Materials and assessment of literacy level for the recognition of social bots in political misinformation contexts

Materiales y evaluación del nivel de alfabetización para el reconocimiento de bots sociales en contextos de desinformación política

Materiais e avaliação do nível de literacia para o reconhecimento dos bots sociais em contextos de desinformação política

Dafne Calvo

*Predoctoral Researcher
(University of Valladolid)*
http://orcid.org/0000-0003-0659-6792
Spain

Lorena Cano-Orón

*Postdoctoral Researcher
(University of Valladolid)*
https://orcid.org/0000-0003-4270-1924
Spain

Almudena E. Abengozar

Researcher
https://orcid.org/0000-0002-7266-6501
Spain

Abstract

The media ecosystem is constantly changing, transforming at a rhythm that educational institutions cannot keep up with. In this media transformation, artificial intelligence (AI) has been introduced, adapted to social networks for various purposes, including political ones. This work focuses on AI in bot *format as an automated tool to publish content on Twitter and on the skills needed to identify them.*

Bots seek to imitate human behaviour to create a climate of concrete opinion, participate in political conversations and interact with real accounts in order to boycott them or increase their relevance. In order to know the level of media literacy of journalism students and increase their skills to identify this type of manipulation of the public sphere, a teaching intervention is designed at university level, applied to a sample of 55 students, consisting of a workshop to identify bots *on Twitter. From the results of the workshop, evaluated through questionnaires, it is observed that the students have previous skills acquired informally to detect* bots *and that after the workshop the ability to identify these accounts slightly increases and, above all, they manifest a more conscious process when they face them. This paper presents the design of the workshop and its evaluation.*

Key Words: *Media literacy; Digital citizenship; Disinformation; Bots; Twitter; Computational campaign*

Resumen

El ecosistema de medios está en constante cambio, transformándose a un ritmo que las instituciones educativas no pueden seguir. En esta transformación mediática se ha introducido la inteligencia artificial (IA), adaptada a las redes sociales con diversos fines, también políticos. Este trabajo se centra en la IA en formato bot *como herramienta automatizada para publicar contenido en Twitter y en las competencias necesarias para identificarlos.*

Los bots *buscan imitar el comportamiento humano para crear un clima de opinión concreto, participar en conversaciones políticas e interaccionar con cuentas reales para boicotearlas o aumentar su relevancia. Con el objetivo de conocer el nivel de alfabetización mediática de alumnado universitario e incrementar sus competencias para identificar este tipo de manipulación de la esfera pública, se diseña una*

intervención docente en ámbito de la educación superior, aplicada en una muestra de 55 alumnos, que consiste en un taller de identificación de bots *en Twitter. A partir de los resultados del taller, evaluados mediante cuestionarios, se observa que el alumnado presenta habilidades previas adquiridas informalmente para detectar* bots *y que, tras el taller, manifiestan un proceso más consciente cuando se enfrentan a ellas. En este trabajo se expone el diseño del taller y su evaluación.*

Palabras clave: *Alfabetización mediática; Ciudadanía digital; Desinformación; Bots; Twitter; Campaña computacional*

Resumo

O ecossistema mediático está em constante mudança, transformando-se a um ritmo que as instituições educativas não conseguem igualar. Nessa transformação da mídia, foi introduzida a inteligência artificial (IA), adaptada às redes sociais para diversos fins, inclusive políticos. Este trabalho centra-se na IA em formato bot *como uma ferramenta automatizada para publicar conteúdos no Twitter e nas habilidades necessárias para identificá-los.*

Os bots *procuram imitar o comportamento humano para gerar um clima de opinião concreta, participar de conversas políticas e interagir com contas reais a fim de boicotá-las ou aumentar sua relevância. A fim de conhecer o nível de literacia mediática dos estudantes de jornalismo e aumentar as suas competências para identificar este tipo de manipulação da esfera pública, é desenvolvida uma intervenção pedagógica a nível universitário, aplicada a uma amostra de 55 estudantes, que consiste num seminário para identificar* bots *no Twitter. A partir dos resultados do workshop, avaliados através de questionários, observa-se que os alunos têm habilidades anteriores adquiridos informalmente para detectar* bots *e que após o seminário a capacidade de identificar essas contas aumenta ligeiramente e, acima de tudo, eles manifestam um processo mais consciente quando eles enfrentam-las. Este artigo apresenta o projeto do workshop e sua avaliação.*

Palavras chave: *Alfabetização midiática; Cidadania digital; Desinformação; Bots; Twitter; Campanha computacional*

1. Introduction

The Internet has transformed the communicative scenario and has articulated new forms of production and distribution of information. These have also permeated electoral campaigns, where different political formations have readapted their formulas and strategies for mobilizing the electorate and disseminating their message to the cyber-connected citizenry (Ciampaglia, 2017; Kreiss & Jasinski, 2016; Nielsen & Vaccari, 2013).

In this sense, the growing use of social networks such as Twitter or Facebook has contributed to the promotion of direct interaction between political representatives and voters and has become a new source and medium for information on the electoral campaign (Allcott & Gentzkow, 2017; Green & Middleton, 2008; Woolley & Howard, 2016).

A recent example is what happened during the referendum on self-determination in Catalonia in 2017 when through associated hashtags the bots sent messages to pro-independence profiles to activate civil disobedience during October 1st (Stella, Ferrara & De Domenico, 2018). Thus, some of the consequences of the implementation of these techniques, such as the use of bots, is the creation of fictitious climates of opinion, the viralization of disinformation, or setting the media agenda in networks.

In this scenario, as Monreal Guerrero et al. (2017) point out, it is essential to strengthen the media and information skills of citizens. To do this, it is useful to offer tools that allow them to reflect on the media and thus be able to develop a critical spirit that promotes autonomous and critical individuals in the face of content consumption and the use of social networks.

