

**H2020-ICT-2018-2 /ICT-28-2018-CSA**  
**SOMA: Social Observatory for Disinformation and Social Media**  
**Analysis**



## D5.1 SOMA Impact Assessment Methodology

<b>Project Reference No</b>	SOMA [825469]
<b>Deliverable</b>	D5.1 SOMA Impact Assessment Methodology
<b>Workpackage</b>	WP5 Assessing the impact of the European Observatory
<b>Type</b>	R = Document, report
<b>Dissemination Level</b>	PU = Public
<b>Date</b>	30/01/2020
<b>Status</b>	V1
<b>Editor(s)</b>	Simona De Rosa, T6 Ecosystems; Andrea Nicolai, T6 Ecosystems
<b>Contributor(s)</b>	Noemi Trino, LUISS Data Lab; Stefano Guarino, LUISS Data Lab
<b>Reviewer(s)</b>	Noemi Trino, LUISS Data Lab; Stefano Guarino, LUISS Data Lab; Gianni Riotta, LUISS Data Lab
<b>Document description</b>	D5.1 describes the methodological framework developed by SOMA for the impact assessment methodology. The deliverable also contains the Source Transparency Index elaborated by the project.

## Document Revision History

Version	Date	Modifications Introduced	
		Modification Reason	Modified by
<b>V0.1</b>	03/07/2019	Table of content and first contents	T6 ECO
<b>V0.2</b>	06/12/2019	Request for contribution to Luiss Data Lab	T6 ECO
<b>V0.3</b>	20/12/2019	Input from Luiss	Luiss Data Lab
<b>V0.4</b>	13/01/2020	Further input from Luiss	Luiss Data Lab
<b>V0.5</b>	16/01/2020	Input integration and finalisation	T6 ECO
<b>V0.6</b>	17/01/2020	Lat input from T6	T6 ECO
<b>V.0.7</b>	20/01/2020	Sent to internal review	T6 ECO
<b>V.0.8</b>	29/01/2020	Review received from LUISS	Luiss Data Lab
<b>V.0.9</b>	29/01/2020	Comments integrated	T6 ECO
<b>V1</b>	30/01/2020	Final version sent to the coordinator for submission	T6 ECO

# Executive Summary

The deliverable 5.1 “SOMA” Impact assessment methodology introduces the work of the Work Package (WP) 5 that will last until the end of the project aimed to develop and test a framework to assess the impacts of disinformation and to provide policy guidelines. Around this topic, further topics related to the impact assessment methodology have been discussed and elaborated.

Following the definition adopted by the High Level Expert Group established by the European Commission in 2018, the definition of disinformation used in the deliverable refers to “all forms of false, inaccurate, or misleading information designed, presented and promoted to intentionally cause public harm or for profit”. In line with the report, we discard references to fake news and misinformation and we concentrate on those false information created intentionally to harm the general public.

According to what stated in the Description of Action (DoA), D5.1 sets up the methodological framework to assess the impact of disinformation and plan the activities for the next months. Indeed, WP5 activities started at the beginning of the project with the identification of current methodological gaps when talking about measuring the impact of disinformation. On such base, we built up a combined and mixed approach that should improve the comprehension of the impacts of disinformation looking at what happens on social media and on the real lives of people.

The deliverable reports the theoretical background and the methodology together with the methods that will be used to collect and analyze data.

The deliverable also describes the proposed approach to quantify the specific impacts produced by the SOMA Observatory.

Then, a relevant part of the methodology is dedicated to the creation of a SOMA Source Transparency Index (STI). This is a list of criteria that should be used to identify when the information circulating on social media is trustworthy or not. The creation of the SOMA STI started from the analysis of existing solutions. Then, SOMA has aggregated some of the already existing indicators and the list has been validated with stakeholders in order to agree on what already proposed by other stakeholders, add what was missing and, finally, create the SOMA index

The deliverable contains also a report about a consultation with data analysts expert on data access and, in particular, about the prioritization of which data should be available for research purposes in order to inform policy makers of the needs and priorities of the research communities.

Finally, the deliverable contains preliminary results of a scientific collaboration established with a research project, [Toffee](#), aimed to investigate the impact of disinformation on the topic of 5G on social media.

## Table of Contents

Executive Summary .....	3
1. Introduction .....	6
1.1 Purpose and Scope .....	7
1.2 Structure of the Deliverable .....	7
2. A methodology to assess the impact of disinformation .....	8
2.1 The framework .....	9
2.2 The process .....	19
2.3 The timeline .....	22
3. Assessing the impacts of the SOMA Observatory .....	23
3.1 The methodological framework to assess the impact of the SOMA Observatory .....	23
3.2 The stakeholder groups and areas of impact .....	24
4. The SOMA Source Transparency Index .....	27
4.1 The SOMA process for establishing a co-designed STI .....	28
4.2 Validation of the SOMA STI index: the process and main results .....	29
5. Data Access for research purposes: what is needed .....	32
5.1 The exercise .....	33
5.2 The results .....	34
6. Conclusions .....	36
7. References .....	38
8. Annexes .....	40
8.1 An initial analysis of BOT presence in the debate of 5G on Social Media by SOMA and TOFFEE projects .....	40
8.2 Use case template .....	43

## List of Figures

Figure 1. SOMA's impact areas .....	11
Figure 2. SOMA's impact areas and related dimensions .....	14
Figure 3. SOMA process for impact assessment .....	21
Figure 4. Impact Value Chain Approach. Source: Adaptation from Impact Measurement working group, 2014 .....	24
Figure 5. STI validation results .....	30

## List of Tables

Table 1. SOMA impact areas with indicators and variable .....	18
Table 2. SOMA activities and related impact areas .....	25
Table 3. SOMA impact areas and related indicators .....	26
Table 4. SOMA Source Transparency Index draft version .....	29
Table 5. SOMA STI validated by the users .....	31

## List of Terms and Abbreviations

Abbreviation	Definition
<b>DOA</b>	Description of Action
<b>EC</b>	European Commission
<b>GA</b>	Grant Agreement
<b>HLEG</b>	High Level Expert Group
<b>STI</b>	Source Transparency Index
<b>UGC</b>	User Generated Content
<b>WP</b>	Work Package

# 1. Introduction

According to the figures provided by Statista (2020), in the world there are almost 3 billion active social media users. Such figures are supposed to increase constantly. In 2018, indeed, the number of estimated users was 2.65 billion, the number will increase to almost 3.1 billion in 2021.

In Europe there are almost 500 million active social media users, and this accounts for an average penetration of over 50% of the European population. This widespread adoption of social media has increased the dynamics of information diffusion. If on the one hand, such a large number of users, and the exponential increase in the use of Social Media platforms, highlights a demand to access knowledge and information, and channels through which to express oneself, on the other hand, the exponential growth and circulation of disinformation “represents risks for our democratic processes, national security, social fabric, and can undermine trust in the information society and confidence in the digital single market” (HLEG report, 2018). Such data seems to be confirmed by a survey conducted by Eurobarometer (2018) covering 26,000 EU citizens in the 28 EU member states. Results show that over 83% considered online disinformation a threat to democracy<sup>1</sup>.

However, even if the debate on disinformation is pressing in everyday life and several researches have been conducted to learn how disinformation can affect different aspects of citizens’ lives, such as impacts on opinions (Quattrocioni, 2016), impacts on rates of vaccination (Aquino et al. (2017) and also impacts on electoral outcomes (Allcott and Gentzkow, 2017; Fournay et al. 2017), still a lot has to be discussed to understand the consequences of the phenomenon.

Indeed, what is still missing in the academic and political debate, is a clear understanding and a tangible measure of the impacts of disinformation on public and personal opinion, culture and policy. Using the words of the Reuters Institute (2018) “We currently know little about the impact that false news has on people’s attitudes and beliefs, which is often the underlying concern”.

This lack means that i) it is unclear which are the tools and methods able to provide better insights on the phenomenon ii) which are the concrete consequences on European populations in terms of social and political behaviours. To fill the gap, research seems to be the only opportunity, though, to understand why the phenomenon is so urgent, what are the effects on the people and what are the tools to use to contrast the spread of disinformation to limit the effects on our democracies.

To foster the discussion around the topic, SOMA has dedicated an entire Work Package (WP) to address this complex issue through the definition of a method to analyse the impacts of disinformation deriving evidence-based recommendations for policy makers. The aim, indeed, is to define a combined framework based on qualitative and quantitative methodology to assess and quantify the impacts of disinformation on the people and test the methodology in order to provide results of the use of the methodology within the project lifetime. Finally, all the results from the methodology development and its validation will be translated into policy recommendations for decision makers at the end of the project. The WP5 has, also, the scope to assess the impacts of the project overall. For this reason, an entire chapter is dedicated to how the impact of SOMA will be measured and analysed.

<sup>1</sup> Eurobarometer 464 (2018) Final results of the Eurobarometer on fake news and online disinformation. <https://ec.europa.eu/digital-single-market/en/news/final-results-eurobarometer-fake-news-and-online-disinformation>

SOMA aim, additionally, is to develop a tool to assess the trustability of information. So, a Source Transparency Index (STI) has been created and validated with users and discussed here in the report.

## 1.1 Purpose and Scope

The purpose of the deliverable is to describe the theoretical framework for the impact assessment of disinformation created by SOMA. The deliverable contains a review of current approaches used to assess the impact of disinformation, mainly on social media, and intends to propose a combined framework able to read social media dynamics with the addition of the impacts on the real world. To do it properly, SOMA relies on a cross-disciplinary approach where different scientific competences work together to improve the complexity of the analysis and its comprehension. Moreover, the deliverable also contains the methodology that will be used to measure the impact of the SOMA Observatory overall. The deliverable contains, also, the results gathered on the creation of a Source Transparency Index co-developed with the users, and a brief information for the EC on which data the platforms should open as soon as possible in order to guarantee to run research activities.

## 1.2 Structure of the Deliverable

The deliverable is structured in four main chapters.

Chapter 2- A methodology to assess the impact of disinformation- reports the description of the methodological framework to assess the impacts of disinformation. First of all, it describes the reason why SOMA is focusing on this topic. Second, it describes the theoretical approach under which the methodological framework has been built. Finally, it provides a description of methods for data gathering and data collection that will be used. An overview of the timeline for activities closes the chapter.

Chapter 3-Assessing the impacts of the SOMA Observatory - describes the process that will be implemented to quantify the impacts of the overall SOMA Observatory, as described in the DoA. One of the WP5 aim is indeed to analyse the impacts reached by the Observatory over the 30 months. For this reason, a methodology to assess the impacts following the impact value chain project has been set up and is here described.

Chapter 4- The SOMA Source Transparency Index- describes the work done to build up the Source Transparency index created by SOMA. The aim was to aggregate already existing variables to create a STI for single news outlets. As the aim of the task was to develop the STI together with users, the Chapter describes also the work done with user to co-develop the index following a participatory approach.

Chapter 5- Data Access for research purposes: what is needed- report some information for the European Commission about the top priorities of researchers when accessing data for research purposes. Information has been collected from the direct engagement of data analysts from research institutions working on disinformation, through a template reported in Annex 8.2 Use case template.

In the Annex 8.1 An initial analysis of BOT presence in the debate of 5G on Social Media by SOMA and TOFFEE projects- the preliminary result of the collaboration established between T6 Ecosystems are reported, on the behalf of the SOMA project, and the Toffee project working on disinformation as well to investigate the relation between bots and the discussion on the 5G on social media.

## 2. A methodology to assess the impact of disinformation

As a matter of fact, disinformation is an urgent topic that needs to be treated carefully. It deserves a lot of attention due to its potential impacts on several aspects of our lives.

To get the real understanding of disinformation, it was stated very clearly by the High Level Expert Group established by the European Commission (EC) on January 2018 the need to “promote continued research on the impact of disinformation in Europe to evaluate the measures taken by different actors and constantly adjust the necessary responses”.

The reason why it is important to promote this kind of research is twofold. First, because only if we understand what is the impact we can provide specific and ad-hoc solutions to respond to the problem. Secondly, because the larger is our comprehension of the phenomenon the higher will be the opportunity to anticipate future challenges opened by the platforms or by other technologies that now or later will reshape our capability to interact with the information.

