashtags associated with three blocks – *#selfharm*, *#shtwt* and *#ouchietwt* – were retrieved (Martínez et al, 2024).

Against this background, the discussion among the articles analysed regarding use of X/Twitter and potential consequences (both positive and negative) for young people at risk of self-harm is particularly relevant. Firstly, we encounter research studies where X/Twitter is perceived as **positive** for people at risk of self-harm (7, 8, 12, 15), which highlight the potential of social media to enable these users to express their emotions and experiences without feeling judged (7), the contingency of using the information contained in the social network to create predictive models that help to prevent episodes of self-harm (12), or the opportunities for prevention-oriented organisations to accomplish advisory, support and awareness-raising activities (8, 15). The same number of studies (1, 2, 5, 13) opt to highlight the ambiguity of the use of social media and the consequences (**both positive and negative**) of NSSH-related posts on young people, with the latter including exposure to negative messages that can encourage future self-harming (5), difficulties in communicating between young people at risk of self-harm and their family environment (2), and understanding and feeling themselves to be considered a valuable individual in the community. These aspects could help to normalise self-harm and perpetuate the behaviour (1, 13). In just two research studies, the results present X/Twitter as a social media network with net **negative** effects by relating negative comments on the network with a greater incidence of depression markers (9) or by concluding that most of the comments on the social network are negatively labelled (29 negative tags compared to 2 positive tags) (3). Finally, there are studies where X/Twitter is not analysed positively or negatively, but from a **neutral** approach (4, 6, 10, 11, 14).

Furthermore, **studies** on self-harming communication or behaviour among young people and teenagers on X/Twitter draw from an abundance of theoretical approaches and analysis methodologies. Among the research, there are a greater number of studies focusing on developing **mathematical and computerised models** (4, 5, 9, 10), including the creation of automated algorithms or algorithms with well-known guidelines such as Machine Based Learning (Samoili et al., 2021). Therefore, and as has been described in the review of the existing literature, the use not only of different APIs but also of big data programs is now frequent when searching for and selecting information on social media. It is paramount to

identify this nascent digital community of young people which is growing around the subject of NSSH on social media, whether using hashtags or keywords.

Analysis of the aims of the analysed studies reveals that just one is directed towards assessing the **actions of a NSSH-related support organisation (15)** through its X/Twitter account, @CrisisTextLine, and it comes to the conclusion that the existence of this account has a **positive effect** on the young people who could perform self-harming behaviours, by being able to advise those young people and also those in their family and/or personal circle. **Prevention with regard to self-harm** is centred on creating models which, based on a series of interactions on X/Twitter, could calculate a platform user's potential risk of self-harm (4, 5, 9, 10). The results of these models are promising, with success rates of over 70% in the model that predicts suicidal tendencies (4). However, they are far from having achieved a degree of effectiveness that enables them to be adopted with assurance by public administrations or by public or private health services. Furthermore, no proposals of action relating to preventing behaviour that tends towards self-harming have been observed, despite express mentions of the importance of public health activities to monitor possible **self-harm content (1, 12)**. The corpus for this review has focused on analysing and interpreting user actions, postponing other types of initiatives such as digital literacy or developing guidelines such as manuals of best practices.

In accordance with Duarte, Noriega and Fregoso (2021), it is noteworthy that there are no studies that address this behaviour as a multi-factor and multi-cause phenomenon nor any that analyse socialisation contexts in academic and family-based environments or which describe risk factors or factors for protection associated with the use of the public space. There is also a complete absence of qualitative works about communication in the digital community. Due to all this, it would be fitting for there to be future research along these lines of work.

5. Conclusions

The main conclusion of the review performed on self-harm content on X/Twitter shared by young people and teenagers is that the scientific production on this specific topic is limited compared to research studies that focus on studying online

self-harm content (SCO-Self-harm Content Online) in different communities and social networks.

