Co-funded by the
Erasmus+ Programme
of the European Union

Device Fingerprinting by News Media Websites in EU: An analysis of data taken in 3 different point of times

MASTERTHESIS
to obtain the Erasmus Mundus Joint Master Degree

in Digital Communication Leadership (DCLead)

of

Faculty of Cultural and Social Sciences
Paris Lodron University of Salzburg

Technical Faculty of IT and Design
Aalborg University in Copenhagen

Submitted by
Fakhira Bano
Matrikel-Nr.: 11728375
fbano18@student.aau.dk, s1047894@stud.sbg.ac.at
Vognporten 14, Albertslund 2620,
Denmark

Internal Supervisor: Reza Tadayoni
External Supervisor: Ursula Maier-Rabler
Tutors: Jannick Sørensen

Department of Communication Studies

Salzburg, 30/09/2019

Acknowledgements

There has been a great and a valuable support from my main thesis supervisor, Ass. Prof. Dr. Reza Tadayoni from Aalborg University Copenhagen from the time of submitting thesis proposal till the final submission. I received his support in the form of regular inputs and advices on each of the milestones of the thesis and content that led this project to this presentable form. Moreover, my secondary supervisor, Ass. Prof. Dr. Ursula Maier-Rabler's valuable and regular inputs and advices are also the main reason this project came into a successful completion.

I would also like to express my utmost gratitude to my tutor and co-supervisor Associate Prof. Jannick Sørensen from Aalborg University for his valuable comments and suggestions to increase the research quality to certain standards of a Masters Degree and Assistant Professor Sokol Kosta from Aalborg University for providing me with the data in a very short period of time. Along with the data provision, Sokol's valuable recommendations for the methods in acquiring results and data analysis helped in making this project successful. As this was the most critical stage and the limited time available for a Master thesis did not allow for a year-long observation from June 2018 till June 2019; however, Sokol and Jannick had previously done data acquisition for a long time period that was very much relevant to my current research purpose.

This research could not be completed without support from my other tutors from AAU and PLUS, friends and family including my husband Nizar Ali whose technical support to process large amount of data without a virtual machine, and suggesting the specific programming approaches to achieve the desired results using python language made it possible to complete the current research. Moreover, I dedicate this research to the European Policy makers, citizens and all the researchers who could get an insight and in-depth knowledge of the methods and the levels to which news media websites capture audience data using device fingerprinting.

Table of Contents

ACKNOWLEDGEMENTS **2**

EXECUTIVE SUMMARY **6**

INTRODUCTION **7**

Background and Motivation 7
Research Focus/Purpose 9
Research Question 10
Intended Audience 11
Scope of the Research 11
Thesis Outline 12

LITERATURE REVIEW **14**

THEORETICAL FRAMEWORK **24**

Audience Commodity/ Digital Labor Theory 24
Selective Exposure Theory 26
Control Theory 26
Liberal theory and Public Interest theory in Journalism 27

RESEARCH DOMAIN **29**

Mass Media 29
Print Media *29*
TV Media *30*
Internet Media *30*
Audience Measurement 31
Cookies and Super-cookies *32*
Device Fingerprinting *33*
Regulation 41
GDPR *41*

RESEARCH METHODOLOGY **44**

Research Strategy 44
Research Approach *44*
Research Method 44
Data Gathering *45*
Available Data *46*
Filtering the Data *47*
Identifying the Device Fingerprint Pattern *47*

RESULTS AND ANALYSIS **49**

Understanding the Results 49
Data Sampling and Coding *49*
Analysis 1: Device fingerprinting methods in each country 52
Selection of limited results *52*
Austrian News Media Websites *53*
Belgian News Media Websites *54*
Danish News Media Websites *56*
Conclusion *58*
Research Sub-Question a *58*
Analysis 2: Levels of capturing unique device fingerprint 58
Data Set from June 2018 ..... *58*
Data Set from February 2019 ..... *59*
Data Set from June 2019 ..... *60*

ANALYSIS 3: COMPARATIVE RESULT OF THE COUNTRIES IN PERCENTAGES 61
Research Sub-Question b *62*
ANALYSIS 4: TRANSITION OF DEVICE FINGERPRINTING LEVELS FROM JUNE 2018, FEBRUARY 2019 AND JUNE 2019 64
Research Sub-Question c: *68*
Transition from June 2018 – June 2019 ..... *69*
Transition from June 2018 – February 2019 ..... *69*
Transition from February – June 2019 ..... *69*

DISCUSSION 71

CONCLUSION 74

REFERENCE LIST 75

APPENDIX A: PYTHON SCRIPTS 80

MYSCRIPT1.PY 80
COMP_COUNTRIES_PIE.PY 82
COMP_COUNTRIES_BAR.PY 85
COLLECTIVE.PY 87

APPENDIX B: RESULTS OF 3 TYPES OF DEVICE FINGERPRINTING DATA FROM EACH COUNTRY IN EU 92

RESULTS FROM JUNE 2018 92
RESULTS FROM FEBRUARY 2019 96
RESULTS FROM JUNE 2019 100

APPENDIX C: COMPARATIVE RESULT OF THE COUNTRIES IN PERCENTAGES 105

APPENDIX D: LIST OF COUNTRIES AND WEBSITES USED FOR CURRENT RESEARCH 106

FIGURE 1: TYPICAL VALUES FROM JAVASCRIPT FILE RUNNING ON A WEBSITE THAT CAPTURES SOME USER'S DATA 34
FIGURE 2: COOKIE PRINTING PROCESS 36
FIGURE 3: EXAMPLE CODE THAT CREATES WEBGL VALUE TO MAKE A CANVAS FINGERPRINT 39
FIGURE 4: PERCENTAGE OF METHODS TO CAPTURE DEVICE FINGERPRINTING BY AUSTRIAN NEWS MEDIA WEBSITES - DATA FROM JUNE 2018 53
FIGURE 5: PERCENTAGE OF METHODS TO CAPTURE DEVICE FINGERPRINTING BY AUSTRIAN NEWS MEDIA WEBSITES - DATA FROM FEBRUARY 2019 53
FIGURE 6: PERCENTAGE OF METHODS TO CAPTURE DEVICE FINGERPRINTING BY AUSTRIAN NEWS WEBSITES – DATA FROM JUNE 2019 54
FIGURE 7: PERCENTAGE OF METHODS TO CAPTURE DEVICE FINGERPRINTING BY BELGIAN NEWS WEBSITES – DATA FROM JUNE 2018 54
FIGURE 8: PERCENTAGE OF METHODS TO CAPTURE DEVICE FINGERPRINTING BY BELGIAN NEWS WEBSITES – DATA FROM FEBRUARY 2019 55
FIGURE 9:PERCENTAGE OF METHODS TO CAPTURE DEVICE FINGERPRINTING BY BELGIAN NEWS WEBSITES – DATA FROM JUNE 2019 55
FIGURE 10: PERCENTAGE OF METHODS TO CAPTURE DEVICE FINGERPRINTING BY DANISH NEWS WEBSITES – DATA FROM JUNE 2018 56
FIGURE 11: PERCENTAGE OF METHODS TO CAPTURE DEVICE FINGERPRINTING BY DANISH NEWS WEBSITES – DATA FROM FEBRUARY 2019 57
FIGURE 12: PERCENTAGE OF METHODS TO CAPTURE DEVICE FINGERPRINTING BY DANISH NEWS WEBSITES – DATA FROM JUNE 2019 57
FIGURE 13: PERCENTAGE LEVELS OF CAPTURING UNIQUE DEVICE FINGERPRINTING BY 28 EUROPEAN COUNTRIES IN EU OBSERVED IN JUNE 2018 59
FIGUR 14: PERCENTAGE LEVELS OF CAPTURING UNIQUE DEVICE FINGERPRINTING BY 28 EUROPEAN COUNTRIES IN EU OBSERVED IN FEBRUARY 2019 60
FIGURE 15: PERCENTAGE LEVELS OF CAPTURING UNIQUE DEVICE FINGERPRINTING BY 28 EUROPEAN COUNTRIES IN EU OBSERVED IN JUNE 2019 61
FIGURE 16: TRANSITION OF PERCENTAGES IN CAPTURING UNIQUE DEVICE FINGERPRINTING FROM JUNE 2018 - FEBRUARY 2019 - JUNE 2019 FOR THE FIRST SET OF SEVEN COUNTRIES IN EU 65
FIGURE 17: TRANSITION OF PERCENTAGES IN CAPTURING UNIQUE DEVICE FINGERPRINTING FROM JUNE 2018 - FEBRUARY 2019 - JUNE 2019 FOR THE SECOND SET OF SEVEN COUNTRIES IN EU 66
FIGURE 18: TRANSITION OF PERCENTAGES IN CAPTURING UNIQUE DEVICE FINGERPRINTING FROM JUNE 2018 - FEBRUARY 2019 - JUNE 2019 FOR THE THIRD SET OF SEVEN COUNTRIES IN EU 67
FIGURE 19: TRANSITION OF PERCENTAGES IN CAPTURING UNIQUE DEVICE FINGERPRINTING FROM JUNE 2018 - FEBRUARY 2019 - JUNE 2019 FOR THE FOURTH AND LAST SET OF SEVEN COUNTRIES IN EU 68

Executive Summary

The current study is inspired from the audience's ability through different devices used to generate the data for the news media companies while browsing their websites. The research intends to answer the main research question 'How do news media companies in EU captured audience data after the implementation of new GDPR' through three sub-questions including 'what are the techniques of audience data gathering through device fingerprinting', 'which countries in EU have an increased percentage of capturing data from audiences using device fingerprinting', and 'what are the transitions in capturing device fingerprints over different time periods after the implementation of GDPR'.

The research then followed a literature review which inspired to apply theory of audience commodity and digital labor theory in the research process. The research was followed with a quantitative research using big-data received from scholars Sørensen & Kosta (2019) who observed and captured data from several websites in three point of times including June 2018, February 2019 and June 2019. This helped this research to refine and retrieve the relevant data form twenty-eight countries and get the results for identifying what kind of data is retrieved through device fingerprinting, the countries with a percentage level of capturing device fingerprints, and a transition over the three time periods in capturing the device fingerprints.

The results showed that after GDPR, news media companies still use device fingerprinting to capture the audience data; however, the percentage levels of identifying unique device fingerprints has been decreased from June 2018 till June 2019 by news media companies in 20 out of 28 European countries. Despite of this drop in capturing device fingerprinting data, it cannot be confirmed whether it is due to GDPR or there are other reasons as the intermediary transitions are still not convincing to agree that GDPR is influencing the news media companies to capture least data from the audience.

Introduction

Communication has been a very basic, important and natural activity in any creature's life for centuries and it is evolving in various ways. For human beings, it was initially words combined with actions which would give and understanding to another human being. A word or action alone would have never made any sense of understanding in communication, therefore, speech has always been accompanied by an action to make sense of the message being disseminated. Languages started to form, in various parts of the world and the way of communication also evolved over centuries. There have been particular modes of communication for certain purposes including sending confidential letters between kingdoms, publishing books, painting on rocks, leaves, etc. However, with the developments in science and technology, we have been communicating with a vast audience in much reduced time and resources. Television and internet have brought a very fast and reliable way of communication among individuals and companies around the world. Initially, television industry used antennas and cables for broadcasting the information which has now evolved in such a way that it is combined with the sophisticated technologies such as internet and the TVs' ability to operate with camera, speakers and microphone that makes the virtual communication and interactivity with the world possible.

Background and Motivation

Television industry has been evolving in different areas since the rapid development in digital technologies. From a traditional broadcasting to smart TV and now websites of these TV channels have become very popular. The news, sports and entertainment TV channels have websites to reach to an increased number of audiences. But the relevance of media content to the intended audience has been a priority task for TV media companies both on the broadcast TV and through their content on their websites. A further step to enhance the quality and relevance of media content and advertisements, based on the audience's data was required. This audience-related data is termed as Audience Measurement while more specific terminologies used in audience measurement is Rating Point which is to measure the viewership of specific TV programs. Coffey (2001) explains the purposes of audience measurement as:

> "Today, Internet audience measurement is used for three main purposes. The first may be called self-promotion. It is important for organizations to be able to make claims about the size and growth of their audiences or technologies. […] The second purpose, which was the driver behind the author's efforts in launching Media Metrix, is to support advertising planning, buying, selling and posting. […] This is the same role that television ratings, radio ratings, and magazine audience estimates play for their respective media. […] The third application of Internet audience measurement data is in strategic planning. […] Knowing the patterns of consumer behavior, how consumers interact with a particular site or group of sites, can help site managers make decisions that improve the traffic flow and objective of the site tremendously." (Coffey, 2001, p. 11)

The audience measurement is done for several purposes including the organizational reports on its growth in terms of audience reach, the advertisements and also for improving the content based on the data received through different techniques in audience measurement. The important purpose which is the basis for the current research is the third reason mentioned here – the improvement and customization of media content based on audience data. For the other two reasons, the media company or any other organization is the directly or indirectly targeted but the third purpose clearly has the audience as target to provide some content based on their data.

In making the content more relevant to the audience, different strategies had been used which include audience measurement. Green (2017) has described Audience measurement as it evolved over time in 5 versions. The fifth and current status of audience measurement in media industry is explained as:

> "Audience Measurement 5.0 is all about […] a Total Understanding of audiences across media and across platforms. There are at least five core principles underlying the evolution of audience measurement into its fifth age. […]
>
> The first is that it will be platform-neutral. Audiences will be measured from wherever they consume media content, whenever they are exposed to it. […]
>
> A second requirement for Audience Measurement 5.0 is that it should be respondent-friendly. […]

> The third component of Audience Measurement 5.0. is Big Data. For television, we can access details of the video content households tune into second-by-second through their set-top boxes (which they need in order to receive satellite or cable signals). Using router meters and other methods, we can also look at all kinds of internet usage, including access to streamed video or audio, as well as requests for text or imagery online. […]
>
> The fourth feature of Audience Measurement 5.0 is that it will be hybrid. In other words, it will increasingly comprise a mix of information from external sources (much of it Big Data) alongside sample-based information which turns data on devices into data about people using the devices. […]
>
> The final major feature of Audience Measurement 5.0 – Data Science, in short, is the glue that joins together the four other key requirements of Audience Measurement 5.0 - allowing us to deliver cross-platform insights, with reduced burden on respondents, potentially drawing on Big Data from multiple sources". (Green, 2017, pp. 6–9)

This new era of audience measurement is more sophisticated and there have been different strategies to achieve the best practice in audience measurement by different companies.

Research Focus/Purpose

Modern technologies have been introduced to let companies put less efforts in doing so, therefore, the current research will focus on a latest and less understood technology for audience measurement which could be used either for enhanced and targeted advertising or customized content production by the television media companies. However, **the current research will be dedicated to identifying the latest technology used for audience measurement in web-based news media content in EU**.

The main purpose of this research would be to seek what data is captured when an audience watches a content on news media websites. Because this subject area is of most importance as using the audience data has several effects and outcomes in different ways. One of them has been highlighted by Napoli (2011) as:

> "Questions related to the utilization of audience research have mapped onto this persistent tension in such a way that journalists and editors have often expressed concern that reliance upon audience research in the production of news inevitably

> undermines established news values, as well as the subjective news judgments of journalists." (Napoli, 2011, p. 43)

Several approaches and technological solutions are also in use to gather the data about the audience which include the use of cookies and browsers' capabilities to access device and viewer's information and much more. Cookie synchronization - the practice of third-party domains sharing pseudonymous user IDs typically stored in cookies - provides the potential for more effective tracking, especially when coupled with technologies such as evercookies (Acar et al., 2014). However, there are a lot of limitations in cookies-based audience tracking such as limited data capture and new ways of disabling cookies from audience side disables the trackers' ability to get audience information. "What is becoming increasingly clear at this point is that within the new media environment, there is a wider array of analytical tools for media organizations to employ in their efforts to understand audiences" (Napoli, 2011, p. 42). There have been new approaches introduced over time such as the smartphone's ability to listen to the audience, video camera with facial recognition, and fingerprints are used to identify user and the different devices used by a specific user.

Fingerprinting is defined by Desmond et. al. as "fingerprinting is a process by which a machine, driver or the software the machine is running can be uniquely identified due to its externally observable characteristics" (Desmond, Yuan, Pheng, & Lee, 2008). All these methods provide the trackers to identify the demographics and, in some cases, retargeting the audience by identifying behavioral preferences of web content by a single audience. This way, companies provide different audience with specialized web-content and advertisements. These are, in one way or the other, taking the audience's information and in some cases disseminating it to third parties. Recently, in 2018, new GDPR has been enforced by the European Union that provides the public to have full control over giving consent to share their personal data to the companies and how companies are obliged to process that data. Moreover, if any company captures and stores the user data, they need to be informed about the data capture in a plain language.

Research Question

The research questions to be answered in the current research are:

How do news media companies in EU capture audience data through their websites?

a) **What are the techniques used in device fingerprinting to gather audience data through the websites of news media companies?**
b) **To which extent the countries in EU capture data from audiences using device fingerprinting technique?**
c) **What are the transitions in capturing device fingerprints over three different time periods?**

Intended Audience

The study primarily intends to assist the stakeholders in the policy making and regulation of data protection in EU. However, the stakeholders are divided into two groups: the policy makers and general public being affected by these policies. This study provides a technological understanding which is used by policy makers to define laws in the future so that the gap between what is understood outside of a technological domain and what technology makes possible should be mitigated. The general public in the EU with some basic knowledge about the research domain could also learn about the backend processes occurring due to their activities and interactions with online news media. Moreover, this study also leaves future researchers to think further and work in the same domain to take such studies to the next levels.

Scope of the Research

1. The current research will only focus on identifying the device fingerprinting data being captured by the European news media companies through their websites. It will be done only for identifying the type of the data captured for the audience measurement through device fingerprinting. There is no such intention of current research to assess the purpose of data capture by any news media website. However, relating the results with the existing theories will help the researcher to draw conclusions.
2. Audience measurement provides the basis of the current research; however, it has evolved in so many ways that several companies use different methods for audience measurement. As it is too broad and unrealistic to cover all the approaches and methods of audience measurement, cookies and super cookies will not be part of the research

analysis; however, an understanding of these terms will be briefly described in the chapter 'Research Domain'.

3. Moreover, the new video-based audience measurement will not be discussed as this is a method used with the TV sets at homes. Therefore, the current research will completely focus on Device Fingerprinting for the data acquisition and analysis.
4. The behavioral characteristics of audiences will not be included as device fingerprinting only provides information of the devices being used to access the media content
5. The data for the current research was received from a previous research done by the scholars Sørensen & Kosta (2019) using a virtual machine and not a human being visiting different websites. Therefore, the results of the scripts that could react based on a human being's interactions with the websites might be different than in the study done by Sørensen & Kosta (2019) who used a virtual machine to crawl through thousands of websites.
6. Moreover, the results and analysis of the current research depends on the genuineness of the data collected by Sørensen & Kosta (2019) and its validity could not be verified during current research due to large amount of data.

Thesis Outline

The next chapter of "**Literature Review**" is presented as an initial research method to understand different scholars' point of views in the current research topic. Moreover, this chapter helped in understanding several critics, problems and opportunities for further research that gave a basis to current research topic.

After the literature review, a specific interest was developed in the current research domain so the theories to base the research were needed to explained. The next chapter "**Theoretical Framework**" comprised of some theories in the specific research domain that led the current research towards data collection and to prove or disapprove the theories in the current research topic of device fingerprinting for audience measurement. During the theoretical framework and literature review, the researcher found out several new concepts about the topic and the intentions to proceed with data collection and analysis.

Therefore, a thorough understanding of research domain and specifically the focus of the research, that is, device fingerprinting needed to be understood and briefly described. The chapter "**Research Domain**" serves as a knowledge base for understanding the current research domain.

Along with a literature review, a theoretical framework and a research domain knowledge, a proper research methodology is also needed to be followed. This is explained in the chapter "**Research Methodology**". It will also explain the particular method used for data acquisition and analysis of the data and results.

The '**Results and Analysis**' will be the next chapter with more clear idea of how the methods of device fingerprinting are used to gather audience data.

The analysis will not include any details on why the data is being captured by the specified media companies. Therefore, further discussions will be provided in the chapter "**Discussions".** At the end, a conclusion will be drawn out from the study of previous works until the results and analysis of the current research which will be part of the chapter "**Conclusion**".

Keywords: News Media websites, audience measurement, device fingerprinting, GDPR and audience tracking, EU websites

Literature Review

Since several new types of devices have gained a great attention in the twenty first century, media has become accessible throughout all devices which a person uses now. It gradually became common that people use different devices that they find feasible in terms of the devices' capabilities, functionalities and it also depends upon person's personal preferences. "Audiovisual content can still be enjoyed live via television sets but now we can also watch it on the screens of other devices such as computers, tablets or mobile phones" (Portilla, 2015). That has greatly influenced how the media companies provide the audience with the content. "The media market is flooded with traditional media outlets, newer media outlets such as cable and satellite television, and countless web sites on the Internet, to name a few" (Kim, 2016).

Normally, media companies assess the performance of their content using several techniques, among which program rating is widely used. "A program's rating refers to the percentage of households tuned to the program compared to the number of television homes that could be tuned to the program" (Kinney, 2011). This rating is further analyzed by the media companies as Napoli (2011) refers to rating analysis in his book 'Audience Evolution: New technologies and the transformation of media audiences' as:

> "Ratings analysis may be more usefully defined as the analysis of the data (whatever their orientation) used by media industry stakeholders to assess performance and success in the audience marketplace. Such a definitional approach imbues the field with the flexibility necessary to adapt to various stages of audience evolution." (Napoli, 2011, p. 171).

In the end of twentieth century, the scholars have already forecasted the way media content is put forwards to the audience by mentioning the future possibilities of audience involvement in what will be presented to the audience as media the content. This notion of putting media content to the audience has a greater importance and influence on the whole journalism industry. Negroponte (1996) predicted this as:

> "Being digital will change the nature of mass media from a process of pushing bits at people to one of allowing people (or their computers) to pull at them. This is a radical change, because our entire concept of media is one of successive layers of filtering,

> which reduce information and entertainment to a collection of "top stories" or "best-sellers" to be thrown at different "audiences".” (Negroponte, 1996, p. 84)

The journalists now know that it is not them who can decide on what to provide as a content and their audience has turned into an influencer of the content being provided to them. This was a notion adopted by media companies a century ago by involving audience’s feedback in several different ways. However, this change is intensified in the digital era when several different devices enabled the audience to watch content on their desired device. At the same time, several different opportunities of watching a television content on several different devices has also opened up the possibilities for the interactions among the audiences and when that occurs online, the data flows from both direction through the medium of the website.

