CUADERNOS.INFO Nº 49
Versión electrónica: ISSN 0719-367x
http://www.cuadernos.info
https://doi.org/10.7764/cdi.49.27809

Received: 11-26-2020 / Accepted: 04-19-2021

Scattered Images: scrapers and the representation of cancer on Instagram

Imágenes desgarradas: el uso de *scrapers* en investigación social en Instagram sobre cáncer

Imagens rasgadas: o uso de scrapers na pesquisa social sobre câncer no Instagram

Miguel Varela-Rodríguez, Universidad de Valladolid, Valladolid, España (miguel.varela@uva.es)

Miguel Vicente-Mariño, Universidad de Valladolid, Valladolid, España (miguel.vicente@uva.es)

ABSTRACT | The use of visual social networks like Instagram in health communications is well-documented, especially when it comes to the analysis of specific contents to study images. However, this methodology poses a challenge due to the growing difficulties in access and a very limited legal and action framework. Based on the assumptions of visual sociology, this article explores a methodology to obtain data from Instagram using scrapers, offering a revision of its technical and ethical implications. Through this method, we analyse the distribution of Instagram posts carrying the hashtag #SacaPecho (a Spanish expression that refers to standing up for or showing support to a cause), created by the Spanish Society Against Cancer (Asociación Española Contra el Cáncer) on the occasion of the International Day Against Cancer (October 19, 2020). We obtain data from more than 7,000 images in a matter of minutes. Our work sheds light on the tools available for researchers to access data from Instagram and proposes a debate on the ethical possibilities to implement them.

KEYWORDS: social media research; health communication; visual sociology; cancer; computational methods in social science.

HOW TO CITE

Varela-Rodríguez, M. & Vicente-Mariño, M. (2021). Imágenes desgarradas: el uso de *scrapers* en investigación social en Instagram sobre cáncer. *Cuadernos.info, (49)*, 72-97. https://doi.org/10.7764/cdi.49.27809

RESUMEN | *El uso de redes de contenido visual como Instagram está bien documentado en la comunicación en salud, especialmente el análisis de contenido para estudiar las imágenes. Sin embargo, esta metodología supone un reto ante las crecientes dificultades en el acceso y un marco legal y de actuación muy limitados. Basado en los postulados de la sociología visual, este artículo explora una metodología para obtener datos de Instagram mediante el uso de scrapers, revisando las necesidades técnicas y las implicaciones éticas en el uso de este tipo de herramientas. Se analiza la distribución de imágenes acompañadas por la etiqueta #SacaPecho, creada por la Asociación Española Contra el Cáncer con ocasión del Día Internacional de la Lucha Contra el Cáncer (19 de octubre de 2020). El uso de scrapers permite obtener referencias de más de 7000 imágenes en poco tiempo. El trabajo permite entender las herramientas al alcance de la investigación social para acceder a datos relevantes en Instagram y propone un debate sobre las posibilidades éticas en este ámbito.*

PALABRAS CLAVE: *investigación en redes sociales; comunicación en salud; sociología visual; cáncer; métodos computacionales en investigación social.*

RESUMO | A utilização de Instagram e de redes semelhantes de conteúdo visual na comunicação sobre saúde está bem documentada, especialmente na análise de conteúdo. Contudo, as crescentes limitações no acesso a esta informação representam um desafio. Baseado na sociologia visual, este artigo explora uma metodologia para obter dados da Instagram através da utilização de scrapers. Revemos as necessidades técnicas e as implicações éticas da utilização deste tipo de ferramenta. Através desta metodologia, analisamos a distribuição de imagens acompanhadas pela hashtag #SacaPecho ("mostre o seu apoio") na Instagram. A campanha foi criada pela Associação Espanhola Contra o Cancro por ocasião do Dia Internacional contra o Câncer (19 de outubro de 2020). A utilização de scrapers permite a obtenção de referências de mais de 7.000 imagens num curto espaço de tempo. O trabalho permite uma compreensão dos mecanismos à disposição da investigação social para acessar a dados relevantes no Instagram e propõe um debate sobre as possibilidades éticas nesta área.

PALAVRAS CHAVE: investigação em redes sociais; comunicação em saúde; sociologia visual; câncer; métodos computacionais na investigação social.

INTRODUCTION

Despite major advances in the study of social networks over the past decade, automated image retrieval and analysis is, in many cases, a challenge for social research. Since Facebook and other providers shut down access to their platforms following the Cambridge Analytica case in 2018 (Bruns, 2019; Rogers, 2018), finding and using network images for research has become more difficult due to technical difficulty and various ethical considerations. Often, researchers are involved in a manual process of identification, selection, and processing, investing large amounts of time in cropping, copying, and pasting images into their computers. While some seek to establish partnerships with Facebook through contests –a costly endeavor, as discussed below– others are faced with using tools that speed up their work, but in many cases break the Terms of Use of platforms such as Instagram. Some advocate abandoning this type of work until the scenario is more favorable. In the meantime, images continue to be a fundamental part of the modes of expression and exchange in contemporary society.

Considering the above, this article reviews the use of a simple technique to search and obtain images on Instagram as a viable alternative after the closure of public APIs in 2018 using two scraping tools known as Instaloader and Instalooter. To do this, it uses a case study: the Spanish Association Against Cancer's (AECC, by its Spanish acronym) campaign around World Cancer Day with the hashtag #SacaPecho. This type of campaign generates large amounts of visual online content. We explore the specific steps to design a search for images uploaded to Instagram with a hashtag on specific dates, we analyze ethical considerations, and propose alternatives.

The objectives are to demonstrate the opportunities available to social researchers to access network images and to stimulate a debate on the difficulties (technical and ethical) linked to the use of such tools and the role of platform providers such as Instagram.

THEORETICAL FRAMEWORK: WAYS OF SEEING IN THE DIGITAL WORLD

The image as a source of sociological information

The study of images and of the social sciences have a long and complex relationship, greatly facilitated by anthropology in the early 20th century. The democratization in access to cameras towards the end of the 1960s, the cheapening of color printing, and the pivotal role of visual media in the Vietnam War (Berger, 2016; Sontag, 2017) impacted the crossover between the study of images and sociology. This opened a space for some sociologists who "rejected the dominant paradigms in both research and theory" (Harper, 1988, p. 59).

Gradually, the visual would become a mode of storytelling, but also a key source of information to (re)construct them. Both are basic functions explored, since 1981, by the International Visual Sociology Association through visual sociology, and permeate much of the work with images in social sciences (Harper, 1996; Pauwels, 2015), especially from computational analysis.

The image in networks and Instagram, beyond content analysis

With the advent of democratized photography, the latter became an activity common to millions of individuals and families, to the point of becoming "image junkies" (Sontag, 2017, p. 33). The image was transformed into a mass-produced and mass-consumed good. Early visual sociologists were aware of this situation and seized it as an opportunity. They worked with a multitude of found photographs, family archives, and other image banks to review the meanings and dynamics contained in the image[1]. The importance of the image in social networks goes beyond the represented and includes representation (Pauwels, 2015). While the analysis of the represented refers to reviewing what the images show, the study of representation has to do with the decisions at the time of producing the images and their modes of circulation.

Rose (2019) refers to this differentiation as the four domains of the image study. In that approach, the image can be studied from its contents, its mode of production, its mode of circulation, or its mode of visualization. Beyond looking for what is represented in the image, Rose is concerned with who produced it, how, why, how it is shared, and under what circumstances consumers approach that image.