The majority of the studies found during the search process refer to NSSH as a factor associated with depression prior to suicidal thoughts and attempts, but never as a behaviour in itself. This seems to have stalled the study of self-harm, its prevention and its emergence among children and teenagers. Understanding why young people begin the practice of self-harming is key to addressing the need that triggers this behaviour in other ways which are not self-destructive. Technological tools are required to detect and locate the messages, which is only possible through combining the insights from different scientific disciplines. In fact, the scientific discipline which has most advanced the detection of online self-harm content is computing, through training artificial intelligence to process written language and interpret it contextually, which makes it possible to distinguish between literal and figurative use, for example. In this respect, an interdisciplinary approach which integrates proven methodologies that can be adapted to overcome the limitations found thus far is suggested.

At the same time, it must be taken into account that social media is evolving and young people and teenagers migrate to more attractive and apparently novel content platforms (like TikTok). Thus, the information contained in this study can serve as the basis for continuing to search the content shared in the digital environment in order to reach young people and teenagers and provide them with the help they need.

Authors' contribution

Esther Martínez-Pastor: Conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, visualization, writing - original draft, writing - review and editing. **María Ascensión Miralles González-Conde:** Conceptualization, formal analysis, investigation, methodology, visualization, writing - original draft, writing - review and editing. **Pablo Sánchez-López:** Conceptualization, formal analysis, investigation, methodology, visualization, writing - original draft, writing - review and editing. All the authors have read and accepted the published version of the manuscript.

Conflicts of interest

The authors have no conflicts of interest to declare.

Funding

This work has been funded by the National Plan from the Ministry of Science and Innovation (Spain), under which the project “*Representación mediática de las autolesiones de los menores en los medios de comunicación y Redes*” [“Media representation of self-harming by children in the media and social networks”] (PID2021-1245500B-I00) is being developed.

References

- Abi-Jaoude, Elia; Treurnicht Naylor, Karline; & Pignatiello, Antonio (2020). Smartphones, social media use and youth mental health. *Canadian Medical Association Journal (CMAJ)*, 192(6), E136–E141. <https://doi.org/10.1503/cmaj.190434>
- Alhassan, Muhammad Abubakar; Inuwa-Dutse, Isa; Bello, Shehu; & Pennington, Diane (2021). Self-harm: Detection and support on Twitter. En *ECSM 2021: 8th European Conference on Social Media*. <https://doi.org/10.48550/arXiv.2104.00174>
- Alhassan, Muhammad Abubakar; & Pennington, Diane (2021). Investigating non-suicidal self-injury discussions on Twitter. En *Proceedings of the International Conference on Social Media and Data Mining (ICSMDM)*. https://pure.strath.ac.uk/ws/portalfiles/portal/120117781/Alhassan_Pennington_ICSMDM_2021_Investigating_non_suicidal_self_injury_discussions_on_twitter.pdf
- American Psychiatric Association (2022). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR)*. <https://www.mignair/images/docs/files/000058/nf00058253-2.pdf>

- Anderson, Monica; Faverio, Michelle; & Gottfried, Jeffrey (2023). *Teens, Social Media and Technology 2023*. Pew Research Center. <https://www.pewresearch.org/internet/2023/12/11/teens-social-media-and-technology-2023/>
- Asociación Española de Pediatría (AEP) (2022). *Posicionamiento sobre el aumento de autolesiones y suicidio en niños y adolescentes*. Comunicado de la Asociación Española de Pediatría (AEP). <https://www.aeped.es/noticias/posicionamiento-sobre-aumento-autolesiones-y-suicidio-en-ninos-y-adolescentes>
- Asociación Española de Pediatría de Atención Primaria (AEPAP) (s.f.). <https://www.aepap.org/>
- Brown, Rebeca; Fischer, Tin; Goldwich, David; Keller, Ferdinand; Young, Robert; & Plener, Paul (2018). #cutting: non-suicidal self-injury (NSSI) on Instagram. *Psychological medicine*, 48(2), 337-346. <https://doi.org/10.1017/S0033291717001751>
- Cipriano, Annarosa; Cella, Stefania; & Cotrufo, Paolo (2017). Nonsuicidal self-injury: A systematic review. *Frontiers in Psychology*, 8:1946. <https://doi.org/10.3389/fpsyg.2017.01946>
- Coady, Alanna; Lainchbury, Keeley; Godard, Rebecca; & Holtzman, Susan (2022). What Twitter can tell us about user experiences of crisis text lines: A qualitative study. *Internet Interventions*, 28, 100526. <https://doi.org/10.1016/j.invent.2022.100526>
- Das Gollapalli, Sujatha; Zagatti, Guilherme Augusto; & Ng, See-Kiong (2021). Suicide risk prediction by tracking self-harm aspects in tweets: NUS-IDS at the CLPsych 2021 shared task. En *Proceedings of the Seventh Workshop on Computational Linguistics and Clinical Psychology: Improving Access* (pp. 93–98). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2021.clpsych-1.10>