> “Many television shows offer corresponding websites or mobile applications. Digital technologies increase the ease with which audience members can engage with both broadcast content and each other, as well as provide a means of tracking, and potentially the ability to quantify, these engagement activities”. (Smith, 2015)

These enable the television websites gather the audience data to process further within the company for the several company’s strategies. One of the most important strategy that companies develop is to growing the economic model and their business so the data in digital age serves as a very important factor in creating business opportunities. As the audiences of media websites use several devices, the turnover from one screen to other also increases an opportunity for the television content providers to provide with more advertisement resulting in an effect – the more the devices a person uses for watching a media content, the more opportunities for the organizations to grow economically. This is explained by Smith (2015) as “advertisements from the primary screen will hold the same audience numbers that the television show does, while the potential grows for advertisements online to actually engage audience members”. (Smith, 2015)

Moreover, the advertisement and marketing started to follow a different approach than earlier while focusing not only on the demographics of the audience or preferences mentioned by the audience but also their behavioral attitudes towards the media content and advertisements. Jenkins (2006) identified these changes as:

> “New models of marketing seek to expand consumer's emotional, social, and intellectual investments with the goal of shaping consumption patterns. In the past, media producers spoke of "impressions." Now, they are exploring the concept of audience "expressions," trying to understand how and why audiences react to the content.” (Jenkins, 2006, p. 63)

These behavioral patterns of the audience are not only used for reshaping the media content but also for the advertisements as this increases the business opportunities for both media companies as well as the advertisers. The media companies and always put the economic factor forwards and to achieve the targeted milestones in their revenues, these companies work together with the advertisers. Behavioral advertising enables both the advertisers and the news publishers to increase their revenues by capturing more and more audience attention by targeting the content that is based on their previously observed behavior and response through several different methods. This

> “online behavioral advertising (OBA) which refers to the collection of data from a particular computer or device regarding Web viewing behaviors for the purpose of using such data to predict user preferences or interests and to deliver advertising to that computer or device, based on what has been inferred from such Web viewing behaviors.” (Boddewyn, 2015, pp. 204–205)

With the development of telecommunication technologies, ability of faster and greater data transfers, the broadcasting system has also evolved to great developments and changes. The main technological influences in the digital period were the development of digital production technology and the Internet, as websites became a staple for television shows and digital production technology allowed additional content to be uploaded to these websites. (Smith, 2015) Along with the internet, “cable television systems have moved far beyond simple delivery of television programming to include high-speed data services, voice telephony, networking, transactional delivery of digital video under the interactive control of customers, and targeted advertising delivery, to name a few” (Large & Farmer, 2009). All these new approaches in the delivery of television programs demand much sophisticated developments in the audience information systems. “Not only must the dynamics of media consumption be changing in ways that undermine established approaches to audiences, but new audience information systems must be capable of capturing alternative approaches to

audiences" (Napoli, 2011, p. 150). These audience information systems also contribute to identifying the behavioral patterns of the audience. Moreover, while explaining the relation between the online behavioral advertising and the audiences's responses, Napoli (2011) explains:

> "Online behavioral advertising involves the delivery of targeted advertising to different members of the audience based on their demonstrated patterns of media consumption or behavioral responses such as information requests and other possible advertisement responses, such as click-throughs or product purchases." (Napoli, 2011, p. 111)

As mentioned earlier that media companies have several options to retrieve audience-related data for several organizational purposes, one of the challenging but very effective approach is video based audience monitoring. This is very efficient in some scenarios where the physical movements, the sophisticated body language, eyes movement and tracking are considered for shaping content and displaying targeted advertisement especially in the wider spaces such as screens on the streets, shopping malls, transport areas etc. This video-based monitoring is explained by Testori (2014) describes the possibilities of video-based monitoring as "count the passers-by in a given area (shopping malls, transportation hubs, stores…), Measure the dwell time, Measure the number of viewers, Split the data by gender and age groups." (Testori, 2014, p. 3). However, this video-based audience monitoring is also available in multi-device web browsing of the tv program content. An audience information system is responsible to capture, process and make decision for content display, which "some research suggests that innovation in content production is inhibited when audience information plays a prominent role in decision-making" (Napoli, 2011, p. 158). However, it is a big question whether this decision-making is something audience is happy about or even whether they are aware about it or not. Whalley (2011) in their paper 'Advertising' explains the way marketers monitor audiences' activities which consumers will not be happy to know about. "Many consumers are not pleased to learn their activities are being monitored when they engage a media outlet" (Whalley, 2011, p. 8).

It has been a long time when all these three approaches - Television viewing, Television recording, and Internet spyware that capture and monitor the audience without them being aware about it became common while Kinney (2011) points out that,

> "This problem (capturing all the activity of the sampled population, short of the subject's willingness to have all of his or her digital devices monitored) is compounded by the number of media companies that are active across a broad array of media platforms. For example, ESPN programs several cable television channels, syndicates radio shows, operates a national radio network, publishes a magazine, produces a national Website, and is moving into the local Website market." (Kinney, 2011, p. 4)

However, some studies describe that these audience data have helped in reshaping the media content in a more audience-focused. Many companies and even people believe that the only way to provide relevant and interesting content by the publishers is to know the audience and therefore the audience data capturing is advocated in a supportive way. Napoli (2011) describes the importance of audience tracking for media companies to reach to the targeted audience which does not end up in providing a content to an irrelevant audience:

> "the ability to gather and analyze more granular data about media audiences allowed for more targeted approaches to identifying desirable audiences for advertisers, and thus provided an important impetus for magazines to reorient themselves in ways that served more narrowly targeted audiences." (Napoli, 2011, p. 28)

Earlier in this chapter it has been made clear that there are several approaches to the audience measurement or audience tracking. These techniques have evolved to a certain level in the digital era that companies do not have to rely on different resources to understand their audiences. Among several new techniques for audience tracking is the cookie-based audience tracking. These are small pieces of information that are saved on a user's or audience's device and that could be retrieved by these websites every time a user browses through webpages to uniquely identify the user. Most of the companies and advertisers today still rely on the cookie-based audience tracking but advertisers and media producers faced several challenges such as audience could easily block them and/or remove them after every session of browsing; hence making it difficult to track back the same audience. But as explained by Ring (2015),

> "super-cookies are like cookies on steroids: they are designed to be permanently stored on a user's computer, and are typically more difficult for users to detect and

> remove from their devices because they cannot be deleted in the same way as regular cookies." (Ring, 2015, p. 5)

With some critics and security-related issues, the super-cookies never became very famous, which could have several reasons. Also, there is very less literature available on super-cookies. However, the most challenging tracking method which an audience is simply unable to detect and avoid is the device fingerprinting. Device fingerprinting is the least understood method of audience tracking where the websites capture different kind of information and combine them in such a way that creates a unique fingerprint for each user across the internet.

> "By using browser fingerprinting to piece together information about your browser and your actions online, trackers can covertly identify users over time, track them across websites, and build an advertising profile of them. The information that browser fingerprinting reveals typically includes a mixture of HTTP headers (which are delivered as a normal part of every web request) and properties that can be learned about the browser using JavaScript code: your time zone, system fonts, screen resolution, which plugins you have installed, and what platform your browser is running on." (Budington, 2018)

The philosophy behind the device fingerprinting is that every device has some system properties, display properties, browser properties, etc. When combined together, the whole set of information is varied from every other device which increases the chances of uniquely identifying each user across multiple devices. This technique helps media companies and advertisers to get the several different type of audience's browsing data including the number of different devices an audience is using to watch a particular content. There is a wide variety of information available in different studies about the consequences and effects of audience being monitored and their data being captured by media companies and other third-party software; however, it has been clarified in some studies that this data and process of audience monitoring has large economic and business effects as Curran et. al. precisely explains:

> "What is being sold here is our profile, our consumption habits and our search history in precisely the way that Garnham argued that the main commodity in the cultural industries is the audience as it is sold, over and over again, to advertisers." (Curran, Fanton, & Freedman, 2016, p. 82)

The point that how the audience are becoming a commodity instead of some kind of consumers in the mass media, Mansell (2012) quotes in her book 'Imagining the Internet: Communication, Innovation, and Governance:

> "Dallas W. Smythe, a political economist who studied the economic features of the communication system, examined the 'audience commodity' at the time when the mass media were predominant. He argues that 'readers and audience members of advertising-supported mass media are a commodity produced and sold to advertisers because they perform a valuable service for the advertisers'." (Mansell, 2012, p. 58)

There is also another reason, the technological developments have enabled the broadcasters and advertisers with a great opportunity of using audience related statistics to keep track of their progress. Today it has become one of the most important focus of the broadcasters to use these statistics.

> The importance of measuring audiences both online and for television can be seen in the ability of broadcasters to use audience numbers, interactions, and demographic information to acquire sponsors or 65 advertisers to be featured online or during televised content. (Smith, 2015, pp. 64–65)

Comparing the traditional advertising where advertisement were based on surveys, market analysis and purchase patterns from the markets and the new ways of advertising where audience has become a commodity, "the difference between the audience commodity of traditional mass media and of the internet is that on the internet the users are also content producers" (Curran et al., 2016, p. 129) where the content produced by the audience is used in a hidden way through computer programs that create audience-specific content and advertisement. There are several companies solely working on audience traffic flow for several different media companies and their programs in real time. Kinney (2011) also pointed out about one of the most popular companies in the US working on TV rating:

> "Nielsen's television ratings data is sold to any person or company interested in television ratings, including sports leagues, television and radio broadcasters, advertising agencies, and brand marketers…. Nielsen uses its peoplemeter technology in 25 U.S. markets to project ratings and demographic profiles." (Kinney, 2011, p. 3).

While doing their work on monitoring audience on several websites, some of these companies also provide new definitions to what has been simply called Audience Measurement. Jenkins

in his book ‘Convergence Culture: Where old and new media collide’ gives an example of company named ‘Initiative media’:

> “Initiative Media, a company that advises many of the Fortune 500 companies about their advertisement placements, advocates an alternative approach to audience measurement they call "expression." Expression charts attentiveness to programming and advertising, time spent with the program, and the degree of viewer loyalty and affinity to the program and its sponsors.” (Jenkins, 2006, pp. 67–68)

Smith (2015) explains the whole process of cookie-based audience monitoring and how it is influencing both the audience, advertisers and the content providers.

> “Blattberg and Deighton argue that audience tracking, already twenty years ago, was a key advantage that computer-mediated systems offer marketing communicators (1991). This tracking, typically found through the use of "cookies," or small encrypted text files that allow web developers to help users navigate their websites, can allow advertisers to see what users clicked on, interacted with, or even where users went after using a website (Allaboutcookies.org, n.d.).” (Smith, 2015, pp. 23–24)

Audience is not playing a role of consumer anymore but they are source of generating value by simply watching and clicking on different media websites. Their activities contribute to the businesses and hence becoming commodity in media industry especially in advertising. Deuze (2012) in his book ‘Media Life’ explains:

> “As users of all kinds of cards, as people under constant surveillance while being in public, and as online shoppers, we provide value-generating labor for the business and corporations that collect, record, mine or sell data about us. In doing so, we extend the work we were already doing for companies simply by watching television or listening to the radio: we were making ourselves available to be sold as audiences to advertisers.” (Deuze, 2012, pp. 113–114)

When it comes to audience awareness about their data being captured and processed for further decision making in shaping the audience-specific content and advertisement, the law and regulation comes into the scenario.

> “fundamental questions they raised about the appropriate regulatory treatment of audience measurement services, particularly whether ratings data should be afforded free speech protection under the First Amendment, and therefore be free from any

> potential government intrusions. The answer to this question has a direct impact on whether policymakers and the courts have the right to influence the operation of audience information systems." (Napoli, 2011, pp. 137–138)

In EU, the new GDPR is already introduced in the year 2018 with several new changes in which it (GDPR) "aims to give users more control over the collection and distribution of their personal information, also on websites" (Sørensen & Kosta, 2019). On the other hand, those who do not follow any advertisement strategy face typical challenges particularly in identifying their audiences' preferences and the content to be created so that the efforts of the media company do not go in vain. Napoli (2011) explained this problem by exemplifying the public broadcasting:

> "A number of accounts of the operation of public service media (e.g., public broadcasting) have highlighted the extent to which professionals within such organizations have resisted—although in most cases, eventually succumbed to—the pressures to conduct and rely upon sophisticated audience research." (Napoli, 2011, p. 44)

With the introduction of new changes in EU's GDPR in 2018, there have been several changes observed in user/audience consent agreements by several companies. A very recent research conducted by Sørensen and Kosta (2019) during an eight months study in 2018 to identify the Third Party (trackers/websites - TPs) on several EU and non-EU public and private websites including news media websites. "[…] these data suggest that the private websites present a slight decrease of the number TPs after the commencement of GDPR, while the public ones are hardly affected." (Sørensen & Kosta, 2019, p. 1596) However, they mentioned that the reason of decrease in TPs after the GDPR might be because "advertisers prefer to reduce the amount of TPs to ease the obstacle of obtaining consent" or it declined due to "a technological change in the advertising industry" (Sørensen & Kosta, 2019, p. 1599). Further research is needed to see what kind of audience specific information is gathered by these media companies.

The literature review has provided much knowledge and insight to the topic of current research which will help in defining the methodology and specifically sampling the research

in further data gathering and analysis. The next chapter will provide an in-depth information about the method to be followed for achieving the data in the current research topic.

Theoretical Framework

Several different theories are related to the current research topic and that provide a basis for the research process and conclusions. However, the research has categorized these theories into a way that provides a multi-angular view to the current research focus. For example, media companies, audiences and regulators need to understand the latest forms of data and how it can be transformed into a useful piece of information, its value and the source where it has originated. Keeping in view the fact that the modern era is an information era, each piece of data that contributes to information has an economic value. The audience's and the data generation from audience's side make them a commodity rather than just a mere consumer. This is explained in the Audience Commodity Theory and Digital Labor Theory.

While researching on theories in audience measurement, some other theories were found very interesting and they might not be critically analyzed based on results of the current research, but for the future reference to the researchers, these theories will definitely provide the basis and to critically analyze the theories in the light of data analysis in other related directions. These other theories that were studied during this research included Selective Exposure theory which is focused on the audiences that have some basic personality traits, attributes and characteristics that shape their decision making which cannot be ignored in audience research. Moreover, how media companies try and practically following strategies to control the behavior of watching the media content is based on Control theory. At the end, two of the most important theories in audience measurement and tracking are the Liberal and Public Interest theories in Journalism. These will be briefly discussed in the current chapter.

Audience Commodity/ Digital Labor Theory

Every individual from a teenage in a developed world has at least one personal device with an ability to provide the media content available through the internet. As long as the individual browses through the internet, especially when the individual browses to view the media content online, the traces of the individual's device data are left for the website owners. Due to the audience being the source of that data, this data is generally identified in the academia as audience data. As today, the audience data is used for reshaping the media content which has an economic value and particularly for the advertisements, the audience has become more

than just the viewers of media content. Every information in the digital age has an economic value and considering the fact that it has been produced from an audience end, "the audience commodity theory has always entailed a contentious proposition that the so-called work of audiences has been effectively subsumed within the capitalist logic of accumulation" (Caraway, 2011, p. 694). Here, the so-called work is not a traditional work that the workers know as a job that they wish to do to get an income. However, this so-called work is a work done by an audience while browsing a website and due to the 'browsing' and 'using' a device to browsing, an audience work by leaving some of the data from their device to the media companies. It is a so-called work because it takes time and the activity produces a good in terms of data that is valuable for a company. And therefore, the audience commodity theory by Dallas Smythe comes into the discussion.

> The origin of the audience commodity stems in part from a legitimate critique of traditional Marxist formulations of what constitutes productive labor under capital. Marx, falling in line with Smith and Ricardo, argued that productive labor is that which produces new value for capital. (Caraway, 2011, p. 694)

For every activity and visits on the websites, the audience leave behind the data for those media websites to be stored. This data can be turned into meaningful form programmatically which becomes an information. When a work is done by an individual, it is a form of labor and in the digital era when the product is a data then digital labor theory needs to be considered.

Fuchs (2015) explained the labor theory as productive labor and defined by Marx in some briefly that the productive labor is the work that produces values and it is the labor that produces capital and surplus-value for the purpose of accumulation (Fuchs, 2015). McQuail in his book 'Reconsidering Value and Labour in the Digital Age' explains further in a clearer way that:

> "In an innovative and sophisticated move, the Canadian Dallas Smythe (1977) gave birth to the theory that audiences actually *work* for advertisers (thus, for their ultimate oppressors) by giving their free time to watch media, which labor is then packaged and sold by the media to advertisers as a new kind of 'commodity'." (McQuail, 1997, pp. 2–3)

Based on the definitions and discussions about audience commodity theory and digital labor theory, the current research will follow a certain direction to retrieve a data that could be analyzed and represented to support or criticize these theories.

Selective Exposure Theory

In further researching in the domain of audience measurement, Selective exposure theory seemed interesting and relevant to the current research domain. Selective exposure is based on the psychological tendency of individuals to make decisions based on their assumptions or beliefs to choose further information. This theory was first presented by Joseph Klapper (1960). He mentioned these three basic concepts of the theory:

- Selective exposure - people keep away from communication of opposite hue.
- Selective Perception - If people are confronting unsympathetic material, they do not perceive it, or make it fit for their existing opinion.
- Selective retention - Furthermore, they just simply forget the unsympathetic material.

When choosing any communication medium, the audience have several options including tv, newspaper, and the websites. According to the theory, an individual has several assumptions or beliefs such as availability of time to watch news on a TV or on website where the same and much more content is available to be watched later on at any time, the belief an individual has regarding the availability of more information on the website than on the TV, the availability of resources to the individual, and many other beliefs an individual has which motivates that person to choose one medium over the other. These multiple choices related to a person's beliefs impacts the generation of data from several devices and captured by the media companies.

Control Theory

In certain times in history, people needed the news to stay tuned to the certain stories, situations etc. but as the news production incorporated the economic factor in it, it has started depending on the audience as the main determinant of economic growth in the news agency. McQuail (1997) explains the need of audience to the media agencies as "media need their audience more than audiences need their media, and there is also reason to view audience research as primarily a tool for the close control and management (call it manipulation) of media audiences." (McQuail, 1997, p. 3) In the efforts of controlling the attention of

audience, the media companies work really hard in understanding their audience through several ways and ultimately control them so that they watch their content and stay tuned to their content. Smith and Russel (2014) explain the control theory which is based on cybernetic theory by Norbert Wiener as:

> "The foundation of control theory is the negative feedback loop, which is based on the discrepancy between the perception of a present and a desired ideal state. It is through this negative feedback loop that perceptions of present situations are compared with reference values that form ideal situations." ('Control Theory', 2014, p. 1)

In his article McQuail (1997) again criticizes the way media influencing in creating least understanding and an increasing dependence on the media as a player of monopoly, providing the audience what the media companies want instead of providing a complete freedom to the audience to watch certain media content. "The media were attributed the power to create extreme dependence in respect of basic psychic needs for identity and self-realization. The way they were organized made it virtually impossible to answer back and the media could impose a 'psychological illiteracy'" (McQuail, 1997, p. 1).

This both (large media corporations or wealthy individuals) reflects a degree of monopoly and also opens the way for influence on news content that favors big business interests. Such influence is not unrestrained and is not often easy to demonstrate, but it does exist, with little effective counterweight. (Whitney, Sumpter, & McQuail, 2004, p. 13)

Liberal theory and Public Interest theory in Journalism

These theories are related to journalism and are important to mention in the current research as it is pointed out by many scholars that the news media content based on audience data will shape the journalistic approaches and hence can challenge the certain theories in the field of journalism.

> "Liberal theory is one of the four main variants of normative theory. It exalts individual freedom of expression and publication over all other goals and forms of relationship to the wider society. Journalism has no set purpose and should accept no limits to its autonomy, a principle going far beyond the outlawing of censorship. […] Particular journalists will choose their own purpose or market and orient themselves

> accordingly without any overarching purposes and without imposed responsibilities from the society. Moreover, public interest theory covers all branches of theory that assign some positive social purpose to the work of journalists, on behalf of some higher general interest and wider public good (beyond simply pleasing clients or the immediate audience and making money).” (Carpentier, 2008, pp. 50–51)

The current research does not aim to support or criticize these theories based on the data analysis of current research. However, news media and the importance and power of society in shaping the news is influencing the journalism.

Research Domain

As the current research domain is too broad and without a thorough understanding of specific concepts and processes in the research domain, device fingerprinting is difficult to understand, therefore, this chapter will address those concepts and terms in detail for a better understanding of the research domain. Moreover, it will also provide a detailed understanding on device fingerprinting as it is the core focus of the current research. Without an understanding of these concepts and processes in media and audience measurement, it is almost impossible to understand the data acquisition and analysis done in the later chapters.

Mass Media

Media is a collection of information that is communicated through different sources including publishing on papers in different forms including newspapers, magazines etc., broadcasting over TV and radio, digital and still billboards, and in modern era it is communicated over the internet. With the advent of internet and availability of different devices in a human's personal life, media companies have put their focus on their media presence as platform independent and over the internet which provides greater audience reach.

Even though we have a modern and very efficient way of the information capture through internet, TV and print media still has its own huge market and importance. However, the media that is available on internet is accessible anytime and anywhere with least discrimination of the audience watching it. However, internet users still face a lot of issues which will be discussed further in this chapter. First, the popular media types that were mentioned above need to be briefly discussed.

Print Media

Print media involves all the media that is available on newspapers, magazines, brochures, company catalogues, etc. That is, all media that is available in a print form on any kind of paper is referred to as print media.

> "The print media evolved through three historical discoveries or milestones. These are the evolution of language, the evolution of writing and the invention of the printing machine. […] The print media can be classified into books, newspapers, magazine and pamphlets." (Odorume, 2012, pp. 1–2)

The modern era is called an era of the digitalization and that has enabled print media to transform into a model where they not only keep the old model of printing the media in different forms such as newspaper, magazines, brochures etc. but also their publication in the digital form which is available on the websites. This happens due to the digital disruptions in normally most of the organizations in different fields.

TV Media

Television was introduced to be able to broadcast a set of information especially moving images to a wide audience at the same time. Since the introduction of TVs in mid-late twentieth century, different kinds of media have been introduced to be broadcasted over TV. "It is a mode of (tele) communication typically used for transmitting moving (color) images (and potentially sound) to audiences" (Wherry & Schor, 2015). There have been significant developments during the twentieth century in the television industry where it was initially black-and-white television content was delivered. Then with certain technological developments, colorful images and videos could be broadcasted. While the advent of recorded cassettes, discs and players for these types of media storage, television was used to watch the content that has been recorded previously. Moreover, the quality of content kept improving over time from resolution, frames per second in recorded content and the colors of the media content being broadcasted or delivered using television.

Today, with the internet being the most dominating technology in the world, it has influenced the television industry as well. There is a new concept that became popular in twenty-first century, that is, Internet TV. "Recently, there has been a rise in Internet TV services, with services such as BBC iPlayer, Hulu, and Netflix broadcasting their material onto TV screens via the Internet " (Wherry & Schor, 2015).

Internet Media

The internet media here specifically refers to the online presence of television media companies in the form of official websites, social media and mobile applications. "Most, if not all, broadcasters have embraced the Internet and have developed websites for audience members to visit and digest information about the broadcast content" (Smith, 2015). Today,

the increasing internet penetration rates have attracted television media companies to ensure their presence in various forms where internet can be used by viewers to watch the media content. Along with the media websites and applications, advertisers also have been attracted by the internet advertisement and they try to make sure of more audience capture in the process where audience watch content on these websites. These television media websites offer various opportunities that are found much more flexible to watch than on traditional television at homes as the content can be watched anytime and repeatedly whenever the audience has internet and a device supporting these websites content to be displayed. "Viewers could go online to find more information about what they were watching, whether it was from an official site provided by the broadcaster, or through other options available such as forums or fan pages" (Smith, 2015).

With these media websites, there have been many efforts in understanding audiences so that the relevant advertisement would be displayed. Since, there have been numerous types of media content, the categories of advertisements need to be sorted based on the type of media content displayed on the websites.

Audience Measurement

The audience measurement has been a practice since mass media was introduced for public. The certain categories of audiences have been measured in different ways throughout the media history. All the different kinds of mass media including, print, digital, and internet-based media have different strategies to gather audience information and shape the advertisements as well as media content according to this audience data.

> "It took 38 years for the radio to attract 50 million listeners, 13 years for TV to gain the attention of 50 million viewers. The Internet took only 4 years to attract 50 million participants, and Facebook reach 50 million participants in only one and a half years." (Nair, 2011, p.46 as cited in; Smith, 2015, p. 24).

For the audience measurement in the era of print media only, there were limited ways to collect the audience information and customization of advertisement and media content which included traditional surveys and interviews; however, the technologies and methodologies became sophisticated in the era of television. Different companies used surveys, interviews

and later on some companies developed personalized devices to gather television audience information.