The volume of available images in social networks is overwhelming, with more than 95 million uploaded every day to Instagram (Abutaleb, 2016); thus, the study of the modes of production and circulation becomes more relevant and complex. Hand (2016) asserts that "it could be said that the numbers of 'found' images mean we have to abandon any idea of analyzing individual ones" (p. 216). Authors such as Manovich (2018) and followers of cultural analysis easily adapt to this situation. In addition to the image's content and form, they are interested in its means of production and, especially, its means of circulation. In their proposal for a cultural analytics, they understand that the volume of images is such that an individual categorization loses meaning and limits the researcher's analytical capacity, but today we as academic researchers "live in the 'shadow' of a world of social networks, recommendations, apps, and interfaces that all use media analytics" (Manovich, 2018, p. 2).

---

1. . See, as an example in social networks, Manovich (2016), González-Requena (1988, 2013), with their cinema examples, or Abril-Curto (2013) on political communication.

This way of approaching networks, favoring an analysis of production and circulation modes, without intervening on the individual texts, has had particular importance in networks with open programming interfaces, such as Twitter, but its use with Instagram is limited. Partly, as Rose explains, because the images allow the identification of the individuals represented, which raises ethical issues (Rose, 2019), and partly, however, because of economic issues.

On Twitter or YouTube, studies have been particularly fruitful when trying to analyze volume and content in political communication (KliglerVilenchik et al., 2020), in gender studies (Giraldo-Luque et al.,2018 ; Regueira, Ferreiro, & Vila, 2020) –applying web scraping techniques to the study of women's representation on YouTube or screen scraping to analyze more than 100 thousand tweets in the #NiUnamenos movement–, as well as during the COVID-19 pandemic (Babvey et al., 2020; Cinelli et al., 2020; Orduña-Malea et al., 2020). In fields such as marketing, computational network analysis is an established practice (Harrigan et al., 2020; Purba et al., 2020; Tafesse & Wood, 2020).

These methods have a high level of development and free and open-source options (Arcila-Calderón et al. 2019), and reduce the workload and time required to obtain and study social network data.

On Instagram, however, the research has not followed this evolution, but has a markedly qualitative approach. For example, in the analysis of hand-extracted selfies (Abidin, 2016; Couture Bue, 2020), or through the then available Instagram API (Jang et al., 2015). In 2016, Munk and colleagues claimed that "social research on Instagram is still in its infancy" (2016, p. 2), and encouraged the use of Instagram beyond the specific image and toward social network analysis. While there are papers that develop methods to automatically obtain information from Instagram, they either predate the closure of the social network's APIs in 2018 (Highfield & Leaver, 2014) or do not explore the ethical implications (Dewi et al., 2019). Nor do they respond to the emergence of algorithms in social networks as a defining moment of our era (Schie, Westra, & Schäfer, 2017).

Despite these dynamics, social research on images in social networks continues to be limited by commercial dynamics and ethical debates, by the blocking of access, and by the lack of definition and uncertainty about what is and is not acceptable when studying them.

Scraping as an image-retrieval method and the ethical debate it raises

Scraping is one of the techniques for obtaining images from online sources emerged from the advent of social networks and the explosion of image production since the early 2000s. The term refers to the way these tools operate: based on

a set of parameters, they scrape all the content of a webpage that matches those parameters and make it available to the user.

Scraping methods are described in detail by Marres and Weltevrede (2013), Hirschey (2014) and, in the context of YouTube, Regueira and colleagues (2020). Scraping tools can access a given web page (such as the Instagram page dedicated to the #cancer tag) and download all its content, including images, accompanying text and their metadata, via an HTML connection. This process is known as web scraping, where the tool connects directly to the provider's database (Instagram, in our study). An alternative method is screen scraping: instead of connecting to the database, the tool captures all the content loaded on the user's screen. For browsers such as Google Chrome, Google's own extension store includes tools to perform this type of extraction.

In general, scraping is a technique with a questionable reputation, often linked to illegal practices (Marres & Weltevrede, 2013). Not for nothing is it a technique banned on some platforms, including Instagram according to its terms of use (Instagram, 2018), and which has been severely curtailed following the Cambridge Analytica case in 2018. There are several reasons for this, ascribed primarily to two risk areas of risk: for the user, and for the platform provider.

The former has to do with key issues in social studies with images, already recognized by Becker (1974) in his proposal for a visual sociology: privacy, consent, and inclusion. Anthropologists, ethnographers, and sociologists are no strangers to these debates. Many papers explore the ethical implications of using images in social research (Markham & Buchanan, 2012, 2017; Prosser et al., 2008; Rose, 2019). Several explore this issue in relation to Instagram and social media specifically (Henricks, 2017; Marres & Weltevrede, 2013; Walsh & Baker, 2017; White & Boatwright, 2020).

In addition to ethical considerations, Prosser and colleagues (2008) invite researchers to review the legal framework for their research. For Instagram, it includes reviewing the Terms of Use and making sure that the use of these techniques is permitted (it is not). Schie and colleagues (2017) lay out the research process in three phases, in which they call for "a moment of ethical reasoning" (p. 196): research design, in which questions related to potential subjects (of the images, in our case) are addressed; safe exploration of the data, with clear documentation of the steps followed and storing the data in a protected environment, and the final retrieval process.

This is where the second area of risk comes in: the ownership of images, the intervention of commercial entities such as Facebook, Google and other large

companies, and the protection of their business model. The debate around the role of these companies in delimiting what is and what is not searchable is heated. This issue is captured in the ethical frameworks of Townsend and Wallace (2016) and Prosser and colleagues (2008), who start their model with the inclusion of legal considerations.

Beyond the ethical argument, the legal environment in scraping images from social networks complicates the researchers' work, especially since providers such as Facebook shut down their APIs in the wake of the Cambridge Analytica case. Hirschey details the variety of options available to large companies, such as "overwhelm the legal resources of scrapers and quickly drive them out of business" (2014, p. 925). It thus becomes a legal and business battle that has little to do with research ethics, as denounced by Rogers (2018) or Bruns (2019), and in which researchers must deal with economic logic.

Cancer and health education as an example of the importance of online images

Image has played a particularly important role in the field of health communication and education; more specifically, communication around cancer. As early as 1978, Susan Sontag reflected on the impact breast cancer patients' images (Sontag, 1978). Similarly, Ehrenreich (2001) makes a fierce criticism of the cult of pink that she encountered after her breast cancer diagnosis. The image of cancer has been studied in specialized magazines and media (Grant & Hundley, 2008), in awareness campaigns (Cartwright, 1998) and, more recently, in social networks such as Instagram (Basch & MacLean, 2019). There is a predomination of work devoted to discourse analysis (Döbrössy et al., 2020; Gibson et al., 2016; Sutton et al., 2018; Zhang et al., 2019), viralization dynamics (Noar et al., 2018; Wang et al., 2019), or physiological response to images (Chou et al., 2020).

These works show the growing importance of social networks as an area of consumption of information related to cancer, to the point of having become the second medium of information for patients (Blanch Hartigan & Viswanath, 2015). Pardo (2019) highlights the role of networks in the construction of a more human image that is less dominated by medicine and depersonalization, and reflects on the areas of production and circulation already highlighted by Rose.

Cancer is the subject of important photographic and video documentaries, such as those by Jo Spence (1986), Angel Merendino[2] and Jeremy Nicholl[3].