Formal education has the possibility of offering such specific knowledge "making it necessary to incorporate strategies that help to search, select and interpret the most faithful information available in the virtual environment" (Mayor Buzón et al., 2017:63). In this sense, formal education must deepen the knowledge that allows for the development of a trained citizenry capable of identifying and disrupting disinformation strategies on the web. This makes an increase in democratic quality possible inasmuch as it contributes to discerning between truthful information and that generated for partisan purposes.

The general objective of this work is to generate teaching materials for the technological empowerment in the recognition of bots, so that civil society itself has a multitude of resources to facilitate the verification of the information received from social networks and the implications of this in the political processes in which it is involved (Levi et al., 2019; Tucker et al., 2018; Woolley, 2016).

In particular, the specific objectives that articulate this research are:

- SO1: Assess the degree of media competence in identifying automated fake accounts (bots) on social networks (Grimme et al., 2017).

- SO2: To contribute to cognitive resources for conscious recognition of disinformation flows on the Net (Levi et al., 2019) by creating and disseminating teaching materials that enable a workshop for educational intervention in the classroom.

2. Theoretical Approach

2.1. The computational campaign and the threat of misinformation

In the Spanish context, which has so far followed North American trend, computational strategies for political communication—i.e., public segmentation / microtargeting—were consolidated in the 2016 general election, while political parties experimented with new practices—i.e., the use of bots or virtual assistants based on artificial intelligence—(Campos-Domínguez & García-Orosa, 2018). An outstanding example of a machine learning application to political communication is the Spanish Calisto software, which articulates political discourse from Partido Popular officials extracting information from the Internet and prioritizing it according to how it benefits the party's interests (Redondo, Calvo & Díez-Garrido, 2017).

Bots also serve the strategies of political actors contesting electoral campaigns. They can be defined as software that automatically simulate human behavior through artificial intelligence and machine learning (Persily, 2017; Vosoughi, Roy, & Aral, 2018; Woolley & Howard, 2016). They are therefore capable of performing different tasks with diverse sophistication, such as extracting information from webpages, chatting with humans and/or sending basic messages (Gorwa & Guil-

beault, 2018; Howard, Woolley & Calo, 2018; Powers & Kounalakis, 2017).

Howard et al. (2018) suggest that bots can both serve legitimate interests such as running news feeds and malicious purposes such as sending spam messages. The versatility of bots' results in a wide range of names that describe them according to certain functions [Figure 1]. Specifically, social bots are software that has a fully or partially automated presence on social networks (Murthy et al., 2016: 4955). Political bots belong to the last category, dealing with political aims (Bessi & Ferrara, 2016; Woolley, 2016). Political bots perform three main functions, as Powers & Kounalakis (2017) noted:

1. Boycotting campaigns and the creation of critical discourses on political opponents' networks (roadblock bots).

2. Following specific accounts of party representatives to simulate a greater relevance on the Internet (follower bots).

3. Promoting the ideas and discourses of political parties through the publication of content that is favorable to them and their ideas (propaganda bots).
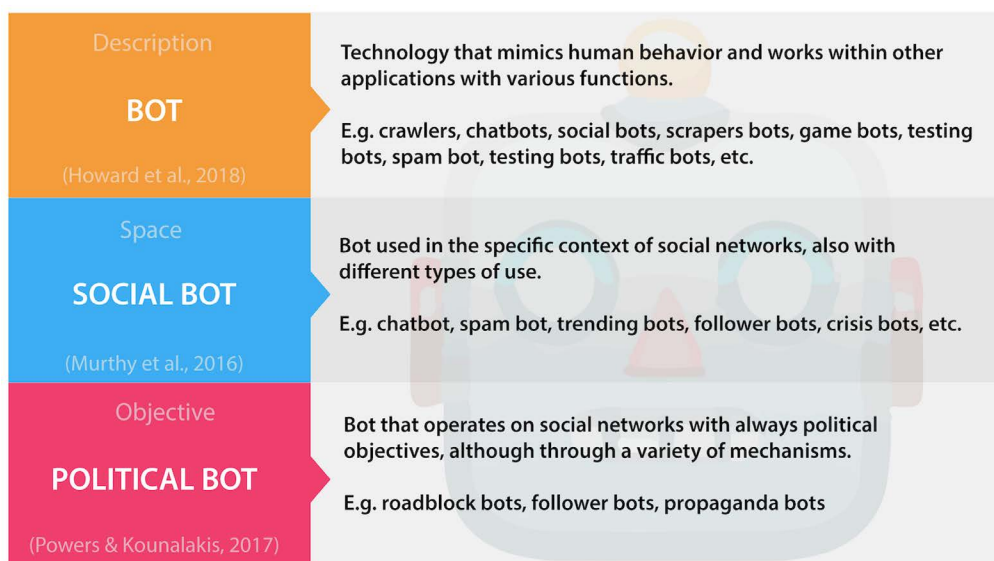

Description **BOT** (Howard et al., 2018)	Technology that mimics human behavior and works within other applications with various functions. E.g. crawlers, chatbots, social bots, scrapers bots, game bots, testing bots, spam bot, testing bots, traffic bots, etc.
Space **SOCIAL BOT** (Murthy et al., 2016)	Bot used in the specific context of social networks, also with different types of use. E.g. chatbot, spam bot, trending bots, follower bots, crisis bots, etc.
Objective **POLITICAL BOT** (Powers & Kounalakis, 2017)	Bot that operates on social networks with always political objectives, although through a variety of mechanisms. E.g. roadblock bots, follower bots, propaganda bots

Figure 1: Typologies of bots.
Source: Authors' own.

One of the best-known strategies is astroturfing, which consists of the massive publication of coordinated messages from different accounts that pretend not to be related to the beneficiaries of the campaign that configures and manages them (Howard et al., 2018; Levi et al., 2019). With these capabilities, these human behavior mimicking algorithms are likely to contribute not only to computational propaganda but also to disinformation flows on the Internet (Bradshaw & Howard, 2017; Powers & Kounalakis, 2017).