> "Additionally, if audiences have to sign up for use of a website or application, advertisers then potentially have access to knowledge about the demographic who uses the website or application in question, which allows advertisers to have more refined." (K. Babin, personal communication, January 8, 2015, as cited in; Smith, 2015, p. 24)

These and several other facts of gathering audience information is also due to the different kinds of devices a user uses to watch television content. Moreover, not only the specific television content watched by a certain user was important to be recorded but the behavior of the viewer based on the content that the viewer watched has a great influence on how advertisements and media content need to be customized or broadcasted to the user.

> "Of the adult cell phone owners that use Internet, email or apps on their phone, 35% used their phone to visit a website that was mentioned on television, 20% used their phone to see other audience members' comments online, and 19% used their phone to post a comment online 25 about a program they were watching." (Boyles & Smith, 2012, p. 4, as cited in; Smith, 2015, pp. 24–25)

Some of the popular methods and technologies that are used by online media websites to collect audience information will be discussed further in this chapter. However, based on the current research topic, the newest and least understood method, that is, device fingerprinting will be discussed in detail.

Cookies and Super-cookies

Websites globally have been using a technique to capture some data from a user's browser called cookie. This technique has been in use for several purposes including the planning and dissemination of content and advertisement for an audience. "The most common way to track web browsers (by 'track' we mean associate the browser's activities at different times and with different websites) is via HTTP cookies, often set by with 3rd party analytics and advertising domains" (Krishnamurthy, B., Wills, C., 2006, as cited in; Eckersley, 2010). However, due to its popularity and several web users being able to identify and stop them to capture or store their data from browsers, cookies are now considered a weak method for

audience measurement. Therefore, more sophisticated technologies were discovered to gather data from audience and used by media and advertising companies.

Another tracking mechanism called super-cookies was introduced which is much harder to control from a user end. "[super-cookies] are typically installed by vendors on behalf of advertising networks, so the vendor can gather and sell valuable browsing information that helps the advertisers deliver 'targeted' ads to users, based on their past online behavior and preferences" (Ring, 2015). It is not only difficult to identify super-cookies and stop them from tracking but it is a great security risk for the users. Hackers and malicious software programs can use these sets of data against individuals and companies. However, due to a lot of complaints and issues raised from super-cookies, different tech-companies introduced "Do Not Track" feature which would override the browser's capability to track any user's activity and information.

Device Fingerprinting

Another more advanced and least understood method of audience measurement and tracking is device fingerprinting. The more devices a person uses the more opportunities media companies and advertisers have as Acer et al. (2013) mentions "with the advent of smartphones and tablets, fingerprinting allows advertisers to augment previously gathered user-data and track the user across devices" (Acar et al., 2013). This is a technique initially introduced to tackle security issues such as fraud etc.; however, due to its audience data gathering technique, it is considered to be a breach of privacy as most of the users over the web are not aware of this method. Device fingerprinting is done using a JavaScript file running while browsing and that gets certain data from user's device. The following captured image is from a website while the user checked the JavaScript running during the browsing and the values JavaScript was capturing were found to be the browser, platform and other details. These in combination with some other values constitute to form a unique print called browser fingerprint.

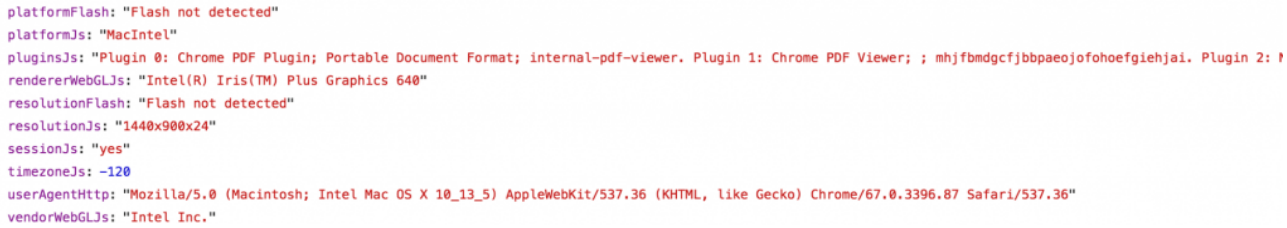


Figure 1: Typical values from JavaScript file running on a website that captures some user's data

The basic purpose for which device fingerprinting was introduced is explained by Van Goethem et. al. (2016) as:

> "This information (gathered during multi-factor authentication), which can consist of the user's IP address, behavioral and contextual information, or a fingerprint of the browser he is using to authenticate, is then compared against the user's typical behavior. In case some elements from this information deviate from what is expected, the user is either denied access, or is required to use a stronger authentication method for verification." (Van Goethem, Scheepers, Preuveneers, & Joosen, 2016, p. 107)

There are several ways a device fingerprinting is done including, browser fingerprinting, canvas fingerprinting and cookie printing. Several websites use different combination of the specific data in these three categories to create a unique signature to identify a user across the web.

In browser fingerprinting, some specific data including, browser type, version, the platform it is running on and other details are captured. "Fingerprinting user devices through the browser is an increasingly common practice used of advertising and antifraud companies" (Acar et al., 2013).

Time elapsed: 338.64ms

timezone	-120
screenSize	1440,900
availSize	1440,822
colorDepth	24
pixelRatio	2
userAgent	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_14_6) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/76.0.3809.132 Safari/537.36
cookiesEnabled	true
mathtan	-1.4214488238747245
dateFormat	01/01/1970, 01:00:00

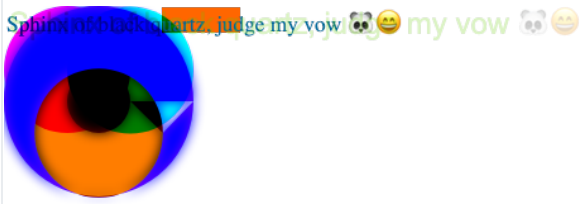
touchCompatibility	0,false,false
languages	en-GB,en-GB,en-US,en,,,
localStorage	true
sessionStorage	true
userData	false
indexedDB	true
doNotTrack	1
hardwareConcurrency	4
cpuClass	undefined
platform	MacIntel
plugins	>Plugin 0: Chrome PDF Plugin, internal-pdf-viewer, Portable Document Format, application/x-google-chrome-pdf, pdf;;>Plugin 1: Chrome PDF Viewer, mhjfbmdgcfjbbpaeojofohoefgiehjai, , application/pdf, pdf;;>Plugin 2: Native Client, internal-nacl-plugin, , application/x-nacl,
iePlugins	empty
webGLVendor	Intel Inc.
webGLRenderer	Intel(R) Iris(TM) Plus Graphics 640
adBlock	true
installedFontsJs	Andale Mono;Arial;Arial Black;Arial Hebrew;Arial Narrow;Arial Rounded MT Bold;Arial Unicode MS;Comic Sans MS;Courier;Courier New;Geneva;Georgia;Helvetica;Helvetica Neue;Impact;LUCIDA GRANDE;Microsoft Sans Serif;Monaco;Palatino;Tahoma;Times;Times New Roman;Trebuchet MS;Verdana;Wingdings;Wingdings 2;Wingdings 3;
canvasFp	f4f5187ab709b42b0aeed033095c4e5d 
audio	124.90863108361373

Table 1: A sample browser and canvas fingerprinting captured from the source http://fp.virpo.sk/

A canvas fingerprinting is normally taken using pixel information of a system and encodes it to a cryptographic hash or using another method.

> “To obtain this (canvas) fingerprint, a website renders text and WebGL scenes to a <canvas> element, then examines the pixels produced. Different systems produce different output, and therefore different fingerprints. Even very simple tests— such as rendering a

single sentence in a widely distributed system font— produce surprising variation". (Mowery & Shacham, 2012, p. 1)

In order to understand cookie printing it is important to understand how cookies are being used by the websites to gather audience information. Following illustration ('Browser Fingerprinting', n.d.) describes how a cookie is stored and then used by the websites for the required information including a user's device details and activities.

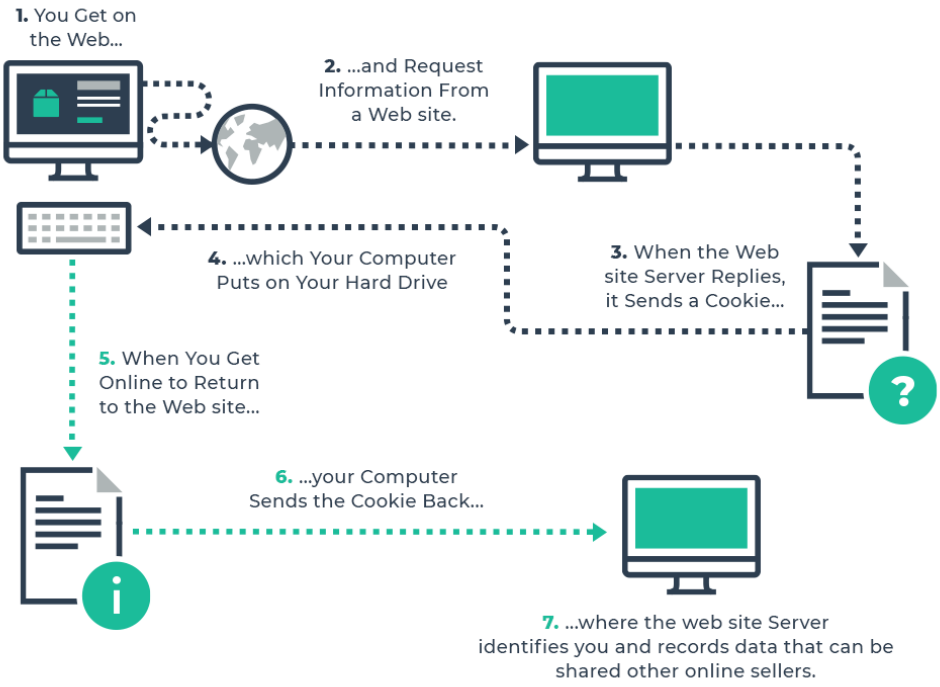


Figure 2: Cookie printing process

There are different details that are captured during browser, canvas and cookie fingerprinting which are retrieved by running a script in JavaScript language on each website/webpage that a user visits. That script takes certain properties which make a browser, canvas and/or cookie fingerprinting. These properties are briefly described below.

a. Display properties – Canvas FP

There are several values through which a system's display properties are retrieved. Display properties include but not limited to screen's dimensions in width and height with resolution in pixels, the window size that is available, the color depth, etc. Following is the details extracted for getting display properties and other values from a system that (Hraška, 2018) has discussed in his thesis.

	Value example

Display properties	
Screen size	1440x900
Available size	1440x827
Color depth	24
Pixel ratio	2

Table 2: List of all the features of display properties with example values. (Hraška, 2018)

All these properties are retrieved through JavaScript functions window.screen.width returns screen width, window.screen.height returns screen height, window.screen.availWidth is the screen width where a web browser can be displayed, window.screen.availHeight is the screen height where a web browser can be displayed, window.screen.colorDepth, window.devicePixelRatio and these values return values that are in the format mentioned in table 2.

b. Browser details – Browser FP

Among various details collected from browser, the most important are plugins and user-agent string. The Plugins parameter provide enough information to store in a fingerprint so that next time a user visits a website or a webpage, fingerprint is identified by a comparison of similar values of these features. Browser plugin feature provides all the details of installed plugins on a browser that is used to browse the website/webpage.

	Value example
Browser features	
AdBlock	true
Cookies enabled	true
Do Not Track (DNT)	false
Plugins	{name: Chrome PDF Plugin, fileName: internal-pdf-viewer, description: Portable Document Format, mimeType: . . . }, {. . . }
IE plugins	empty
Indexed database	true
Local storage	true
Session storage	true
Binary Behaviors	false
User-agent	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_13_3) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/65.0.3325.181 . . .

Table 3: List of all the features of browser details with example values. (Hraška, 2018)

window.navigator.plugins One of the most important string values here is retrieved from user-agent parameter which returns the browser name, its version, a set of values

from the platform it is running on, e.g., the hardware, operating system, and a combination of other values that could uniquely identify a user's browser. Browser name/version number (Machine name, Operating system with Hardware platform and version) Browser Engine/development build number

c. System properties – Canvas FP

Some of the fingerprints also retrieved from properties more specific to the system from where a website or webpage is accessed. CPU, timezome, languages being in use in the system, fonts and date format are retrieved using specific JavaScript functions such as getDateFormat, window.navigator.language, window.navigator.cpuClass, etc.

	Value example
System properties	
CPU class	undefined
Timezone	-120
Languages	en-US, sk-SK, sk, en
Installed fonts	Andale Mono; Arial; Arial Black; Arial Hebrew; Arial Narrow; Arial Rounded MT Bold; Arial Unicode MS; Comic Sans MS;
Date format	01/01/1970, 01:00:00
Tanh	-1.4214488238747245

Table 4: List of all the features of system properties with example values. (Hraška, 2018)

d. Hardware Properties – Canvas FP

To get hardware properties from a system through browser fingerprinting, certain JavaScript methods are used. For example, window.navigator.platform, window.navigator.maxTouchPoints, window.navigator.hardwareConcurrency etc.

	Value example
Hardware properties	
Hardware concurrency	4
Touch compatibility	0,false,false
WebGL vendor	Intel Inc.
WebGL renderer	Intel(R) Iris(TM) Graphics 540
Platform	MacIntel

Table 5: List of all the features of hardware properties with example values. (Hraška, 2018)

Following is a screenshot taken from an experiment of browser fingerprinting by the researcher where the JavaScript program takes WebGL vendor and renderer details.

```
var getWebGLVendor = function () {
  try {
    var canvas = document.createElement("canvas");
    var ctx = canvas.getContext("webgl") || canvas.getContext("experimental-webgl");
    return ctx.getParameter(ctx.getExtension('WEBGL_debug_renderer_info').UNMASKED_VENDOR_WEBGL);
  } catch (e) {
    return 'not supported'
  }
}

var getWebGLRenderer = function () {
  try {
    var canvas = document.createElement("canvas");
    var ctx = canvas.getContext("webgl") || canvas.getContext("experimental-webgl");
    return ctx.getParameter(ctx.getExtension('WEBGL_debug_renderer_info').UNMASKED_RENDERER_WEBGL);
  } catch (e) {
    return 'not supported'
  }
}
```

Figure 3: Example code that creates WebGL value to make a canvas fingerprint

"WebGL is a JavaScript API that extends the HTML 5 canvas API to render 3D objects from the browser. […] The first attribute reports the name of the GPU, for example ANGLE (VMware SVGA 3D Direct3D11 vs 4 0 ps 4 0). […] The second WebGL attribute (vendor) is expected to provide the name of the GPU vendor, whose value actually depends on the OS." (Vastel, Laperdrix, Rudametkin, & Rouvoy, 2018, p. 5)

e. HTTP Headers

Another part of a browser fingerprint is HTTP Headers which contributes to maximum level of uniqueness of the fingerprint generated.

	Value example
HTTP headers	
Accept	application/json, text/plain, */*
Accept encoding	gzip, deflate, br
Accept language	en-US,en;q=0.9,sk;q=0.8
User-agent	Mozilla/5.0 (Macintosh; Intel Mac OS X 10_13_3) Apple. . .

Table 6: List of all the features of HTTP Headers with example values. (Hraška, 2018)

f. Canvas Properties – Canvas FP

Some typical values that are retrieved by running a fingerprint script return values such as shown in the Table 7 below.

	Value example
Orthogonal features	
Canvas	875f14dcfa55c0f534b7809b0b5109d1
Audio	124.94877783898846

Table 7: List of all the features of canvas properties with example values. (Hraška, 2018)

The actual code that extracts and forms a final value of a fingerprint is retrieved from (‘Fingerprint’, n.d.) and it looks like the following:

```
var canvas =
document.createElement('canvas')
var getCanvasFp = function() {
  var result = []
  // Very simple now, need to
make it more complex (geo shapes
etc)
  canvas.width = 2000
  canvas.height = 200
  canvas.style.display = 'inline'
  var ctx =
canvas.getContext('2d')
  // detect browser support of
canvas winding
  //
http://blogs.adobe.com/webplatfor
m/2013/01/30/winding-rules-in-
canvas/
  //
https://github.com/Modernizr/Mode
rnizr/blob/master/feature-
detects/canvas/winding.js
  ctx.rect(0, 0, 10, 10)
  ctx.rect(2, 2, 6, 6)
  result.push('canvas winding:' +
((ctx.isPointInPath(5, 5,
'evenodd') === false) ? 'yes' :
'no'))

  ctx.textBaseline = 'alphabetic'
  ctx.fillStyle = '#f60'
  ctx.fillRect(125, 1, 62, 20)
  ctx.fillStyle = '#069'
  ctx.font = '13pt no-real-font-
123'
  ctx.fillText('Sphinx of black
quartz, judge my vow
\ud83d\udc3c\ud83d\ude04', 2, 20)
  ctx.fillStyle = 'rgba(102, 204,
0, 0.23456789)'
  ctx.font = '18pt Arial'
  ctx.fillText('Sphinx of black
quartz, judge my vow
\ud83d\udc3c\ud83d\ude04', 4, 22)

  // canvas blending
  //
http://blogs.adobe.com/webplatfor
m/2013/01/28/blending-features-
in-canvas/
  //
http://jsfiddle.net/NDYV8/16/
  ctx.globalCompositeOperation =
'multiply'
  ctx.fillStyle =
'rgb(255,0,255)'
  ctx.beginPath()
  ctx.arc(50.123456789, 50, 50,
0, Math.PI * 2, true)
  ctx.closePath()
  ctx.fill()
  ctx.fillStyle =
'rgb(0,255,255)'
  ctx.beginPath()
  ctx.arc(100, 50.456, 50, 0,
Math.PI * 2, true)
  ctx.closePath()
  ctx.fill()
  ctx.fillStyle =
'rgb(255,125,0)'
  ctx.beginPath()
  ctx.arc(75, 100, 50.789, 0,
Math.PI * 2, true)
  ctx.closePath()
  ctx.fill()
  ctx.fillStyle =
'rgb(125,0,255)'
  // canvas winding
  //
http://blogs.adobe.com/webplatfor
m/2013/01/30/winding-rules-in-
canvas/
  //
http://jsfiddle.net/NDYV8/19/$
  ctx.arc(75, 75, 75, 0, Math.PI
* 2, true)
  ctx.shadowBlur = 10;
  ctx.shadowColor = "blue";
  ctx.arc(75, 75, 25, 0, Math.PI
* 2, true)
  ctx.fill('evenodd')

  if (canvas.toDataURL) {
result.push('canvas fp:' +
canvas.toDataURL()) }
  window.canvasThing = canvas;
  return md5(result.join('~'))
```

At the end of the above code, a function called md5 is executed which is a hashing technique to get the results in a format like the one mentioned for the value Canvas in table 7.

Regulation

In this section, the study regarding regulation in EU will be discussed and specifically GDPR will be briefly discussed for particular parts when it is related to the audience data gathering using any method e.g. cookies or any other method. The researcher aims to present the statements and discuss whether they particularly address or defines any regulation for data capture done in a method such as device fingerprinting. For this purpose, GDPR which was implemented in May 2018 in EU will be studied.

GDPR

The General Data Protection Regulation in the EU had been implemented and kept modifying from time to time. As digital era brings more sophisticated opportunities for organizations to capture and process personal data, the regulation needed to be revised once again. The purpose of the EU's GDPR is to reserve the rights of individuals to capture, use and process the data by certain public, private and international organizations. According to the EU Commission's website about the Data Protection in the EU, the GDPR was revised and adopted in May 2016; however, the companies and organizations had two years of timeline to revise their strategies and comply with the new GDPR by May 2018 ('Data protection in the EU', n.d.).

It is important to first of all understand what does personal data mean and how does the EU's GDPR perceives as a meaning of personal data. The same website of EU Commission's webpage entitled 'What is personal data' answers to it completely and precisely as:

> Personal data is any information that relates to an **identified or identifiable living individual**. Different pieces of information, which collected together can lead to the identification of a particular person, also constitute personal data.
>
> Personal data that has been de-identified, encrypted or **pseudonymised** but can be used to re-identify a person remains personal data and falls within the scope of the GDPR.

> Personal data that has been rendered **anonymous** in such a way that the individual is not or no longer identifiable is no longer considered personal data. For data to be truly anonymised, the anonymisation must be irreversible. ('What is personal data?', n.d.)

While further explaining the examples of personal data, the same webpage of EU Commission explains that the following will be considered as personal data:

- a name and surname;
- a home address;
- an email address such as name.surname@company.com;
- an identification card number;
- location data (for example the location data function on a mobile phone)*;
- an Internet Protocol (IP) address;
- a cookie ID*;
- the advertising identifier of your phone;
- data held by a hospital or doctor, which could be a symbol that uniquely identifies a person.

**Note that in some cases, there is a specific sectoral legislation regulating for instance the use of location data or the use of cookies – the ePrivacy Directive (Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 (OJ L 201, 31.7.2002, p. 37) and Regulation (EC) No 2006/2004) of the European Parliament and of the Council of 27 October 2004 (OJ L 364, 9.12.2004, p. 1)*

According to the specified note, the regulation provided by the GDPR regarding a cookie ID explained in the mention ePrivacy Directive as:

> Where such devices, for instance cookies, are intended for a legitimate purpose, such as to facilitate the provision of information society services, their use should be allowed on condition that users are provided with clear and precise information in accordance with Directive 95/46/EC about the purposes of cookies or similar devices so as to ensure that users are made aware of information being placed on the terminal equipment they are using. […] Access to specific website content may still be made conditional on the well-informed acceptance of a cookie or similar device, if it is used for a legitimate purpose.

Moreover, the examples that will not be considered personal data includes a company registration number, an email address such as info@company.com, anonymised data.

Research Methodology

Research Strategy

The research strategy involves choosing a right approach towards conducting the current research, the data gathering/collection process, procedures to refine the data, and to explain the ways of getting results from the data and perform data analysis.

Research Approach

The current research aims to have a **deductive** approach with the process of collection and refinement of the data and analysis of the results by building connections with the **theories** to the findings as to support the theories or to mention the critics if the findings go against the theories. The theories that will be used for building the conclusions include selective exposure theory, digital labor theory along with audience commodity theory, control theory and liberal and public interest theory in journalism. However, the research will focus on connecting only the **digital labor theory and audience commodity theory** in the analysis section. The control theory, selective exposure theory and theories in journalism only serve as a guideline and motivation for the current research which also opens up the opportunities of further research to connect it to the control theory, selective exposure theory and theories in journalism.

The research approach also involves a study of the **literature** focusing on evolution of digital media, new models of audience measurement, critics and the effects of audience measurement on journalistic approaches, audience commodity and the legal aspects of audience tracking. Moreover, the current research focus is on a specific technology within the certain media category and without its complete understanding, the results and analysis would make least sense to anyone. Therefore, a study of the current **research domain** is already done after the theoretical framework that helps readers understand the umbrella of the domain in which the current research focus of device fingerprinting resides.

Research Method

The current research follows a **quantitative method** to generate statistical results and to analyze them in a quantitative way. In a quantitative way, the data is required to be in

numbers mostly and that the data can be countable during analysis. The current research questions allow the research to have the data available in a format that could be used to generate results so that these results are quantifiable and comparable in numbers. Statistical methods can be used directly such as creating averages, percentages and comparing instances of the results with the each other to draw conclusions.

Data Gathering

There has been **data** available from several websites including news and entertainment ones, the researcher aims to use the specific data received from the websites of public and private news channels of the countries within EU. Therefore, for the current research purpose the data and its nature are very important to find some meaningful results. The current research needs to collect data for a year to observe the data capture by the European news media websites which is not possible in a standard time of 6 months for a Master project. Therefore, it was found out that Sørensen & Kosta (2019) could provide the data from their research that could be refined and normalized to the relevant data including news media websites in EU and the data that could provide information regarding device fingerprinting values. They had retrieved this huge amount of data using a Virtual Machine that crawls through the websites and captures JavaScripts running on them, HTTP headers, HTTP responses etc. Their data was huge considering their research focus which included third-party trackers as an important factor.However, for the current research, specific data including only JavaScript table was required as device fingerprinting can easily be found out from JavaScripts running on these websites.