2. https://www.mywifesfightwithbreastcancer.com/

3. http://mybeautifulcancer.com/

Documentary and analytical works share an approach: that of individual image analysis. Through discourse and content analysis, they develop cancer stories in an almost ethnographic effort, adding narrative value. However, in preliminary scans conducted by the authors of this article in 2019, more than 40 million images with keywords associated with cancer were detected on Instagram alone. Such a volume of cancer-related images can overwhelm any attempt at this type of social media analysis.

METHODOLOGY

For the purpose of this article, we explored images accompanied by the hashtag #SacaPecho on Instagram during October 19, 2020, International Breast Cancer Awareness Day. Said tag was launched by the Spanish Association Against Cancer (Asociación Española Contra el Cáncer, 2020), one of the largest non-profit organizations in the country and with a great importance in the social image and research funding against this disease. #SacaPecho, part of their campaign to raise awareness about breast cancer that year, achieved thousands of impressions in the first few hours. We will use the scraping technique to analyze the dynamics of publication around this hashtag and its key date, detailing each step followed. The Python programming language was used to obtain the data.

Tools used and installation

Python (Miniconda) and Pip

Python is a programming language that many social researchers are familiar with as an alternative to R for statistical analysis. It can be downloaded as a complete package using Miniconda, a simple installer that includes the latest version available, plus other useful tools. Although there are other ways to install Python, Miniconda is arguably the most intuitive and straightforward. It is available, free of charge, from its website[4]. On a PC, Miniconda will appear installed as Anaconda Prompt (miniconda 3)[5].

Pip is a package installer for Python –a tool used to add components that use this language and that add functionalities such as the production of graphs from the data obtained or the definition of search parameters. Pip allows to download codes developed by third parties and install them with a simple line of text through the previously installed Anaconda Prompt (Miniconda 3). It is integrated in Miniconda and we used it to install our scraping tools.

4. https://docs.conda.io/en/latest/miniconda.html

5. Python 3.9 version to the date of this article's writing.

To install both tools, simply download the installation package from the webpage and follow the steps indicated.

Jupyter Notebook

The Python programming environment can be implemented from the command prompt or from a dedicated word processor, a simpler and more convenient option. In this case, we will use Jupyter Notebook, a tool developed specifically to facilitate working with Python. It is an intuitive text editor. To install it, you just need to:

1. Open Anaconda Prompt (Miniconda 3), previously installed. It will open a black screen in which we will be able to introduce text.
2. Enter this code:

   conda install -c conda-forge jupyterlab

With this, Jupyter Notebook will be installed and ready to use. See below for details on how to start it.

Instaloader e Instalooter

Instaloader and Instalooter are scraping tools that connect directly to the Instagram database, without the need to use an official interface. Despite breaking Instagram's Terms of Use, both are, to the date of this article's writing, available through Pip, and updated regularly. The steps for installation are simple:

1. Open Anaconda Prompt (Miniconda 3).
2. Introduce the following codes:

   pip install instaloader[6]pip install instalooter

With this, the scraping tools are installed. More information is available on the tools' pages:

https://instaloader.github.io/installation.html

https://instalooter.readthedocs.io/en/stable/install.html

6. Instaloader often results in errors if used in a browser from which we have recently accessed Instagram. Most errors are solved by upgrading Instaloader to its latest version (pip install -upgrade), updating Python (conda update python), avoiding massive downloads, and simply waiting a few hours in case of a crash.

Step 1. Building a folder to work

Once the working tools have been installed, the first step is simply to create a folder in the desired directory. All the data obtained will be saved in this folder. In our example:

C:/Users/Test/Imágenes/SacaPecho

Step 2. Search design

Using the research design process detailed by van Schie and colleagues (2017), we propose to design a tag search employing the installed scraping tools in three steps:

2.1. Define the hashtag to be studied: we are interested in images accompanied by #SacaPecho due to the importance of the Spanish Association Against Cancer and the high volume of images related to breast cancer observed in exploratory stages.

2.2. Define search dates: October 19, 2020, International Breast Cancer Awareness Day, a key date in the #SacaPecho campaign.

2.3. Get an idea of how many publications have that tag. Access the Explore page for #SacaPecho selected to understand the overall volume of images: sacapecho/. 12,792 posts as of October 30, 2020.

That much is the information provided by Instagram in its search engine. For any additional steps we will need a scraping tool such as Instaloader or Instalooter.

Step 3. Launch Jupyter Notebook in the desired folder

With the search parameters defined, in this third step we will be able to obtain a list with the metadata of all the publications (or posts) made between two specific dates and carrying a specific tag. To do this, we will use the text editor Jupyter Lab. We launch Jupyter-Lab through Anaconda Prompt:

[Open Anaconda Prompt (Miniconda)] cd c:/Users/Test/Images/Instaloader
jupyter-lab

Entering both commands will open the default web browser with the Jupyter Lab editor. Once inside, select Notebook > Python and we are ready to program the search.

Step 4. Download metadata

To obtain the metadata of all images with the hashtag #SacaPecho uploaded on October 19, 2020, we developed a code using both scraping tools. The complete code is available upon request to the authors and is composed of two parts: a first performs the metadata search and records them. The second part downloads a sample of images. Thus, the code allows defining parameters such as the search

tag, the date range, and a file in which to save the data obtained. In addition, all the necessary Python packages have already been incorporated. It is important to note that the developed code only downloads data from public profiles[7].

```
#We define the hashtag (desired_tag) desired_tag="sacapecho [or another hashtag]"

#We define the sample_size. It Will be used to later download an images' sample. In no case should more than 100 be downloaded in a test, to avoid overloading the system or being blocked by Instagram.

sample_size=50[or another number]

#We define the dates (SINCE and TO).

SINCE = datetime(2020, 10, 19) # Farther from today, inclusive

TO = datetime(2020, 10, 20) # Closer to today, not inclusive # Define the destination folder.

destination="[DIRECTORIO COMPLETO DE LA CARPETA DE TRABAJO]"

# We define the file where the results will be saved with the extension .txt. It is important not to forget to change it if another job is to be launched, to avoid overwriting the results.

file_to_record="[ DESIRED FILE NAME].txt"
```

It is also possible to limit the search to videos, stories, images, or reels (short-format videos). It is also possible to modify the metadata to be downloaded:

```
else: # the data to be requested are indicated below

print(post.shortcode,",",post.date,",",post.profile,",",post.likes)[8]
```

Step 5. Downloading images

Once the metadata has been obtained, we can download a sample of images to check that the desired results are being collected. This fifth step is contained in the second code cell accompanying the article, and uses the Instalooter tool,

7. Downloading private profiles is possible if the user accesses its own Instagram account via Python, something that is not explored in this article.

8. These parameters are listed in the Instaloader documentation, but the most common for a simple search will be post.shortcode (a number identifying the image) and post.date (upload date). Any other parameter will slow down the search and should be edited to avoid exposing details of the accounts studied.

which allows us to set a maximum number of downloads between two dates. We suggest a maximum of 50, to avoid overloading the system.

With Jupyter, we can launch both codes in a single session and start getting results quickly. The complete code is documented and contains simple parameters to change to modify the search.

RESULTS

The implementation of the developed code yielded data from 7556 images uploaded to Instagram on October 19, 2020, with the hashtag #SacaPecho. A first search, limited to the image ID and the date and time of upload, took 12 minutes to produce a text file with the desired information. A more extensive search, including the number of Likes and the profile that posted it, took more than five hours. We assume this discrepancy is due to Instaloader spacing out requests for this type of data to prevent Instagram from blocking access to it.