Thanks to a wide scientific literature it has been possible to prove that the quick and huge spread of disinformation through the web and social media platforms could generate some effects also on the real world and on real users reached by disinformation (Kumar and Shas, 2018). But it is also true that “we currently know little about the impact that false news has on people’s attitudes and beliefs, which is often the underlying concern” (Reuters Institute, 2018).

Several studies have been trying to understand the impact of disinformation on elections, however, as stated by Cresci et al. (2018) the results of these experiments have often been inconclusive or in contrast with previous results (O’Connor et al. 2010; Gayo-Avello 2012). The main problem emerged in relation to the fact that social data are, by their nature, constantly evolving and noisy. In other words, useful information for an analysis is often buried within a multitude of other data that are not interesting, or at worst, misleading.

So, still a lot of research must be undertaken to get robust evidence of the major impacts of disinformation and of the extent to which users are compromised by disinformation. Several research institutes and researchers are claiming for better focus on analysing the impact of disinformation. Just to mention some of them, Reuters Institute (2018) states that “Further research is needed to understand the reach and influence of online disinformation in these and other countries”. While Lazer et al. (2018) write “How common is fake news, and what is its impact on individuals? Solutions should be commensurate to the magnitude of the problem, but there are surprisingly few scientific answers to these basic questions”.

As a matter of fact, finding scientific evidence replying to this question is not easy. However, such a goal can be achieved only by taking into account that further reflections are needed on methods and approaches used to assess and quantify disinformation.

So far, several studies have investigated the impact of disinformation using the user engagement metrics, mainly counting the number of likes, shares, retweets, etc. As reported by Kumar and Shas, (2018) “On



web and social media, the impact is measured as the engagement it produces via its readers, using statistics such as number of reads, number of days it survived without being removed, or number of people it reached via reshares.”

To mention some of the studies conducted using the user engagement metrics, not pretending to provide an exhaustive list, we can cite, Frigerri et al. (2014) who analyse the spread of rumours on Facebook; Del Vicario et al. (2015) using quantitative analysis to understand how facebook users consume information related to conspiracy theories and scientific evidences; Bessi et al (2015) investigating the role of polarization on the web; Kumar et al. (2016), analysing the impact of hoax articles on Wikipedia; Silverman (2016), who focuses on the engagement of election fake news articles on Facebook.

Even if quantitative approaches such as complex networks and machine learning can provide relevant and deep understanding of social media data, making visible, for example, communities and common interests or determining the sentiment of some texts, some shortages are evident and it is fair to say that the comprehension of the phenomenon is still limited also due to the research approaches in use.

As written by Lazer et al. (2018) “knowing how many individuals encountered or shared a piece of fake news is not the same as knowing how many people read or were affected by it. Evaluating medium to long run impact on political behaviour of exposure to fake news (for example, whether and how to vote) is essentially non-existent in the literature. The impact might be small evidence suggests that efforts by political campaigns to persuade individuals may have limited effects. However, mediation of much fake news via social media might accentuate its effect because of the implicit endorsement that comes with sharing. Beyond electoral impacts, what we know about the effects of media more generally suggests many potential pathways of influence, from increasing cynicism and apathy to encouraging extremism. There exists little evaluation of the impacts of fake news in these regards”.

On the other hand, at the moment of writing, only few papers have utilized methods that are different from the quantitative methodologies to collect an evaluation of impacts. One of those is Allcott et al. (2019) who has provided an evaluation of the facebook’s impacts on a randomized sample of users to understand what happens when users deactivate their accounts. It is fair to say that also the qualitative approach has its own limits and failures. For example, it is difficult to create a representative sample or to collect a sufficient number of data to analyse.

Taking into account the pros and cons of both approaches and following the need to foster interdisciplinary research aimed at creating a sounding evaluation framework, SOMA has elaborated a mixed approach to evaluate the impacts of disinformation.

## 2.1 The framework

As reported by Reuters (2018) “Against the backdrop of increased online news use, and growth in the use of social media to find news (Newman et al. 2017), governments, policymakers, and other stakeholders have started to take formal steps towards assessing and tackling this issue”. However, there are two emerging issues.

First of all, as said by Lazer et al. (2018) “Evaluations of the medium-to-long-run impact on political behaviour of exposure to fake news (for example, whether and how to vote) are essentially non-existent in the literature”. This means that we still need additional knowledge on the impact that disinformation has on people and on their choices. Second, more academic research is needed to enlarge the base of evidence.

However, to find methods and approaches to map the impact of disinformation is not easy. As reported by Napoli (2014), due to the technological changes, the field of media impact assessment is in a period of rapid innovation and change, investigating new approaches to defining and measuring impact.

In particular, what Napoli stresses is the need to provide a comprehensive overview and assessment of media impact. In Napoli's word "In recent years, funders of public interest media initiatives have been working to improve the assessment strategies and tools that guide their work (see, e.g., Brock, Buteau, & Herring, 2012). Their goal is to have a more concrete and more comprehensive sense of the overall impact that these media initiatives are having, both within their intended audiences and beyond. Such impact can extend into realms such as politics, policy, education, public health, and the economy, depending on the focus and goals of the particular public interest media initiative".

Due to the relevancy of media, it is important to focus on the social value perspective. This means that it is important to go beyond the economic and financial impact to look more at the impact in regard to the well-being of individuals and communities.

Such an approach can be easily translated on the impact of social media and, as a consequence, to the impact of disinformation spread through social media. However, what is missing is a deep understanding on how this impact on people in their every day-life in their social sphere of action.

To fill the gap, the SOMA methodology has been developed taking into account three dimensions: social, economic and political.

First of all, as stated by Whiteman (2004), the provision of a methodological approach for the impact of media is a complex process. For this reason, there is the need to use and deploy multiple methods to collect data and to then analyse a wider range of aspects that a single approach cannot cover. For example, questionnaires, interviews, participant observation and content analysis are some of the methods that could be used in parallel to get a complete picture of the context and a deep insight on the data collected.

In order to properly enlarge the data collection and the investigation, SOMA proposes a framework where quantitative and qualitative methods are both used. Both methods, indeed, are not conflicting as they look at different data and different aspects. For this reason, we think that a combination of both methods can create an added value in the comprehension of the phenomenon.

Quantitative methodologies such as complex networks analysis and machine learning will be used at the beginning of the process to map the impact of disinformation spreading on social media (more information about techniques are reported in Paragraph 2.2). For example, how long disinformation survives, how often disinformation is viewed, and how frequently news outlets containing disinformation are cited on the Web will be crucial to understand the impact on social media users. However, even if these aspects are very relevant, they are not the only ones to consider. Once the results of the quantitative analysis will be ready, data gathering will be organized through questionnaires, focus groups and interviews to understand the impact on real users according to their beliefs and choices (more information are reported in Paragraph 2.2).

As reported by Napoli (2014) as a starting point it is important to differentiate the concept of media impact that is the focus here from the more traditional notion of media effects. "First, while there is certainly a substantial amount of overlap (see, e.g., Leiserowitz, 2004), the field of media effects can be

characterized as having a strong micro orientation, in that the unit of analysis is typically the individual media user, and the focus is on the relatively narrow question of whether exposure to a particular media message impacted that user's attitudes, beliefs, cognitions or behaviours. When we talk about media impact, on the other hand, the orientation can be characterized as a bit more macro, in that the concerns extend beyond whether individual media users had their attitudes, beliefs, cognitions, or behaviours affected, to also include broader systemic changes at the levels of organizations and institutions" (Napoli, 2014).

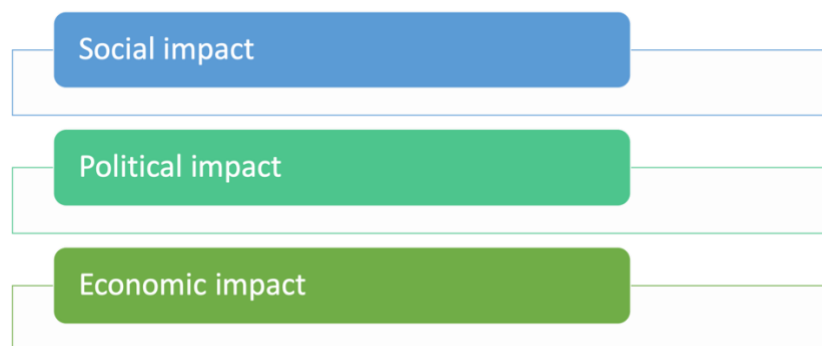
It is in the interest of SOMA to look at what are the impacts on the individuals (at micro level) in order to then link which are the impact on institutions (macro level). Essentially, the SOMA methodology will allow data collection at individual level exploring the link to political and scientific institutions. Through this data gathering it will be possible to understand the impacts on institutions, also, in terms of trust.

The methodology for impact assessment is built on a robust methodology called Sequoia (Passani et al. 2014), developed and validated in previous European projects, including SEQUOIA, MAXICULTURE and especially IA4SI, and redefined for the scope of SOMA.

The Sequoia methodology has been used as a starting point for the definition of the SOMA methodology in order to build up specific indicators and variables fitting for the scope of the analysis.

In line with this, the framework used for SOMA aims to map three different areas of impact (Figure 1):

- Social
- Political
- Economic



*Figure 1. SOMA's impact areas*

### **Social impact**

The social impact has been declined in terms of a significant change that addresses a pressing social challenge.

The increase of information available on the web has certainly opened a lot of opportunities for users to access news and information as well as to share contents and to distribute user generated content (UGC). For this reason, the aim of the social index is to understand the changes that have occurred on individuals at micro level. The social impact index is composed of 7 sub-categories focused on the six crucial aspects that link information to opinions, social inclusion, digital skills and media literacy skills:

1. Impact on news awareness and bias on information spread on social media
2. Impact on information quality
3. Impact on news media access
4. Impact on social inclusion
5. Impact on opinions and behaviours change through the use of social media

6. Impact on digital skills
7. Impact on media literacy and digital media literacy

### **Political impact**

As stated by the Sequoia methodology, the political impact index looks at how a phenomenon influences “participatory democratic processes” having an impact on the users and on European citizens political participation.

Political impact is particularly relevant for the scope of the SOMA project since the role of digital spaces and social media platforms in influencing the functioning of democracies worldwide is currently highly discussed (e.g. Brexit, Trump’s election), (Gaumont et al., 2018).

Moreover, as reported by the European Commission (2018), “Disinformation erodes trust in institutions and in digital and traditional media and harms our democracies by hampering the ability of citizens to take informed decisions<sup>2</sup>”.

Indeed, several studies are focused on political activities using Twitter or data coming from social media. Garimella et al (2016) focused on the analysis of Twitter to understand the political debate, as well as Connover et al. (2011) analysed the political situation in the USA. Allcott and Gentzkow (2017) analysed the role of fake news in the Trump victory in the 2016 US presidential elections. Fournery et al. (2017) analysed the correlation between the spread of fake news across US states and voting outcomes in the US. Also, polarization has been widely studied thanks to the works Quattrociocchi et al. (2016) who has analysed the creation of echo chambers on Facebook to measure communities’ polarization, and other scholars (Morales et al., 2015). Other examples of network theory applied to social media are provided by Cherepnalkoski et al. (2016; 2017), who determined the political affiliation of the Members of the European Parliament according to retweets.

However, these relevant studies are all focused on the analysis of patterns and connections on social media. This means that the evidence retrieved by the analysis explains what the processes on the social media world are and it provides some hypotheses on how this is reflected on the social interactions outside the platforms.

For this reason, the SOMA methodology aims to add 6 sub-dimensions dedicated to analysing the political impact on disinformation on real users and on their everyday life.

In addition to the impact on democratic participation, the SOMA project has enlarged the range of investigation mapping the impacts in relation to the following dimensions extrapolating indicators, where possible, from already well-known approaches active in the field of media studies:

1. Impact on digital democracy
2. Impact on trust in institution
3. Impact on electoral votes
4. Impact on media freedom and pluralism
5. Impact on access to information and trustability of news outlet
6. Impact on polarization

### **Economic impact**

Another relevant aspect of disinformation is related to the economic sphere.

<sup>2</sup> Available at <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0236&from=EN>

Platforms, in fact, are changing ways of consumption and users' economic behaviours also dramatically influencing the media value chain.