- Duarte Tánori, Karen Guadalupe; Vera Noriega, José Ángel; & Fregoso Borrego, Daniel (2021). Factores contextuales relacionados a las conductas de autolesión no suicidas: una revisión sistemática. *Psicología y Salud*, 31(1), 51–59. <https://doi.org/10.25009/pys.v31i1.2675>
- Emparanza, José Ignacio, & Urreta, Iratxe (2005). Revisión sistemática y meta-análisis. *Anales de pediatría continuada*, 3(6), 1-9.
- Faura-Garcia, Juan; Orue, Izaskun; & Calvete, Esther (2021). Clinical assessment of non-suicidal self-injury: A systematic review of instruments. *Clinical alpsychology&psychotherapy*, 28(4), 739-765. <https://doi.org/10.1002/cpp.2537>
- Faura-Garcia, Juan; Orue, Izaskun; & Calvete, Esther (2022). Nonsuicidal Self-Injury Thoughts and Behavior in Adolescents: Validation of SITBI-NSSI. *Psicothema*, 34(4), 582-592. <https://doi.org/10.7334/psicothema2022.13>
- Fernández, Rosa (2024, 23 October). Número de usuarios mensuales activos (MAU) de X (Twitter) a nivel mundial desde 2010 hasta 2023. *Statista*. <https://acortar.link/GhuHyD>
- Fundación ANAR (2023). Informe anual. Teléfono chat ANAR. <https://consejoaudiovisualdeandalucia.es/2024/06/04/informe-anual-del-telefono-chat-anar/>
- Gámez-Guadix, Manuel; Almendros, Carmen; Rodríguez-Mondragón, Laura; & Mateos-Pérez, Estíbaliz (2020). Autolesiones online entre adolescentes españoles: análisis de la prevalencia y de las motivaciones. *Revista de Psicología Clínica con Niños y Adolescentes*, 7(1), 9-15. <https://doi.org/10.21134/rpc-na.2020.07.1.1>
- Guccini, Federica; & McKinley, Gerald (2022). “How deep do I have to cut?”: Non-suicidal self-injury and imagined communities of practice on Tumblr. *Social Science & Medicine*, 296, 114760. <https://doi.org/10.1016/j.socscimed.2022.114760>

- Gustina Sari, Genny; & Wirman, Welly (2022). Communication Patterns of Adolescent Self-Harm Suffering in Interpersonal Relationships. *Journal Kajian Komunikasi*, 10(29). <https://doi.org/10.24198/jkk.v10i1.29384>
- Hilton, Charlotte (2017). Unveiling self-harm behaviour: what can social media site Twitter tell us about self-harm? A qualitative exploration. *Journal of Clinical Nursing*, 26(11-12), 1690-1704. <https://doi.org/10.1111/jocn.13575>
- Jashinsky, Jared; Burton, Scott H.; Hanson, Carl L.; West, Josh; Giraud-Carrier, Christophe; Barnes, Michael; & Argyle, Trenton (2014). Tracking suicide risk factors through Twitter in the US. *Crisis*, 35(1), 1-9. <https://doi.org/10.1027/0227-5910/a000234>
- Khasawneh, Amro; Madathil, Kapil Chalil; Zinzow, Heidi; Wisniewski, Pamela; Ponathil, Amal; Hunter, Rogers; Agnisarman, Struthy; Roth, Rebecca; & Narasimhan, Meera (2021). An investigation of the portrayal of social media challenges on YouTube and Twitter. *ACM Transactions on Social Computing*, 4(1), 1-23. <https://doi.org/10.1145/3444961>
- Klonsky, Eric David (2011). Non-suicidal self-injury in United States adults: prevalence, sociodemographics, topography and functions. *Psychological Medicine*, 41(9), 1981-1986. <https://doi.org/10.1017/S0033291710002497>
- Kuric, Stribor; Sanmartín, Anna; Ballesteros, José Carlos; & Gómez Miguel, Alejandro (2023). *Barómetro Juventud, Salud y Bienestar 2023*. Centro Reina Sofía de Fado Juventud. <https://doi.org/10.52810.5281/zenodo.8170910>
- Leiva, Oscar (2023, 9 October). *Autolesiones. Es importante resaltar que la mayoría de personas que se provocan autolesiones no quieren terminar con sus vidas*. UNICEF. <https://acortar.link/h3QjMb>
- Lerman, Kristina; Karnati, Aryan; Shuchan, Zhou; Chen, Siyi; Kumar, Sudesh; He, Zihao; Yau, Joanna; & Horn, Abigail (2023). *Radicalized by Thinness: Using a*