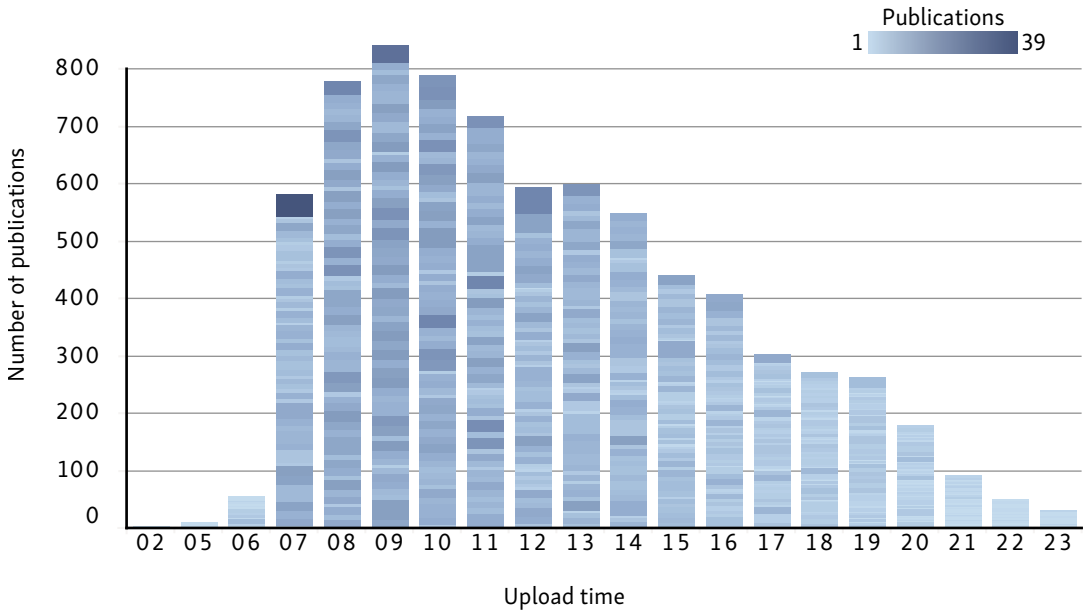
From the visual representation of the data obtained we can extract the hourly distribution of the publications.

Figure 1 shows the number of posts in one-hour slots, detailing the minute of upload. Publications with the hashtag #SacaPecho are maintained throughout the day, with the highest peak between 09:00 and 10:00. Most posts are observed at sharp hours (07:00, 08:00, 09:00, etc.), with the highest number of simultaneous posts at 07:00, suggesting that a significant volume may be posts scheduled to be uploaded at specific times, either by organizations or individuals. It is not in vain that we observe here publications of the different profiles (national and provincial) of the Spanish Association Against Cancer itself, possibly programmed as part of the campaign.

Using the additional metadata, we obtained the name of the account that uploaded each image, with a total of 6840 different accounts. It is relevant to note the distribution of posts by account (figure 1), contrasting the volume of posts by accounts associated with the AECC (containing AECC in their name) and by external accounts. This suggests a great capacity for campaign outreach.

Also noteworthy is the number of likes accumulated by the images depending on the time of publication and, especially, on the account from which the publication originates (figure 2).

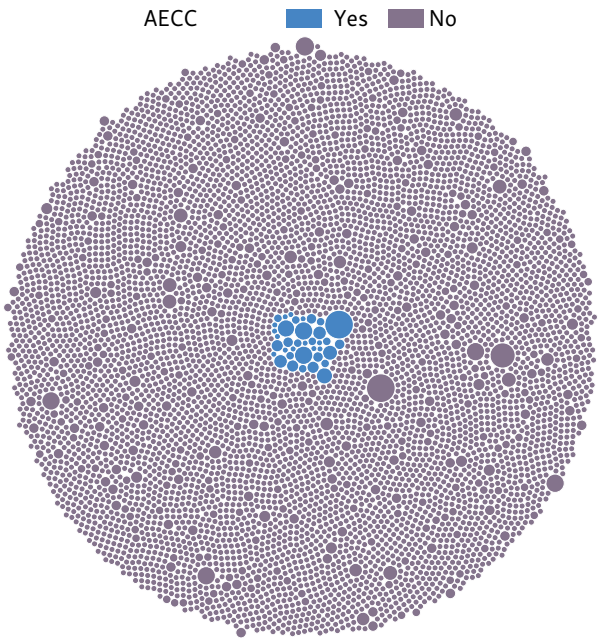
In this regard, the presence of what could be classified as super influencers is undeniable: accounts that reach an extremely high level of visibility and Likes, especially compared to those obtained by the publications of average users and the AECC itself.



The subdivisions indicate the distribution per minute of upload. A more intense color and a larger size indicates a greater number of images uploaded in the indicated minute, starting from minute 00 and ending at minute 59 along the X-axis.

Chart 1. Time and minute of posts uploading to Instagram with the hashtag #SacaPecho on October 19, 2020

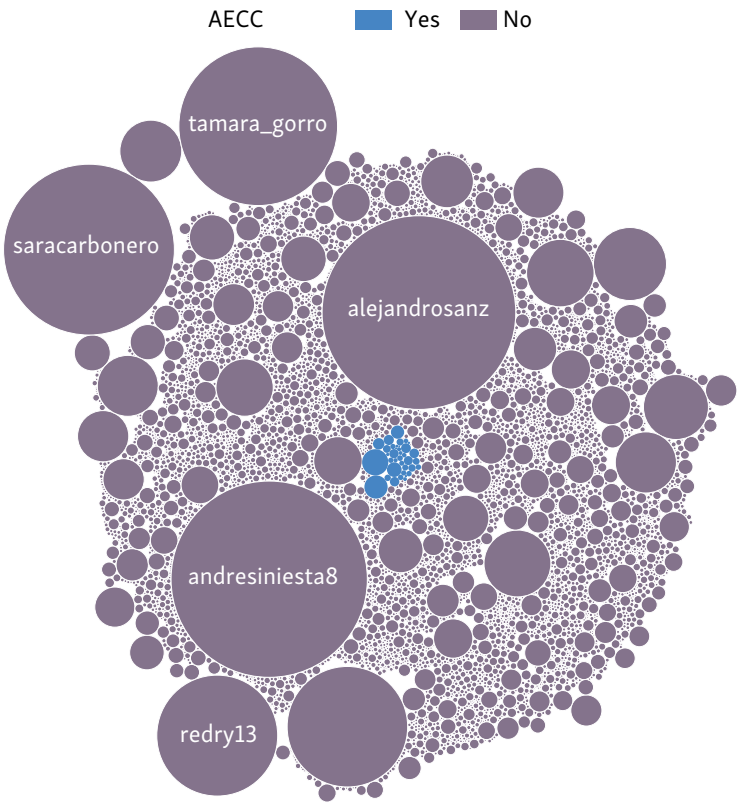
Source: Own elaboration with data from Instagram.



Profiles containing the word AECC in their name are shown grouped in a different color. A larger bubble indicates a greater number of publications.

Figure 1. Distribution by number of publications on Instagram on October 19, 2020, with the hashtag #SacaPecho, grouped by profile

Source: Own elaboration.



AECC refers to profiles that contain this acronym in their name.