Agnihotri et al., (2016) stress the positive effects of the platformisation stating that “Social media is changing our world. Through the advent of smartphones and social media, accessibility of information is higher than it ever has been before. Customers are frequently asked to “like” companies on Facebook, to “follow” companies on Twitter, or to “connect” via LinkedIn. As a result, customers are becoming better connected to companies, more knowledgeable about product selections, and more powerful in buyer-seller relationships”.

This aspect is certainly true, however, the numbers related to the economic losses caused by disinformation are striking.

The UK Royal Economic Society (RES, 2018) recently stated that analysing 49 million tweets mentioning the names of 55 large FTSE100 companies, social media bots have influenced stock market performance, while other findings highlight that also small investors fall victim to bot tweets. On the other hand, “Forbes reported (Rapoza, 2018) that \$130 billion in stock value was lost when the Associated Press (AP) Twitter account was hacked and a post was made claiming there was an “explosion” that injured Barack Obama. Although stock prices recovered, this points to how disinformation on social media can impact high-frequency trading algorithms that rely on text signals to make investment calls. Similarly, the French construction company Vinci became a target in 2016 of a false press release (Agnew, 2016) which caused share price to go down by 19% (a loss of more than 6 billion euro) before recovering by close of trading” (EPRS, 2019).

It is also true that few people can take advantage of social media interaction and algorithms to reach a wider public, monetizing visibility by publishing click-bait stories or using politically driven content (Flore et al., 2019). In this case, the use of headlines to prompt a person into clicking to view an article typically aims to generate advertisement-based revenues for the publication. The more clicks an article receives, the more revenues it earns. For this reason, clickbaiting is designed to attract a large number of readers and rarely uses quality guidelines and researched sources. In fact, digital advertising seems to be one of the possible ways to spread disinformation<sup>3</sup>. Digital advertisements are often “click-based, which rewards sensational and viral content. Particularly dangerous is the functioning of the so called “High risk” advertising. This typology can have a harmful impact on consumers as they can be designed to confuse consumers and defraud Brands. These High Risk ad sectors include ads with adult content, malware and those making use of fraudulent techniques, such as system fraud, click generator fraud or prize fraud<sup>4</sup>” (EUIPO, 2016).

In addition, also the way of news consumption has dramatically changed. As reported by Forbes “Social media has become the main source of news online with more than 2.4 billion internet users, nearly 64.5 percent receive breaking news from Facebook, Twitter, YouTube, Snapchat and Instagram instead of traditional medias”. However, identifying the impacts of new business models on news consumption is still a topic of investigation. The same holds for the relation between news consumption and transparency and media ownership awareness.

<sup>3</sup> Available at <https://eur-lex.europa.eu/legal-content/EN/TEXT/PDF/?uri=CELEX:52018DC0236&from=EN>

<sup>4</sup> Available at

<https://euiipo.europa.eu/ohimportal/documents/%2011370/80606/Digital+Advertising+on+Suspected+Infringing+Websites>

<sup>5</sup> Available at <https://www.forbes.com/sites/nicolemartin1/2018/11/30/how-social-media-has-changed-how-we-consume-news/>

To clarify those aspects in relation to the other vertical impact area, the SOMA methodology will investigate the following 4 sub-dimensions:

1. Impact on news consumption patterns
2. Impact on users subscription to online journals
3. Impact on online advertising
4. Impact on transparency of media ownership, as an element to map the changes on impact media value chain

The following image reports the three impact areas with all related dimensions (Figure 2).

Social Impact	Political Impact	Economic Impact
<ul style="list-style-type: none"> <li>•Impact on awareness and bias on information on social media</li> <li>•Impact on information quality</li> <li>•Impact on news media access</li> <li>•Impact on opinions and behaviours change through the use of social media</li> <li>•Impact on social inclusion</li> <li>•Impact on media literacy and digital media literacy</li> </ul>	<ul style="list-style-type: none"> <li>•Impact on civic and democratic participation</li> <li>•Impact on digital democracy</li> <li>•Impact on trust in political institutions (politicians, government, parliament)</li> <li>•Impact on citizens' behaviours in relation to electoral vote</li> <li>•Impact on media freedom pluralism</li> <li>•Impact on news knowledge and access to information</li> <li>•Impact on media reliability</li> <li>•Impact on polarisation</li> </ul>	<ul style="list-style-type: none"> <li>•Impact on news consumption patterns</li> <li>•Impact on users subscription to online journals</li> <li>•Impact on online advertising</li> <li>•Impact on transparency of media ownership, as an element to map the changes on impact media value chain</li> </ul>

*Figure 2. SOMA's impact areas and related dimensions*

The Table 1 illustrates the impact areas, the indicators and the variables that will be used for the identification of impact of disinformation. However, even if preliminary questions have been already inserted in order to provide a clear idea of the process, the questionnaire will be redefined according to the false news outlet selected to run the validation.

At the moment of writing, the three impact areas and related variables and indicators have been defined as reported hereafter. Indicators are synthetic descriptions of a sub-dimension and can be composed of one variable (simple indicator) or of two or more variables (complex indicator). Variables are characteristic of a unit being observed that may assume a numerical measure or a qualitative definition. Variables will be transformed into one or more questions in the questionnaire that SOMA will develop in order to gather the data needed for the identification and description of impacts.

Area of impact	Sub-dimension	Indicator	Variable
Social impact	<ul style="list-style-type: none"> <li>Impact on news awareness and bias on information spread on social media</li> </ul>	<ul style="list-style-type: none"> <li>Level of information neutrality perceived</li> </ul>	<ul style="list-style-type: none"> <li>How do you assess news neutrality shared on social media? (Likert scale)</li> </ul>
	<ul style="list-style-type: none"> <li>Impact on information quality</li> </ul>	<ul style="list-style-type: none"> <li>Perceived quality of the information</li> </ul>	<ul style="list-style-type: none"> <li>Assess the quality of the information received via social media (Likert scale)</li> <li>Self-Perceived ability to recognize disinformation</li> </ul>
	<ul style="list-style-type: none"> <li>Impact on news media access</li> </ul>	<ul style="list-style-type: none"> <li>Ways and processes for news consumption</li> </ul>	<ul style="list-style-type: none"> <li>How did you get into the news? Please describe the process (Consider closed options for data collection. Eg:1) A friend shared the news on SN 2) Media public page on a social 3) Search engine 4) Online newspaper 5) Link on other website 6) Other (specify)</li> <li>From which sources do you generally get information (local newspaper, national newspaper, tv and radio, social media)</li> </ul>
	<ul style="list-style-type: none"> <li>Impact on social inclusion</li> </ul>	<ul style="list-style-type: none"> <li>Perception on level of access and availability of media for vulnerable groups</li> </ul>	<ul style="list-style-type: none"> <li>Assess social media inclusiveness for vulnerable groups (Likert scale)</li> </ul>
	<ul style="list-style-type: none"> <li>Impact on opinions and behaviors change through the use of social media</li> </ul>	<ul style="list-style-type: none"> <li>Change behaviours according to the news</li> </ul>	<ul style="list-style-type: none"> <li>Would you change your behavior according to the news? Y/N</li> </ul>
	<ul style="list-style-type: none"> <li>Impact on media literacy and digital media literacy</li> </ul>	<ul style="list-style-type: none"> <li>Level of expertise in digital media skills</li> </ul>	<ul style="list-style-type: none"> <li>Ability to interpret and evaluate the various codes of representation and the function they perform within a message.</li> <li>Understanding the role played by the information and communication</li> </ul>

			<p>technologies in society, and their possible effects.</p> <ul style="list-style-type: none"> <li>Ability to discern why certain media, products or content are popular and why they are successful individually or collectively: the wants and needs that satisfy the senses, emotions, and stimulate the cognitive, aesthetic and cultural interest, etc., of audiences</li> </ul>
Area of impact	Sub-dimension	Indicator	Variable
Political impact	<ul style="list-style-type: none"> <li>Impact on civic and democratic participation</li> </ul>	<ul style="list-style-type: none"> <li>Level of civic participation</li> </ul>	<ul style="list-style-type: none"> <li>How much are you engaged in civic participation (Likert scale)</li> </ul>
	<ul style="list-style-type: none"> <li>Impact on digital democracy</li> </ul>	<ul style="list-style-type: none"> <li>Level of participation to debate (online and offline)</li> </ul>	<ul style="list-style-type: none"> <li>How closely are you engaged in public debate? (Likert scale)</li> <li>Do you follow politics on social media?</li> </ul>
	<ul style="list-style-type: none"> <li>Impact on trust in institutions (politicians, government, parliament, international organization, scientists)</li> </ul>	<ul style="list-style-type: none"> <li>Level of trust in institutions</li> </ul>	<ul style="list-style-type: none"> <li>Do you trust political institutions? Y/N</li> <li>Consider adding another question: on a scale from 0 to 5, how much do you trust the following institutions: a) parliament b) government c) politicians d) EU</li> <li>Assess the trustability of news shared by scientific stakeholders (Likert scale)</li> </ul>
	<ul style="list-style-type: none"> <li>Impact on citizens' behaviours in relation to electoral vote</li> </ul>	<ul style="list-style-type: none"> <li>Change the vote attitude according to the news</li> </ul>	<ul style="list-style-type: none"> <li>Would you change your vote according to the news? Y/N</li> </ul>
	<ul style="list-style-type: none"> <li>Impact on media freedom pluralism</li> </ul>	<ul style="list-style-type: none"> <li>Access to different sources of information</li> </ul>	<ul style="list-style-type: none"> <li>How many sources do you have access to read a news? (Likert scale)</li> </ul>



	<ul style="list-style-type: none"> <li>Impact on news knowledge and access to information</li> </ul>	<ul style="list-style-type: none"> <li>Level of engagement of politics</li> </ul>	<ul style="list-style-type: none"> <li>How closely do you follow politics (Likert scale)</li> </ul>
	<ul style="list-style-type: none"> <li>Impact on media reliability</li> </ul>	<ul style="list-style-type: none"> <li>Trust in media outlets</li> </ul>	<ul style="list-style-type: none"> <li>Do you trust traditional media outlets) (e.g newspaper)? Y/N</li> <li>Do you trust information shared via social media? Y/N</li> </ul>
	<ul style="list-style-type: none"> <li>Impact on polarisation</li> </ul>	<ul style="list-style-type: none"> <li>Perception to be included in a eco-chambre</li> </ul>	<ul style="list-style-type: none"> <li>Do you think are embedded in a echo-chamber? Y/N</li> <li>Consider adding another question to generalize the phenomenon: do you think people tend to be involved in discussion mostly with people who share their views?</li> </ul>
Area of impact	Sub-dimension	Indicator	Variable
Economic impact	<ul style="list-style-type: none"> <li>Impact on news consumption patters</li> </ul>	<ul style="list-style-type: none"> <li>Influence news consumption patterns</li> </ul>	<ul style="list-style-type: none"> <li>How do you get access to information? (newspaper, online journal, open journals, social media)</li> </ul>
	<ul style="list-style-type: none"> <li>Impact on users subscription to online journals</li> </ul>	<ul style="list-style-type: none"> <li>Subscription to online journals</li> </ul>	<ul style="list-style-type: none"> <li>Do you have an active subscription to online journals? Y/N</li> <li>Are willing to pay a fee to access online information?</li> </ul>
	<ul style="list-style-type: none"> <li>Impact on online advertising</li> </ul>	<ul style="list-style-type: none"> <li>Awareness about online advertising</li> </ul>	<ul style="list-style-type: none"> <li>Are you used to check the transparency of AD included in the article?</li> <li>Are you aware about the potential risks of (ads with malware, fraudulent and adult content?</li> </ul>
	<ul style="list-style-type: none"> <li>Impact on transparency of media ownership</li> </ul>	<ul style="list-style-type: none"> <li>Awareness on media ownership</li> </ul>	<ul style="list-style-type: none"> <li>Do you check the ownership of media outlet? Y/N</li> </ul>

			<ul style="list-style-type: none"><li>• Which of the following information help you to establish trust on media outlet (editorial board, code of conduct, contacts and details, etc.)</li></ul>
--	--	--	---

*Table 1. SOMA impact areas with indicators and variable*

## 2.2 The process

According to this framework, the methodology will be based on three main steps (Figure 3). In order to guarantee an interdisciplinary collaboration in the setting and testing of the methodology, researchers coming from different backgrounds and fact-checkers will be engaged in the process.