Model of Radicalization to Understand Pro-Anorexia Communities on Twitter. Cornell University. <https://arxiv.org/abs/2305.11316>

López Vázquez, Paula; Armero Pedreira, Paula; Martínez-Sánchez, Lidia; García Cruz, Jose Manuel; Bonet de Luna, Concha; Notario Herrero, Félix; Sánchez Vázquez, Ana Rosa; Rodríguez Hernández, Pedro Javier; & Díez Suárez, Azucena (2023). Autolesiones y conducta suicida en niños y adolescentes. Lo que la pandemia nos ha desvelado. *Anales de pediatría*, 98(3), 204-212. <https://doi.org/10.1016/j.anpedi.2022.11.006>

Martínez-Pastor, Esther; Atauri-Mezquida, David; Nicolás-Ojeda, Miguel Ángel; & Blanco-Ruiz, Marian (2023). Visualización e interpretación de las interacciones en los mensajes de autolesiones no suicidas (ANS) en Twitter. *Redes. Revista hispana para el análisis de redes sociales*, 34(2), 238-253. <https://doi.org/10.5565/rev/redes.996>

Martínez-Pastor, Esther; & Gaete-Salgado, Catalina (2023). Young content creators around self-harm: identifying metalanguages on X (Twitter). *Revista panamericana de comunicación*, 5(2), 55-70. <https://doi.org/10.21555/rpc.v5i2.2984>

Martínez-Pastor, Esther; Nicolás-Ojeda, Miguel Ángel; Blanco-Ruiz, Marian; & Moraleda, Raquel (2024). *Guía abre los ojos ante las autolesiones en redes sociales*. <https://burjcdigital.urjc.es/handle/10115/39148>

Ministerio de Sanidad (2021). *Estrategia de Salud Mental del SNS, período 2022-2026*. Ministerio de Sanidad. Secretaría General Técnica. Centro de publicaciones. <https://acortar.link/2FPpOV>

Newman, Nic; Fletcher, Richard; Eddy, Kirsten; Robertson T., Craig; & Kleis Nielsen, Rasmus (2023). *Digital News Report 2023*. Reuters Institute. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2023-06/Digital_News_Report_2023.pdf