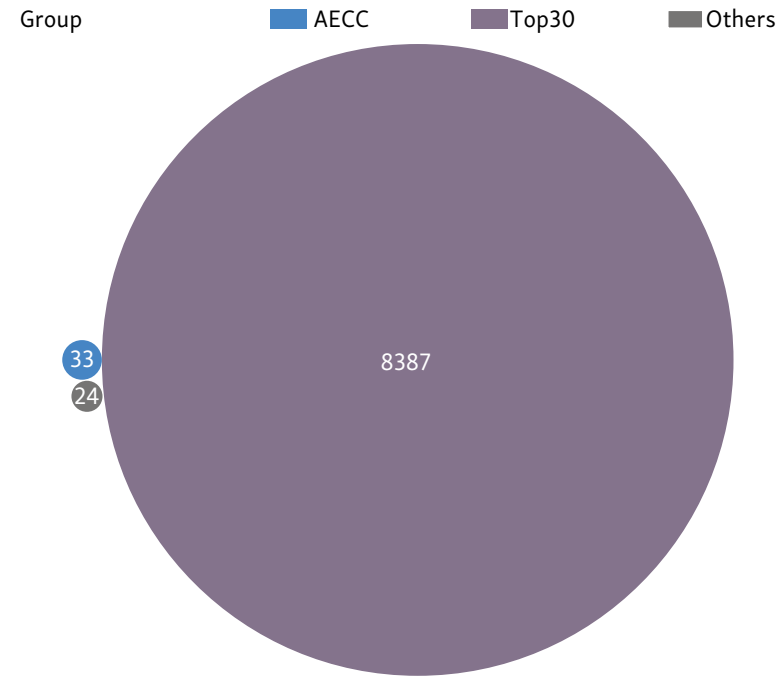
Figure 2. Number of likes received by Instagram posts with the hashtag #SacaPecho made on October 19, 2020, grouped by profile

Source: Own elaboration.

With this data, a list of the top 30 accounts was obtained by adding the Likes, which was compared with the other two groups (figure 3).

With a few institutional exceptions such as soccer teams, most of these superinfluencers are accounts of famous or influential people: singers, writers, athletes, or actors. Some of them have suffered from some form of cancer and represent the personalization of the disease's image explored above. Along, they multiply the reach of the accounts associated with the AECC by a factor of 75.

Using the second part of the developed code, we obtained a sample of publications, allowing to draw a general picture of their content. To illustrate this step, we downloaded a sample of 50 randomly selected images from the original list. All of them include pink as the dominant color, constructing a color distribution in accordance with Ehrenreich's (2001) impressions already pointed out. Figure 4 shows the color distribution, so that general characteristics can be appreciated but individual images cannot be identified.



The top 30 refers to the 30 profiles that received the most Likes, while AECC refers to profiles that contain that acronym in their name.

Figure 3. Median Likes received on posts with the hashtag #SacaPecho made on October 19, 2020, grouped into three profiles

Source: Own elaboration.

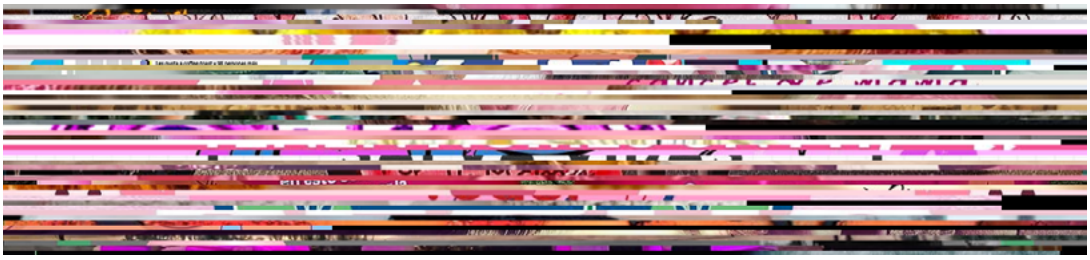


Figure 4. Composite of 50 sample images posted on Instagram with the hashtag #SacaPecho on October 19, 2020

Source: Own elaboration.

A content analysis reveals that the sample publications prioritize contents traditionally related to breast cancer in popular culture, highlighting images in everyday situations: pink ribbons, scarves and masks; diagrams on the possible manifestations of cancer that can be found in a self-examination (to detect possible warning signs); the promotion of sports and physical activity in general, or women in fighting pose (with the arm flexed as a sign of strength) are some of the most repeated elements. It is worth mentioning that the content analysis of photographic material can be the object of application of other computational methods, capable

of processing the information of these files in an automatic way to deepen the collective sense of the massive flow of circulation of visual messages.

DISCUSSION: ALTERNATIVES TO SCRAPING

Beyond the technical limitations, the use of these tools raises significant ethical issues. There are alternatives to scraping, certainly. The first is, as described by Bruns (2019), to retreat. Accept that access to images found on Instagram is hostile and difficult, and engage in networks with easier access (Flickr, Twitter, or Reddit). This would deny the evidence that Instagram is, today, the largest network of images, with a great influence on social dynamics, marketing or, in our case, communication and health education.

The second is manual work. Many of the works in social research focused on the represented follow this method. Performing manual searches and extracting screenshots is a mechanism used by Abidi (2016), Cinelli and colleagues (2020), or Couture Bue (2020). It entails a problem: among Instagram's multiple sorting variables (such as location, hashtags, username, or date and time), only hashtags can be explored with some ease. From Instagram's Explore search page, researchers can browse the various tags associated with images. For example, by searching for the hashtag #cancer we can see all the images with this hashtag (at the time of writing this article, more than 17 million). This, moreover, has two limitations: the first is that Instagram prioritizes the featured content, i.e., the content that has received the most attention on the social network in recent days, based on a series of algorithms that are difficult to interpret. The second is that the images are displayed on an infinite page, without any categorization and unaccompanied by metadata, as well as in thumbnails, so the researcher will have to go image by image making screenshots. In this case, collecting data in volume may help eliminate the bias of Instagram's algorithms, but it brings us back to a technical-ethical debate.

In response, some researchers, such as Giraldo-Luque and colleagues (2018) advocate using screen scraping tools, such as extensions for browsers like Google Chrome. With this, any additional information about the image would be lost, and its legality remains unclear.

A third alternative is to work through Instagram's official API. APIs (Application Programming Interface) are programming applications provided by the platforms, official interfaces through which researchers can formally request the platform to obtain the necessary data. Highfield and Leaver (2014) provide a detailed methodology for this process, albeit written in 2014, before Instagram shut down its API. Since then, the process for obtaining permission to use Instagram's API

has been confusing. Rogers (2018) and Bruns (2019) point to constant delays in processing requests, changes in API formats, and public contests for access tied to elusive dates. In fact, Rogers (2018) notes that the requirements for working with official APIs such as Facebook's are unreasonable, with high transaction costs, and points to a return to web scraping.

However, the work of Zarei and colleagues (2020) is a promising example of the possibilities of establishing a partnership with Instagram to obtain data through its official API. They develop a database of (meta)data obtained from Instagram and related to COVID-19. Although they do not detail how they achieved this, it is reasonable to assume that they obtained explicit permission from Facebook (as the owner of Instagram) to implement their search and crawling tool (a method similar to scraping) thanks to the social relevance of their research topic and the pressure to facilitate pandemic research. In principle, it could open the door to other potential works.