The following steps will be carried out to clean and prepare the data to find the results for the current research:

a) To find existing software(s) to process the collected data from previous research done by Sørensen & Kosta (2019) and clean it up for duplicates and anomalies to get the specific information regarding audience that is gathered by the news media.
b) If there is a lack of existing software tools to conduct the data analysis, an environment is needed to be built for data mining. In that case, python script (s) will be developed and the data from JavaScript files retrieved from these news media

websites will be analyzed for specific patterns that represent data for device fingerprint.

c) The results will be generated using the python script as graphs or tables to show the device fingerprinting features, comparing the results among the countries and the levels of device fingerprinting done in percentages individually by the websites and by the countries in EU

Although the raw data in a SQLite database is available for the current research from a research done by Sørensen & Kosta (2019) but as it is too large and contains more data than required for the current research objectives, the methods are specifically technology-based to clean the data for the specific purpose of current research. The next step is to identify the nature of the raw data and define the processes for filtering and cleaning it up, identify the device fingerprint patterns and features using a script that needs to be developed.

Available Data

The data (raw data for current research) received from Sokol and Sørensen (2019) was very large as they have included more than a thousand media websites in different categories based on the type of media – e.g. News, Entertainment, Sports etc. - as well as based on different geographical regions – e.g. EU, Non-EU and not-EU. Moreover, a single set of data among three sets in SQLite was more than 15 gigabytes including millions of rows with data from third party websites running on main websites. The objectives of the current research only allow to use limited data in categories of news as media type and choosing countries from EU. Also, the research objective is to find out whether the news media websites specifically in EU use device fingerprinting and if yes, it is to find out what kind of fingerprinting type is being used. Therefore, the first step is to run certain SQL queries and extract the required rows and column from three data sets (June 2018, February 2019 an dJune 2019) into a file that could be easily loaded into a python script.

The researcher received the raw data captured from all these media websites for three periods of time, that is, first from June 2018, second February 2019 and the third one from June 2019. This information is identified from the ‘javascript’ table in the databases.

Analyzing the data for a possibility of any information that might exist and contribute to identifying device fingerprint patterns is a challenge from a huge database. However, by

studying the patterns and ways of processing the data for a device fingerprint can help in this process. The researcher has studied the patterns and values if device fingerprinting and presented in a previous chapter ‘Research Domain’.

Filtering the Data

The most important step is to narrow down to the problem practically, that is, to reduce the amount of data without losing any information that might contribute in identifying the device fingerprint by any news media website. Therefore, the researcher used SQL queries to select each record of websites of all the media companies that have values existing in the columns named ‘symbol’ and ‘value’ as these two columns contain the information to detect device fingerprint. This preliminary result for cleaning up data is also taken based on unique ‘script_url’ that is the website links to avoid the duplicate records as the script_url field contained not only the websites but also the webpages, so, for each media website, there could be at least one record. The researcher studied in detail and tested the raw data by performing the following query to make sure that by choosing only unique URLs do not result in loss of any required data to identify fingerprint against each website. The following query was run on each of the data sets available in SQLite format:

> *select * from ‘javascript’ where ‘symbol’ IS NOT NULL AND ‘value’ IS NOT NULL GROUP BY ‘script_url’*

The results for all three raw data sets were exported to CSV (Comma Separated Values) that could be processed in a python script. But before performing this step, the researcher needed to normalize all the three databases to reduce the size by resolving the duplicate records. Along with the raw data, Sørensen & Kosta (2019) also provided a CSV file with only top-level domains for specifically the news media websites and only those from EU with two hundred and eighty-two records.

Identifying the Device Fingerprint Pattern

So far, the data available to proceed further is available in the four files including ‘june2018.csv’, ‘feb2019.csv’, ‘june2019.csv’ with the data that contained fingerprint patterns, and ‘News-in-EU-PublicPrivate.csv’ containing all top-level domains for each news media website in countries within the EU. These files are loaded in four python scripts

developed during this research (and listed in Appendix A) to identify the device fingerprint patterns. The patterns were studied and understood in the chapter Research Domain under the heading of Device Fingerprinting.

The first python script that is developed to identify the browser fingerprint, canvas fingerprint and cookie print is named as ‘myscript.py’. This script gives results in the form of a bar graph for each country against the top-level domains in the EU.

The second script is developed, named as ‘comp_countries_pie.py’, to identify the same fingerprints but creating graphical results in the form of a donut pie chart comparing the cumulative percentage levels of device fingerprinting used by all the websites in each country.

The third script named as ‘comp_countries_bar.py’ is developed to identify the same fingerprints but creating graphical results in the form of a horizontal bar graph comparing and representing the percentage levels of device fingerprinting used by all the websites in each country.

The fourth script is developed for the results to be retrieved named as ‘collective.py’ to get the last result for a transition each country’s news media website in capturing unique fingerprint.

Results and Analysis

The results are generated using string comparisons for device fingerprint and using different libraries to generate the graphical results by writing different code snippets in the python script. There are some results in which some data is missing for certain websites. As the data for the current research was received from the scholars Sørensen & Kosta (2019) and the current results were retrieved using the javascript database, as the javascript can have those certain methods (mentioned in the 'Research Domain' chapter) that create a device fingerprint, the websites being visited needed to have JavaScript enabled and working. Therefore, the missing data could be because of two reasons: either the website did not have JavaScript enabled or the website did not use any device fingerprinting method at all. As the current research scope is not to find out whether these websites allow JavaScript or not, this missing data will not be considered in the data analysis to stay neutral.

Understanding the Results

The current research topic and the research questions can be answered and analyzed through a deep understanding of research in data and processing it to get the desired results. It is important to understand the sampling and coding methods used to retrieve the results and provide the analysis.

Data Sampling and Coding

For the identification of device fingerprint, three variables were defined for browser fingerprinting, canvas fingerprinting and cookie printing and initialized them as dictionaries with the values against the 'symbol' field of JavaScript functions which run and detect the fingerprint values from 'symbol' column such as window.navigator.userAgent etc. A percentage level is assigned to each of these values that makes a total of 100 for each dictionary of browser fingerprint, canvas fingerprint and cookie print. The domain study provided enough knowledge to assign a relevant percentage level to each of these values in the dictionary **based on their contribution in making a unique fingerprint**.

For example, in a browser fingerprinting (also mentioned in the following table), a symbol function of window.navigator.userAgent returned a combination of values (examples of these values can be seen in the chapter 'Research Domain' under the topic of 'Device fingerprinting'), hence, contributing more to the uniqueness of device fingerprint; whereas a

symbol function of window.sessionStorage would return a single value that contributes less to make a device fingerprint unique.
Similarly, in canvas fingerprinting, the symbol value of the function CanvasRenderingContext2D returned 6 values in most cases which proved that this method contributes the most in making a unique device fingerprint.
The following levels were assigned for each symbol value based on their value contribution to make the whole device fingerprint as unique as possible:

Dictionary	**Symbol Value and assigned percentage level**
Browser fingerprinting	window.navigator.userAgent = 20, window.sessionStorage= 10, window.navigator.platform= 10, window.navigator.language= 10, window.localStorage= 10, window.navigator.plugins= 20, window.navigator.doNotTrack= 10, window.navigator.cookieEnabled= 10 Altogether 100 therefore, the above will be considered as percentages
Canvas Fingerprinting	window.screen.colorDepth=10, window.screen.pixelDepth=10, HTMLCanvasElement= 20, CanvasRenderingContext2D=60 Altogether 100 therefore, the above will be considered as percentages
Cookie printing	window.document.cookie = 100

Table 8: Device Fingerprinting categories and corresponding symbol values

These values were assigned to efficiently generate results that could be understandable and comparable among countries within EU and the news media websites operating in those countries to see the rates of device fingerprinting data being captured. In the next chapter these results will be presented and analyzed in detail.

As soon as the study of data is done using the python scripts, the results were retrieved and analyzed in the categories such as:

1. **Device fingerprinting methods in each country in EU:** The percentage value of each method (browser fingerprinting, canvas fingerprinting and cookie printing) mentioned in above table will be calculated for each news media website in each of the twenty-eight countries in EU.

   Result Generation: The result generated using a python script "myscript.py" presented in Appendix A which was written soon after defining methodology. This is the first result to identify type of device fingerprinting used for each of the individual countries in EU observed in 3 different point of times – June 2018, February 2019 and June 2019
2. The **Levels of capturing unique device fingerprint by the news media websites in each of the 28 countries in EU** technique in percentages.

   Result Generation: This result in three bar graphs for three point of times – June 2018, February 2019 and June 2019 - is generated for each country using the python script 'comp_countries_bar.py' presented in Appendix A. This result represents the percentage of uniqueness of the device fingerprinting done by websites in each country. This is calculated using mean method of statistics as all the websites were considered equally.
3. A **Comparative Result of the Countries in Percentages** in each of the twenty-eight countries as a cumulative of hundred percent.

   Result Generation: The result in three pie charts is generated using the python script 'comp_countries_pie.py' presented in Appendix A. This result is generated by calculating the average of individual percentages of unique device fingerprinting done by each website in a country. This average of all the media websites in each country is then calculated as percentage of unique device fingerprinting generated by each country to see the comparison among EU countries.
4. Analysis of device fingerprinting done in these 28 countries for three time periods – **Comparisons of the changes in device fingerprinting from June 2018, February 2019 and June 2019**.

   Results Generation: This result as a spaghetti plot or a mixed line graph is generated with a python script 'collective.py' which represents the overall changes from June

2018 till February 2019 and then from February 2019 till June 2019 in capturing device fingerprinting by these 28 websites as a cumulative result.

This result will help readers understand that whether capturing unique device fingerprinting by news media websites in certain countries has increased or decreased over the three specified periods of time.

All these results are for each of the three data sets retrieved in June 2018, February 2019 and June 2019.

Analysis 1: Device fingerprinting methods in each country

Selection of limited results

The first results will be analyzed in a selective manner choosing only three countries for each of three time periods – June 2018, February 2019 and June 2019 for Austria, Belgium and Denmark. This is because this type of result contained 28(countries)*3(data sets) = 84 graphs as below in figure 4; however, presenting all the 84 results in the current research content is not possible. Moreover, the second sub-question of the current research to identify **what are the techniques used in device fingerprinting to gather audience data through the websites of news media companies**. Therefore, it is not necessary to include all the countries but it is important to find out specifically those techniques used for capturing audience data by these websites. However, the rest of the results are included in the Appendix B. Also, the percentages calculated in the third analysis will also be presented in the form of donut pie graphs in Appendix C. However, an analysis will be provided in the current chapter.

The second and fourth results will be included and analyzed in this chapter.

Also, during the process of capturing results, the data was null for some websites in in the raw data. This might be because these websites did not capture any kind of audience data that could contribute to make a device fingerprint. But this is an assumption, further tools and techniques of research can be used to find out the exact reason of missing data for these websites. This could be done to ensure that these websites do not use any higher level of technologies to avoid data capture as the scholars Sørensen & Kosta (2019) did for a year. However, these websites will be identified during the following analysis for the specific countries including Austria, Belgium and Denmark.

To answer the second sub-question "how do the news media websites in EU capture data from audience using device fingerprinting" will be answered here.

The results were retrieved for each of the twenty-eight countries in EU and for each of the three data sets using a python script where each type of fingerprinting data was analyzed.

Austrian News Media Websites

The data set of **June 2018** and for Austrian news media websites revealed that:

Cookie printing: All of them retrieved cookie printing. Therefore, it has 100 percent contribution in making a unique device fingerprint.

Browser fingerprinting: All of these websites also used browser fingerprinting to some extent that is equal to or below 50 percent.

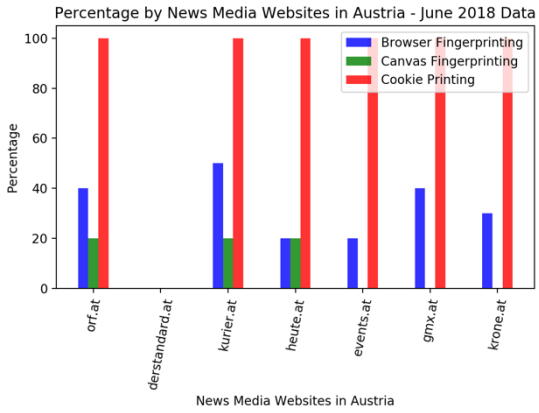


Figure 4: Percentage of methods to capture device fingerprinting by Austrian News Media Websites - Data from June 2018

Canvas fingerprinting: only three of these websites retrieved data that make a canvas fingerprint at a rate equal to or less than 20 percent. There was no data captured for one website derstandard.at in the raw data set received from the scholars Sørensen & Kosta (2019) and the possible reasons are mentioned earlier in this chapter, this missing data will not be analyzed.

The data for Austrian news media websites retrieved in **February 2019** showed slightly different results as shown in figure 5.

Cookie fingerprinting: There is no difference than the data captured in June 2018.

Browser fingerprinting: It is also the same as for browser fingerprinting values in June 2018.

Canvas fingerprinting: However, the canvas fingerprinting is changed from the data captured in June 2018. The website orf.at and heute.at captured 0 percent of data for canvas fingerprinting in February 2019 whereas both of the websites captured around 20 percent data

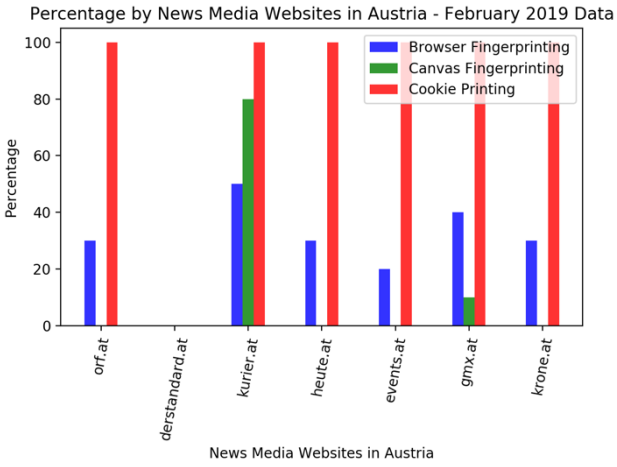


Figure 5: Percentage of methods to capture device fingerprinting by Austrian News Media Websites - Data from February 2019

in June 2018. Instead, gmx.at captured 10 percent of canvas fingerprinting data in second data set whereas 0 percent in the June 2018.

However, the rate of capturing data for canvas fingerprinting is increased in February 2019 by the website kurier.at from 20 percent in June 2018 to 80 percent in February 2019. It will be interesting to see the results for these websites in June 2019.

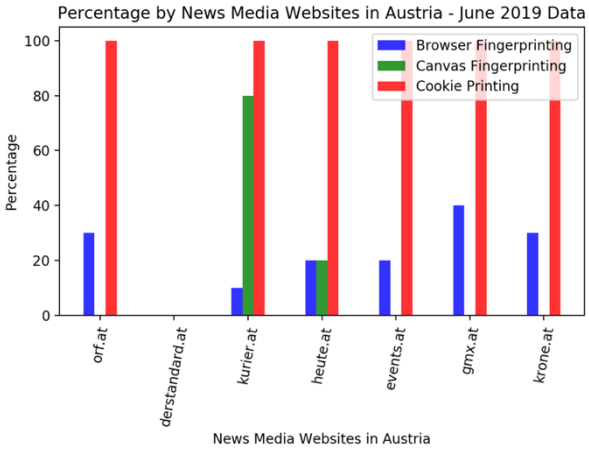


Figure 6: Percentage of methods to capture device fingerprinting by Austrian News Websites – Data from June 2019

In the data set of **June 2019**, the website kurier.at captured same amount of data for canvas and cookie fingerprinting; however, the browser fingerprinting levels decreased from 50 percent to 10 percent as compared to the data captured in February 2019. Whereas, this time heute.at captured canvas fingerprinting data instead of gmx.at which is observed to be 20 percent.

Conclusion:

For Austrian news media websites, the mostly used method for audience data capture is the cookie, second most used method is browser fingerprint and the least used method is canvas fingerprint.

Belgian News Media Websites

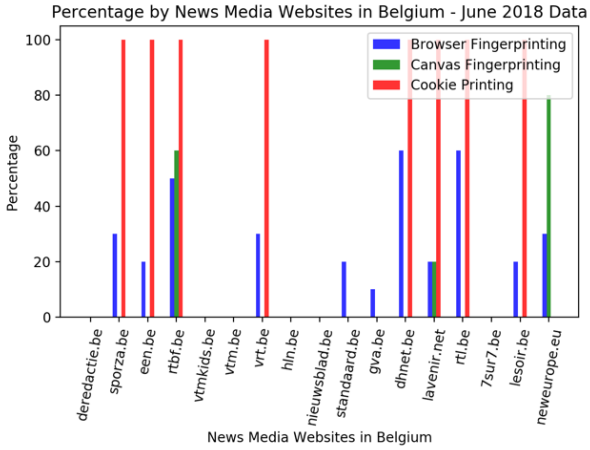


Figure 7: Percentage of methods to capture device fingerprinting by Belgian News Websites – Data from June 2018

The raw data received from Sørensen & Kosta (2019) was missing data for five websites including deredactie.be, vtmkids.be, vtm.be, hln.be, and 7sur7.be and the possible reasons have been mentioned earlier in this chapter. However, the data for the other twelve websites were available to be analyzed.

For the data gathered in **June 2018** the following result was found:

Cookie Fingerprint: Eight out of twelve websites captured cookie print; standard.be, gva.be and neweurope.be did not capture the cookie print.

Browser Fingerprint: Eleven out of twelve websites captured browser fingerprinting on an average of 31 percent of browser fingerprinting data capture. Only two of them exceeded 50 percent in gathering data that can make a browser fingerprint.

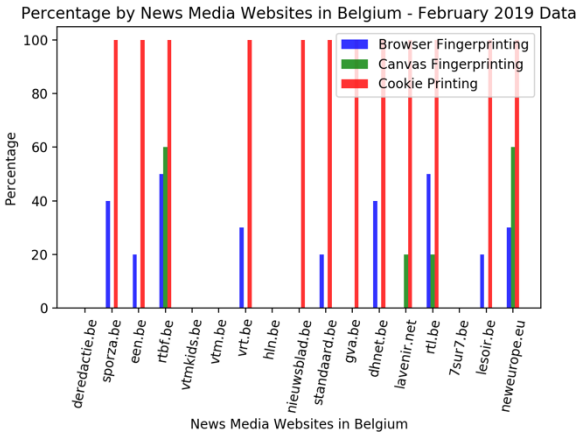


Figure 8: Percentage of methods to capture device fingerprinting by Belgian News Websites – Data from February 2019

Canvas Fingerprinting: Only three websites among these twelve websites collected data for canvas fingerprinting on an average of 53 percent.

In **February 2019**, the following data was observed for all three categories of device fingerprint:

Cookie Fingerprinting: Cookie fingerprint was found to be the same as in June 2018 but in February 2019, nieuwsblad.be, standard.be, gva.be and neweurope.be also captured cookie print.

Browser Fingerprinting: In comparison to the data retrieved in June 2018, browser fingerprinting decreased for gva.be from ten percent to 0, dhnet.be from sixty to forty percent and rtl.be from sixty to fifty percent. For the rest of the website it was the same.

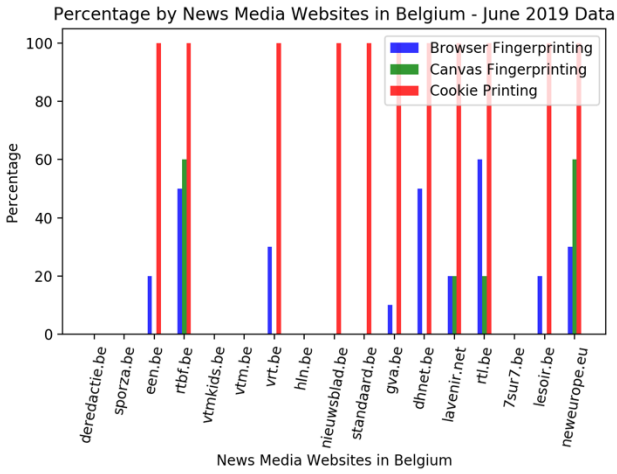


Figure 9:Percentage of methods to capture device fingerprinting by Belgian News Websites – Data from June 2019

Canvas Fingerprinting: The average of canvas fingerprinting is found to be the same in June 2018 and February 2018 with a decrease from 80 percent to 60 percent by neweurope.be and an increase from 0 to 20 percent by rtl.be.

In **June 2019**, the result was found to be as following for Belgian news websites:

Cookie Fingerprint: The data capture of cookie fingerprint was same as February 2019 for all the websites except for sporza.be which did not capture cookie data at all.

Browser Fingerprint: The browser fingerprinting data was not captured by sporza.be and standaard.be at all; however, gva.be increased by ten percent, dhnet.be ten percent, lavenir.be twenty percent and rtl.be increased 10 percent of the data capture for browser fingerprint.

Canvas Fingerprinting: Canvas fingerprint was found to be the same – captured by four of the websites on an average of 40 percent of contribution to make a device fingerprint.

Conclusion

The data capture through cookies by Belgian news media websites remained highest as compared to browser and canvas fingerprinting. The second most popular method of device fingerprinting involved browser fingerprinting. The third one was found to be canvas fingerprinting.

Danish News Media Websites

In **June 2018** the following results were found for the Danish news media websites in capturing device fingerprinting data:

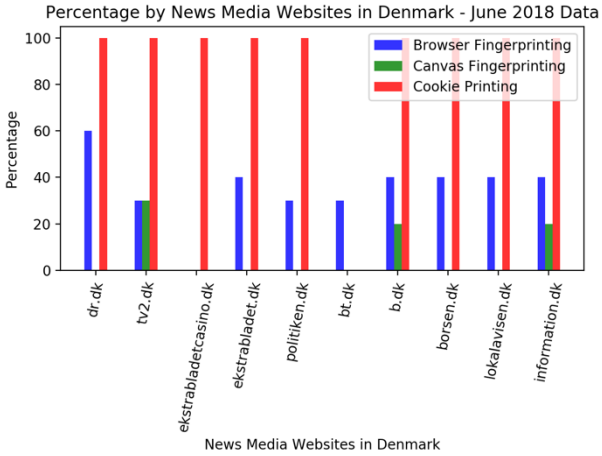


Figure 10: Percentage of methods to capture device fingerprinting by Danish News Websites – Data from June 2018

Cookie Fingerprint: All the news media websites except bt.dk captured cookie fingerprint.

Browser Fingerprint: All the news media websites except ekstrabladetcasino.dk captured browser fingerprint data. Rest of the websites captured on an average of 38 percent of data that could contribute to make a unique fingerprint.

Canvas Fingerprinting: Only three of the websites captured canvas fingerprint data on an average of 23 percent of data that could contribute to a unique device fingerprint.

In **February 2019** the following data was found to be analyzed:

Cookie Fingerprint: The number of websites that *did not capture* cookies was *increased* from one to three websites in February 2019.

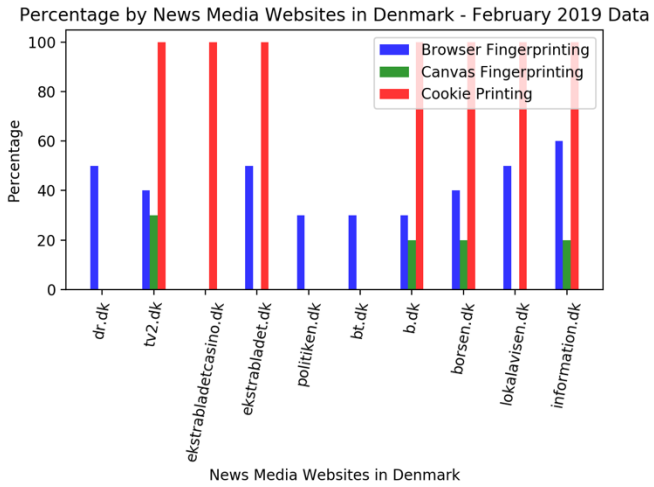


Figure 11: Percentage of methods to capture device fingerprinting by Danish News Websites – Data from February 2019

Browser Fingerprint: Capturing the browser fingerprint data was slightly *increased* by tv2.dk from thirty to forty percent, ekstrabladet.dk from forty to fifty percent, b.dk from thirty to forty percent, lokalavisen.dk from forty to fifty and information.dk from forty to sixty percent with a total average of 42 percent of data capture.