Once T6 will define the methodology, Pagella Politica will be asked to support T6 in selecting false news outlets on the following topics: health and migration. The criteria will be i)-a false news in the radar of PP ii) a false news that, according to PP knowledge, had a certain relevance on the internet and, possibly, on traditional media.

Then, LUISS will conduct quantitative analysis on the selected news to understand their impact in online social networks using algorithms and automated tools.

From a computational perspective, there are a number of features that will be considered in assessing (dis)information cascades. As a matter of fact, Soma's toolbox, as developed by LUISS Data Lab, is structured upon an integrated framework, that simultaneously takes into account (i) message-based features (ii) topic-based features (iii) user-based features and (iv) propagation-based features (cfr. Castillo, 2013).

In particular, the work will start with an initial phase during which data will be automatically collected from selected social media. It has to be stressed, however, that data collection is limited by the available current data exchange solutions (as highlighted in D2.2 by Møller and Bechmann, 2019). Møller and Bechmann (2019) use the expression "data exchange solutions" to include different methods of data resources and access offered by social media platform (eg. freely available databases, data access via tools and data access via APIs.). Data exchange also considers restrictions on data access implemented by social media platforms, that could undermine both the scope and the degree of completeness and replicability of research with social media data.

For the purpose of this task, LUISS team envisages to make use of data collected from the following three sources:

- Facebook data provided by Social Science One, and namely CrowdTangle API, considering both the potentialities and the limitations of the initiative (cfr again Møller and Bechmann, 2019).
- Twitter APIs, used by a considerable volume of academic literature because of their straightforward availability and their accessibility to the general public (Gaumont et al. 2018). In the case of Twitter, in particular, it has to be taken into account the limited representativeness of its users and the fact that current Twitter APIs, being designed for content-based queries, do not allow direct investigation on patterns of interactions among users.
- Reddit posts, thanks to the complete dumps of the platform often made available by its users, of relying on the free API, available for research purposes.

Due to different rules and privacy policies enforced by these three platforms, we expect to obtain very variable data in terms of both content type and source. In particular, only Twitter users profiles are public so the chance of reconstructing accurate socio-demographic indicators will be limited.

Data collection will start on a fixed date and data will be collected for an observation and monitoring time-frame, in agreement with Pagella Politica and T6. The Subject Finder module of the SOMA disinformation toolbox will filter the dataset and return information about the prevalence of selected themes or news pieces, using keyword-based queries to generate a restricted dataset and plotting both the temporal and spatial distribution of the query-matching records. In a second phase, the Classifier module, based on a semi-automatic “self-training” process, will provide a deeper understanding of the relation between polarization and disinformation in our dataset.

Network-oriented analysis, in parallel, will be fundamental to understand the prevalence of the selected disinformation stories on social media. The methodology will adopt a tool for extracting a graph representation from Twitter data, that can be oriented either at hashtags/keywords or at users/accounts. In the case of hashtags, the tool outputs a network of words, connected based on their pattern of co-occurrence in the corpus. Vice versa, when dealing with users (accounts), the tool extracts an interaction graph between social media accounts, considering only the type(s) of interactions specified by the consumer (e.g., retweets and/or replies and/or mentions).

In this way, the community around disinformation will be reconstructed, thanks to:

- a community-detection tool that returns the community structure of an input graph (possibly, the output graph of the previous tool);
- instruments for identifying the key-players of an input graph, based on well-known centrality metrics (degree, PageRank, Betweenness, Closeness).

The quantitative approach will provide evidence about the patterns of production/propagation of relevant disinformation pieces:

1. by allowing to track relevant news stories and reconstruct their prevalence over time and space;
2. by detecting central debating communities and capture their distinctive narrative in our dataset;
3. by identifying the role of authoritative accounts (“influencers”) and bots in driving the production and sharing of news stories both globally and in specific “disinformation networks”.

Once the quantitative analysis of the impact of disinformation on the social network will be done, the qualitative data gathering will start. T6 will conduct focus groups, questionnaires or interviews to get the replies to the indicators and variables defined in the methodology to understand the direct impact on people using qualitative methods. The methods will be selected to allow the widest participation possible.

With reference to the SEQUOIA methodology aimed at providing information on results and estimates of impact on a specific activity or project, evaluation methods have to be chosen in relation to the scope of the study. Literature in the field identifies a broad range of techniques to be deployed including:

- surveys of intended beneficiaries;
- control and comparison groups;
- participatory methods including workshops and focus groups;
- case studies.



*Figure 3. SOMA process for impact assessment*

At the moment of writing T6 envisages to organize interviews or questionnaires to reply to the indicators and variables established in the methodology to assess the impact of disinformation. The aim is to get 50 stakeholders engaged for each topic in order to have 100 stakeholders in total.

To properly compose the sample of users to participate to the assessment, and to compare the data collected online with what observed out of the social media, we will have two options. The first one is the following. After having gained evidence that a user shared a piece of disinformation, we will get in contact with that user through a SOMA public profile. We will inform the user that we are part of a EU project studying the prevalence and impact of disinformation in social media and that we gained evidence that him/her shared an allegedly false piece of information. We will then ask the user to help us carrying on a qualitative analysis by filling an anonymous survey. This approach will allow us to obtain a well-defined, albeit possibly small, sample of users for which we have an exact matching of social-media-derived data and direct user-provided data.

Should this be impossible, or should the number of participants be too low, we will build a sample of users trying to respect as much as possible the characteristics of the profile emerged from the quantitative analyses. To this end, we will compose a sample of users using standard techniques for guaranteeing completeness and representativeness of the sample, keeping into consideration all relevant socio-demographic indicators. Concurrently, we will extract from our social media data subsets of data/users that are, in some sense, equally representative. This may possibly mean using different sets of data for different research questions, based on what aspects of a user profile can be inferred with reasonable accuracy. In any case, we will be very careful in specifying what outcomes of the two analyses are directly comparable.

In the latter case, we can already foresee some shortcomings, but we will take care to include in the analysis a paragraph containing some corrections for researchers that would like to test the SOMA methodology in the future overcoming the issue of the sample. In addition, such shortcoming would be of course considered directly by SOMA researchers in future work of research to improve and better define the framework.

Results from the validation of the methodology will be used to provide information on impact assessment to the European Commission in the form of policy recommendations and will be also used to assess the impact of the Observatory, as described in Chapter 3.

## **2.3 The timeline**

As said, the process for impact assessment is mainly based on three main steps: creation of the methodology, the test of the methodology and then the creation of a whitepaper containing main results and recommendations for policy makers.

According to the timing reported in the GA, the first phase started in M1 (October 2018) and can be considered closed with the submission of the current deliverable D5.1.

The second phase started in M16 (February 2020) and will last until M25 (November 2020). This stage will be dedicated to the data collection and reporting, as described in the previous paragraphs.

Finally, from M25 to M30 (April 2021) the whitepaper will be elaborated and submitted.

### **3. Assessing the impacts of the SOMA Observatory**

In addition to the test of the methodology developed by SOMA to assess the impact of disinformation, as reported in the DoA, WP5 will also have an additional and relevant objective. This will be the analysis of the impact of the SOMA Observatory during the project lifetime.

As reported by the International Association for Impact Assessment (IAIA) “Impact Assessment simply defined is the process of identifying the future consequences of a current or proposed action. The “impact” is the difference between what would happen with the action and what would happen without it”.

The EC INFOREGIO Unit (European Commission, 2012: 119) defines impact as “a consequence affecting direct beneficiaries following the end of their participation in an intervention or after the completion of public facilities, or else an indirect consequence affecting other beneficiaries who may be winners or losers. Certain impacts (specific impacts) can be observed among direct beneficiaries after a few months and others only in the longer term (e.g. the monitoring of assisted firms). In the field of development support, these longer-term impacts are usually referred to as sustainable results. Some impacts appear indirectly (e.g. turnover generated for the suppliers of assisted firms). Others can be observed at the macro-economic or macro-social level (e.g. improvement of the image of the assisted region); these are global impacts. Evaluation is frequently used to examine one or more intermediate impacts, between specific and global impacts. Impacts may be positive or negative, expected or unexpected”.

Accordingly, the aim of this deliverable (and more specifically of Task 5.2) SOMA is to establish a methodological framework to assess the impacts of the project to reply to the following question: “what is the difference the SOMA project makes?”. This is a crucial question to provide, at the end of funding period, an evidence on how project activities and related investments, have impacted different areas and several stakeholders.

#### **3.1 The methodological framework to assess the impact of the SOMA Observatory**

The impact of SOMA overall will be based on the impact value chain approach (Figure 4) which is the de facto standard for many international bodies, including the European Commission. This means that the methodology analyses project inputs, outputs, outcomes and impacts.

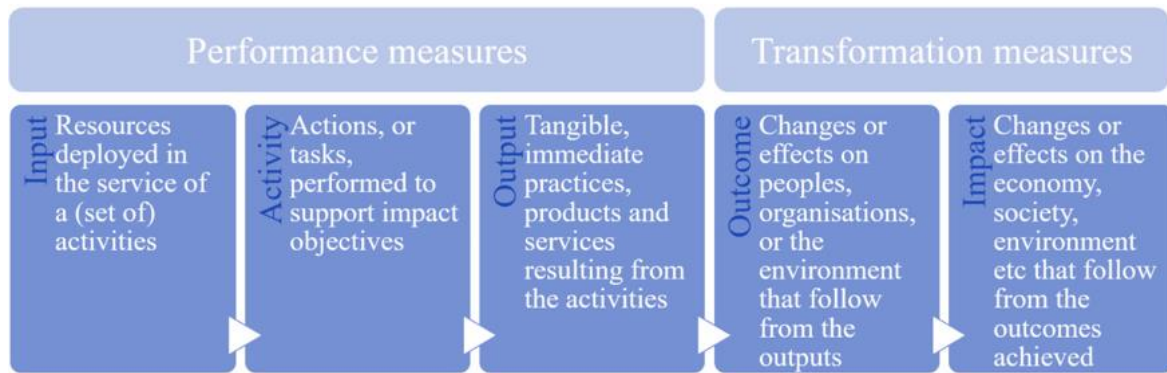


Figure 4. Impact Value Chain Approach. Source: Adaptation from Impact Measurement working group, 2014

As reported in the figure, the assessment of the overall impacts starts at the beginning of the project through a constant monitoring of the project's activities allowing to derive the so called "outputs", namely tangible results that are related to the activities put in place during the project lifetime. The collection and analysis of the outputs is the preliminary step to derive the outcomes of the project activities. This means to evaluate the changes on the stakeholders interested by the project activities. On such bases, the impact of the project overall is defined assessing the measure of the difference made by SOMA not only on the stakeholders engaged directly in the project's activities, but also taking into account a much broader perspectives on the effects of the project on the society.

The SOMA impact assessment framework follows a mixed method approach including both qualitative and quantitative socio-economic methods, reflected in the data gathering process that will include secondary data analysis, web-ethnography, in-depth interviews, questionnaires, focus groups and, where possible, data analytics.

### 3.2 The stakeholder groups and areas of impact

The defined methodology will be based on a defined target of stakeholders, these are:

- Researchers, who have applied and are members of SOMA platform
- Fact-checkers, who have applied and are members of SOMA platform
- General public who has participated to the test of the methodological framework for impact assessment

The reason why the target is enlarged also to stakeholders is because the overall impacts will be analysed according to the main activities putted in place by the project, for this reason we have considered the main stakeholders that will be touched by the activities.

The main activities considered as relevant enough to produce impacts are the following:

- tools and algorithms developed for data analytics;
- SOMA platform;
- networking and research activities to foster collaboration between researchers and fact-checkers;
- media literacy sessions;
- methodology developed for impact assessment.



Then, we have linked the main impact areas of interest of the different stakeholders also stressing regarding on how data will be collected (Table 2). It is important to stress that almost all activities will be assessed through dedicated surveys, or interviews, but the methodology for impact assessment will be assessed directly analysing the results coming from the methodology test performed, as reported in Chapter 2, and also the STI, described in Chapter 4 will be of help to measure certain indicators.