The fourth option is to hope that the use of the of these techniques goes unnoticed. Bruns (2019) and Rogers (2018) discuss this possibility, not without highlighting the difficulties involved. We have seen it used in Himawan and colleagues (2020), Dewi and colleagues (2019), Nobles and colleagues (2020), or Harrigan and colleagues (2020). However, as Bruns himself points out, the more researchers use these tools, the higher their visibility and, consequently, the higher the probability that Facebook will detect and block them.

This brings us to what Hirschey calls a pragmatic acceptance of scraping. He believes that providers should recognize the benefit of allowing it from an economic standpoint, and only take legal action if "when (1) the scraper presents a threat to the data host's core business and (2) the data host has a strong enough claim to prevail legally against the scraper" (2014, p. 901).

Hirschey made these recommendations in 2014, aware that the legal environment was not entirely favorable to providers, as there was no specific regulation for the practice of scraping. Naturally, the author makes no reference to the use of data to advance social research and development, as his argument is purely economic. This leaves us, the social researchers, in the position of being even more subject to economic logic in research if we want to collaborate with providers such as Facebook.

Limitations and conclusions of the study

Scraping is a method with advantages and disadvantages and, in general, a difficult area for social research. In a very simple way, with several lines of code and using open source and completely free tools, we have been able to obtain key

information to understand the distribution of cancer in social networks. While it is true that research on Instagram is not yet at the level of work on networks such as Twitter or Reddit (Lama et al., 2019; Record et al., 2018; Schrading et al., 2015) –mainly due to Facebook's corporate opacity– there are tools available to researchers to continue learning about this network.

From a technical and ethical point of view, Instaloader and Instalooter are tools with severe limitations. Although they facilitate data collection, scraping as such does not provide an easy answer to the complex process of determining which tags to explore nor how to maximize the representativeness of the data. This is solved by a design process as described by Schie and colleagues (2017), making the technique a support element rather than an answer.

Likewise, the design of these tools makes complex searches, where we can cross variables such as different hashtags or different date ranges, difficult. Nor is it possible to search by location, as Instagram does not provide that data. It is a frustrating method, as blockages by Instagram are common, as is the intentional slowing down of results.

Finally, although well documented, the tools require a certain level of programming knowledge in the Python language, and are exposed to a great risk of disappearing, in view of the changes and limitations imposed by Facebook in recent years. Similar tools, such as Instagram PHP Scraper, Django-Instagram, or Instagram-Scraper face similar challenges. Commercial options such as Meta Eyes or Vurku have not been studied in an academic environment.

The use of these resources allows the scalability of the analysis, since the volume of information that can be processed reaches large dimensions. However, the scraping process must seek connections with other computational methods, which deepen the automated analysis of the images obtained. The mass production and circulation of images in the digital ecosystem means that research teams must be trained in the combination of various computational methods that allow not only to collect and store, in this case, photographic archives, but also to advance in the coding, analysis and statistical processing of the most objectifiable aspects of visual communication. The dialogue between the possibilities offered by computer science and the long path completed by visual sociology or visual communication studies during the last decades appears as a fruitful space for research during the next decade.

Our results seem to indicate that the use of scrapers can help social researchers keep their finger on the networks' pulse, accelerating their work. The variety of users who have used the hashtag #SacaPecho and, especially, the super

influencers' individual character, suggests that the dynamics of personalization and appropriation of the discourse of disease in social networks –as opposed, as Pardo points out, to an exclusively medical-professional discourse– is very present on Instagram. This only underscores the usefulness of this network for understanding social discourse.

Obtaining this type of data at a much higher speed than we would have achieved manually allows us to make a classification of the circulation place and audience place (Rose, 2019) of cancer-related images. In the case of #SacaPecho, the information gains in relevance as it is a clearly delimited campaign, with a specific hashtag and search date. However, it is also possible to obtain similar results for more encompassing tags, such as #CancerdeMama, with wider time ranges. From there, we could make a distribution by types of cancer, visualize the results, and look for correlations between their visibility, their visual characteristics, and their epidemiological data.

The paths open for empirical research to analyze the image in digital social networks are as abundant as they are attractive, but ethical dilemmas and technical difficulties will coexist as long as the economic dimension of the platforms that host them does not find a mutually beneficial relationship with academic research and its necessarily critical nature.

REFERENCIAS

Abidin, C. (2016). "Aren't These Just Young, Rich Women Doing Vain Things Online?": Influencer Selfies as Subversive Frivolity. *Social Media + Society*, *2*(2), 2056305116641342. https://doi.org/10.1177/2056305116641342

Abril Curto, G. (2013). *Cultura visual, de la semiótica a la política* (Visual culture, from semiotics to politics). Plaza y Valdés.

Abutaleb, Y. (2016, June 21). Instagram's user base grows to more than 500 million. *Reuters*. Retrieved from https://www.reuters.com/article/us-facebook-instagram-users-idUSKCN0Z71LN

Arcila-Calderón, C., Ortega-Mohedano, F., Álvarez, M., & Vicente-Mariño, M. (2019). Análisis distribuido y supervisado de sentimientos en Twitter: Integrando aprendizaje automático y analítica en tiempo real para retos de dimensión big data en investigación de comunicación y audiencias (Distributed Supervised Sentiment Analysis of Tweets: Integrating Machine Learning and Streaming Analytics for Big Data Challenges in Communication and Audience Research). *Empiria. Revista de metodología de ciencias sociales*, (42), 113-136. https://doi.org/10.5944/empiria.42.2019.23254

Asociación Española Contra el Cáncer (AECC). (2020, October 9). La AECC anima a "sacar pecho" contra el cáncer de mama. *AECC Noticias*. Retrieved from https://www.aecc.es/es/actualidad/noticias/aecc-anima-sacar-pecho-contra-cancer-mama

Babvey, P., Capela, F., Cappa, C., Lipizzi, C., Petrowski, N., & Ramirez-Marquez, J. (2020). Using social media data for assessing children's exposure to violence during the COVID-19 pandemic. *Child Abuse & Neglect*, 104747. https://doi.org/10.1016/j.chiabu.2020.104747

Basch, C. H. & MacLean, S. A. (2019). Breast Cancer on Instagram: A Descriptive Study. *International Journal of Preventive Medicine*, *10*(1), 166. https://doi.org/10.4103/ijpvm.IJPVM_36_19

Becker, H. S. (1974). Photography and Sociology. *Studies in the Anthropology of Visual Communication*, *1*(1), 3-26. Retrieved from https://repository.upenn.edu/cgi/viewcontent.cgi?article=1003&context=svc

Berger, J. (2016). *Modos de ver: Basado en la serie de la televisión BBC en colaboración con John Berger* (Ways of Seeing: Based on the BBC Television Series). Barcelona, Spain: Gustavo Gili.