Thus, bots facilitate the spread of fake news, rumors, and misleading information of various kinds as they can viralize it by sharing (Allcott & Gentzkow, 2017; Bessi & Ferrara, 2016; Shao et al., 2018; Vosoughi et al., 2018). The massive dissemination of a message together with data extraction and segmentation practices contribute not only to republishing messages but also to directing their scope toward the targeting of profiles likely to react to them. In this way, they contribute to the fragmentation of public space around opposing perspectives on specific issues in an increasingly complex society (Bradshaw & Howard, 2017; Stella et al., 2018).

The automatic production and viralization of content can set the political agenda, generate specific opinion flows, attack dissident voices, modify the online debate on a given issue and, in short, generate a fictitious scenario that influences the public sphere of citizens connected to the Internet (Powers & Kounalakis, 2017; Tucker et al., 2018; Tufekci, 2014).

Auditing this type of partisan practice is complex since understanding them requires a high level of technological expertise (Grimme et al., 2017). The detection of bots requires ever more sophisticated techniques to facilitate a more indepth and detailed identification of them, as well as to avoid confusion between human-operated and automated accounts (Bessi & Ferrara, 2016; Chu et al., 2012; Grimme et al., 2017; Murthy et al., 2016).

Tools such as Botometer (formerly known as BotOrNot), for example, allow you to audit particular accounts on Twitter, but do not cover coordinated attacks, nor do they address the farms of these types of accounts that pretend to be real peo-

ple, nor are they able to distinguish hybrid bots, which have automated processes but human-generated content (Grimme et al., 2017; Powers & Kounalakis, 2017).

When academics, journalists or citizens facing recognition of such practices lack the expertise to extract sensitive information such as the IP address or account identifiers of an account, this task is even more complex (Acker, 2018; Gorwa & Guilbeault, 2018).

2.2. Media Literacy as Technological Empowerment

Media Literacy (ML) in this context of identifying bots serves as a source of citizen empowerment as it acts as a backbone between users and technologies and makes it possible to reduce the digital divide with respect to the types of use (OECD, 2008). The development of ML reduces the distance between users and the media ecosystem by teaching technological and media skills.

As early as 1982, UNESCO expressed the need to improve media skills in order to foster critical and participatory citizenship, capable of understanding the media, its messages, and languages (UNESCO, 1982). The need to understand the medium and its transmedia narratives in relation to competencies is still valid today (Scolari, 2016), and is a focus of interest for educators, researchers and educational institutions (European Commission, 2009; Ferrés, 2007; Vuorikari et al., 2016; Mayor Buzón et al., 2019).

It is a matter of forming conscious individuals who are able to access a large volume of information, know how to freely decide what content is relevant and appropriate for them, as well as being able to adopt a responsible option among the multiple alternatives offered (Valerio-Ureña & Valenzuela-González, 2011; Monreal Guerrero et al., 2017).

In this sense, the young and university students, who are the object of this work, all share the fact that they have grown up in a global digital environment. Some of these young people are highly trained participatory users and inventively connect to the Internet and social networks. Technology is an organic part of their

lives for them whereas their adaptation to and learning of the continuous technological change has been a challenge so far (Alcoceba, 2017).

Media skills acquired outside formal institutions are no longer invisible, several authors point out that informal and peer learning is becoming essential in developing the skills that schools do not offer (Fundación Telefónica, 2019; Pereira, Fillol & Moura, 2019). This is why formal, non-formal, and informal learning must permeate and intertwine the skills acquired and the issues faced by users in this ecosystem of environments.

The European reference framework for digital skills, DigiComp (Kluzer & Pujol, 2018), presents a set of indicators of competences that aims to describe those that are needed today to use digital technologies with confidence, in a critical, collaborative and creative way to face the challenges related to work, learning, leisure and participation in our digital society. Kluzer & Pujol (2018) highlight that, in order to consider a digitally competent citizenry, it is necessary to present and develop digital skills in five areas:

- Information and data literacy
- Communication and collaboration
- Digital content creation
- Safety
- Problem solving

To improve the identification of bots, the emphasis is placed on the first of the skill blocks proposed in the DigiComp report, deepening the subsections, which promote skills for searching, filtering, evaluating, and managing data and digital content (Kluzer & Pujol, 2018). The promotion of critical competencies of users is presented as one of the fundamental axes of the ML and of this work, which seeks that people develop critical reading and autonomous thinking (Alvarado, 2012; Monreal Guerrero et al., 2017).

Media Literacy has a critical dimension: it includes analysis, evaluation, and critical reflection (Buckingham, 2007; Portalés-Oliva, 2019). Specifically, the identification of bots requires entering classrooms and non-formal or informal learning to develop such reflective processes and extrapolating learning related to bots on Twitter to other platforms or automated tools.

According to Kellner (2001), literacy is a necessary condition for enabling people to participate in the economy, culture, and politics—at local, national, or global levels—and emphasizes the crucial links between literacy, democracy, empowerment, and participation. Along these lines, by re-reading texts and audiovisual messages and identifying them, it is possible to develop personal strategies to highlight the authentic opinions and messages of those generated intentionally to change public opinion.

In this regard, the previous initiatives that have been carried out in the field of media literacy on issues of digital disinformation come from foundations such as Maldita.es (2019), which contributes to the identification of hoaxes or false news on social networks and which offers training and content search engines so that users can consult suspicious information. Along these lines, initiatives in the public sphere such as those carried out by INCIBE (2019) or OSI (2019) in the area of communications security, aimed at companies and citizens respectively, are also noteworthy. An example of this is the IS4K (2019) website dedicated to ML information and training for children and young people.

3. Material and methods

The objectives of this work are to provide a proposal for the creation of pedagogical strategies to influence a media scenario in constant change and transformation (Cobo & Moravec, 2011; Scolari, 2016). To this end, a teaching intervention was designed with a participatory methodology that allows the integration of student participation in the classroom to achieve active and cooperative learning (Imbernón & Medina, 2008; Pérez-Pérez, 2014). This action also seeks to articulate research with empirical data of both scientific and social interest (Villasante, Montañés & Martí, 2002).