Canvas Fingerprinting: It was *increased* from 0 to 20 percent by borsen.dk while it was same for the rest of the websites as in June 2018. The average was 22 percent of data capture.

In **June 2019** the following result was retrieved:

Cookie Fingerprint: It was found to be the *same* as in February 2019, so the data capture remained decrease as compared to June 2018.

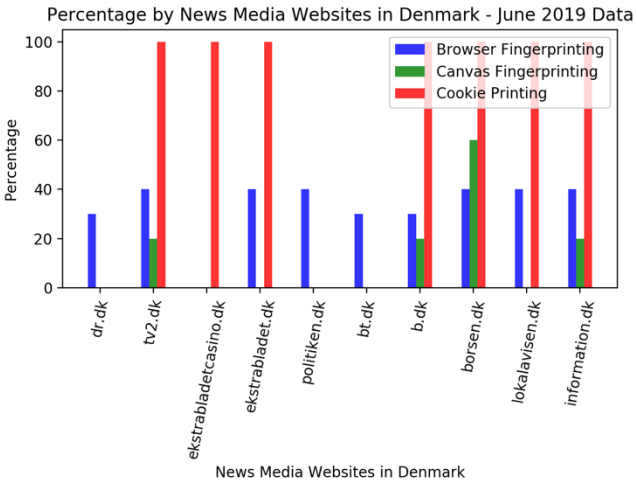


Figure 12: Percentage of methods to capture device fingerprinting by Danish News Websites – Data from June 2019

Browser Fingerprint: It was found to be slightly *decreased* as compared to the data in February 2019 from an average of 42 percent to 36 percent of data capture that contributes to unique device fingerprint.

Canvas Fingerprinting: The overall canvas capture was *increased* from 22 percent to 30 percent in June 2019.

Conclusion

The cookie fingerprinting and browser fingerprinting were done equally but the cookie fingerprinting trend decreased slightly in the data retrieved in February 2019 and June 2019. Canvas fingerprinting done the least by the Danish news media websites as well.

Research Sub-Question a

This analysis shows the results that answers our first sub-question that "**What are the techniques used in device fingerprinting to gather audience data through the websites of news media companies"**. These news media companies capture data that represent cookie print, browser fingerprinting, and canvas fingerprinting from a device used to browse the website. This data had been captured using the code myscript.py in Appendix A.

This result and analysis support the **digital labor theory and audience commodity** theory by proving that the data generated from audience through their devices is captured and used by the media companies which depends on each media company's strategies to use for particular purposes.

Analysis 2: Levels of capturing unique device fingerprint

The purpose of analyzing the data collectively including browser fingerprinting, cookie fingerprint and canvas fingerprint as an average and also taking an average for all the news websites in each of these 28 European countries, is to have an overview of the uniqueness levels of device fingerprinting.

Data Set from June 2018

In June 2018, news media websites in Estonia had the most unique device fingerprint of around 55% distinctiveness. This was followed by Luxembourg with around 54% and Portugal with around 53% uniqueness.

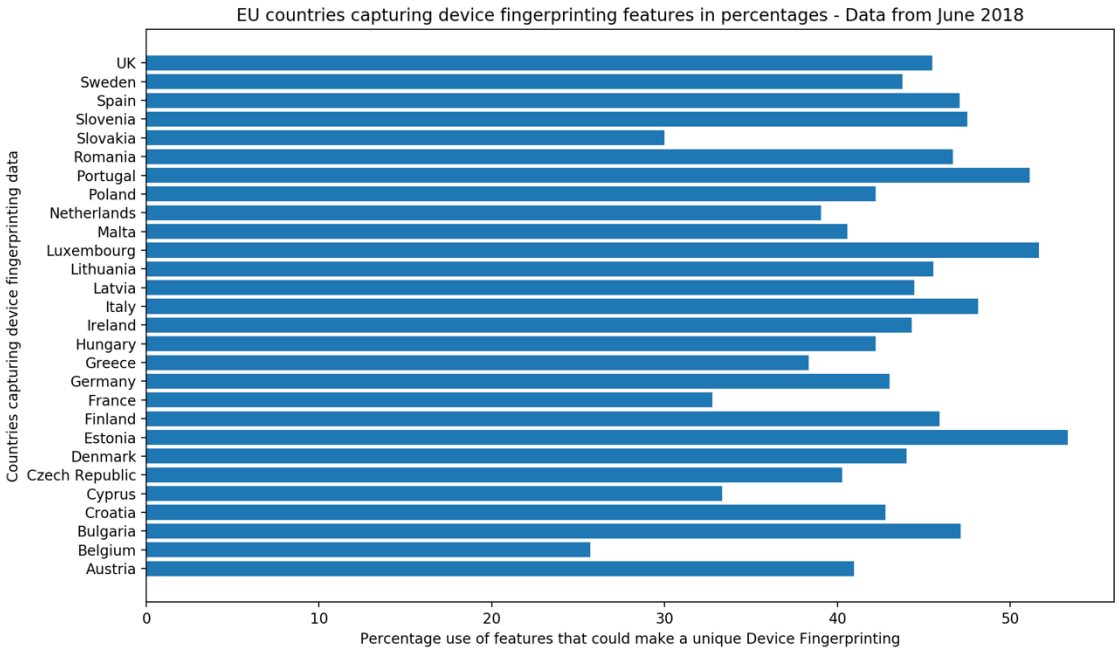


Figure 13: Percentage levels of capturing unique device fingerprinting by 28 European countries in EU observed in June 2018

The lowest uniqueness of fingerprinting was done by Belgium with around 27% uniqueness of the device fingerprinting; whereas, websites from Slovakia and France captured data for device fingerprinting with second and third lowest uniqueness.

Data Set from February 2019

In February 2019, news media websites in Portugal and Lithuania captured the data that could make most unique device fingerprint. Followed by Estonia and Luxembourg with almost the same level of around 53% uniqueness.

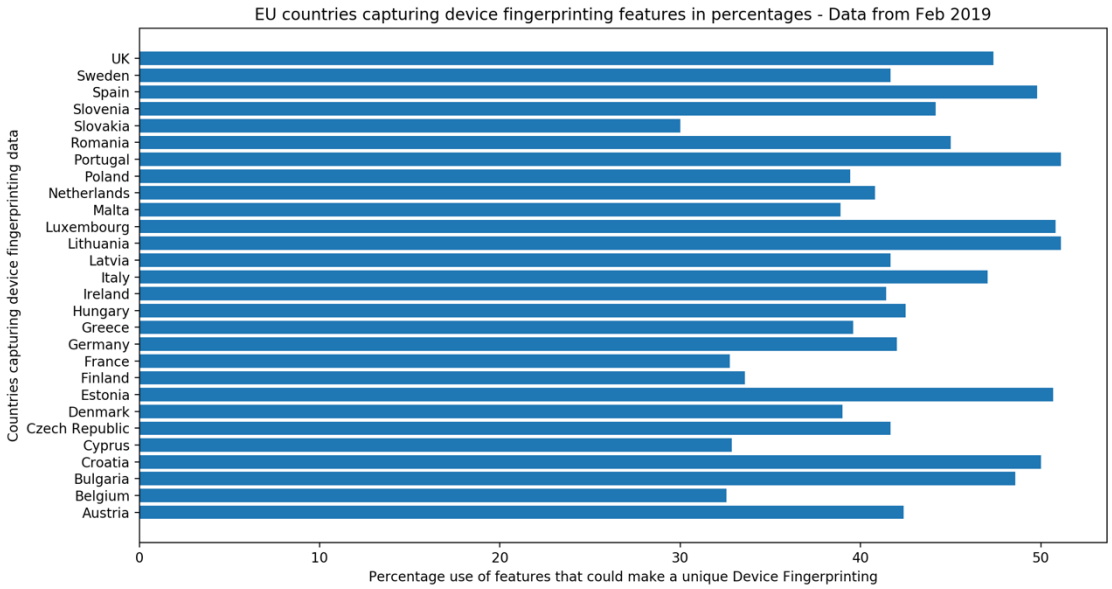


Figure 14: Percentage levels of capturing unique device fingerprinting by 28 European countries in EU observed in February 2019

However, the website from Slovakia captured data that could create least unique device fingerprinting data with around 32% uniqueness. Belgium the second lowest with around 33% whereas, Cyprus and France with captured third lowest with around 34% uniqueness.

Data Set from June 2019

In June 2019 – the last time when the websites were observed – the news media websites from Spain and Portugal captured data that could create a unique device fingerprint with around 52% uniqueness followed by Bulgaria and then Estonia with around 50% uniqueness. However, the websites from Slovakia captured least unique fingerprinting with around 26% uniqueness and Belgium with second lowest with around 30% uniqueness of the device fingerprint.

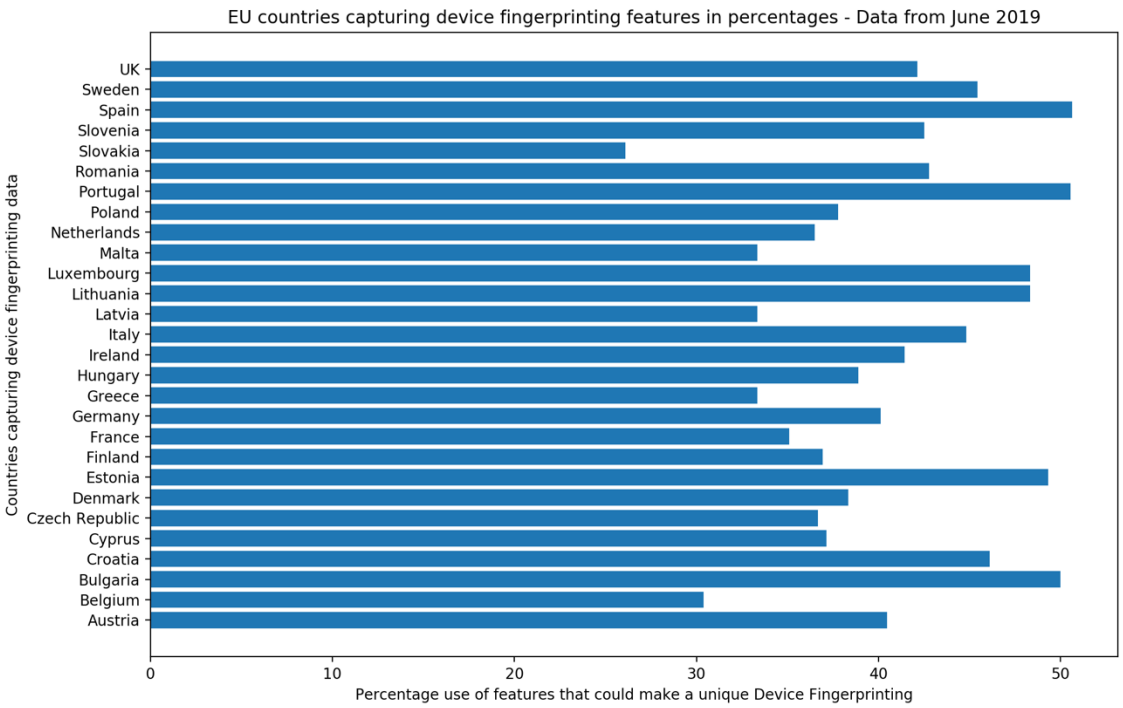


Figure 15: Percentage levels of capturing unique device fingerprinting by 28 European countries in EU observed in June 2019

Analysis 3: Comparative Result of the Countries in Percentages

The results were generated in the form of pie-charts and presented in Appendix C that helped in understanding the results to analyze in the current topic.

The following table summarizes the comparison of data capture in percentages by news media websites in each country in the EU for all three data sets including June 2018, February 2019, and June 2019.

	News media Websites in Country	**Percentage Data Capture in June 2018**	**Percentage Data Capture February 2019**	**Percentage Data Capture June 2019**
1	Austria	3.4	3.6	3.6
2	Belgium	2.1	2.7	2.7
3	Bulgaria	3.9	4.1	4.4
4	Croatia	3.6	4.2	4.1
5	Cyprus	2.8	2.8	3.3

6	Czech Republic	3.4	3.5	3.2
7	Denmark	3.7	3.3	3.4
8	Estonia	4.5	4.3	4.3
9	Finland	3.8	2.8	3.2
10	France	2.7	2.8	3.1
11	Germany	3.6	3.5	3.5
12	Greece	3.2	3.3	2.9
13	Hungary	3.5	3.6	3.4
14	Ireland	3.7	3.5	3.6
15	Italy	4.0	4.0	3.9
16	Latvia	3.7	3.5	2.9
17	Lithuania	3.8	4.3	4.3
18	Luxembourg	4.3	4.3	4.3
19	Malta	3.4	3.3	2.9
20	Netherlands	3.3	3.4	3.2
21	Poland	3.5	3.3	3.3
22	Portugal	4.3	4.3	4.4
23	Romania	3.9	3.8	3.8
24	Slovakia	2.5	2.5	2.3
25	Slovenia	4.0	3.7	3.7
26	Spain	3.9	4.2	4.5
27	Sweden	3.7	3.5	4.0
28	UK	3.8	4.0	3.7

Table 9: Comparative result of the 28 European countries in cumulative percentages that capture unique device fingerprinting - Data from June 2018, February 2019, June 2019

Research Sub-Question b

To which extent the countries in EU capture data from audiences using device fingerprinting technique?

In **June 2018**, the news media websites in *Belgium* captured the **minimum** levels of data for device fingerprinting with a score of 2.1/100 percent of the total device fingerprinting by all the twenty-eight countries in EU mentioned below in the pie chart. The second lowest score

was from the news media websites in *Slovakia* with 2.5% and third lowest score was from the news media websites in France with 2.7%.
In the same year, *Estonia* scored the **highest** in capturing data that could identify a device uniquely with a score of 4.5/100 percent in comparison to all the other twenty-seven countries in EU. News media websites in *Portugal* and *Luxembourg* captured data on second highest levels with 4.3% of total data capture that could make a device fingerprint to uniquely identify devices used by the audiences.

In the data set of **February 2019**, the news media websites in *Slovakia* captured the **least** data that could make a device fingerprint with a percentage of 2.5/100 in comparison to all the 28 countries in EU. Websites in *Belgium* scored 2.7% while those in *Cyprus*, *Finland France* scored 2.8% each.
The **highest** data captured in February 2019 by the news media websites were from *Portugal*, *Luxembourg*, *Lithuania*, and *Estonia* with a score of 4.3% in data capture that makes a device a fingerprint.

In the data set retrieved in **June 2019**, the news media websites in *Slovakia* captured **least** data that can contribute to a device fingerprint with a score of only 2.3 % in comparison to all the other 27 countries in EU. The websites in *Belgium* captured second lowest data with a score of 2.7% while the websites in *Greece*, *Latvia* and *Malta* scored third lowest with 2.9% of data capture to make a device fingerprint.
In the same data captured in June 2019 for all the news media websites in 28 countries in EU, the **highest** amount of data captured by the news media websites in *Spain* with a score of 4.5% whereas, the websites in *Portugal* and *Bulgaria* captured second highest levels of data to make a device fingerprint with a score of 4.4 %.

The analysis number 2 and 3 show that the device fingerprinting is done mostly by the news media websites in countries such as Estonia, Portugal, Lithuania, Luxembourg, Spain and Bulgaria. These countries that scored highest in capturing data of device fingerprinting, varied in three point of times, that is, June 2018, February 2019 and June 2019, with different country leading in each time period.

Whereas, news media websites in countries such as Belgium, Slovakia, France and Cyprus had done the least device fingerprinting while Greece, Latvia and Malta decreased the levels of device fingerprinting to 5% in the last data observed in June 2019. The decrease and increase in levels of capturing device fingerprinting data over the specified periods of time will be analyzed in Analysis 4.

Analysis 4: Transition of Device Fingerprinting levels from June 2018, February 2019 and June 2019

In this analysis the third research sub-question will be answered. Therefore, percentage levels of device fingerprinting as a cumulative for all the news media websites in each country will be represented on the y-axis of the graphs shown below. These levels will be identified for each country over the three time periods in order to conclude whether the levels have decreased or increased over time after the implementation of GDPR in the EU.
As there were 28 countries, the cumulative line graphs – also called spaghetti plot – would be unreadable for the readers of this research. Therefore, these graphs have been split into four only to make it readable for the readers. However, the data representation from all the countries will be analyzed without any discrimination. Moreover, this will also help in analyzing the transitions in four groups of seven countries at one time.

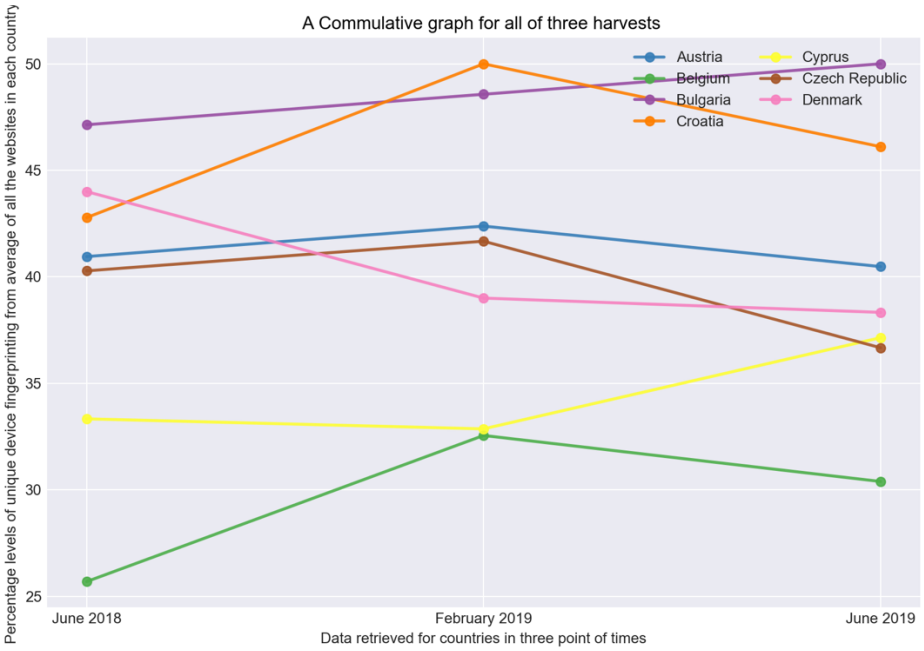


Figure 16: Transition of percentages in capturing unique Device Fingerprinting from June 2018 - February 2019 - June 2019 for the first set of seven countries in EU

The news media websites in countries such as Belgium, Croatia, Austria, Czech Republic, and Bulgaria, the percentage of device fingerprinting average increased in February 2019 than in June 2018; whereas, the news media websites in Cyprus and Denmark the percentage of device fingerprinting average decreased in February 2019.

From February 2019 till June 2019, the news media websites including Austria, Belgium, Croatia, Czech Republic and Denmark captured less data for device fingerprinting. However, news websites from Bulgaria and Cyprus increased the device fingerprinting levels from February 2019 till June 2019.

Overall, the news websites that kept decreasing from June 2018 till June 2019 observed in three point of times included Denmark only. **If the levels of device fingerprinting from June 2018 and June 2019 are compared only, then the countries including Austria, Denmark and Czech Republic (3 countries) have decreased level of data captured for device fingerprinting.**

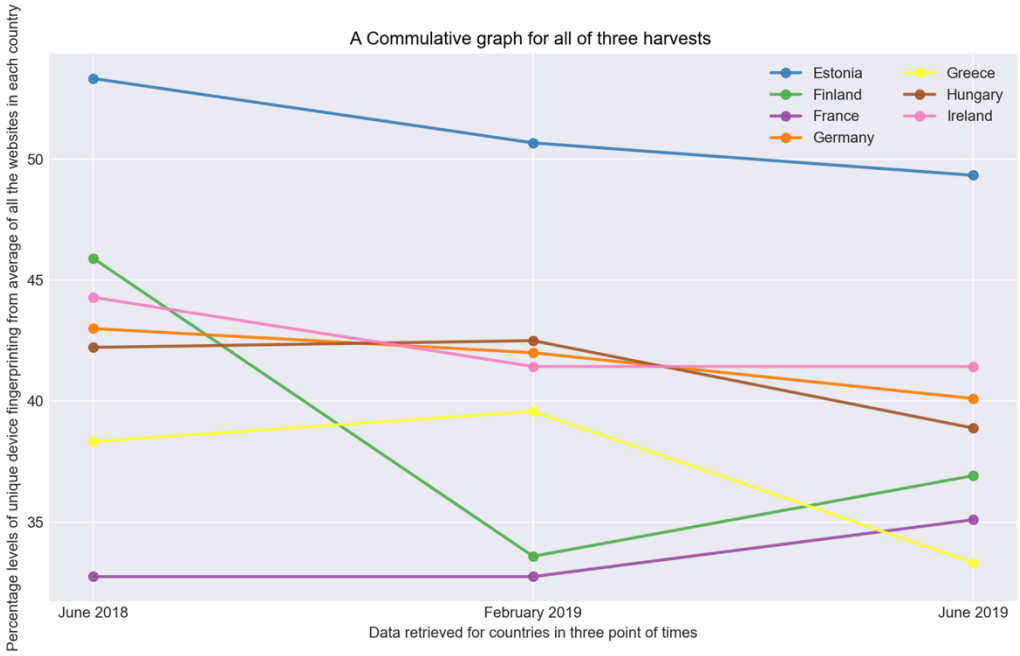


Figure 17: Transition of percentages in capturing unique Device Fingerprinting from June 2018 - February 2019 - June 2019 for the second set of seven countries in EU

For the second section of countries, the news media websites including Estonia, Finland, Ireland, and Germany decreased the levels of data capture for device fingerprinting observed in February 2019 than in June 2018. Whereas, news media websites from countries such as France, Greece and Hungary have an increased levels of device fingerprinting in February 2019.

Observing the transition from February 2019 till June 2019, it is observed that news media websites from Estonia, Germany, Greece, Hungary and Ireland have decreased level of capturing device fingerprinting data; whereas, news websites from countries like Finland and France increased.

Overall from June 2018 till June 2019, the news websites from countries such as Estonia, Finland, Germany, Greece, Hungary and Ireland (6 countries) decreased the level of device fingerprinting; whereas, France had an increased level of device fingerprinting.