Activity	Areas of impact	Stakeholders	Data collected
Tools and algorithms developed for data analytics	Scientific impact Technological impact	Researchers Fact checkers	Survey
SOMA platform	Scientific impact Technological impact	Researchers Fact checkers	Survey
Networking and research activities to foster collaboration between researchers and fact-checkers	Scientific impact Social impact	Researchers Fact checkers	Survey or interview
Media literacy sessions	Social impact Democratic impact	General public	Survey
Methodology developed for impact assessment	Scientific impact Social impact Political impact	General public	Secondary data from impact assessment exercises (as per Chapter 2)

*Table 2. SOMA activities and related impact areas*

To provide an overview about areas of impact and indicators, Table 3 reports main elements that will be analysed to map the SOMA impacts.

Social impact	Indicators
<ul style="list-style-type: none"> <li>Impact on skills and human capital</li> <li>Impact on social capital (bounding and bridging)</li> <li>Impact on media literacy</li> <li>Impact on skills and human capital</li> </ul>	<ul style="list-style-type: none"> <li>Number of people attended the media literacy session</li> <li>Number of media literacy sessions</li> <li>Degree of satisfaction for the media literacy session</li> <li>Percentage of collaboration among the SOMA members</li> </ul>
Democratic impact	
<ul style="list-style-type: none"> <li>Impact on media freedom and pluralism</li> <li>Impact on quality journalism</li> <li>Impact on digital democracy</li> <li>Impact on civic and political participation</li> <li>Impact on responsiveness and trust in democratic institutions</li> </ul>	<ul style="list-style-type: none"> <li>Percentage of trust in online journals</li> <li>Degree of trust in institutions</li> <li>Degree of participation in the political debate</li> </ul>

Technological impact	
<ul style="list-style-type: none"> <li>• Impact on tools usability and accessibility for content verification</li> <li>• Impact on technology readiness level for content verification</li> <li>• Impact on data access, sharing, re-using</li> <li>• Impact on the quality of data sets exchanged and its diversity</li> </ul>	<ul style="list-style-type: none"> <li>• Degree of satisfaction of the SOMA members</li> <li>• Degree of satisfaction of the use</li> <li>• Number of new algorithms developed</li> </ul>
Political impact	
<ul style="list-style-type: none"> <li>• Impact on data access</li> <li>• Impact on platforms regulations</li> </ul>	<ul style="list-style-type: none"> <li>• Number of communication with the platforms</li> <li>• Data access for SOMA researchers to the data collected by the platforms</li> </ul>
Scientific impact	
<ul style="list-style-type: none"> <li>• Impact on publications</li> <li>• New knowledge and datasets that would be otherwise be impossible to generate</li> <li>• New research questions emerged thanks to the collaboration between fact checkers and researchers</li> <li>• New research methodologies, methods and techniques to study disinformation</li> </ul>	<ul style="list-style-type: none"> <li>• Number of scientific publications on SOMA</li> <li>• Number of analysis produced using SOMA tools</li> <li>• Number of conferences or scientific events attended</li> <li>• Number of stakeholders in the SOMA network</li> <li>• Data access for SOMA researchers to the data collected by the platforms</li> </ul>

*Table 3. SOMA impact areas and related indicators*

## 4. The SOMA Source Transparency Index

As reported by the HLEG (EC, 2018), among the actions to increase transparency of online news sources there is the suggestion to develop Source Transparency Indicators. According to the report (2018, pp. 23) “Source transparency Indicators should be developed based on professional codes of conduct and users’ expectations for what constitutes credible and trustworthy content, adjusted to cater for online journalistic practices. Part of the solution to the spread of disinformation is to ensure that quality and trusted news is substantially available and easily findable in the digital environment”.

Accordingly, some initiatives are already working. Among others, the Trust Project<sup>6</sup> and News Guard<sup>7</sup> are indeed working on the definition and implementation of indicators for credible sources in order to let the reader quickly understand what kind of content is reading.

As stated by the HLEG (pp.23), in fact, “All relevant information should be made available next to the content displayed online on whatever platform in order to empower readers to judge content trustworthiness. This will make users aware as to how to assess different sources of information.”.

Following the indications of the HLEG and as reported in the project description of action, SOMA is also committed to develop a Source Transparency Index validated within the SOMA activities through users engagement to assess that the indicators would be valuable for the community of users.

To build up the index, T6 started from a literature review on source indicators, and in particular from the analysis of the indicators developed by the Trust project and by News Guard. In particular, as reported on the Trust project website, out of an initial set of 37 indicators, a core set of eight Trust Indicators has been selected to be first implemented. They are:

- Best Practices: What are the news outlet’s standards? Who funds it? What is the outlet’s mission? Plus, commitments to ethics, diverse voices, accuracy, making corrections and other standards.
- Author/Reporter Expertise: Who made this? Details about the journalist, including their expertise and other stories they have worked on.
- Type of Work: What is this? Labels to distinguish opinion, analysis and advertiser (or sponsored) content from news reports.
- Citations and References: What’s the source? For investigative or in-depth stories, access to the sources behind the facts and assertions.
- Methods: How was it built? Also, for in-depth stories, information about why reporters chose to pursue a story and how they went about the process.
- Locally Sourced? Was the reporting done on the scene, with deep knowledge about the local situation or community? Let’s you know when the story has local origin or expertise.
- Diverse Voices: What are the newsroom’s efforts and commitments to bringing in diverse perspectives? Readers noticed when certain voices, ethnicities, or political persuasions were missing.

<sup>6</sup> <https://thetrustproject.org/>

<sup>7</sup> <https://www.newsguardtech.com/>

- **Actionable Feedback:** Can we participate? A newsroom's efforts to engage the public's help in setting coverage priorities, contributing to the reporting process, ensuring accuracy and other areas. Readers want to participate and provide feedback that might alter or expand a story.

On the other hand, NewsGuard has created its index based on 9 factors divided in credibility and transparency, these are:

- **Does not repeatedly publish false content:** In the last three years the site has not produced multiple stories that have been found—either by journalists at NewsGuard or elsewhere—to be clearly false, and which have not been quickly and prominently corrected.
- **Gathers and presents information responsibly:** Content on the site is created by reporters, writers, videographers, researchers, or other information providers who generally seek to be accurate and fair in gathering, reporting, and interpreting information, even if they approach their work from a strong point of view. They do this by referencing multiple sources, preferably those that present direct, first-hand information on a subject or event.
- **Regularly corrects or clarifies errors:** The site makes clear how to contact those in charge and has effective practices for publishing clarifications and corrections.
- **Handles the difference between news and opinion responsibly:** Content providers who convey the impression that they report news or a mix of news and opinion distinguish opinion from news reporting, and when reporting news, they do not regularly or egregiously misstate, distort, or cherry pick facts, or egregiously cherry pick stories, to advance opinions. Content providers whose clearly expressed purpose is to advance a particular point of view do not regularly and egregiously misstate or distort facts to make their case.
- **Avoids deceptive headlines:** The site generally does not publish headlines that include false information, significantly sensationalize, or otherwise do not reflect what is actually in the story.
- **Website discloses ownership and financing:** The site discloses its ownership and/or financing, as well as any notable ideological or political positions held by those with a significant financial interest in the site, in a user-friendly manner.
- **Clearly labels advertising:** The site makes clear which content is paid for and which is not.
- **Reveals who's in charge, including any possible conflicts of interest:** Information about those in charge of the content is made accessible on the site, including any possible conflicts of interest.
- **Site provides the names of content creators, along with either contact information or biographical information:** Information about those producing the content is made accessible on the site.

Looking at what already exist in the literature and starting from what has been already developed (e.g. News Guard, The Trust Project), T6 has elaborated its own index, discussed hereafter.

## 4.1 The SOMA process for establishing a co-designed STI

As anticipated, starting from a review of existing solutions, T6 has selected twelve potential indicators to build up the SOMA Transparency Index. The dimensions identified, as the most relevant ones to give a quick overview about crucial aspects related to the transparency and trustworthiness of news outlets, are six: headline, author, sources, contents, wording and advertisement. Each dimension has been classified in two variables. The suggested indicators are summarized in Table 4.

It is important to stress that, according to the review, the selected dimensions are already present both in the indicators selected by the Trust project and by NewsGuard. However, there are two characteristics

which differentiate SOMA from the other two examples. First of all, even if some indicators are replicated, the dimensions are not aggregated in the same way as SOMA does. Secondly, as a fundamental difference, SOMA believes that the index should be referred to the article, and not to the journal, in order to guarantee an improved transparency and a more accurate information. For this reason, we suggest to apply the SOMA STI index to individual news items, according to the following dimensions that have been selected, yet not validated with the users.

Table 4 also reports a tentative description of the variables to build up the indicators<sup>8</sup>. Variables should be created in order to be automatized.

Dimensions	Indicators	Variables
<b>Headline</b>	Clickbait headline	Yes/No
	Correlation between Headline and Text	Yes/No
<b>Author</b>	Biography	Clear definition of the Author/authors and its biography Y/N
	Contacts	Easy access to author's details or journal's contacts Y/N
<b>Sources</b>	Links	Sources cited within the articles Y/N
	Evidences	Scientific evidence referred in the article Y/N
<b>Contents</b>	Speculation	Using wording related to speculation, rumors and false claims "Hearsays" (e.g "si dice che") Y/N
	Opinion	Containing personal opinion or providing different elements to let the people to decide what they believe Y/N
<b>Wording</b>	Emotional wording	Using of sensational "emotional" wording Y/N
	Hate speech	Using wording feeding hate speech Y/N
<b>Advertising</b>	Visible the advertisement	Advertising clearly visible in the page Y/N
	Info on advertisement	Transparent access to more information about the advertising Y/N

*Table 4. SOMA Source Transparency Index draft version*

## 4.2 Validation of the SOMA STI index: the process and main results

As reported in the SOMA project description, before its finalization the SOMA STI has to be validated by a group of users to understand if the index is aligned to what the users need. T6 performed the validation of the above-mentioned variables during the media literacy workshops organized in Rome on September 9th, 2019.

Basically, after a theoretical presentation of what a STI is and how it could work in practice, participants were asked to judge how valuable could be each indicator provided in the table 1 if displayed near the

<sup>8</sup> This is just a proposal and should be discussed with developers in case the STI should be integrated in existing platforms.

news. Participants were asked to assess a value from 1 to 10 where 1 was no important and 10 very important.

Out of 66 participants, we collected 57 replies.

In order to homogenize the replies, where there wasn't provided answer, a 0 is inserted. All the field filled by a "0" are defined as "not applicable".

Then, due to the high fragmentation of the replies and to make the results more significant, the values assigned from 1 to 5 have been grouped as "less relevant" and the values from 6 to 10 as "more relevant".

Values collected and the results are reported in Figure 5.

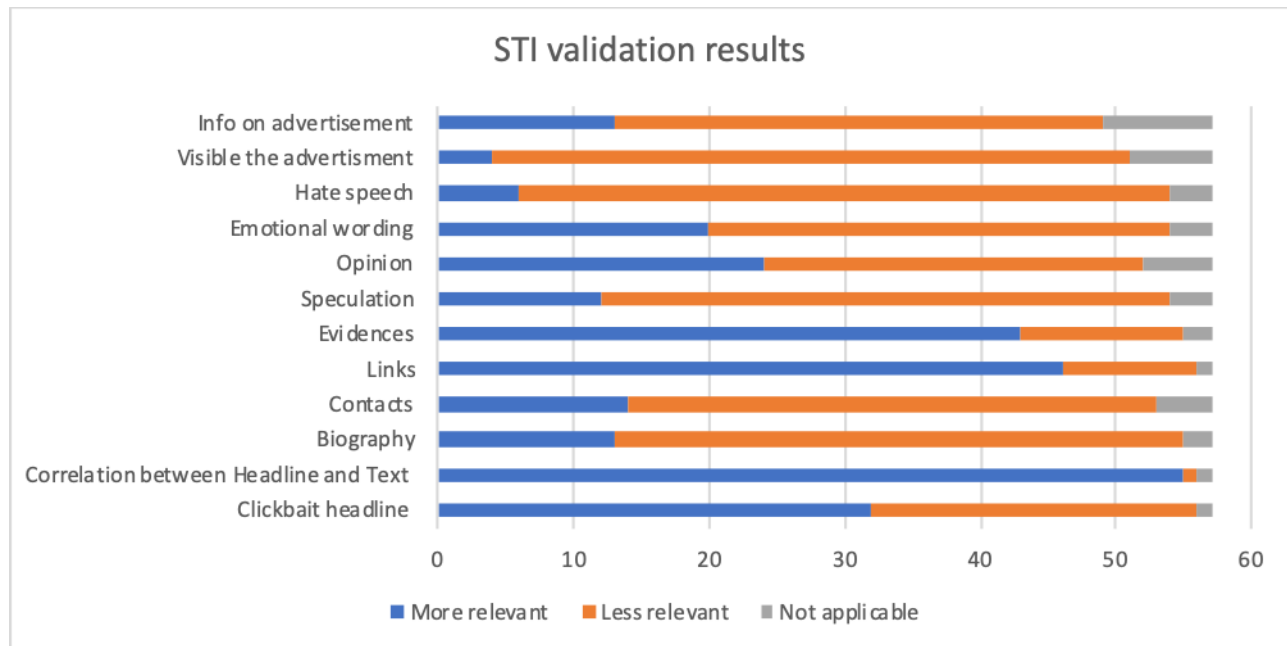


Figure 5. STI validation results

What emerged from the analysis it is that the most relevant dimensions to consider in building the STI are: headline and reference to the sources. Referring to the headline, it is worth to be mentioned that almost all the participants agreed that the indicator of the alignment between the headline and the text it's of crucial importance. However, also an indication about the title is important to help the reader in understanding if he or she is looking at a clickbait headline.