- Page, Matthew J.; McKenzie, Joanne E.; Bossuyt, Patrick M.; Boutron, Isabelle; Hoffmann, Tammy C.; Mulrow, Cynthia D., et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372(71). <https://doi.org/10.1136/bmj.n71>
- Pérez, Sandra; García-Alandete, Joaquín; Gallego, Blanca; & Marco, José H. (2021). Characteristics and Unidimensionality of Non-suicidal Self-injury in a Community Sample of Spanish Adolescents. *Psicothema*, 33(2), 251–258. <https://doi.org/10.7334/psicothema2020.249>
- Perry, I'Sis; & Park, Albert (2021). Exploring Suicidality on Social Media: Qualitative Analysis of Twitter. En *IEEE 9th International Conference on Healthcare Informatics (ICHI)* 362-371. IEEEComputerSociety. <https://doi.org/10.1109/ICHI52183.2021.00061>
- Sánchez-Meca, Julio (2022). Revisiones sistemáticas y meta-análisis en Educación: un tutorial. *RiiTE, Revista interuniversitaria de investigación en Tecnología Educativa*, 13, 5-40. <https://doi.org/10.6018/riite.545451>
- Samoili, Sofia; López Cobo, Monserrat; Delipetrev, Blagoj; Martinez-Plumed, Fernando; Gómez, Emilia; & De Prato, Giuditta (2021). *AI Watch. Defining Artificial Intelligence 2.0. Towards an operational definition and taxonomy of AI for the AI landscape*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2760/019901>
- Seko, Yukari; & Kikuchi, Minako (2022). Mentally Ill and cute as hell: Menhera girls and portrayals of self-injury in Japanese popular culture. *Frontiers in Communication*, 7: 737761. <https://doi.org/10.3389/fcomm.2022.737761>
- Silva, Aline Conceição; Giacchero Vedana, Kelly Graziani; Pereira dos Santos, José Carlos; Cristina Pillon, Sandra; Arena Ventura, Carla Aparecida; & Inocenti Miaso, Adriana (2021). Analysis of nonsuicidal self-injury posts on Twitter: A

quantitative and qualitative research. *Research, Society and Development*, 10(4), e40410413017. <https://doi.org/10.33448/rsd-v10i4.13017>

Simms, Drew (2020). Peer responses to trans youth tweeting about self-harm and suicidality. *Creative nursing*, 26(2), 135-142. <https://doi.org/10.1891/CRNR-D-19-00089>

Smith, Hayden; & Cipolli, William. (2022). The Instagram/Facebook ban on graphic self-harm imagery: A sentiment analysis and topic modeling approach. *Policy & Internet*, 14(1), 170-185. <https://doi.org/10.1002/poi3.272>

Twitter (2023, 2 February). Twitter's Sensitive Media Policy. <https://changewatch.co.uk/twitter/sensitive-media-policy/2023-02-02/>

UNICEF (2021a, 4 March). *Por lo menos 1 de cada 7 niños y jóvenes ha vivido confinado en el hogar durante gran parte del año, lo que supone un riesgo para su salud mental y su bienestar*. [Comunicado de prensa]. <https://acortar.link/p5ddej>

UNICEF (2021b). *Estado Mundial de la infancia 2021: En mi mente: Promover, proteger y cuidar la salud mental de la infancia*. [Informe principal]. <https://acortar.link/Q244Vj>

Wang, Yilin; Tang, Jiliang; Li, Jundong; Boaxini, Li; Wan, Yali; Mellina, Clayton, O'Hare, Neil; & Chang, Yi (2017). *Understanding and discovering deliberate self-harm content in social media*. In Proceedings of the 26th international conference on World Wide Web. 93-102. <https://doi.org/10.1145/3038912.3052555>

Yadav, Shweta; Chauhan, Jainish; Sain, Joy Prakash; Thirunarayan, Krishnaprasad; Sheth, Amit; Schumm, Jeremiah (2020). Identifying depressive symptoms from tweets: Figurative language enabled multitask learning framework. *Proceedings of the 28th International Conference on Computational Linguistics*, 696-709. <https://doi.org/10.18653/v1/2020.coling-main.61>

- Yazdavar, Amir Hossein; Al-Olimat, Hussein Saif; Ebrahimi, Monireh; Bajaj, Goonmeet; Banerjee, Tanvi; Thirunarayan, Krishnaprasad; Pathak, Jyotishman; & Sheth, Amit (2017). Semi-supervised approach to monitoring clinical depressive symptoms in social media. En *Proceedings of the 2017 IEEE/ACM international conference on advances in social networks analysis and mining 2017*, 1191-1198. <https://doi.org/10.1145/3110025.3123028>
- Yepes-Nuñez, Juan José; Urrútia, Gerard; Romero-García, Marta; & Alonso-Fernández, Sergio (Trad.) (2021). Declaración PRISMA 2020: una guía actualizada para la publicación de revisiones sistemáticas. *Revista Española de Cardiología*, 74(9), 790-799. <https://doi.org/10.1016/j.rec.2021.07.010>