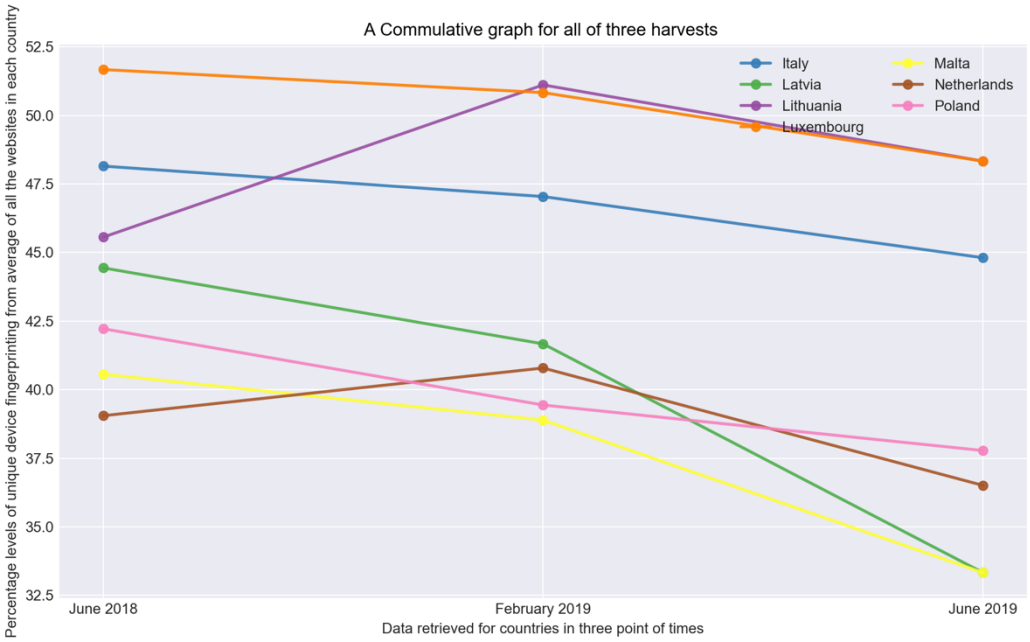


Figure 18: Transition of percentages in capturing unique Device Fingerprinting from June 2018 - February 2019 - June 2019 for the third set of seven countries in EU

In this third section of countries with Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands and Poland, it is observed that the news websites from all the countries mentioned above have decreased the level of device fingerprinting from June 2018 till February 2019 except Lithuania and Netherlands. However, when compared the results from February 2019 and June 2019, all these websites decreased the level of device fingerprinting. **Overall, all these above-mentioned websites (6 countries) decreased the level of device fingerprinting from June 2018 till June 2019 except the websites from Lithuania**.

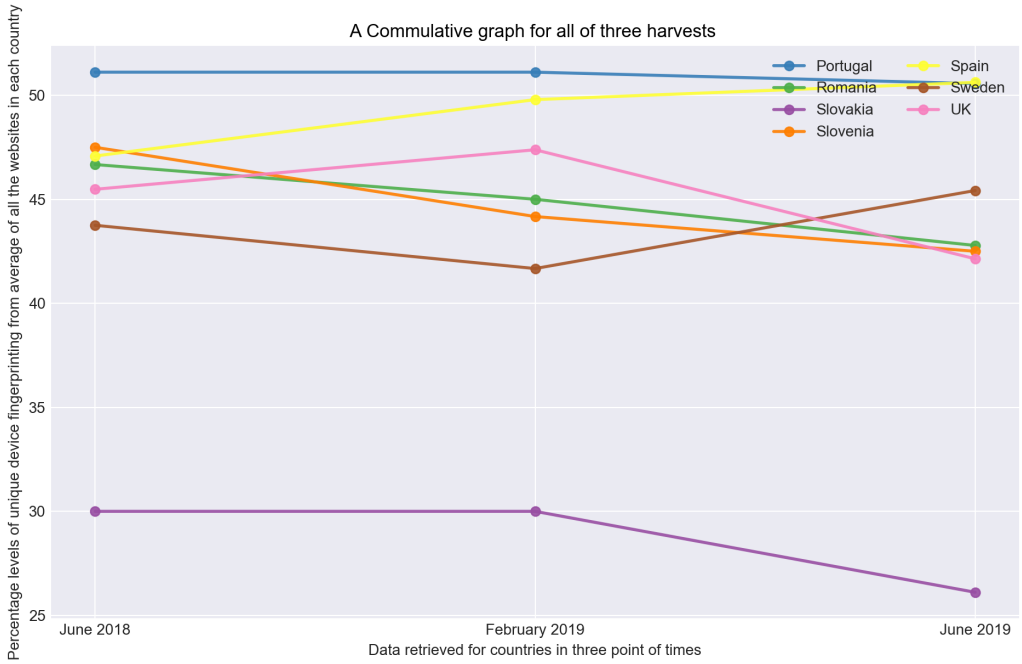


Figure 19: Transition of percentages in capturing unique Device Fingerprinting from June 2018 - February 2019 - June 2019 for the fourth and last set of seven countries in EU

In the last section of seven countries from a total of 28 countries, the news websites from countries such as Slovenia, Romania and Sweden decreased the level of fingerprinting, news websites from Slovakia and Portugal remained the same; whereas news websites from Spain and UK increased the levels of device fingerprinting from June 2018 till February 2019. However, from February 2019 till June 2019, news websites from Portugal decreased the level of device fingerprinting with a minor difference, news websites from UK, Romania, Slovenia and Slovakia decreased, whereas, the news websites in Spain and Sweden increased in the levels of data captured for device fingerprinting.

Overall, these levels decreased in all the countries (5 countries) mentioned above except for Sweden and Spain which increased from June 2018 till June 2019.

Research Sub-Question c:

What are the transitions in capturing device fingerprints over three different time periods?

Transition from June 2018 – June 2019

The data analysis shows that the news media websites from **20 countries among 28 European countries decreased the level of capturing data from audiences' devices from June 2018 to June 2019**. This data is particularly the data that contributes to make a unique device fingerprint which is explained in previous chapters. This decrease includes all the slight changes to major drops and including a rise from some countries in between, that is in February 2019.

List of Countries with Decreased Level of Device Fingerprinting:

These 20 countries include Austria, Denmark, Czech Republic, Estonia, Finland, Germany, Greece, Hungary, Ireland, Italy, Latvia, Luxembourg, Malta, Netherlands, Poland, UK, Romania, Slovenia, Slovakia and Portugal with a very minor decrease in capturing data for device Fingerprinting.

List of Countries with Increased Level of Device Fingerprinting:

Cyprus, Croatia, Belgium, Bulgaria, France, Lithuania, Sweden and Spain have increased levels of capturing data for unique device fingerprinting from June 2018 till June 2019.

Transition from June 2018 – February 2019

List of Countries with Decreased Level of Device Fingerprinting:

Cyprus, Denmark, France, Greece, Hungary, Italy, Latvia, Luxembourg, Malta, Poland, Slovenia, Romania and Sweden

List of Countries with Increased Level of Device Fingerprinting:

Belgium, Croatia, Austria, Czech Republic, Bulgaria, Estonia, Finland, Ireland, Germany, Lithuania, Netherlands, Spain, UK

List of Countries with Same Level of Device Fingerprinting:

Sweden, Portugal

Transition from February – June 2019

List of Countries with Decreased Level of Device Fingerprinting:

Austria, Belgium, Croatia, Czech Republic, Denmark, Estonia, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, UK, Romania, Slovenia and Slovakia

List of Countries with Increased Level of Device Fingerprinting:

Cyprus, Bulgaria, Finland, France, Spain and Sweden

Discussion

The motivation behind this research was to find out the new approaches by media companies in gathering audience data for audience measurement. During the initial research approach, it was found that the newest and least understood method is the device fingerprinting that includes several methods in a JavaScript file that runs on the user's/audience's device while browsing the content on a website. Therefore, the current research motivation was to find out the patterns of device fingerprinting done by the European news media websites. However, to identify the device fingerprinting patterns by the websites of news media companies in EU, a certain observation needed to be done.

This observation could be done during the current research but that could give results from a limited time period; however, the device fingerprinting patterns needed to be observed over a long period of time. Moreover, observing these news media websites involve visiting each news media website and check for the JavaScript running on those websites to see the values from audience device being captured by the JavaScript functions. Doing this for two hundred and eighty-two websites was impossible for a Master thesis in a six-month time period and the data would not give results from a greater time span to provide a general analysis and draw conclusions. Therefore, it was found out during the current research that such data has already been captured by researchers Sørensen & Kosta (2019) in June 2018, February 2019 and June 2019. This data was captured for their own research focus explained in their paper (cited and provided reference in the bibliography); however, for the current research this data was used after filtering out the relevant information required for identifying device fingerprinting.

The new GDPR was implemented introduced in May 2016 and two years were given to follow it by companies and allowing the companies to abide by the rules and regulations mentioned in it. After May 2018, several companies in EU had revised privacy policies and policies regarding data being taken and processed from the users/consumers/audiences by the companies.

Our observation in this analysis shows a post GDPR changes in news media websites specifically in gathering audience device data that contributes to create a unique device fingerprint. The data was taken in June 2018 (right after GDPR's implementation in EU), February 2019, and June 2019 to see the transitions and changes over three time periods.

Using the theory of audience commodity, the research focused on how audience's data is being gathered by news media companies through their websites because every data has an economic value and the audience generated data makes an audience a commodity more than just a consumer.

After developing the python scripts, the data was loaded in the scripts and generated four types of results discussed in detail in the Results and Analysis section.

One of these results provided insights in three sets for June 2018, February 2019 and June 2019 that showed what kind of fingerprinting type – browser fingerprinting, canvas fingerprinting and/or cookie printing is used by each news media website in each of the twenty-eight EU countries. Most of these websites use cookie printing at a maximum level, then browser fingerprinting and canvas fingerprinting is done the least.

The second result provided an insight of a cumulative taken for all the websites in each country and a percentage usage of the overall device fingerprinting by 28 countries in the form three bar-graphs each for data taken in June 2018, February 2019 and June 2019.

The third result was almost similar to the second result with three donut-pie charts represented as a table showing the percentage usage of unique device fingerprinting done in each of the 28 European countries. This result is different from the second one in terms of cumulative percentage as 100 for each of the data harvests including June 2018, February 2019, and June 2019. The data in a tabular form helps the reader in observing the rise and fall of capturing device fingerprinting for each country as a cumulative percentage taken for all countries in each harvest of June 2018, February 2019 and June 2019.

The last and most important result shows the transitions in capturing unique device fingerprinting for every country from June 2018 – February 2019 and from February 2019 – June 2019 in one single plot of line graphs.

This result of the current research showed that the overall device fingerprinting has decreased from June 2018 – June 2019 which can be seen in the results and analysis chapter under the heading "Transition from June 2018 – June 2019". However, the changes from June 2018 till February 2019 showed that there is an equal number of countries that increased or decreased the levels of capturing unique device fingerprinting. Whereas, from the analysis of the transition from February 2019 till June 2019, it is evident that more countries have decreased in the levels of capturing unique device fingerprinting. However, it is still unclear that

whether the decrease in the capture of unique device fingerprinting by the news websites in each of these 20 among 28 European countries is due to the implementation of GDPR as it is understandable that in May 2018, the new GDPR was implemented. It is because, the transitions are not uniform for each country, that is, some countries did not bring any major or easily identifiable changes such as Portugal. While some countries brought changes in increased or decreased levels of device fingerprinting from the first data harvest and the second harvest, which is also changed from the second harvest and third harvest.

Conclusion

The results of this research showed that the news media websites in EU have been using particular device fingerprinting methods such as browser fingerprinting, canvas fingerprinting and cookie fingerprinting. Overall, the levels of the data captured by these websites decreased in most of the countries from June 2018 till June 2019. To relate the results of current research with the effects of new GDPR, further research is required to carry out. Company policies are also needed to be studied and the sampling can be narrowed down to the draw conclusions specific to some countries.

Moreover, **the theory of audience commodity is supported** after the analysis of the results as these results show that even after the implementation of GDPR and regardless of decrease in the levels of capturing device fingerprints, news media websites capture the data from audience's devices. The data that is captured through device fingerprinting is to uniquely identify each user across platforms for several reasons including ensuring security. This audience data gathering makes an audience a commodity given that the data in today's digital world contributes to companies' economic value and especially if the data is given a meaning in terms of identifying the user.

The current research does not conclude in identifying the audience related data captured through device fingerprinting is used for a specific purpose by the news media companies that were included in the current study, but it concludes that the data capture from audience's devices make an audience a commodity while being a consumer of the services and news content from these media companies.

Reference List

Acar, G., Eubank, C., Englehardt, S., Juarez, M., Narayanan, A., & Diaz, C. (2014). The Web Never Forgets: Persistent Tracking Mechanisms in the Wild. *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, 674–689. https://doi.org/10.1145/2660267.2660347

Acar, G., Juarez, M., Nikiforakis, N., Diaz, C., Gürses, S., Piessens, F., & Preneel, B. (2013). FPDetective: Dusting the Web for Fingerprinters. *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security*, 1129–1140. https://doi.org/10.1145/2508859.2516674

Boddewyn, J. (2015). Advertising, Control of. In J. D. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)* (pp. 201–207). https://doi.org/10.1016/B978-0-08-097086-8.95095-5

Browser Fingerprinting: What Is It and What Should You Do About It? (n.d.). Retrieved 14 September 2019, from Pixel Privacy website: https://pixelprivacy.com/resources/browser-fingerprinting/

Budington, K. S. and B. (2018, June 19). The GDPR and Browser Fingerprinting: How It Changes the Game for the Sneakiest Web Trackers. Retrieved 13 August 2019, from Electronic Frontier Foundation website: https://www.eff.org/deeplinks/2018/06/gdpr-and-browser-fingerprinting-how-it-changes-game-sneakiest-web-trackers

Caraway, B. (2011). Audience labor in the new media environment: A Marxian revisiting of the audience commodity. *Media, Culture & Society*, *33*(5), 693–708. https://doi.org/10.1177/0163443711404463

Carpentier, N. (2008). *Democracy, Journalism and Technology: New Developments in an Enlarged Europe: The Intellectual Work of ECREA's 2008 European Media and Communication Doctoral Summer School*. Tartu University Press.

Coffey, S. (2001). Internet Audience Measurement. *Journal of Interactive Advertising*, *1*(2), 10–17. https://doi.org/10.1080/15252019.2001.10722047

Control Theory. (2014). In T. Thompson, *Encyclopedia of Health Communication*. https://doi.org/10.4135/9781483346427.n100

Curran, J., Fanton, N., & Freedman, D. (2016). *Misunderstanding the Internet*. London and New York: Routledge.

Data protection in the EU [Text]. (n.d.). Retrieved 24 September 2019, from European Commission—European Commission website: https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en

Desmond, L. C. C., Yuan, C. C., Pheng, T. C., & Lee, R. S. (2008). Identifying Unique Devices Through Wireless Fingerprinting. *Proceedings of the First ACM Conference on Wireless Network Security*, 46–55. https://doi.org/10.1145/1352533.1352542

Deuze, M. (2012). *Media Life*. Polity.

Eckersley, P. (2010). How Unique is Your Web Browser? *Proceedings of the 10th International Conference on Privacy Enhancing Technologies*, 1–18. Retrieved from http://dl.acm.org/citation.cfm?id=1881151.1881152

Fingerprint. (n.d.). Retrieved 14 September 2019, from http://fp.virpo.sk/

Fuchs, C. (2015). The Digital Labour Theory of Value and Karl Marx in the Age of Facebook, YouTube, Twitter, and Weibo. In E. Fisher & C. Fuchs (Eds.),

Reconsidering Value and Labour in the Digital Age (pp. 26–41). https://doi.org/10.1057/9781137478573_2

Green, A. (2017). Audience Measurement 5.0—Pushing the Boundaries. *Ipsos*. Presented at the GameChangers Ipsos. Retrieved from https://www.ipsos.com/en/audience-measurement-50-pushing-boundaries

Hraška, P. (2018). *Browser fingerprinting* (Comenius University, Faculty of Mathematics, Physics and Informatics). Retrieved from http://virpo.sk/browser-fingerprinting-hraska-diploma-thesis.pdf

Jenkins, H. (2006). *Convergence Culture: Where Old and New Media Collide*. New York University Press.

Kim, S. J. (2016). A repertoire approach to cross-platform media use behavior. *New Media & Society*, *18*(3), 353–372. https://doi.org/10.1177/1461444814543162

Kinney, L. (2011). Audience Measurement. In L. Swayne & M. Dodds, *Encyclopedia of Sports Management and Marketing*. https://doi.org/10.4135/9781412994156.n44

Klapper, J. T. (1960). The effects of mass communication. New York: Free Press.

Large, D., & Farmer, J. (2009). Chapter 1—Linear Broadband Distribution Systems. In D. Large & J. Farmer (Eds.), *Broadband Cable Access Networks* (pp. 1–4). https://doi.org/10.1016/B978-0-12-374401-2.00001-2

Mansell, R. (2012). *Imagining the Internet: Communication, Innovation, and Governance*. Oxford, UK: Oxford University Press.

McQuail, D. (1997). The Audience in Communication Theory and Research. In *Audience Analysis* (pp. 12–24). https://doi.org/10.4135/9781452233406

Mowery, K., & Shacham, H. (2012). *Pixel Perfect: Fingerprinting Canvas in HTML 5*.

Napoli, P. (2011). *Audience Evolution: New Technologies and the Transformation of Media Audiences*. New York: Columbia University Press.

Negroponte, N. (1996). *Being Digital* (1 edition). New York: Vintage.

Odorume, A. (2012). Historiography of the Print Media: A global-cum-Nigerian perspective. *Mgbakoigba: Journal of African Studies*, *1*(0). Retrieved from https://www.ajol.info/index.php/mjas/article/view/117186

Portilla, I. (2015). Television Audience Measurement: Proposals of the Industry in the Era of Digitalization. *Trípodos*, *0*(36), 75-92–92. Retrieved from http://www.tripodos.com/index.php/Facultat_Comunicacio_Blanquerna/article/view/243

Ring, T. (2015). Keeping tabs on tracking technology. *Network Security*, *2015*(6), 5–8. https://doi.org/10.1016/S1353-4858(15)30047-7

Smith, H. (2015). *From bunny ears to smart phones: The development of broadcast technology and policy, audience viewing trends and measurement methods throughout the history of television in canada*. Retrieved from https://trepo.tuni.fi/handle/10024/97985

Sørensen, J. K., & Kosta, S. (2019). Before and After GDPR: The Changes in Third Party Presence at Public and Private European Websites. *The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019*, 1590–1600. https://doi.org/10.1145/3308558.3313524

Testori, M. (2014). The Applications of Video Analytics in Media Planning, Trade and Shopper Marketing. In C. Distante, S. Battiato, & A. Cavallaro (Eds.), *Video Analytics for Audience Measurement* (pp. 3–20). Springer International Publishing.

Van Goethem, T., Scheepers, W., Preuveneers, D., & Joosen, W. (2016). Accelerometer-Based Device Fingerprinting for Multi-factor Mobile Authentication. In J. Caballero, E. Bodden, & E. Athanasopoulos (Eds.), *Engineering Secure Software and Systems* (pp. 106–121). Springer International Publishing.

Vastel, A., Laperdrix, P., Rudametkin, W., & Rouvoy, R. (2018). *Fp-Scanner: The Privacy Implications of Browser Fingerprint Inconsistencies*. 135–150. Retrieved from https://www.usenix.org/conference/usenixsecurity18/presentation/vastel

Whalley, A. J. (2011). Advertising. In *Encyclopedia of Sports Management and Marketing* (Vols 1–4, pp. 10–16). https://doi.org/10.4135/9781412994156

What is personal data? [Text]. (n.d.). Retrieved 24 September 2019, from European Commission—European Commission website: https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_en

Wherry, F. F., & Schor, J. B. (2015). Television. In *The SAGE Encyclopedia of Economics and Society*. https://doi.org/10.4135/9781452206905.n670

Whitney, D., Sumpter, R., & McQuail, D. (2004). News Media Production: Individuals, Organizations, and Institutions. In *The SAGE Handbook of Media Studies* (pp. 393–410). https://doi.org/10.4135/9781412976077

Appendix A: Python Scripts

myscript1.py

```
import os
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

# All the top level domains coming from csv file
df = pd.read_csv('./News-in-EU-PublicPrivate.csv')
# Uncomment each of the following three lines to generate the results for each
harvest
# df_harvest1 = pd.read_csv('../june2018.csv', encoding = "ISO-8859-1")
# df_harvest1 = pd.read_csv('../feb2019.csv', encoding = "ISO-8859-1")
df_harvest1 = pd.read_csv('../june2019.csv', encoding = "ISO-8859-1")

browser_finger_printing_variables = {
                  'window.navigator.userAgent': 20,'window.sessionStorage': 10,'win
dow.navigator.platform':10,
                  'window.navigator.language': 10,'window.localStorage': 10,'window
.navigator.plugins': 20,
                  'window.navigator.doNotTrack': 10,'window.navigator.cookieEnabled
': 10,
}

canvas_finger_printing_variables = {
                    'window.screen.colorDepth': 10 ,'window.screen.pixelDepth': 10,
 'HTMLCanvasElement': 20, 'CanvasRenderingContext2D': 60,
}

cookie_finger_printing_variables = {'window.document.cookie': 100}

# countires dictionary
countries = {}
domains = []
for index, row in df.iterrows():
    country = row['Country']
    top_level_domain = row['TopLevelDomainLookUp']

    matching = df_harvest1[df_harvest1['script_url'].str.contains(top_level_domain)
 == True]
    # Browser finger printing matching and preparing data
    b_f_p_rows = {}
    b_f_p_percentage = 0
    for key, value in browser_finger_printing_variables.items():
```

```
        rows = matching[matching['symbol'].str.contains(key) == True]
        if rows.empty:
            b_f_p_rows[key]= 'No'
        else:
            b_f_p_rows[key]= 'Yes'
            b_f_p_percentage +=value
    # Canvas finger printing matching and preparing data
    c_f_p_rows = {}
    c_f_p_percentage = 0
    for key, value in canvas_finger_printing_variables.items():
        rows = matching[matching['symbol'].str.contains(key) == True]
        if rows.empty:
            c_f_p_rows[key]= 'No'
        else:
            c_f_p_rows[key]= 'Yes'
            c_f_p_percentage +=value
    # Cookies finger printing matching and preparing data
    c_p_rows = {}
    c_p_percentage = 0
    for  key, value in cookie_finger_printing_variables.items():
        rows = matching[matching['symbol'].str.contains(key) == True]
        if rows.empty:
            c_p_rows[key]= 'No'
        else:
            c_p_rows[key]= 'Yes'
            c_p_percentage +=value

    if country in countries:
        countries[country].append({top_level_domain: {'b_f_p': b_f_p_percentage, 'c
_f_p': c_f_p_percentage, 'c_p': c_p_percentage } })
    else:
        countries[country] = []
        countries[country].append({top_level_domain: {'b_f_p': b_f_p_percentage, 'c
_f_p': c_f_p_percentage, 'c_p': c_p_percentage } })

for country_key, domains in countries.items():
    n_groups = len(domains)
    country_bfps = []
    country_cfps = []
    country_cps = []
    # create plot
    fig, ax = plt.subplots()
    index = np.arange(n_groups)
    bar_width = 0.15
    opacity = 0.8
    domain_names = []
    # preparing bars data
```

```
    for item in domains:
        for domain_name, domain in item.items():
            country_bfps.append(domain['b_f_p'])
            country_cfps.append(domain['c_f_p'])
            country_cps.append(domain['c_p'])
            domain_names.append(domain_name)
    rects1 = plt.bar(index, country_bfps, bar_width, alpha=opacity, color='b', label='Browser Fingerprinting')

    rects2 = plt.bar(index + bar_width, country_cfps, bar_width, alpha=opacity, color='g', label='Canvas Fingerprinting')

    rects3 = plt.bar(index + bar_width + bar_width, country_cps, bar_width, alpha=opacity, color='r', label='Cookie Printing')

    plt.xlabel('News Media Websites in '+country_key)
    plt.ylabel('Percentage')
    plt.title('Percentage by News Media Websites in '+country_key+ ' - June 2019 Data')
    plt.xticks(index + bar_width, domain_names)
    plt.xticks(rotation=80)
    plt.legend(loc=1)
    plt.tight_layout()
    plt.show()
```