Another crucial dimension to consider is the one related to sources. The majority of the participants, in fact, agreed on the importance of both the creation of an indicator about scientific evidence contained in the text and whether links are inserted to other references.

The indicator related to wording appears less relevant. In particular, most of the participants did not judge as useful the presence of an indicator defining if there is a wording supporting hate speech in the text. Finally, also the information if the advertising is visible did not seem useful.

The indicators that were less understood from the audience are the one on information on advertisement, not surprisingly this category reached the higher value of "not applicable".

According to the findings, comparing Table 4 with Table 5, we removed two indicators from the Table 4 that seem to be not relevant for the participants. Moreover, we reorganized the grid in order to have the

most important dimensions on the top and the less relevant on the bottom. This prioritization should help in pointing out the most relevant information to users.

For this reason, we suggest developing a STI based on the following dimensions and indicators, that should be visible and organised in the following order.

Dimensions	Indicators
<b>Headline</b>	Correlation between Headline and Text
	Clickbait headline
<b>Sources</b>	Links
	Evidences
<b>Author</b>	Biography
	Contacts
<b>Contents</b>	Speculation
	Opinion
<b>Wording</b>	Emotional wording
<b>Advertising</b>	Info on advertisement

*Table 5. SOMA STI validated by the users*

Furthermore, based on the conversation with the participants, we collected the suggestion of an additional indicator that could clearly express the category the news belongs to, in order to let the user be more conscious about what he/she is reading and so putting more attention on some aspects instead of others according to the typology of news outlet.

## 5. Data Access for research purposes: what is needed

As a matter of fact, all the activities described in the deliverable, namely the impact assessment and the evaluation of news outlet, suffer from an important shortcoming: the limited access to data. This is not a new issue but rather a hot topic in the debate with the social media platforms, owners of data, when discussing the need to open data access at least for academic research purposes. The limited access to data, indeed, hinders greatly the capability of researchers to investigate what happens on social media.

Due to the relevancy of the topic, EC is working to improve the conditions of the researchers to access data increasingly bringing up the dialogue with the platforms to achieve an agreement for a framework of cooperation on this issue.

In particular, SOMA has co-organised together with Social Science One and the European Commission a meeting held in Brussels on October 18<sup>th</sup> 2019 with representatives from, platforms, academia, media and governments to discuss about moving forward on data access points for researchers. Main point of reflection was: what kind of data should the platforms let be available for the benefit of the research and for the entire community of researchers?

It is evident that the more data the researchers have, the more in depth research they can produce for the benefit of our societies and knowledge. On the other hand, platforms argue that due to GDPR and EU Data privacy regulation they cannot open their datasets. In particular, during the meeting it was discussed the implications of the General Data Protection Regulation (GDPR) for data sharing practices. Facebook's lawyers argue that GDPR lacks clarity regarding whether and how companies might share data with academics. The lawyers contend that without further clarifications, the company's liability risks are too great. Various policymakers present at the workshop replied that GDPR does not present such a barrier. GDPR is designed to serve the public interest and academic research that serves the public interest is exempt from many GDPR requirements. Further, companies are themselves afforded leeway to weigh public—and even corporate—interest in their data sharing approaches. The key to sharing data with researchers, policymakers argued, is clearly documenting that risks have been considered, that reasonable actions have been taken to mitigate those risks, and that it is in the public interest for independent experts to undertake the research for which the data are shared. Researchers agreed that official guidance from European and member state-level authorities is urgently needed.

Such an *impasse* implies a situation in which European researchers have small or no data access to properly perform social media analysis with important effects on the poor understanding of the digital environment and its consequences on the society. In particular advanced research, for example, is limited by the inability to recreate real world platform conditions and user experiences. Work in collaboration with Twitter, Facebook, Google, and others—and with the informed consent of users—is then needed to implement advanced research directly on the platforms. And critically, the results of any such research must be published publicly, not embargoed by non-disclosure agreements.

In this sense, asking the platforms to release all data *a priori* could be too ambitious limiting the power of the request from European institutions. Following the meeting discussion, the use of Sandboxes or Safe harbour approach has been considered as one of the possible ways forward. The sandbox concept “originates with financial regulatory services. In brief, a regulatory body establishes and oversees a forum where organisations can experiment and test innovative products, services, business models and delivery



mechanisms in a live market with real consumers. It is a supervised “safe haven” which encourages and supports innovation in ways which will comply with legislative and other requirements. Such a supervised space can provide organisations with reduced time-to-market for new products and services, combined with assurance that they have built in appropriate safeguards”<sup>9</sup>. Such an approach can be realized through spaces (whether physical or virtual) within which scholars would directly access and analyse data. In such spaces clear and robust limits on the type and amount of data researchers could access will put in place, as well as the means of analysis researchers could undertake. Researchers then will take the responsibility of researchers to ethically receive and analyse any data., taking all the needed steps to preserve platform users’ privacy and other digital rights. Moreover, while GDPR does not explicitly contemplate any Sandbox mechanism, its risk-based approach is broadly consistent with the approach<sup>10</sup>. Researchers accessing sandbox would be able nevertheless to engage in processing because their participation will have required appropriate mitigations and voluntary regulatory oversight. However, a still open question would be whether companies providing access via a sandbox would be themselves fully immune from liability for sharing data under a safe harbor model.

To support the European Commission in the discussion on the topic, SOMA has launched a short exercise conducted within its network to provide a quick feedback to the European Commission on what data are a priority for research. Results are of course limited, as it was the timing for collecting feedback, but can anyway provide some useful insights for further thoughts.

## 5.1 The exercise

To understand which kind of data are a priority for researchers, SOMA has launched a small round of consultations engaging three highly experienced institutions working on data science.

Replies to the use cases have been provided by SOMA project partners (Aarhus University and LUISS Data Lab) and from Institute collaborating with SOMA, IMT School of advanced studies also working on online disinformation with the research project Toffee<sup>11</sup>.

With the three institutions, we have shared a use case template (Annex 8.2) aimed at collecting information on past or undergoing researches and on which kind of data researchers needed or still need to achieve their research’s objectives. We asked also to explain the process of analysis and its aims, together with the team that should potentially contribute to the research.

The aim of the exercise was to collect, from different research institutions focused on different research topics an overview of what kind of data are needed for their investigations and without it would be impossible to proceed.

<sup>9</sup> Regulatory Sandboxes in Data Protection: Constructive Engagement and Innovative Regulation in Practice, Regulatory Sandbox White Paper 2019, Centre for Information Policy Leadership, UK

<sup>10</sup> See 8

<sup>11</sup> Available at <https://toffee.imtlucca.it/home>

## 5.2 The results

As anticipated, three institutions have provided four different use cases. Results of the data gathering is reported hereafter.

In terms of research topic, the cases address different themes. In particular, one analysis focuses on the spread of disinformation in social media networks. The second one focuses on filter bubbles. The other two are more similar and are related to the link between social media and political debate. The first one focuses on the interference of the spreading of fake content on online social networks in the online political debate and the other one on the role that social media had in the dynamics of the *Yellow Vests* movement and in defining its relationship with politics and society at large.

Even if the researches are dealing with different topics and hypothesis, the four studies agree on the fact that data used for research will be safely stored at the university and not shared with other parties. Results, in aggregated way, will be used to be presented in scientific conferences and published in scientific journal articles. Also, all the cases studied see the participation only of researchers, in most of the case data scientists.

Analysing the data collected it is possible to assume that, even if the focus of the research is different, the typologies of data requested are the same.

In particular, the data requested in all the four use cases are: user data, document data (also defined content data) and historical data.

Looking at the breakdown of the category user data, this means: location, timing, network, groups, shares of specific users or classes of users. Newsfeed data and unique ids of participants (first degree). It seems also crucial to get the information regarding the social networks (i.e. the networks of contacts of every user in the dataset), the data of creation of the accounts. Whenever individual private data are unavailable, aggregated and/or anonymised data may be sufficient. For instance, for membership to private groups it may be enough to have it characterized in terms of time-space distribution, age and sex. However, identifying influencers and leaders requires access to as much non-anonymized information as possible, especially with regard to public figures.

In relation to document data, it is possible to refer to: content, location, timing, comments, replies of specific or classes of posts, tweets, messages, exposure data in the newsfeed (likes, shares, comments, posts, links, pictures, videos etc. Regarding the content itself (i.e. the text of the shared message), it would be important to have access to the text of the messages, the time they were posted, the author (anonymized) and the sources (i.e. the URLs) cited in the messages. If, for privacy reasons, the text cannot be provided, it could be provided the sentiment of the message and the number of selected keywords contained. For instance, sentiment analysis of messages in a group of users may be performed even on anonymous data, provided that co-variates such as timing, location, age and sex are available -- at least, to some level of granularity (e.g., differential privacy may be enforced to some extent).

Finally, also historical data, such as real-time stream data, URL metadata, aggregated data, panel data, statistics, samples are important. However, it seems that this category has been less discussed by the participant to the data gathering. Some of the participant also stressed that to perform a complete comparison, even the results of the bot detection used by the different platform would be needed.

Looking at how documents have been provided it seems that user data, network data, and data about contents are definitely needed. Aggregated data can be a solution even if more information could provide much more knowledge for the research to be performed.

According to the results of this little exercise, SOMA suggests the European Commission, during its talk with the platforms, to ask for the following.

It is evident that to guarantee scientific research a safe access would be needed to different social media data creating rules and procedures to let the researchers work with data respecting specific and strict regulations. However, while such safe space will be built, it is requested to create dedicated APIs for researchers working at the universities.

First, the API should make accessible historical data in a predefined sufficient period of time for aggregated data of users, network and contents. If it is not possible to provide personal information due to GDPR it shouldn't be an issue to provide aggregated information. It should be noted that, in this case, there's a specific scope for Sandbox use by companies that might engage together with a Code of Conduct under Article 40 of the GDPR. Concerns over the appropriate measures to be adopted in order to promote compliance with the rules applying to the processing of sensitive social information could be allayed by the existence of an approved and detailed code. Such a code could outline in a fair and transparent manner the following<sup>12</sup>:

- the relevant safeguards to be applied regarding the information to be provided to data subjects;
- relevant safeguards to be applied in respect of the data collected from third parties;
- communication or dissemination of the data;
- the criteria to be implemented to ensure respect for the principle of data minimisation;
- the specific security measures;
- appropriate retention schedules;
- and the mechanisms to manage the data as a result of the exercise of data subjects' rights (As per Articles 32 and 89 of the GDPR).

Second, provide such data in a controlled platform (sandbox) managed by SOMA will guarantee the correct use of data only for research and fact checking purposes. Finally, establish a working group representing the platforms to start a continuous collaboration with SOMA project to enlarge the data access as well as the guarantee that data will be used for fair and trustable purposes.