Thus, this research has a double aspect: on one hand, it presents the educational intervention designed as a workshop to improve the detection of bots on social networks. On the other hand, it collects both qualitative and quantitative data on the media literacy level of the students and enables us to test the effectiveness of the teaching tool and its effect on students.

3.1. Bots identification workshop design

The workshop was structured with a participatory approach responding to the objectives of this research. This intervention was divided into three phases [Figure 2]; phases 1 and 2 allow for the collection of information to evaluate the degree of media competence; phases 2 and 3 provide resources for the conscious recognition of bots. Each of the proposed phases was divided into two distinct parts, always in the same order, in which information was first provided and then put into practice.
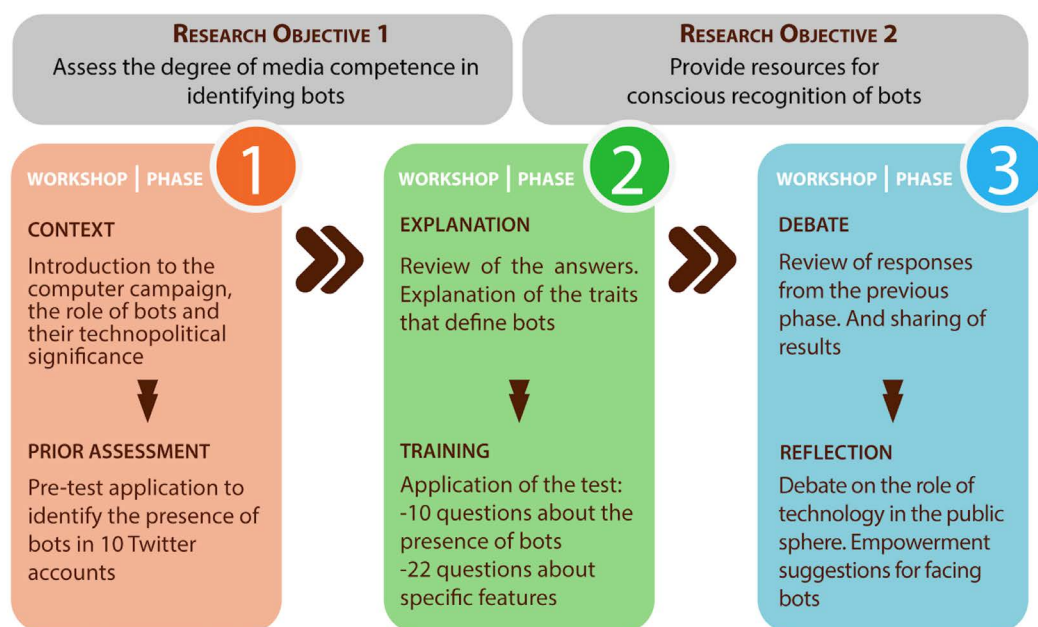

Figure 2: Objectives and phases of the workshop.
Source: Authors' own.

Phase 1: Context and prior assessment

The first phase of the workshop began with an explanation of political communication on the Internet to contextualize where the sophistication of the techniques derived from the computational campaign were developed (Nickerson & Rogers, 2014; Tufekci, 2014). It was also explained the conflicts that this type of strategy presented for the manipulation of the electoral public sphere and the inequality between online citizens and political subjects with greater human, technical and economic resources (Powers & Kounalakis, 2017; Tucker et al., 2018).

Later, with the aim of observing the initial level of competence of the students, the pretest was carried out, which consisted of reflecting on 10 Twitter accounts to discern whether or not they were automated fake profiles. This pretest took the form of an online questionnaire with simple answer questions, which evaluated the media competence level from which they started and whose quantitative results are explored in the results section.

Phase 2: Explanation and training

The next phase of the intervention consisted of sharing and formalizing the intuitive knowledge that the students previously showed in the pretest. Thus, first, the results of the previous questionnaire were shared and then the elements that facilitated the recognition of the bots on the Internet were explained in order to contribute to a critical reading of the contents on social networks.

After the theoretical explanation that started with the characteristics pointed out by Acker (2018) and with the experimentation with tools that contribute to the detection of bots from the information they had analyzed, such as Botometer, Twittonomy, Foller.me or the Twitter API, the self-assessment test was carried out.

This test was again constituted as a questionnaire, with 10 single-answer questions similar to those of the pretest and, additionally, with a specific section where the students tried to identify the bots from images showing specific characteristics of this profile type, which also had single-answer questions. Finally, to check which features were more easily identified, the test ended with an open comment section, which provided specific qualitative data on the process of identifying the bots.

Phase 3: Discussion and reflection

The last part of the workshop began with a collective exhibition of the results of the previous test. Once the characteristics that facilitated the identification of the recognition of bots were recognized, as well as the tools that contributed to the identification of bots, the last step was to discuss the mechanisms and strategies to perform when one wants to carry out a political action beyond the conscious knowledge of these disinformation strategies on social networks.

The theoretical presentation in this section consisted of a debate on the mechanisms offered by Twitter to denounce this type of account, the ways of publicizing these opaque techniques and disseminating their identification, as well as the cases that have transcended to the media, where the mobilization on social networks of Internet users has meant specific accounts being suspended.

3.2. Tools and materials for implementation

To measure the level of media literacy with respect to the detection of automated accounts on Twitter and the evolution of the student body after the teacher's intervention, online questionnaires were used as an interactive basis. In addition to facilitating the extraction of quantitative and qualitative data, this instrument allowed the assessment of media literacy in educational contexts.