Blanch-Hartigan, D. & Viswanath, K. (2015). Socioeconomic and sociodemographic predictors of cancer-related information sources used by cancer survivors. *Journal of Health Communication*, *20*(2), 204-210. https://doi.org/10.1080/10810730.2014.921742

Bruns, A. (2019). After the 'APIcalypse': Social media platforms and their fight against critical scholarly research. *Information, Communication & Society*, *22*(11), 1544-1566. https://doi.org/10.1080/1369118X.2019.1637447

Cartwright, L. (1998). Community and the Public Body in Breast Cancer Media Activism. *Cultural Studies*, *12*(2), 117-138. https://doi.org/10.1080/095023898335500

Chou, W.-Y. S., Trivedi, N., Peterson, E., Gaysynsky, A., Krakow, M., & Vraga, E. (2020). How do social media users process cancer prevention messages on Facebook? An eye-tracking study. *Patient Education and Counseling*, *103*(6), 1161-1167. https://doi.org/10.1016/j.pec.2020.01.013

Cinelli, M., Quattrociocchi, W., Galeazzi, A., Valensise, C. M., Brugnoli, E., Schmidt, A. L., Zola, P., Zollo, F., & Scala, A. (2020). The COVID-19 Social Media Infodemic. *Sci Rep,* (10), 16598. https://doi.org/10.1038/s41598-020-73510-5

Couture Bue, A. C. (2020). The looking glass selfie: Instagram use frequency predicts visual attention to high-anxiety body regions in young women. *Computers in Human Behavior*, *108*, 106329. https://doi.org/10.1016/j.chb.2020.106329

Dewi, L.C., Meiliana, & Chandra, A. (2019). Social Media Web Scraping using Social Media Developers API and Regex. *Procedia Computer Science*, *157*, 444-449. https://doi.org/10.1016/j.procs.2019.08.237

Döbrössy, B., Girasek, E., Susánszky, A., Koncz, Z., Győrffy, Z., & Bognár, V. K. (2020). «Clicks, likes, shares and comments» a systematic review of breast cancer screening discourse in social media. *PLOS ONE*, *15*(4), e0231422. https://doi.org/10.1371/journal.pone.0231422

Ehrenreich, B. (2001, November). Welcome to Cancerland: A Mamogram Leads to a Cult of Pink Kitsch. *Harper's Magazine*. Retrieved from https://harpers.org/archive/2001/11/welcome-to-cancerland/

Gibson, F., Hibbins, S., Grew, T., Morgan, S., Pearce, S., Stark, D., & Fern, L. A. (2016). How young people describe the impact of living with and beyond a cancer diagnosis: Feasibility of using social media as a research method. *Psycho-Oncology*, *25*(11), 1317-1323. https://doi.org/10.1002/pon.4061

Giraldo-Luque, S., Fernández-García, N., & Pérez-Arce, J. C. (2018). La centralidad temática de la movilización #NiUnaMenos en Twitter (Thematic centrality on Twitter: The case of# NiUnaMenos mobilisation). *El profesional de la información*, *27*(1), 96-105. Retrieved from http://profesionaldelainformacion.com/contenidos/2018/ene/09.pdf

González Requena, J. (1988). *El discurso televisivo: Espectáculo de la posmodernidad* (The television discourse: Spectacle of postmodernity). Madrid, Spain: Cátedra.

González Requena, J. (2013). El ser de las imágenes: De la teoría al análisis de imagen (The Being of Images: From Theory to Image Analysis). *www.gonzalezrequena.com*. Retrieved from http://gonzalezrequena.com/textos-en-linea-0-2/libros-en-linea/el-ser-de-las-imagenes/

Grant, J. A. & Hundley, H. (2008). Fighting the Battle or Running the Race? *Visual Communication Quarterly*, *15*(3), 180-195. https://doi.org/10.1080/15551390802235578

Hand, M. (2016). Visuality in Social Media: Researching Images, Circulations and Practices. In L. Sloan & A. Quan-Haase (Eds.), *The SAGE Handbook of Social Media Research Methods* (pp. 215-231). SAGE Publications Ltd.

Harper, D. (1988). Visual Sociology: Expanding Sociological Vision. *The American Sociologist*, (*19*), 54-70. https://doi.org/10.1007/BF02692374

Harper, D. (1996). Seeing Sociology. *The American Sociologist*, *27*(3), 69-78. Retrieved from https://www.jstor.org/stable/27698785

Harrigan, P., Daly, T. M., Coussement, K., Lee, J. A., Soutar, G. N., & Evers, U. (2020). Identifying influencers on social media. *International Journal of Information Management*, (*56*), 102246. https://doi.org/10.1016/j.ijinfomgt.2020.102246

Henricks, S. C. (2017). Social Media, Publicly Available Information, and the Intelligence Community. *American Intelligence Journal, 34*(1), 21-31. Retrieved from https://www.jstor.org/stable/26497113

Highfield, T. & Leaver, T. (2014). A methodology for mapping Instagram hashtags. *First Monday, 20*(1). https://doi.org/10.5210/fm.v20i1.5563

Himawan, A., Priadana, A., & Murdiyanto, A. (2020). Implementation of Web Scraping to Build a Web-Based Instagram Account Data Downloader Application. *IJID (International Journal on Informatics for Development)*, *9*(2), 59-65. https://doi.org/10.14421/ijid.2020.09201

Hirschey, J. (2014). Symbiotic Relationships: Pragmatic Acceptance of Data Scraping. *SSRN Electronic Journal,* (29). https://doi.org/10.2139/ssrn.2419167

Instagram. (2018). *Condiciones de uso* (Terms of use). Retrieved from https://help.instagram.com/581066165581870?ref=dp

Jang, J. Y., Han, K., Shih, P. C., & Lee, D. (2015). Generation Like: Comparative Characteristics in Instagram. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 4039-4042. https://doi.org/10.1145/2702123.2702555

Kligler-Vilenchik, N., Baden, C., & Yarchi, M. (2020). Interpretative polarization across platforms: How political disagreement develops over time on Facebook, Twitter, and WhatsApp. *Social Media + Society, 6*(3), 2056305120944393. https://doi.org/10.1177/2056305120944393

Lama, Y., Hu, D., Jamison, A., Quinn, S. C., & Broniatowski, D. A. (2019). Characterizing Trends in Human Papillomavirus Vaccine Discourse on Reddit (2007-2015): An Observational Study. *JMIR Public Health and Surveillance*, *5*(1), e12480. https://doi.org/10.2196/12480

Manovich, L. (2016). Instagram and Contemporary Image. *manovich.* Retrieved from http://manovich.net/content/04-projects/151-instagram-and-contemporary-image/instagram_book_manovich_2017.pdf

Manovich, L. (2018). *AI Aesthetics*. Moscow, Russia: Strelka Press.

Markham, A. N. & Buchanan, E. (2012). Ethical Decision-Making and Internet Research: Recommendations from the AOIR Ethics Working Committee (Version 2.0). Association of Internet Researchers. Retrieved from https://aoir.org/ethics/

Markham, A. N. & Buchanan, E. (2017). Research Ethics in Context: Decision-Making in Digital Research. In M. T. Schäfer & K. Es, van (Eds.), *The Datafied Society. Studying Culture through Data*. Amsterdam, The Netherlands: Amsterdam University Press.