comp_countries_pie.py

```
import os
import pandas as pd

import numpy as np
import matplotlib.pyplot as plt
import statistics

# All the top level domains coming from csv file
df = pd.read_csv('./News-in-EU-PublicPrivate.csv')
# Uncomment each of the following three lines to generate the results for each harvest
# df_harvest1 = pd.read_csv('../june2018.csv', encoding = "ISO-8859-1")
# df_harvest1 = pd.read_csv('../feb2019.csv', encoding = "ISO-8859-1")
df_harvest1 = pd.read_csv('../june2019.csv', encoding = "ISO-8859-1")

browser_finger_printing_variables = {
                  'window.navigator.userAgent': 20,'window.sessionStorage': 10,'window.navigator.platform':10,
```

```
                  'window.navigator.language': 10,'window.localStorage': 10,'window
.navigator.plugins': 20,
                  'window.navigator.doNotTrack': 10,'window.navigator.cookieEnabled
': 10,
}

canvas_finger_printing_variables = {
                   'window.screen.colorDepth': 10 ,'window.screen.pixelDepth': 10,
 'HTMLCanvasElement': 20, 'CanvasRenderingContext2D': 60,
}

cookie_finger_printing_variables = {'window.document.cookie': 100}

# countires dictionary
countries = {}
domains = []
for index, row in df.iterrows():
    country = row['Country']
    top_level_domain = row['TopLevelDomainLookUp']

    matching = df_harvest1[df_harvest1['script_url'].str.contains(top_level_domain)
 == True]

    # Browser finger printing matching and preparing data
    b_f_p_rows = {}
    b_f_p_percentage = 0
    for key, value in browser_finger_printing_variables.items():
        rows = matching[matching['symbol'].str.contains(key) == True]
        if rows.empty:
            b_f_p_rows[key]= 'No'
        else:
            b_f_p_rows[key]= 'Yes'
            b_f_p_percentage +=value
    # Canvas finger printing matching and preparing data
    c_f_p_rows = {}
    c_f_p_percentage = 0
    for key, value in canvas_finger_printing_variables.items():
        rows = matching[matching['symbol'].str.contains(key) == True]
        if rows.empty:
            c_f_p_rows[key]= 'No'
        else:
            c_f_p_rows[key]= 'Yes'
            c_f_p_percentage +=value
    # Cookies finger printing matching and preparing data
    c_p_rows = {}
    c_p_percentage = 0
    for  key, value in cookie_finger_printing_variables.items():
```

```
        rows = matching[matching['symbol'].str.contains(key) == True]
        if rows.empty:
            c_p_rows[key]= 'No'
        else:
            c_p_rows[key]= 'Yes'
            c_p_percentage +=value

    if country in countries:
        countries[country].append({top_level_domain: {'b_f_p': b_f_p_percentage, 'c
_f_p': c_f_p_percentage, 'c_p': c_p_percentage } })
    else:
        countries[country] = []
        countries[country].append({top_level_domain: {'b_f_p': b_f_p_percentage, 'c
_f_p': c_f_p_percentage, 'c_p': c_p_percentage } })
country_list = []
d_f_p_values = []

for country_key, domains in countries.items():
    country_bfps = []
    country_cfps = []
    country_cps = []
    d_f_p_array = []
    domain_names = []
    d_f_p_array = []
    # preparing pie data
    for item in domains:
        for domain_name, domain in item.items():
            country_bfps.append(domain['b_f_p'])
            country_cfps.append(domain['c_f_p'])
            country_cps.append(domain['c_p'])
            domain_names.append(domain_name)

            d_f_p_array.append(country_bfps)
            d_f_p_array.append(country_cfps)
            d_f_p_array.append(country_cps)
    country_list.append(country_key)
    d_f_p_values.append( round(np.mean(d_f_p_array), 2))


# Pie chart, where the slices will be ordered and plotted counter-clockwise:
labels = country_list
fig1, ax1 = plt.subplots()
patches, texts, autotexts = plt.pie(d_f_p_values, labels = country_list, autopct='%
1.1f%%', pctdistance=0.85, startangle=90)

plt.setp(autotexts, size=8, weight="bold")
```

```
centre_circle = plt.Circle((0,0),0.70,fc='white')
fig = plt.gcf()
fig.gca().add_artist(centre_circle)
plt.tight_layout()

plt.title('Comparison of EU countries capturing device fingerprinting - Data from J
une 2019')
ax1.axis('equal')
plt.show()
```

comp_countries_bar.py

```
import os
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import statistics

# All the top level domains coming from csv file
df = pd.read_csv('./News-in-EU-PublicPrivate.csv')
Uncomment each of the following three lines to generate the results for each
harvest
# df_harvest1 = pd.read_csv('../june2018.csv', encoding = "ISO-8859-1")
df_harvest1 = pd.read_csv('../feb2019.csv', encoding = "ISO-8859-1")
# df_harvest1 = pd.read_csv('../june2019.csv', encoding = "ISO-8859-1")

browser_finger_printing_variables = {
                  'window.navigator.userAgent': 20,'window.sessionStorage': 10,'win
dow.navigator.platform':10,
                  'window.navigator.language': 10,'window.localStorage': 10,'window
.navigator.plugins': 20,
                  'window.navigator.doNotTrack': 10,'window.navigator.cookieEnabled
': 10,
}

canvas_finger_printing_variables = {
                    'window.screen.colorDepth': 10 ,'window.screen.pixelDepth': 10,
 'HTMLCanvasElement': 20, 'CanvasRenderingContext2D': 60,
}

cookie_finger_printing_variables = {'window.document.cookie': 100}
# countires dictionary
countries = {}
domains = []
for index, row in df.iterrows():
```

```
    country = row['Country']
    top_level_domain = row['TopLevelDomainLookUp']

    matching = df_harvest1[df_harvest1['script_url'].str.contains(top_level_domain)
 == True]
    # Browser finger printing matching and preparing data
    b_f_p_rows = {}
    b_f_p_percentage = 0
    for key, value in browser_finger_printing_variables.items():
        rows = matching[matching['symbol'].str.contains(key) == True]
        if rows.empty:
            b_f_p_rows[key]= 'No'
        else:
            b_f_p_rows[key]= 'Yes'
            b_f_p_percentage +=value
    # Canvas finger printing matching and preparing data
    c_f_p_rows = {}
    c_f_p_percentage = 0
    for key, value in canvas_finger_printing_variables.items():
        rows = matching[matching['symbol'].str.contains(key) == True]
        if rows.empty:
            c_f_p_rows[key]= 'No'
        else:
            c_f_p_rows[key]= 'Yes'
            c_f_p_percentage +=value
    # Cookies finger printing matching and preparing data
    c_p_rows = {}
    c_p_percentage = 0
    for  key, value in cookie_finger_printing_variables.items():
        rows = matching[matching['symbol'].str.contains(key) == True]
        if rows.empty:
            c_p_rows[key]= 'No'
        else:
            c_p_rows[key]= 'Yes'
            c_p_percentage +=value

    if country in countries:
        countries[country].append({top_level_domain: {'b_f_p': b_f_p_percentage, 'c
_f_p': c_f_p_percentage, 'c_p': c_p_percentage } })
    else:
        countries[country] = []
        countries[country].append({top_level_domain: {'b_f_p': b_f_p_percentage, 'c
_f_p': c_f_p_percentage, 'c_p': c_p_percentage } })


country_list = []
d_f_p_values = []
```

```
for country_key, domains in countries.items():
    country_bfps = []
    country_cfps = []
    country_cps = []
    d_f_p_array = []
    domain_names = []
    d_f_p_array = []
    # preparing bars data
    for item in domains:
        for domain_name, domain in item.items():
            country_bfps.append(domain['b_f_p'])
            country_cfps.append(domain['c_f_p'])
            country_cps.append(domain['c_p'])
            domain_names.append(domain_name)

            d_f_p_array.append(country_bfps)
            d_f_p_array.append(country_cfps)
            d_f_p_array.append(country_cps)
    country_list.append(country_key)
    d_f_p_values.append( round(np.mean(d_f_p_array), 2))

y_pos = np.arange(len(d_f_p_values))
# Create horizontal bars
plt.barh(y_pos, d_f_p_values)
# Create names on the y-axis
plt.yticks(y_pos, country_list)

plt.xlabel('Percentage use of features that could make a unique Device Fingerprinti
ng')
plt.ylabel('Countries capturing device fingerprinting data')
plt.title('EU countries capturing device fingerprinting features in percentages - D
ata from Feb 2019')

# Show graphic
plt.show()
```

collective.py

```
import os
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

df_harvest1 = pd.read_csv('../june2018.csv', encoding = "ISO-8859-1")
df_harvest2 = pd.read_csv('../feb2019.csv', encoding = "ISO-8859-1")
```

```
df_harvest3 = pd.read_csv('../june2019.csv', encoding = "ISO-8859-1")
def prepareData(df_harvest):
    # All the top level domains coming from csv file
    df = pd.read_csv('./News-in-EU-PublicPrivate.csv')
    browser_finger_printing_variables = {
                    'window.navigator.userAgent': 20,'window.sessionStorage': 10,'w
indow.navigator.platform':10,
                    'window.navigator.language': 10,'window.localStorage': 10,'wind
ow.navigator.plugins': 20,
                    'window.navigator.doNotTrack': 10,'window.navigator.cookieEnabl
ed': 10,
    }

    canvas_finger_printing_variables = {
                        'window.screen.colorDepth': 10 ,'window.screen.pixelDepth':
 10, 'HTMLCanvasElement': 20, 'CanvasRenderingContext2D': 60,
    }

    cookie_finger_printing_variables = {'window.document.cookie': 100}

    # countires dictionary
    countries = {}
    domains = []
    for index, row in df.iterrows():
        country = row['Country']
        top_level_domain = row['TopLevelDomainLookUp']

        matching = df_harvest[df_harvest['script_url'].str.contains(top_level_domai
n) == True]
        # Browser finger printing matching and preparing data
        b_f_p_rows = {}
        b_f_p_percentage = 0
        for key, value in browser_finger_printing_variables.items():
            rows = matching[matching['symbol'].str.contains(key) == True]
            if rows.empty:
                b_f_p_rows[key]= 'No'
            else:
                b_f_p_rows[key]= 'Yes'
                b_f_p_percentage +=value
        # Canvas finger printing matching and preparing data
        c_f_p_rows = {}
        c_f_p_percentage = 0
        for key, value in canvas_finger_printing_variables.items():
            rows = matching[matching['symbol'].str.contains(key) == True]
            if rows.empty:
                c_f_p_rows[key]= 'No'
            else:
```

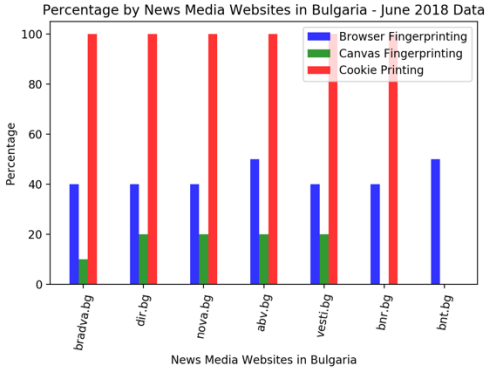
```
                c_f_p_rows[key]= 'Yes'
                c_f_p_percentage +=value
        # Cookies finger printing matching and preparing data
        c_p_rows = {}
        c_p_percentage = 0
        for  key, value in cookie_finger_printing_variables.items():
            rows = matching[matching['symbol'].str.contains(key) == True]
            if rows.empty:
                c_p_rows[key]= 'No'
            else:
                c_p_rows[key]= 'Yes'
                c_p_percentage +=value

        if country in countries:
            countries[country].append({top_level_domain: {'b_f_p': b_f_p_percentage
, 'c_f_p': c_f_p_percentage, 'c_p': c_p_percentage } })
        else:
            countries[country] = []
            countries[country].append({top_level_domain: {'b_f_p': b_f_p_percentage
, 'c_f_p': c_f_p_percentage, 'c_p': c_p_percentage } })
    country_list = []
    d_f_p_values = []

    for country_key, domains in countries.items():

        country_bfps = []
        country_cfps = []
        country_cps = []
        d_f_p_array = []
        domain_names = []
        # preparing bars data
        for item in domains:
            for domain_name, domain in item.items():
                country_bfps.append(domain['b_f_p'])
                country_cfps.append(domain['c_f_p'])
                country_cps.append(domain['c_p'])
                domain_names.append(domain_name)

                d_f_p_array.append(country_bfps)
                d_f_p_array.append(country_cfps)
                d_f_p_array.append(country_cps)
        country_list.append(country_key)
        d_f_p_values.append( round(np.mean(d_f_p_array), 2))

    return {'country_list': country_list, 'dfp_values': d_f_p_values}
harvests = ['June 2018', 'February 2019', 'June 2019']
x_axis_len = len(harvests)
```

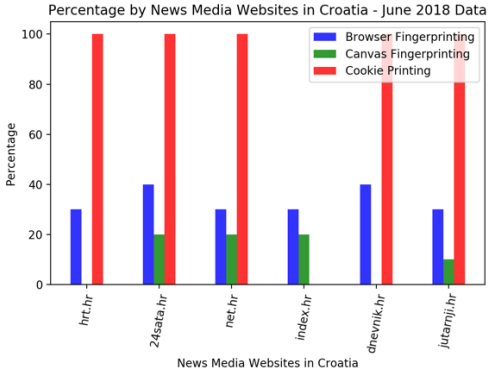
```
#result for harvest 1
result1=prepareData(df_harvest1)

#result for harvest 2
result2=prepareData(df_harvest2)

#result for harvest 3
result3=prepareData(df_harvest3)

all_countries1 = result1['country_list']
all_values1 = result1['dfp_values']
all_values2 = result2['dfp_values']
all_values3 = result3['dfp_values']
# print(len(all_values1))
harvests_dict = {}

for i in range(len(all_values1)):
    all_values = []
    all_values.append(all_values1[i])
    all_values.append(all_values2[i])
    all_values.append(all_values3[i])
    harvests_dict[all_countries1[i]] = all_values
# print (harvests_dict)
# print('^^^^^^^^^^^^^^^^^^^^^^^^^^^^^^')
values_length = len(harvests)
h_length = values_length+1

df1=pd.DataFrame({'x': range(1, h_length), 'Austria':harvests_dict['Austria'], 'Bel
gium':harvests_dict['Belgium'],
                'Bulgaria':harvests_dict['Bulgaria'], 'Croatia':harvests_dict['Croa
tia'], 'Cyprus':harvests_dict['Cyprus'],
                'Czech Republic':harvests_dict['Czech Republic'], 'Denmark':harvest
s_dict['Denmark']})

df2=pd.DataFrame({'x': range(1, h_length), 'Estonia':harvests_dict['Estonia'],'Finl
and':harvests_dict['Finland'],
                'France':harvests_dict['France'], 'Germany':harvests_dict['Germany'
],'Greece':harvests_dict['Greece'],
                'Hungary':harvests_dict['Hungary'], 'Ireland':harvests_dict['Irelan
d']})

df3=pd.DataFrame({'x': range(1, h_length), 'Italy':harvests_dict['Italy'], 'Latvia'
:harvests_dict['Latvia'],
                'Lithuania':harvests_dict['Lithuania'], 'Luxembourg':harvests_dict[
'Luxembourg'], 'Malta':harvests_dict['Malta'],
```

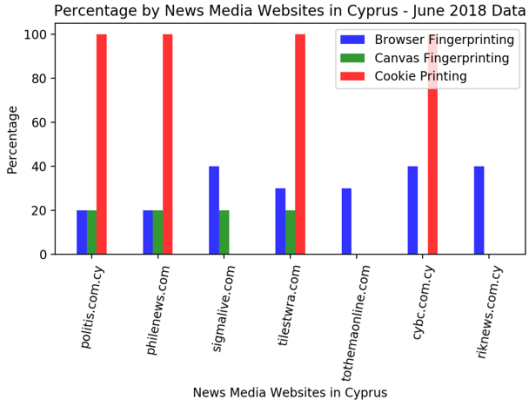
```
                'Netherlands':harvests_dict['Netherlands'], 'Poland':harvests_dict[
'Poland']})

df4=pd.DataFrame({'x': range(1, h_length), 'Portugal':harvests_dict['Portugal'], 'R
omania':harvests_dict['Romania'],
                'Slovakia': harvests_dict['Slovakia'], 'Slovenia': harvests_dict['S
lovenia'],
                'Spain':harvests_dict['Spain'], 'Sweden':harvests_dict['Sweden'], '
UK':harvests_dict['UK']})

def create_plot(df):
    # style
    plt.style.use('seaborn-darkgrid')

    # create a color palette
    palette = plt.get_cmap('Set1')

    # multiple line plot
    num=0
    for column in df.drop('x', axis=1):

        num+=1
        plt.plot(harvests, df[column], marker="o", color=palette(num), linewidth=2,
 alpha=0.9, label=column)

        # Add legend
        plt.legend(loc=1, ncol=2)

        # Add titles
        plt.title("A Commulative graph for all of three harvests ", loc='center', f
ontsize=12, fontweight=0, color='black')
        plt.xlabel("Data retrieved for countries in three point of times")
        plt.ylabel("Percentage levels of unique device fingerprinting from average
of all the websites in each country")

    plt.show()

create_plot(df1)        # create plot for first set of  7 countries in EU
create_plot(df2)        # create plot for second set of  7 countries in EU
create_plot(df3)        # create plot for third set of  7 countries in EU
create_plot(df4)        # create plot for last set of 7 countries in EU
```

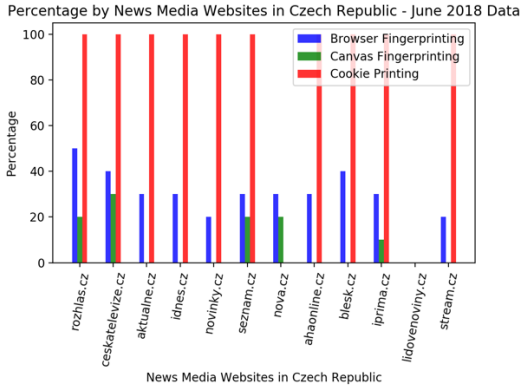
Appendix B: Results of 3 types of Device Fingerprinting data from each country in EU

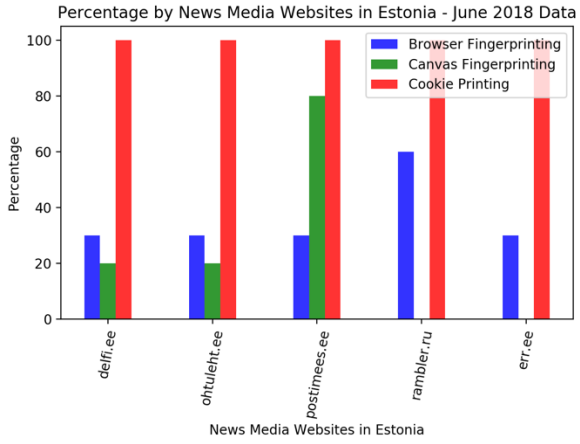
Results From June 2018

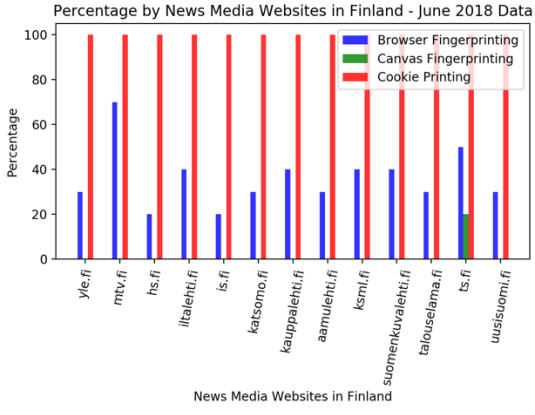


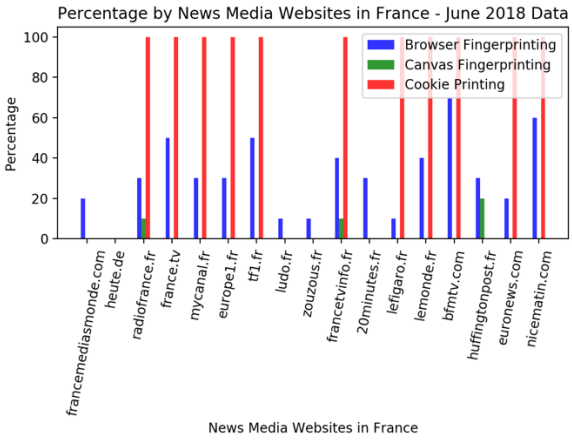


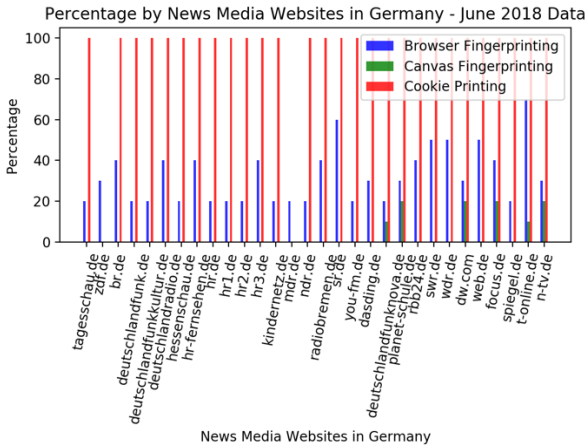


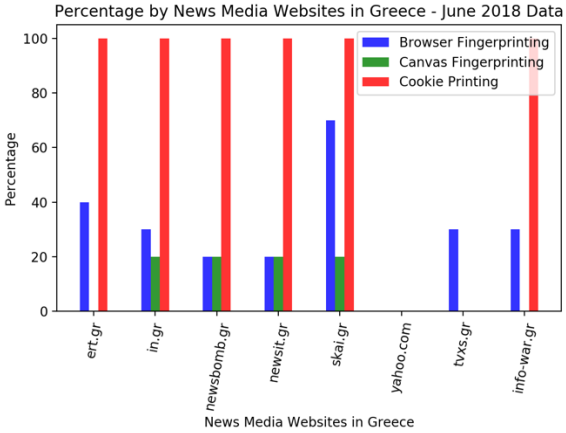


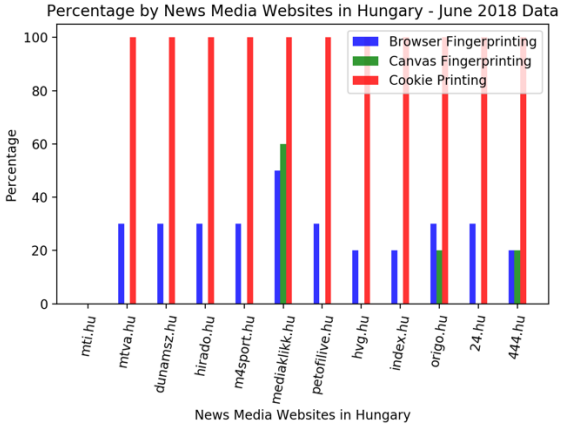


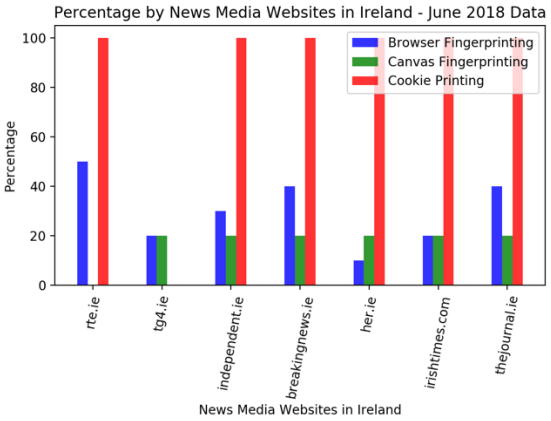


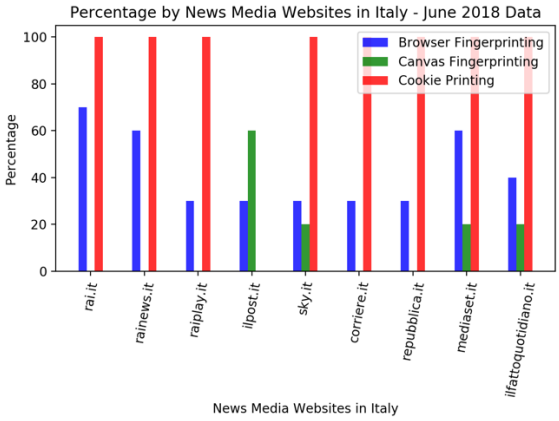

Percentage by News Media Websites in France - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
francemediasmonde.com
heute.de
radiofrance.fr
france.tv
mycanal.fr
europe1.fr
tf1.fr
ludo.fr
zouzous.fr
francetvinfo.fr
20minutes.fr
lefigaro.fr
lemonde.fr
bfmtv.com
huffingtonpost.fr
euronews.com
nicematin.com
News Media Websites in France