The questionnaire as an instrument to measure media competences has been previously used by national and international institutions such as the OECD (2015), the European Commission (2017), the Spanish Ministry of Education (2011), the French Ministry of Education (2019) or ACARA-Australia (2011), among others. Furthermore, this tool continues to be used in recent studies on these competencies (Mayor Buzón et al., 2019; Guzmán-Simón et al., 2017; Jiménez-Cortés et al., 2015). Additionally, this instrument contributes to the replicability of both the study and the teaching proposal, beyond the present article.

The questionnaires were implemented through an online platform, configured as exercises for the recognition of automated fake accounts on Twitter. In the test format, three types of questions were posed:

MONOGRAPH

- Profile identification: Simple answer question for the recognition of false profiles with a single valid answer ("Yes", "No" or "Don't know").

- Focus on features: Simple answer question in which five options were presented (two from bots, two from real accounts and "I don't know"), with screenshots taken that focused on the specific information of the Twitter account (its profile picture, its followers, the date the account was created, for example). This section is multi-responsive and consisted of examining which features belong to bots and which belong to real accounts.

- Open-ended questions: These proposed reflections on the process of identifying bots, the relevance of their analysis, and ways to deal with disinformation strategies on the Internet.

Data Craft: The Manipulation of Social Media Metadata (Acker, 2018) was used to extract the variables that facilitated the identification of bots. This document includes a detailed set of features shared by this type of automated account: "Username, date and time of publication, follower accounts, likes and dislikes, hashtags or location tags".

The set of these traces gives readers clues as to how they were produced (Acker, 2018:8). From this report, a set of public and open metadata was extracted that has guided the questions asked in the test, as well as the theoretical explanation. The transfer of these categories to the online test questions, as well as the rest of the issues included in the test and in the subsequent discussion are included in the annexes to this document [see: research instruments].

The examples of bots with which the tests were built were obtained manually from the findings of the DataPolitik association's report (2019), which points to the existence of suspicious accounts in the political hashtag #SíguemeYTeSigoVox. From that tag, an initial manual search was carried out for accounts that partially or totally fulfilled the characteristics proposed by Acker (2018). After this first extraction, an additional phase was carried out to verify these suspicious accounts using Botometer[1]. This tool reviews account activity and provides a score based on characteristics such as frequency of posting or date of account creation.

Thus, profiles extracted from #SíguemeYTeSigoVox were analyzed with this tool and those with a score of less than three were discarded (where five was the maximum bot-like score). Sixty profiles were selected that had a high probability of being bots and used in the Spanish electoral campaign in November 2019. Their messages and profiles were used as examples in both the theoretical and practical parts of the workshop.

3.3. Study corpus

The workshop was held twice for university students. This type of audience was chosen because it is the segment of the population that most frequently connects to the Internet (INE, 2018) and is the one that most frequently uses social networks to inform itself about political and electoral campaign issues (CIS, 2019). This was applied in two Spanish universities with students of computer engineering, audiovisual communication, and journalism, always in dates close to the context of political communication in which the teaching intervention was framed.

This diversity of study disciplines allowed us to observe the interest and capacity to recognize bots in a knowledge sector with technical training to understand the functioning of these automated mechanisms, on one hand, and in a knowledge area whose theoretical knowledge is related to media literacy and media consumption, on the other. A total of 55 people took part in the participatory methodology proposed. According to the data collected from the survey, the students were between 18 and 29 years old, 63.6% were women and 34.5% were men.

4. Results

4.1. Pre-workshop digital skills

From the pretest data, it can be seen that the students were able to recognize some of the characteristics that effectively defined the bots, mainly those related to the profile name or biography, photo and account details [Table 2]. Some of the characteristics identified in this phase were the number of foreign messages in the profile; the high frequency of publications and repetitive messages; the massive and

reciprocal tracking of other fake accounts; the type of content published; and the type of account name – "Identifiers with many numbers or very random" (Q7761).

At this stage, the existence of informally acquired knowledge is perceived (Pereira, Fillol & Moura, 2019) which likewise contributes to the perception of online misinformation. In one of the questionnaires, it is openly stated that the exercise was done based on "intuition" (Q7145) and, in other cases, this reference is made more implicitly: "If you only do retweets it gives more of a bot feeling, although not necessarily" (Q9407); "I don't ever follow a bot, even if it's followed by someone I trust" (Q0142).

4.2. Results After Teaching Explanation

In the results of the students' initial critical skills for identifying bots, it can be seen that the pretest data at both universities show an average score of 6 out of 10. 58% of the students obtained an initial score ranging from 4 to 6 [Table 1], which shows that they had a medium base knowledge, acquired through informal learning, but which was not infallible.

Score	Between 0 and 3 (n | %)	Between 4 and 6 (n | %)	Between 7 and 10 (n | %)	Total average
Pretest (n=55)	3	5%	32	58%	20	36%	6/10
Test (n=52)	5	10%	29	56%	18	35%	5,7/10

Table 1: Assessment of pretest and test bot account identification questions.
Source: Authors' own.

In fact, at a later stage than the theoretical explanation of the key characteristics for the recognition of bots, the results for the account identification questions are slightly lower (5.7 out of 10) than the pretest. This could mean that, while a priori they had not paid much attention to them, doubts about identifying the profiles subsequently increased. The economy of attention may have been another influential factor.

When we looked closely at the data for the characteristics corresponding to the test part [Table 2], we found that the section of questions regarding the detection of bots from the profile name was the most accurate. However, the set of questions about the specific feature of the follower type that defines a bot generated the highest percentage of errors and doubt.

In this sense, the data show that the more complex the characteristic to be observed, the greater the percentage of doubt and error. In other words, more time and resources are needed to identify a bot accurately, which is probably why bots continue to multiply and have a more complex reality effect to identify. The breakdown of results by question can be found in research instruments.

Question Blocks	% Correct	% Error	% Doubt
1. Profile name	82.0	16.6	1.7
2. Biography, photo and account details	79.6	27.1	2.1
3. Followers and interactions	64.5	40.9	2.4
4. Publications and activity	65.6	36.9	1.6

Table 2: : Percentages of correct, error and doubt in the detection of bots from specific characteristics (n=54).
Source: Authors' own.