SOMA, as an Observatory on Disinformation focused on social media, is committed to support the European Commission in the discussion with the platform and to start a dialogue for the benefit of researchers and fact-checkers but above all for the safe of European values of trust and transparency.

<sup>12</sup> Guidelines 1/2019 on Codes of Conduct and Monitoring Bodies under Regulation 2016/679  
Adopted on 12 February 2019. European Data Protection Board

## 6. Conclusions

The current deliverable discusses in separated chapters transversal topics all aimed to the final objective to report to the European Commission insights on the phenomenon of disinformation at European level providing concrete suggestions and guidance coming from analysis and evidences.

Indeed, the first contribution we are providing to the European Commission is a methodological framework to map the impacts of disinformation. The deliverable, indeed, sets the foundation of the methodological framework that SOMA is developing for this purpose. As mentioned in the document, the development of a such methodology is not an easy task. Several researchers have developed, tested and published different approaches to assess the impact of disinformation. However, at the best of our knowledge, current approaches still suffer of some gaps and need to be improved.

The approach proposed by SOMA is a try to overcome the main shortcoming of approaches that are based only on quantitative methods or on qualitative methods. What we propose is to apply a combined approach that could highlight both the importance of quantitative analysis to collect data from social media and qualitative analysis to collect data from users and their behaviours outside the social networks. The SOMA combined approach intends also to be a call to foster cross-collaboration among different disciplines since the phase of methodology development, and not in a later stage of process development. We strongly support the idea that trans-disciplinary research is needed, and this is also valuable to understand how complex a phenomenon such disinformation is. Accordingly, we have structured a method based on the direct engagement of fact-checkers, data scientists and social scientists that we want to improve and apply as a basic method for the work done by SOMA:

This approach is at the base of the first scientific collaboration that SOMA, and in particular by T6 as WP5 leader, has established with scientific partners. I the elaboration of data extracted from determined use case provides an example to understand what kind of data are extremely urgent to be delivered by the platform to allow researchers to do their job (Chapter 5). Another example of the cross-disciplinary approach is the collaboration established with the Toffee project to investigate the impact of bots on the discussion on 5g (presented in the Annex 8.2).

At the moment of the writing, we have defined the approach and in the next deliverable, in D5.2, we will provide results from the validation of the methodology to prove if it works and reporting about how it should be refined and what is still to be improved. In addition to the methodology validation, the next deliverable will report some evidences of the impact of disinformation on two crucial topics (health and migration) to the European Commission.

The second contribution is related to the assessment of the SOMA Observatory. The deliverable, indeed, also contains the methodology that will be used to map the impact of the Observatory to provide tangible measure of the measure of the investment made by the EC. The results will allow also to inform potential next actions of the EC according to the results obtained during the project lifetime.

SOMA has also described in the deliverable the SOMA Source Transparency index developed with users engaged in a media literacy session. The index could be further developed and integrated in existing tools as a form of the exploitation.

Finally, the deliverable contains preliminary results of conversations that SOMA is establishing with the platforms about data access for scientific purposes. Such first exercise is another relevant piece of work conducted under the WP5 to drive the work of policy on disinformation.

Indeed, almost all topics mentioned so far, and in particular the potential application of the impact assessment methodology, as well as the real impact of the Observatory are very much related to capability of researchers to access data.

For this reason, in addition to theoretical speculation, it is critically important to support the EC in the discussion with the platforms in order to foster a wider and easier access to data for European researchers.

Broadly speaking, the testing of the evaluation framework, the assessment of the SOMA Observatory and the ongoing dialogue with the platforms will allow SOMA to develop a policy document suggesting next steps to the EC in relation to disinformation.

## 7. References

- Agnihotri, R., Dingus, R., Hu, M.Y. and Krush, M.T., (2016). Social media: Influencing customer satisfaction in B2B sales. *Industrial Marketing Management*, 53, pp.172-180.
- Allcott H. and Gentzkow M., (2017). Social media and fake news in 2016 election. *Journal of Economic Perspectives* 31(2) 211-236.
- Allcott, H., Braghieri, L., Eichmeyer, S. and Gentzkow, M. (2019). *The welfare effects of social media* (No. w25514). National Bureau of Economic Research
- Asbjørn Møller L., Bechmann A. (2019) D2.2: Research Data exchange (and transparency) solution with platforms. Available at <https://www.disinfoobservatory.org/resources/>
- Bartl, M., Ernst, H., Füller, J. and Mühlbacher, H. (2004), Community based innovation: A method to utilize the innovative potential of online communities, *Proceedings of the 37th HICSS Conference*, Hawaii.
- Bason, C. (2010) *Leading Public Sector Innovation. Co-creating for a better society*. Bristol: Policy Press.
- Bessi, A., Petroni, F., Del Vicario, M., Zollo, F., Anagnostopoulos, A., Scala, A., Quattrociocchi, W. (2015). Viral Misinformation: The Role of Homophily and Polarization. In *WWW (Companion Volume)* (pp. 355-356).
- Cherepnalkoski D, Mozetič I. (2016) Retweet networks of the European Parliament: evaluation of the community structure. *Applied Network Science*; 1(1):2. <https://doi.org/10.1007/s41109-016-0001-4>
- Cherepnalkoski D, Mozetic I. A retweet network analysis of the European Parliament. In: 2015 11TH International Conference On Signal-Image Technology & Internet-Based Systems (SITIS); 2015. p. 350-357.
- Conover MD, Goncalves B, Ratkiewicz J, Flammini A, Menczer F. (2011) Predicting the political alignment of twitter users. In: *Proceedings<2011 IEEE International Conference on Privacy, Security, Risk and Trust and IEEE International Conference on Social Computing, PASSAT/SocialCom*. p. 192±199
- Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H.E. and Quattrociocchi, W., (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3), pp.554-559.
- Eurobarometer 464 (2018) Final results of the Eurobarometer on fake news and online disinformation. <https://ec.europa.eu/digital-single-market/en/news/final-results-eurobarometer-fake-news-and-online-disinformation>
- Flore, M., Balahur, A., Podavini, A., Verile, M. (2019) *Understanding Citizens' Vulnerabilities to Disinformation and Data-Driven Propaganda*, EUR 29741 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-76-03320-2, doi:10.2760/919835, JRC116009
- Fourney A., Racz M. Z., Ranade G., Mobius M., Horvitz E. (2018). Geographic and Temporal Trends in Fake News Consumption During the 2016 US Presidential Election. *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*, 2071-2074
- Friggeri, A., Adamic, L.A., Eckles, D., and Cheng J. (2014). Rumor cascades. In *Proceedings of the 8th International AAAI Conference on Web and Social Media*.
- Garimella K, De Francisci Morales G, Gionis A, Mathioudakis M. (2016) Quantifying Controversy in Social Media. In: *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining< WSDM'16*. New York, New York, USA: ACM Press. p. 33±42. Available from: <http://dl.acm.org/citation.cfm?doid=2835776.2835792>.

- Gaumont N, Panahi M, Chavalarias D (2018) Reconstruction of the socio-semantic dynamics of political activist Twitter networks: Method and application to the 2017 French presidential election. *PLoS ONE* 13(9): e0201879. <https://doi.org/10.1371/journal.pone.0201879>
- Kumar, S., and Shah, N. (2018), False Information on Web and Social Media: A Survey. 1, 1 (April 2018), <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>
- Kumar, S., West, R., and Leskovec, J. (2016) Disinformation on the web: Impact, characteristics, and detection of wikipedia hoaxes. In *Proceedings of the 25th International Conference on World Wide Web*.
- Lazer, D.M., Baum, M.A., Benkler, Y., Berinsky, A.J., Greenhill, K.M., Menczer, F., Metzger, M.J., Nyhan, B., Pennycook, G., Rothschild, D. and Schudson, M.. (2018). The science of fake news. *Science*, 359(6380), pp.1094-1096.
- Morales AJ, Borondo J, Losada JC, Benito RM. (2014) Measuring political polarization: Twitter shows the two sides of Venezuela. *Chaos*. 2015; 25(3). <https://doi.org/10.1063/1.4913758> PMID: 25833436
- Napoli, P.M. Measuring Media Impact. An overview of the field.
- Passani A., Monacciani F., Van Der Graaf S., Spagnoli F., Bellini F., Debicki M., and Dini P. (2014a), SEQUOIA: A methodology for the socio-economic impact assessment of Software-as-a-Service and Internet of Services research projects, *Research Evaluation*, 2014 23: 133-149
- RES (2018) Social media bots influence stock market performance. Media briefing, March 2018. <http://www.res.org.uk/details/mediabrief/10926361/SOCIAL-MEDIA-BOTS-INFLUENCE-STOCK-MARKET-PERFORMANCE.html>, visited 26 November 2018.
- Reuters Institute. (2018). Measuring the reach of “fake news” and online disinformation in Europe
- Silverman G. (2016) This analysis shows how viral fake election news stories outperformed real news on facebook. *Buzzfeed News*, 16.

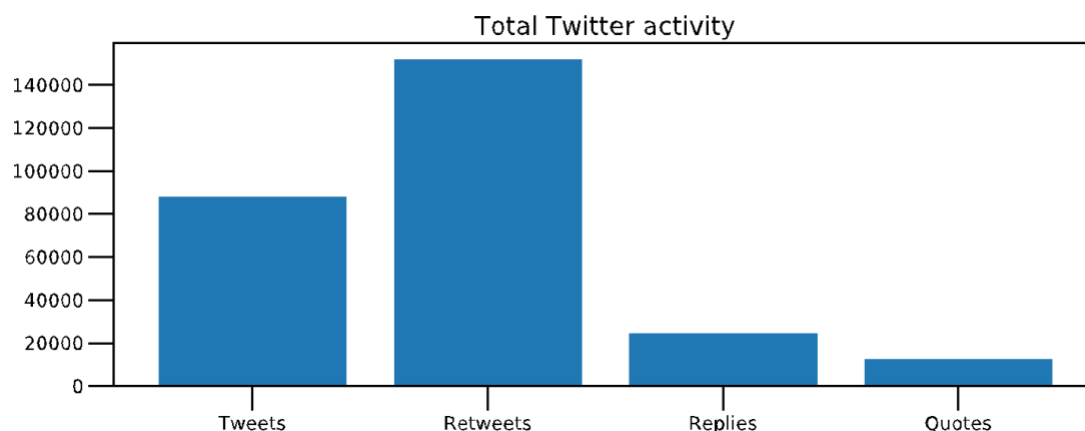
## 8. Annexes

### 8.1 An initial analysis of BOT presence in the debate of 5G on Social Media by SOMA and TOFFEE projects

We present here a preliminary analysis of the traffic on Twitter related to the discussion on 5G, the fifth generation wireless technology for digital cellular networks that has begun wide deployment in 2019. The total activity is composed by 277,688 tweets, collected along 15 days, from November 26<sup>th</sup> to December 10<sup>th</sup> 2019. The data have been collected searching for tweets containing the keyword “5G”. The analysis has been made on different languages to reflect the audience of European Union, with English, French, German, Italian and Spanish data collection.

#### Tweeted sources

Respect to other datasets, the tweeting activity (i.e., the one related neither to replies, nor to retweets nor to quotes) is respectively quite important, see the figure below.



Bot presence so far can be considered marginal (social bots, as detected by Botometer<sup>13</sup>, represent almost the 3.78% of the total number of users). Most of bot activity is done by tweeting posts from a series of websites as '[www.theguardian.com](http://www.theguardian.com)', '[www.youtube.com](http://www.youtube.com)', '[www.politico.eu](http://www.politico.eu)', '[economictimes.indiatimes.com](http://economictimes.indiatimes.com)', '[www.weforum.org](http://www.weforum.org)', '[mass.innovationnights.com](http://mass.innovationnights.com)', '[cio.economictimes.indiatimes.com](http://cio.economictimes.indiatimes.com)', '[www.reuters.com](http://www.reuters.com)', '[www.frandroid.com](http://www.frandroid.com)', '[www.forbes.com](http://www.forbes.com)', '[www.linkedin.com](http://www.linkedin.com)', '[www.techrepublic.com](http://www.techrepublic.com)', '[www.inc.com](http://www.inc.com)', '[www.bloomberg.com](http://www.bloomberg.com)', '[www.ericsson.com](http://www.ericsson.com)'.