Marres, N. & Weltevrede, E. (2013). SCRAPING THE SOCIAL?: Issues in live social research. *Journal of Cultural Economy*, *6*(3), 313-335. https://doi.org/10.1080/17530350.2013.772070

Munk, A. K., Abildgaard, M. S., Birkbak, A., & Petersen, M. K. (2016). (Re-)Appropriating Instagram for Social Research: Three Methods for Studying Obesogenic Environments. *Proceedings of the 7th 2016 International Conference on Social Media & Society*, 1-10. https://doi.org/10.1145/2930971.2930991

Noar, S. M., Leas, E., Althouse, B. M., Dredze, M., Kelley, D., & Ayers, J. W. (2018). Can a selfie promote public engagement with skin cancer? *Preventive Medicine*, (111), 280-283. https://doi.org/10.1016/j.ypmed.2017.10.038

Nobles, A. L., Leas, E. C., Noar, S., Dredze, M., Latkin, C. A., Strathdee, S. A., & Ayers, J. W. (2020). Automated image analysis of instagram posts: Implications for risk perception and communication in public health using a case study of #HIV. *PLOS ONE*, *15*(5), e0231155. https://doi.org/10.1371/journal.pone.0231155

Orduña-Malea, E., Font-Julián, C. I., & Ontalba-Ruipérez, J.-A. (2020). Covid-19: Análisis métrico de vídeos y canales de comunicación en YouTube (Covid-19: metric analysis of videos and communication channels on *YouTube*). *Profesional de la Información*, *29*(4), e290401. https://doi.org/10.3145/epi.2020.jul.01

Pardo, R. (2019). Fotografía y enfermedad: Iconografías en transformación (Photography and disease: Iconographies in transformation). In M. Morcate & R. Pardo (Eds.), *La imagen desvelada: Prácticas fotográficas en la enfermedad, la muerte y el duelo* (The Unveiled Image: Photographic Practices in Illness, Death, and Grief) (pp. 19-60). Vitoria-Gasteiz, Spain: Sans Soleil Ediciones.

Pauwels, L. (2015). *Reframing Visual Social Science: Towards a More Visual Sociology and Anthropology*. Cambridge University Press. https://doi.org/10.1017/CBO9781139017633

Prosser, J., Clark, A., & Wiles, R. (2008). *Visual Research Ethics at the Crossroads*. NCRM Working Paper. Manchester, Reino Unido:Realities, Morgan Centre. Retrieved from http://eprints.ncrm.ac.uk/535/

Purba, K. R., Asirvatham, D., & Murugesan, R. K. (2020). Influence maximization diffusion models based on engagement and activeness on instagram. *Journal of King Saud University - Computer and Information Sciences*. https://doi.org/10.1016/j.jksuci.2020.09.012

Record, R. A., Silberman, W. R., Santiago, J. E., & Ham, T. (2018). I Sought It, I Reddit: Examining Health Information Engagement Behaviors among Reddit Users. *Journal of Health Communication*, *23*(5), 470-476. https://doi.org/10.1080/10810730.2018.1465493

Regueira, U., Ferreiro, A. A., & Vila, S. D. (2020). La mujer en YouTube: Representación y participación a través de la técnica Web Scraping (Women on YouTube: Representation and participation through the Web Scraping technique). *Comunicar: Revista científica iberoamericana de comunicación y educación*, *63*, 31-40. https://doi.org/10.3916/C63-2020-03

Rogers, R. (2018). *Social Media Research After the Fake News Debacle* (Dataset). University of Salento. https://doi.org/10.1285/I20356609V11I2P557

Rose, G. (2019). *Metodologías visuales: Una introducción a la investigación con materiales visuales* (Visual Methodologies. An Introduction to Researching with Visual Materials). Murcia, Spain: CENDEAC.

Schie, van, G., Westra, I., & Schäfer, M. T. (2017). Get Your Hands Dirty: Emerging Data Practices as Challenge for Research Integrity. In M. T. Schäfer & K. Es, van (Eds.), *The Datafied Society. Studying Culture through Data,* 183-200. Amsterdam University Press. https://doi.org/10.5117/9789462981362

Schrading, N., Ovesdotter Alm, C., Ptucha, R., & Homan, C. (2015). An Analysis of Domestic Abuse Discourse on Reddit. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, 2577-2583. https://doi.org/10.18653/v1/D15-1309

Sontag, S. (1978). *Illness As Metaphor.* Farrar, Straus and Giroux.

Sontag, S. (2017). *Sobre la fotografía* (On photography). Debolsillo.

Spence, J. (1986). *Putting myself in the picture: A political, personal, and photographic autobiography*. Camden, United Kingdom: Camden Press.

Sutton, J., Vos, S. C., Olson, M. K., Woods, C., Cohen, E., Gibson, C. B., Phillips, N. E., Studts, J. L., Eberth, J. M., & Butts, C. T. (2018). Lung Cancer Messages on Twitter: Content Analysis and Evaluation. *Journal of the American College of Radiology*, *15*(1), 210-217. https://doi.org/10.1016/j.jacr.2017.09.043

Tafesse, W. & Wood, B. P. (2020). Followers' engagement with instagram influencers: The role of influencers' content and engagement strategy. *Journal of Retailing and Consumer Services*, *(58)*, 102303. https://doi.org/10.1016/j.jretconser.2020.102303

Townsend, L., & Wallace, C. (2016). *Social Media Research: A Guide to Ethics*. The University of Aberdeen. https://www.gla.ac.uk/media/Media_487729_smxx.pdf

Walsh, M. & Baker, S. A. (2017). The selfie and the transformation of the public–private distinction. *Information Communication and Society*, (20), 1185-1203. https://doi.org/10.1080/1369118X.2016.1220969

Wang, X., Chen, L., Shi, J., & Peng, T.-Q. (2019). What makes cancer information viral on social media? *Computers in Human Behavior, (93)*, 149-156. https://doi.org/10.1016/j.chb.2018.12.024

White, C. L. & Boatwright, B. (2020). Social media ethics in the data economy: Issues of social responsibility for using Facebook for public relations. *Public Relations Review*, *46*(5), 101980. https://doi.org/10.1016/j.pubrev.2020.101980

Zarei, K., Farahbakhsh, R., Crespi, N., & Tyson, G. (2020). A First Instagram Dataset on COVID-19. *arXiv:2004.12226 [cs]*.Retrieved from http://arxiv.org/abs/2004.12226

Zhang, J., Le, G., Larochelle, D., Pasick, R., Sawaya, G. F., Sarkar, U., & Centola, D. (2019). Facts or stories? How to use social media for cervical cancer prevention: A multi-method study of the effects of sender type and content type on increased message sharing. *Preventive Medicine*, *126*, 105751. https://doi.org/10.1016/j.ypmed.2019.105751

FUNDING

This article has been written in the framework of the research project *Verdad y ética en las redes sociales. Percepciones e influencias educativas en jóvenes usuarios de Twitter, Instagram y YouTube (Internética)* (Truth and ethics in social networks. Perceptions and educational influences in young users of Twitter, Instagram, and YouTube (Internética), funded by the 2019 call for R+D+i projects of the Spanish Ministry of Science, Innovation and Universities (PID2019-104689RB-100).

ABOUT THE AUTHORS

MIGUEL VARELA, sociologist and adjunct professor in the Department of Sociology and Social Work at the Universidad de Valladolid. His research work focuses on the analysis of visual discourse, with special attention to social networks and digital technology applied to peace work and the sociology of health. Trained in documentary photography, he also holds an MA in Photojournalism and Documentary Photography from the University of the Arts London and an MSc in Peace Work and Development from Linnaeus University and is a member of the International Visual Sociology Association.

https://orcid.org/0000-0002-6443-9667

MIGUEL VICENTE-MARIÑO, associate professor at the Universidad de Valladolid, where he directs the Department of Sociology and Social Work and coordinates the Recognized Research Group in Applied Social Sciences. Ph.D. in Audiovisual Communication from the Universitat Autònoma de Barcelona, his lines of research focus on the analysis of audiences and contents of mass media and digital social networks, with focus on methodological innovations and computer tools.

https://orcid.org/0000-0002-1467-3535