Beyond the numerical data, the qualitative information shows an increase in the cognitive resources to identify the bots. Students reported having again used some of the characteristics of the bots perceived informally during the pretest, such as the repetition of certain messages—"The repetition of patterns" (Q6952), "Tweets repeated" (Q8219)—or the followers and following of the account—"Number of followers" (Q3731), "The difference between the number of followers and following accounts" (Q9860).

However, they also described new features, such as the announcement of a mutual follow-up in the name, biography or messages of the profile: "FollowBack" (Q5423), "Ask for followback" (Q7655), "#síguemeytesigoVOX" (Q2518); the use of photomontages or model bots in their profile image: "Images that do not match

gender" (Q9407), "Pixelated images" (Q7004) or the date of creation close to the political-electoral context: "When it was created" (Q2651), "Age of account" (Q7761).

Some of the participants also referred to features such as the time of publication of messages and the frequency of publication over time and thus expressed, more or less explicitly, the use of the tools suggested during the intervention: "Twitter activity at Foller.me" (Q7464) or "Time of publication" (Q9407). In other cases, moreover, the responses were aimed at identification through a wide set of variables that corresponded to the proposals in the seminar: "The tips of the presentation" (Q0662), "The examples shown" (Q7145), "What was explained in the talk" (Q6817), etc.

5. Discussion and conclusion

This research addresses the use of artificial intelligence to manipulate the public sphere. The use of bots in political communication campaigns is an increasingly widespread technique (Pastor-Galindo et al., 2020; SINC Agency, 2019) and therefore, training in the identification of bots is a current need in electoral contexts. This work has tried to present both an educational proposal for bot detection and an empirical investigation on media competence concerning bots at the university level, given that this educational sector receives information about institutional politics through social networks (INE, 2018; CIS, 2019).

Results on the media competence [first objective] show that more than half of the participants are able to recognize automated fake accounts without previous specific indications. Prior training is not necessary to identify some bots' key characteristics such as a high number of retweets in their account, misspelled messages, and multiple numbers in their profile name. Put simply, they demonstrated high good to identify misinformation trends and an ability to adopt a critical attitude toward the information on social networks (Valerio-Ureña & Valenzuela-González, 2011).

Research also provides insight concerning the provision of resources to recognize bots consciously [second objective]. We understand that the workshop presented in this study offers new analytical tools (Acker, 2018), which goes beyond the informal training of participants, as they stated during the workshop.

Enhancing the recognition of certain characteristics of bots' behavior—asking for following back, being created in recent months—contribute to the university students' capacity of critical questioning. The identification of bots is more accurate when analyzing complete accounts such as from sets of messages and metadata than focusing on single messages. The most difficult challenge comes when fake accounts present more sophisticated and subtle behaviors (Levi et al., 2019).

We also find that the higher the number of messages and accounts that need to be analyzed, the lower the number of bots that the students identify. The slight decrease may respond to a falling off in attention and the over-abundance of bots. This reaction leads us to suggest that it is complex to sustain critical questioning in an Internet environment of polluted information.

Bearing that in mind, future educational workshops should be understood as a useful contribution in dealing with disinformation insofar as they set out critical tools for media consumption on social networks (Monreal Guerrero, et al., 2017). The final phase of the workshop addresses the need for training beyond the sole identification of bots to enable people to report and remove misleading content that potentially modifies political behavior and polarizes citizens in complex public debates.

Participants also point to the responsibility of social networking companies and political parties due to their agency to take action against misinformation on the Internet. These reflections draw attention to the relevance of the political and social analysis of technologies within the complex ecology where current public issues are discussed (Murthy et al., 2016; Shao et al., 2018; Stella et al., 2018).

This study has some limitations namely the long duration of the activity—two hours—that results in a loss of interest or fatigue among participants and the

difficulties in replicating the workshops in diverse contexts and thus being able to count on a wider sample.

Finally, the target of the workshop implies a reflection on the ability of bots and misinformation identification in certain segments of the population.

Differences between media literacy, age, and educational level refer back to social inequalities between those who have resources to information about politics and those who have less training when exposed to flows of misinformation on the Internet (Grimme et al., 2017). Future contributions to the research on misinformation should address the degree of media competence, sociodemographic groups and, in turn, educational initiatives adapted to other profiles should emerge.

Acknowledgments

References

ACARA-Australia (2011). *National Assessment Program*. Australian Curriculum, Assessment and Reporting Authority.

Acker, A. (2018). *Data Craft: The Manipulation of Social Media Metadata*. Data & Society.

Agencia SINC (2019, 20 de diciembre). Al menos 1.500 cuentas *bot* de Twitter apoyaron a Vox en la campaña del 10N, según la Universidad de Murcia. Eldiario.es. Retrieved from: https://www.eldiario.es/tecnologia/cuentas-Twitter-manejadas-bots-campana_0_976052563.html

Alcoceba, J. A. (2017). Juventud, TICs y aprendizaje invisible. El desarrollo generacional de habilidades y talentos digitales. *Revista de Estudios de Juventud, 4*(117), 21-35.

Allcott, H. & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives, 31*(2), 211–236. https://doi.org/10.1257/jep.31.2.211

Alvarado, M. (2012) Lectura crítica de medios: una propuesta metodológica. *Comunicar, 39*, 101-108. https://doi.org/10.3916/C39-2012-02-10

Bessi, A. & Ferrara, E. (2016). Social *bots* distort the 2016 U.S. Presidential election online. *First Monday, 21*(11). http://dx.doi.org/10.5210/fm.v21i11.7090

Bradshaw, S. & Howard, P. N. (2017). *Troops , Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation* (No. 12). Oxford.