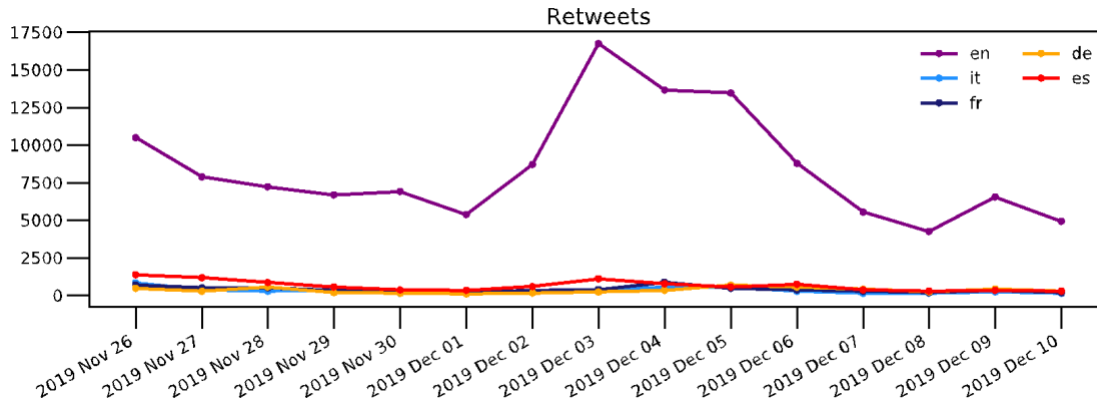
In the following, we will focus on retweeting activity and we will compare the behaviour of bots with respect to all users.

#### Retweeting activity

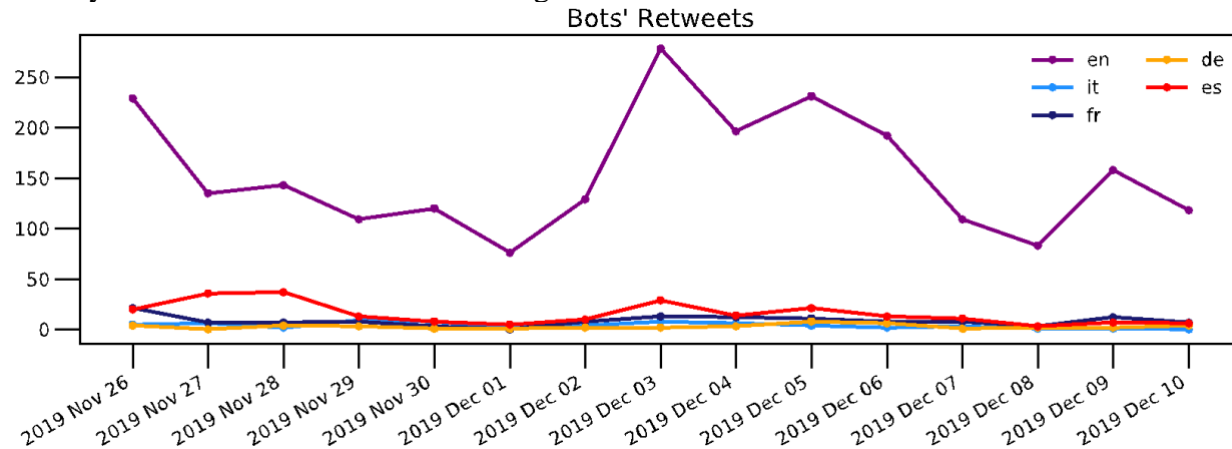
<sup>13</sup> Kai-Cheng Yang, Onur Varol, Clayton A. Davis, Emilio Ferrara, Alessandro Flammini, Filippo Menczer: Arming the public with AI to counter social bots. CoRR abs/1901.00912 (2019)



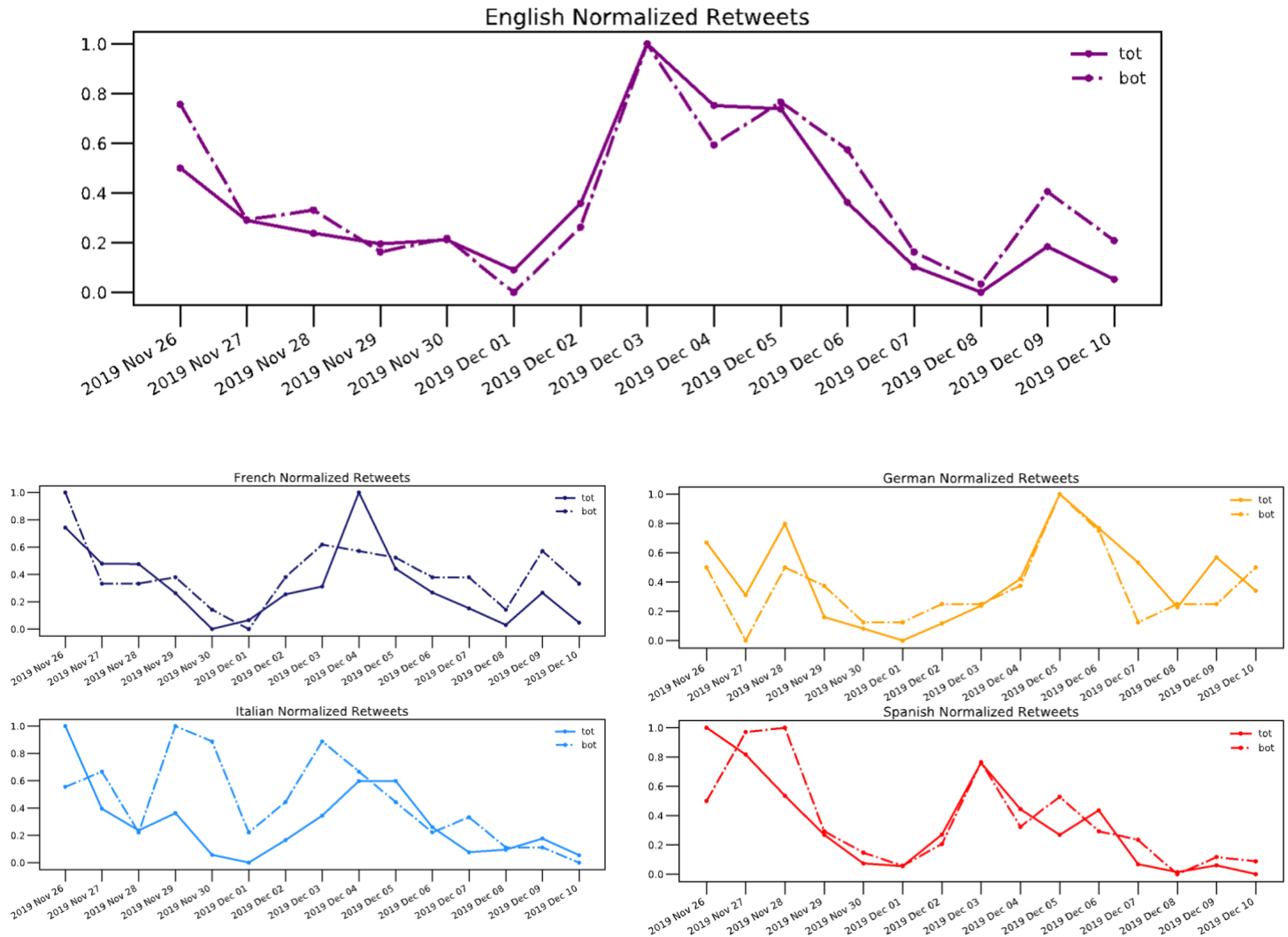
In the figure below, we display the retweeting activity in different languages. English messages are unsurprisingly by far the most numerous in the set.



It is worth to note that retweets from bots also introduce new pieces of information, despite their low number. The total retweeting activity, restricted to social bots, is smaller by two order of magnitude, but actually the trends followed are interesting.



Indeed, the bot retweeting activities in the different languages follow the global ones, but for Italy and France. In the following pictures we used the normalization  $t(x) = \frac{x - \min(x)}{(x) - \min(x)}$ , in order to compare the trends. Using the normalization above, data are stretched in order to have their minimum in 0 and their maximum in 1.



As a general comment it's clear that the tweet volume is still very low, which could mean we are slightly ahead of time for the topic to fly within social media.

In any case this initial recollection will be useful to mark a zero scenario to be used for future analysis.

The most cited sites are mostly reliable (most accessed ones are '[www.ericsson.com](http://www.ericsson.com)', '[www.bloomberg.com](http://www.bloomberg.com)', '[www.inc.com](http://www.inc.com)', '[www.techrepublic.com](http://www.techrepublic.com)', '[www.linkedin.com](http://www.linkedin.com)', '[www.forbes.com](http://www.forbes.com)', '[www.frandroid.com](http://www.frandroid.com)', '[www.reuters.com](http://www.reuters.com)', '[cio.economictimes.indiatimes.com](http://cio.economictimes.indiatimes.com)', '[mass.innovationnights.com](http://mass.innovationnights.com)', '[www.weforum.org](http://www.weforum.org)', '[economictimes.indiatimes.com](http://economictimes.indiatimes.com)', '[www.politico.eu](http://www.politico.eu)', '[www.youtube.com](http://www.youtube.com)', '[www.theguardian.com](http://www.theguardian.com)'), so that identified bots should not be malicious ones.

In the list there are both sites of technological topic (like Ericsson or sites dealing with technology and innovation like techrepublic, frandroid), and economic information sites (forbes, bloomberg, economictimes) and political and news sites (World Economic Forum, Reuters, il Guardian).

Obviously, by selecting the texts in English we recovered data from practically all over the world (the same is true, but to a different extent, also for Spanish: we have many messages that come from all over South America): in particular there are (relatively) many retweets that mention sites of the Indian Times

30/01/2020

group ('cio.economictimes.indiatimes.com', 'economictimes.indiatimes.com'): in this case also said, the bots will have been designed by the magazine to increase visibility of the latest news.

At the moment we are downloading data for a longer period of time (the actual plan is to download 6 months of data). A longer dataset will allow for a more detailed analysis on the role of misinformation in the debate around the 5G technology.

In particular, we plan to detect the social bots present contributing to the discussion, analyse their position in the retweeting network and analyse the origin of the pieces of news shared by them, i.e. if they come from trustworthy sources or not. In particular, we plan to target the debate regarding the effect of the 5G technology on human health, due to the increasing interest on the subject in online social media.

## 8.2 Use case template

<p><b>Title</b> The title that you assign to your project should be brief descriptive, and informative to a nonspecialist audience</p>
<p><b>Abstract</b> Provide a one-paragraph description written for a nonspecialist audience clearly explaining the project's principal focus, questions, and expected results.</p>
<p><b>Datasets</b> Select the datasets for which you are requesting access.</p>
<p><b>Typology of data to get access according to the following categories</b> Report the list of all kind of data you need, or ideally you would like to have, to do your research. Please be more specific as possible</p>
<p>User data (location, timing, network, etc)</p> <hr/> <hr/>
<p>Data from the message (post, tweet, etc)</p> <hr/> <hr/> <hr/>

Other (please, specify)

---

Project description:

A justification and description for the overall project, and the datasets yet available for research access; which available Facebook and other platform data you plan to analyze; which non-available Facebook and other platform data, are needed for your research goal; statistical or other methods to be used and why they are appropriate; expected outcomes, public dissemination plan, potential risks, and their significance to understanding democracy.

---

Short answer questions:

- Practical Importance of Project

- Expected Outcomes and Products

- Ethical and Privacy Consideration

- Data Analysis Experience

<hr/> <hr/> <hr/>
<p><b>List of Participants</b></p> <p>If the proposed project includes any participants other than the Principal Investigator, please select who they are according to the categories reported below.</p> <p>You must clearly identify which participants are requesting data access, as they will be required to agree to the same terms of service that govern data access.</p> <ul style="list-style-type: none"><li><input type="radio"/> Social scientists</li><li><input type="radio"/> Data scientists</li><li><input type="radio"/> Students</li><li><input type="radio"/> Developers</li><li><input type="radio"/> Business consultants</li><li><input type="radio"/> Other</li></ul> <p>Please specify _____</p>