Buckingham, D. (2007). Digital Media Literacies: rethinking media education in the age of the Internet. *Research in Comparative and International Education, 2*(1), 43-55. https://doi.org/10.2304/rcie.2007.2.1.43

Campos-Domínguez, E. & García-Orosa, B. (2018). Comunicación algorítmica en los partidos políticos: automatización de producción y circulación de mensajes. *El Profesional de la Información, 27*(4), 769–777.

Centro de Investigaciones Sociológicas (2019). Macrobarómetro de marzo 2019. Preelectoral elecciones generales 2019. Recuperado el 23 de abril de 2020 en: http://www.cis.es/cis/export/sites/default/-Archivos/Marginales/3240_3259/3242/Cru3242edad.html.

Cobo Romaní, C. & Moravec, J.W. (2011). *Aprendizaje Invisible. Hacia una nueva ecología de la educación*. Publicacions i Edicions de la Universitat de Barcelona.

Chu, Z., Gianvecchio, S., Wang, H. & Jajodia, S. (2012). Detecting Automation of Twitter Accounts. *IEEE Transactions on Dependable and Secure Computing, 9*(6), 811–824.

MONOGRAPH

Ciampaglia, G. L. (2017). Fighting fake news: a role for computational social science in the fight against digital misinformation. *Journal of Computational Social Science, 1*(1), 147–153. https://doi.org/10.1007/s42001-017-0005-6

DataPolitik (2019). *Espejismos en la red. Análisis de la campaña #SiguemeYTeSigoVox*. Asociación Heurística.

European Commission (2017). *Testing and Refining Criteria to Assess Media Literacy Levels in Europe*. DTI/EAVI.

European Commission (2009). *Study on Assessment Criteria for Media Literacy Levels*. European Union.

Ferrés, J. (2007). La competencia en comunicación audiovisual: dimensiones e indicadores. *Comunicar, 29*, 100-107. https://doi.org/10.3916/C29-2007-14

Fundación Telefónica (2019). *Sociedad Digital en Spain 2018*. Taurus.

Gorwa, R., & Guilbeault, D. (2018). Unpacking the Social Media *Bot*: A Typology to Guide Research and Policy. *Policy & Internet*, 1–24. https://doi.org/10.1002/poi3.184

Green, D.P., & Middleton, J.A. (2008). Do Community-Based Voter Mobilization Campaigns Work Even in Battleground States? Evaluating the Effectiveness of MoveOn's 2004 Outreach Campaign. *Quarterly Journal of Political Science, 3*(1), 63–82. https://doi.org/10.1561/100.00007019

Grimme, C., Preuss, M., Adam, L., & Trautmann, H. (2017). Social *Bots*: Human-Like by Means of Human Control? *Big Data, 5*(4), 279–293.

Guzmán-Simón, F., García-Jiménez, E., & López-Cobo, I. (2017). Undergraduate students' perspectives on digital competence and academic literacy in a Spanish University. *Computers in Human Behavior, 74*, 196-204.

Howard, P. N., Woolley, S., & Calo, R. (2018). Algorithms, *bots*, and political communication in the US 2016 election: The challenge of automated political communication for election law and administration administration. *Journal of Information Technology & Politics, 15*(2), 81–93. https://doi.org/10.1080/19331681.2018.1448735

Imbernón, F., & Medina, J. L. M. M. (2008). *Metodología participativa en el aula universitaria*. La participación del alumnado. Cuadernos de docencia universitaria. Octaedro.

Instituto Nacional de Ciberseguridad (INCIBE) (2019). *Kit de concienciación*. Retrieved from: https://www.incibe.es/protege-tu-empresa/kit-concienciacion.

Instituto Nacional de Estadística (INE) (2018). *Población que usa Internet (en los últimos tres meses)*. *Tipo de actividades realizadas por Internet*. Retrieved from: https://www.ine.es/ss/Satellite?L=es1254735110672&pagename=ProductosYServicios%2FPYSLayout

Internet Segura For Kids (IS4K) (2019). *Troles y haters en la Red, ¿cómo actuar?* [online]. Retrieved from: https://www.is4k.es/blog/troles-y-haters-en-la-red-como-actuar

Jiménez-Cortés, R., Rebollo-Catalán, A., García-Pérez, R. & Buzón-García, O. (2015). Motivos de uso de las redes sociales virtuales: Análisis de perfiles de mujeres rurales. *RELIEVE, 21*(1). doi: 10.7203/relieve.21.1.5153

Kellner, D. (2001). New Technologies/New Literacies: Reconstructing Education for the New Millennium. *International Journal of Technology and Design Education, 11*, 67–81.

Kluzer S., & Pujol Priego L. (2018). *DigComp into Action - Get inspired, make it happen. A user guide to the European Digital Competence Framework*. European Commission. https://doi.org/10.2760/112945.

Kreiss, D., & Jasinski, C. (2016). The Tech Industry Meets Presidential Politics: Explaining the Democratic Party's Technological Advantage in Electoral Campaigning, 2004–2012. *Political Communication, 33*(4), 544–562. https://doi.org/10.1080/10584609.2015.1121941

Levi, S., Guixaró, R., Carbonell, M., Palau, G., Bodi, E., García, G., & Alberti, P. (2019). *Informe para la acción estratégica y legislativa. Monopolios de la manipulación informativa y recortes de la libertad de expresión*. Xnet.

Maldita.es (2019). Maldita educa [online]. Retrieved from: https://maldita.es/nosotros/maldita-educa/.

Mayor-Buzón, V., García-Pérez, R., & Rebollo-Catalán, A. (2019). Explorando factores predictores de la competencia digital en las redes sociales virtuales. *Pixel-Bit. Revista de Medios y Educación*, 56, 51-69. https://doi.org/10.12795/pixelbit.2019.i56.03

Ministerio de Educación de Spain (2011). *Competencia mediática. Investigación sobre el grado de competencia de la ciudadanía en Spain*. Ministerio de Educación Cultura y Deporte.

Ministerio de Educación de Francia (2019). *Informer et accompagner les professionnels de l'éducation*. Éduscol.

Monreal Guerrero, I. M., Parejo, J. L., & Cortón de las Heras, M. D. L. O. (2017). Alfabetización mediática y cultura de la participación: retos de la ciudadanía digital en la Sociedad de la Información. *EDMETIC, 6*(2), 148-167.

Murthy, D., Powell, A. B., Tinati, R., Anstead, N., Carr, L., Halford, S. J. & Weal, M. (2016). *Bots* and political influence: A sociotechnical investigation of social network capital. *International Journal of Communication, 10*(20), 4952–4971.

Nickerson, D. & Rogers, T. (2014). Political Campaigns and Big Data. *Journal of Economic Perspectives, 28*(2), 51–74. https://doi.org/10.2139/ssrn.2354474

Nielsen, R. K. & Vaccari, C. (2013). Do People "Like" Politicians on Facebook? Not Really. Large-Scale Direct Candidate-to-Voter Online Communication as an Outlier Phenomenon. *International Journal of Communication, 7*, 2333–2356.

OECD (2008). New Millennium Learners. Initial findings on the effects of digital technologies on school-age learners.CERI.

OECD (2015). *Programme for International Student Assessment*. PISA.

Oficina de Seguridad del Internauta (OSI) (2019). *Guía de privacidad y seguridad en internet* [online]. Retrieved from: https://www.osi.es/es/guia-de-privacidad-y-seguridad-en-internet?origen=tip

Pastor-Galindo, J., Zago, M., Nespoli, P., López Bernal, S., Huertas Celdrán, A., Gil Pérez, M., Ruipérez-Valiente, J. A., Martínez Pérez, G, y Gómez Mármol, F. (2020). Spotting political social *bots* in Twitter: A use case of the 2019 Spanish general election. *arXiv* preprint arXiv:2004.00931.

Pereira, S., Fillol, J. & Moura, P. (2019). El aprendizaje de los jóvenes con medios digitales fuera de la escuela: De lo informal a lo formal. *Comunicar, 58*, 41-50. https://doi.org/10.3916/C58-2019-04

Pérez-Pérez, I. (2014). La metodología participativa en la Educación Superior: una evaluación de los estudiantes de Sociología. *Espacio abierto, 23*(4), 643-660.

Persily, N. (2017). Can democracy survive the Internet? *The Journal of Democracy, 28*(2), 63–75.

Portalés-Oliva, M. (2019). *Nuevos entornos digitales y alfabetización mediática. Dispositivos móviles, jóvenes y el lenguaje audiovisual*. Tesis doctoral. Universitat Autònoma de Barcelona.

Powers, S., Kounalakis, M. (2017). *Can Public Diplomacy Survive the Internet? Bots, Echo Chambers, and Disinformation*. Advisory Commission on Public Diplomacy.

Redondo, M., Calvo, D. & Díez-Garrido, M. (2017). Calisto, un software para la construcción del discurso político/ Calisto, a software for the construction of political discourse. *El profesional de la Información, 26*(4), 756–764. https://doi.org/10.3145/epi.2017.jul.19

Scolari, C.A. (2016). Alfabetismo transmedia: estrategias de aprendizaje informal y competencias mediáticas en la nueva ecología de la comunicación. *Telos, 193*, 13-23.

Shao, C., Ciampaglia, G. L., Varol, O., Yang, K.-C., Flammini, A. & Menczer, F. (2018). The spread of low-credibility content by social *bots*. *Nature Communications, 9*(4787). https://doi.org/10.1038/s41467-018-06930-7

Stella, M., Ferrara, E. & De Domenico, M. (2018). *Bots* sustain and inflate striking opposition in online social systems, 115(49), 12435–12440. https://doi.org/10.1073/pnas.1803470115

Tucker, J. A., Theocharis, Y., Roberts, M. E., Barberá, P., Tucker, J. A., Theocharis, Y. & Roberts, M. E. (2018). From Liberation to Turmoil : Social Media And Democracy. *Journal of Democracy, 28*(4), 46–59.

Tufekci, Z. (2014). Engineering the public: Big data, surveillance and computational politics. *First Monday, 19*(7). http://dx.doi.org/10.5210/fm.v19i7.4901

UNESCO (1982). *Declaración de Grünwald sobre la educación relativa a los medios de comunicación*. Framework for Action.

Valerio-Ureña, G. & Valenzuela-González, R. (2011). Redes sociales y estudiantes universitarios: del nativo digital al informívoro saludable. *El Profesional de la Información, 20*(6), 667-670. https://doi.org/10.3145/epi.2011.nov.10

Villasante, T. R., Montañés, M., & Martí, J. (2002). *La investigación social participativa*. Construyendo ciudadanía. El viejo topo.

Vosoughi, S., Roy, D. & Aral, S. (2018). The spread of true and false news online. *Science, 359*, 1146–1151. https://doi.org/10.1126/science.aap9559

Vuorikari, R., Punie, Y., Carretero Gomez S. & Van den Brande, G. (2016). *DigComp 2.0: The Digital Competence Framework for Citizens*. European Commission. https://doi.org/10.2791/11517

Woolley, S. C. (2016). Automating power: Social *bot* interference in global politics. *First Monday, 21*(4). http://dx.doi.org/10.5210/fm.v21i4.6161

Woolley, S. C., & Howard, P. N. (2016). Political communication, Computational Propaganda, and autonomous agents. *International Journal of Communication, 10*, 4882–4890. https://doi.org/1932–8036/20160005

Notes

[1]   Botometer is an online application developed at the Indiana University Network Science Institute (IUNI) and Center for Complex Networks and Systems Research (CNetS). Software documentation is available at: https://botometer.iuni.iu.edu/#!

Research instruments

Dafne Calvo, Lorena Cano-Orón, & Almudena Esteban. (2020, June 6). Assessment of literacy level for the recognition of social bots in political misinformation contexts. Zenodo. http://doi.org/10.5281/zenodo.3882244