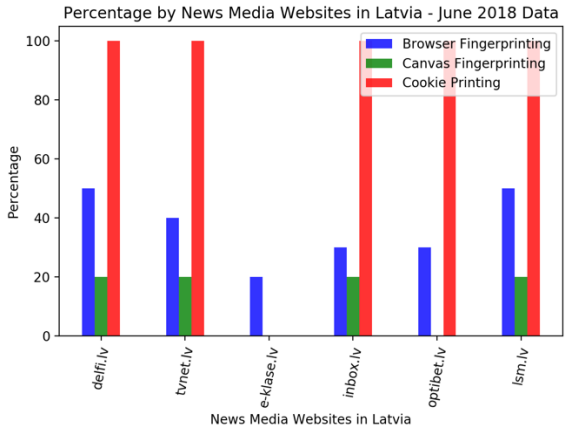

Percentage by News Media Websites in Germany - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
tagesschau.de
zdf.de
br.de
deutschlandfunk.de
deutschlandfunkkultur.de
deutschlandradio.de
hessenschau.de
hr-fernsehen.de
hr.de
hr1.de
hr2.de
hr3.de
kindernetz.de
mdr.de
ndr.de
radiobremen.de
sr.de
you-fm.de
dasding.de
deutschlandfunknova.de
planet-schule.de
rbb24.de
swr.de
wdr.de
dw.com
web.de
focus.de
spiegel.de
t-online.de
n-tv.de
News Media Websites in Germany

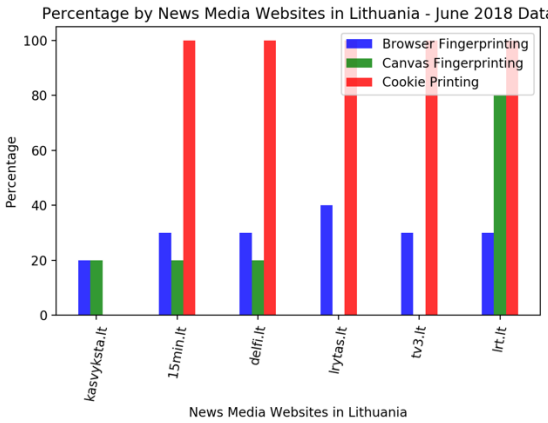

Percentage by News Media Websites in Greece - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
ert.gr
in.gr
newsbomb.gr
newsit.gr
skai.gr
yahoo.com
tvxs.gr
info-war.gr
News Media Websites in Greece

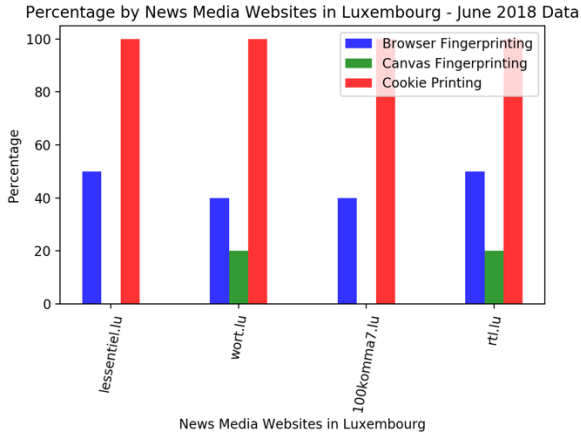

Percentage by News Media Websites in Hungary - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
mti.hu
mtva.hu
dunamsz.hu
hirado.hu
m4sport.hu
mediaklikk.hu
petofilive.hu
hvg.hu
index.hu
origo.hu
24.hu
444.hu
News Media Websites in Hungary

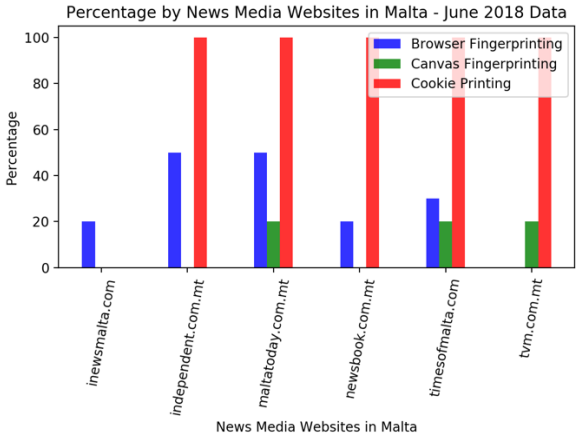

Percentage by News Media Websites in Ireland - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
rte.ie
tg4.ie
independent.ie
breakingnews.ie
her.ie
irishtimes.com
thejournal.ie
News Media Websites in Ireland

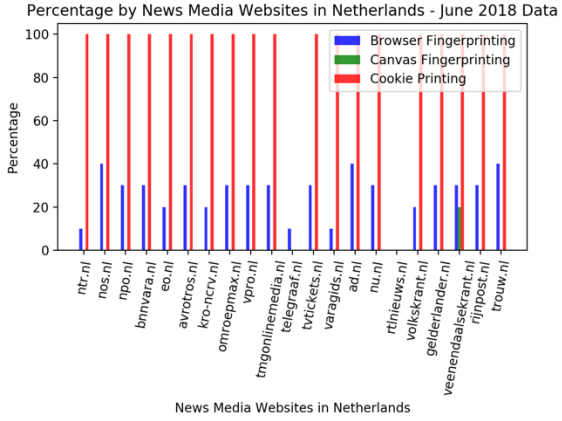

Percentage by News Media Websites in Italy - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
rai.it
rainews.it
raiplay.it
ilpost.it
sky.it
corriere.it
repubblica.it
mediaset.it
ilfattoquotidiano.it
News Media Websites in Italy

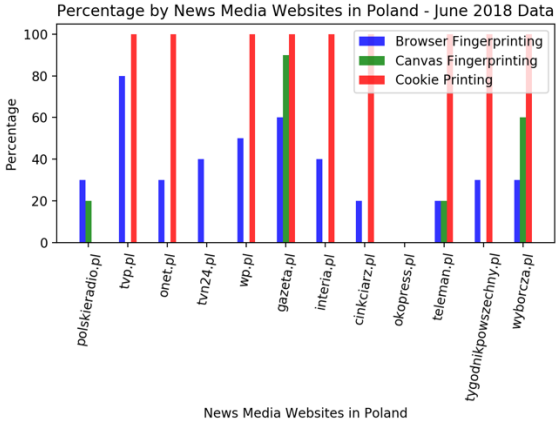

Percentage by News Media Websites in Latvia - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
delfi.lv
tvnet.lv
e-klase.lv
inbox.lv
optibet.lv
lsm.lv
News Media Websites in Latvia

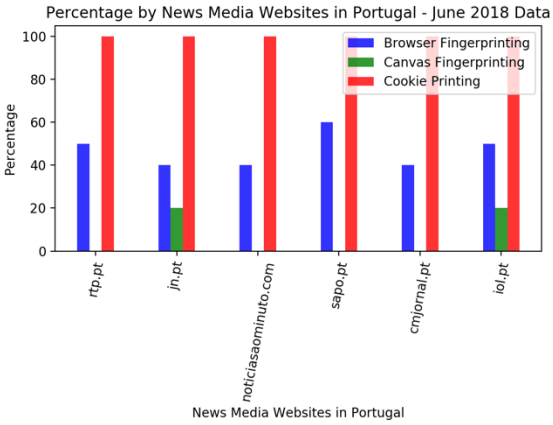

Percentage by News Media Websites in Lithuania - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
kasvyksta.lt
15min.lt
delfi.lt
lrytas.lt
tv3.lt
lrt.lt
News Media Websites in Lithuania

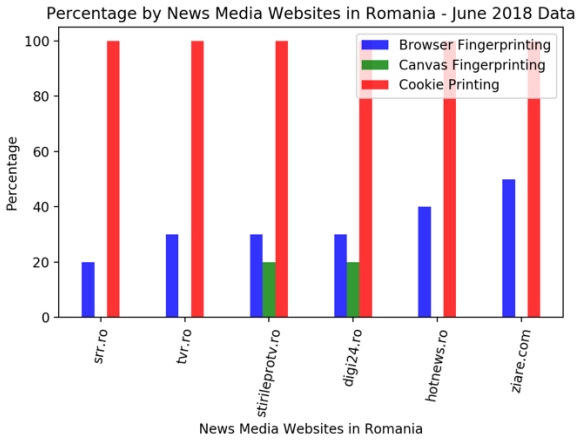

Percentage by News Media Websites in Luxembourg - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
lessentiel.lu
wort.lu
100komma7.lu
rtl.lu
News Media Websites in Luxembourg

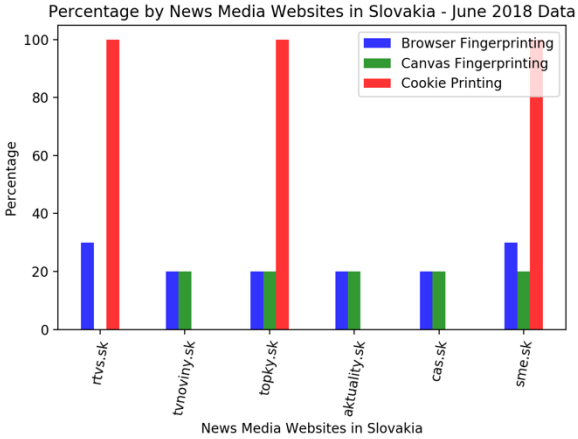

Percentage by News Media Websites in Malta - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
inewsmalta.com
independent.com.mt
maltatoday.com.mt
newsbook.com.mt
timesofmalta.com
tvm.com.mt
News Media Websites in Malta

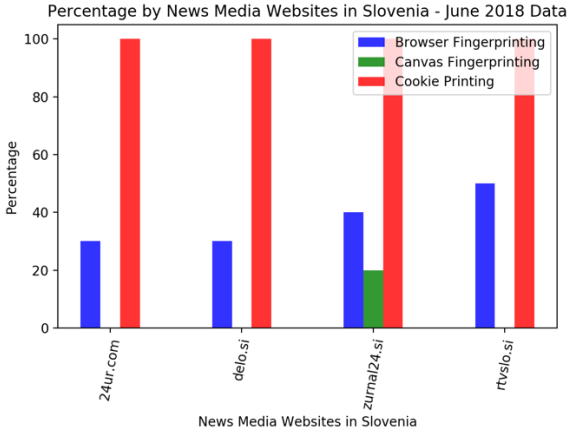

Percentage by News Media Websites in Netherlands - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
ntr.nl
nos.nl
npo.nl
bnnvara.nl
eo.nl
avrotros.nl
kro-ncrv.nl
omroepmax.nl
vpro.nl
tmgonlinemedia.nl
telegraaf.nl
tvtickets.nl
varagids.nl
ad.nl
nu.nl
rtlnieuws.nl
volkskrant.nl
gelderlander.nl
veenendaalsekrant.nl
rijnpost.nl
trouw.nl
News Media Websites in Netherlands

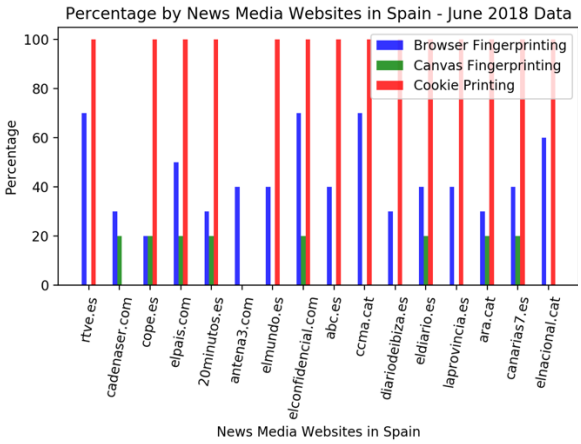

Percentage by News Media Websites in Poland - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
polskieradio.pl
tvp.pl
onet.pl
tvn24.pl
wp.pl
gazeta.pl
interia.pl
cinkciarz.pl
okopress.pl
teleman.pl
tygodnikpowszechny.pl
wyborcza.pl
News Media Websites in Poland

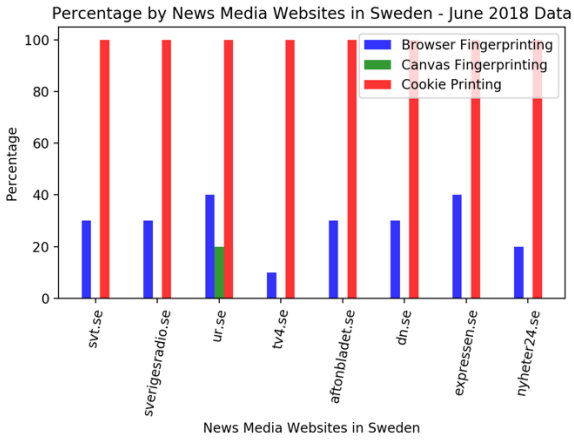

Percentage by News Media Websites in Portugal - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
rtp.pt
jn.pt
noticiasaominuto.com
sapo.pt
cmjornal.pt
iol.pt
News Media Websites in Portugal

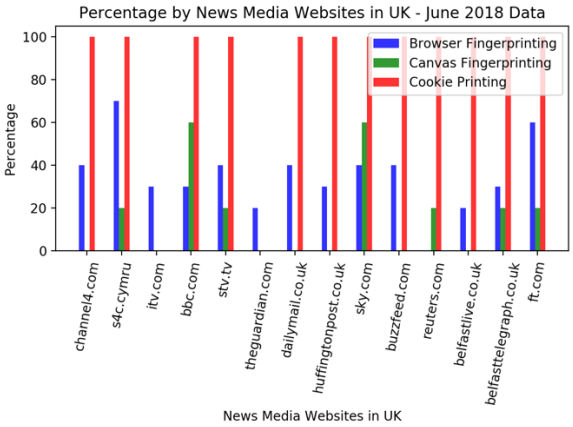

Percentage by News Media Websites in Romania - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
srr.ro
tvr.ro
stirileprotv.ro
digi24.ro
hotnews.ro
ziare.com
News Media Websites in Romania


Percentage by News Media Websites in Slovakia - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
rtvs.sk
tvnoviny.sk
topky.sk
aktuality.sk
cas.sk
sme.sk
News Media Websites in Slovakia

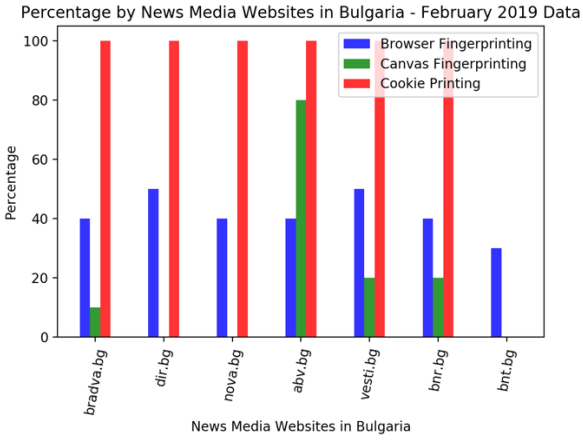

Percentage by News Media Websites in Slovenia - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
24ur.com
delo.si
zurnal24.si
rtvslo.si
News Media Websites in Slovenia

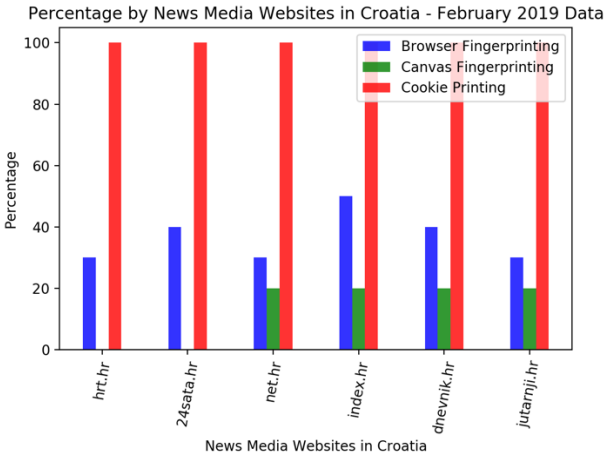

Percentage by News Media Websites in Spain - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
rtve.es
cadenaser.com
cope.es
elpais.com
20minutos.es
antena3.com
elmundo.es
elconfidencial.com
abc.es
ccma.cat
diariodeibiza.es
eldiario.es
laprovincia.es
ara.cat
canarias7.es
elnacional.cat
News Media Websites in Spain

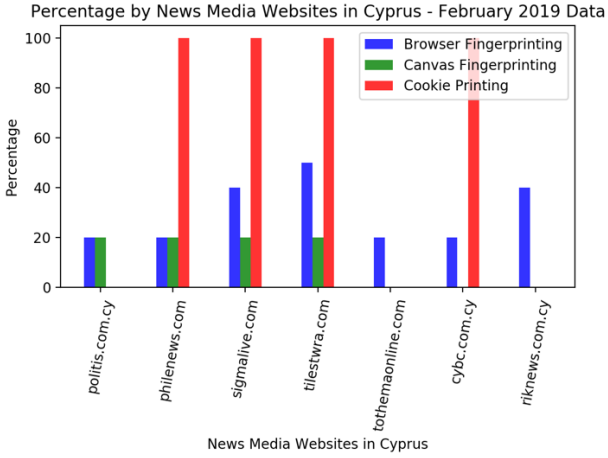

Percentage by News Media Websites in Sweden - June 2018 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
svt.se
sverigesradio.se
ur.se
tv4.se
aftonbladet.se
dn.se
expressen.se
nyheter24.se
News Media Websites in Sweden

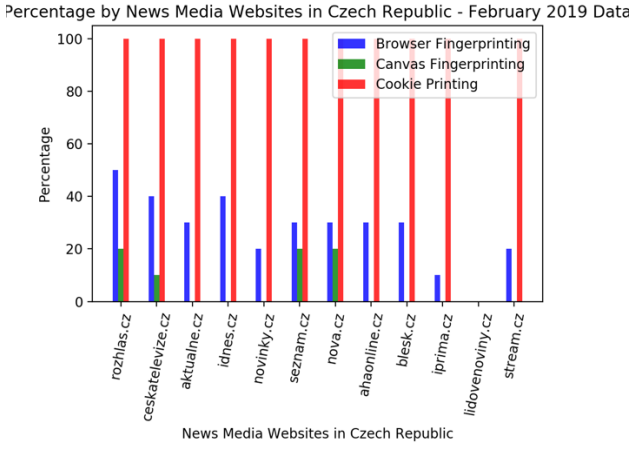


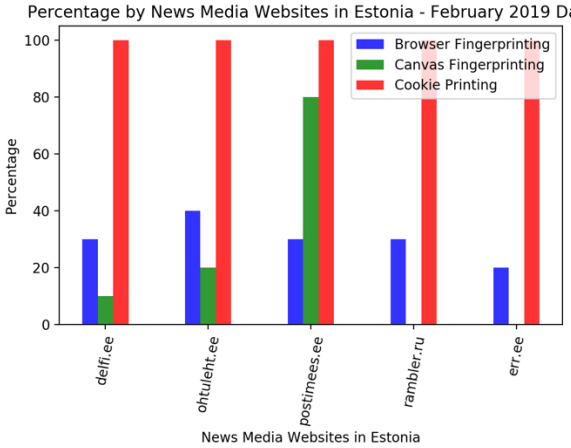
Results from February 2019

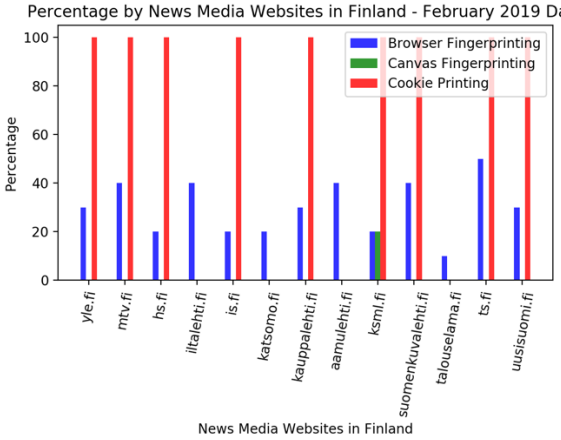


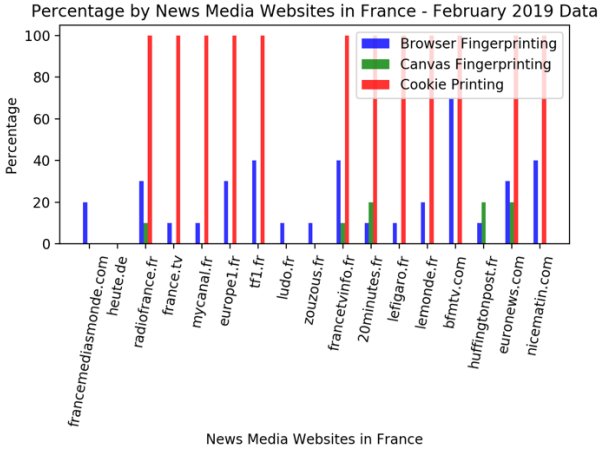


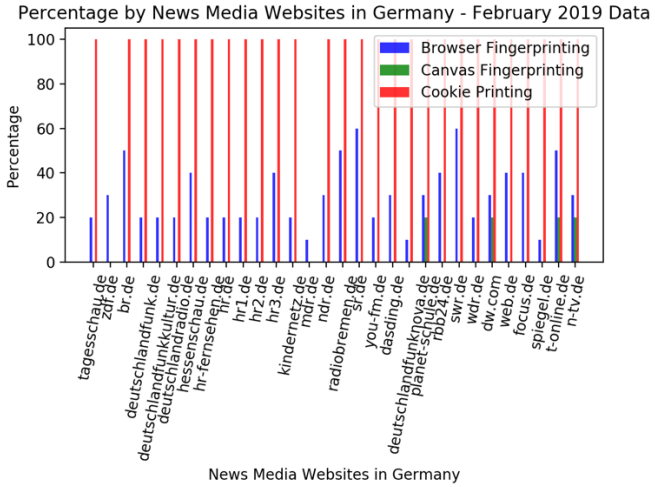


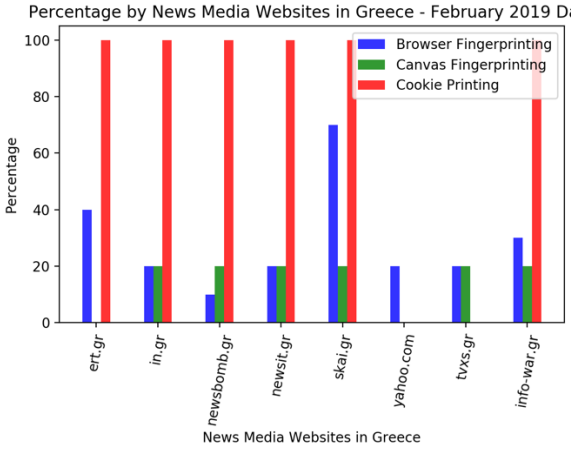


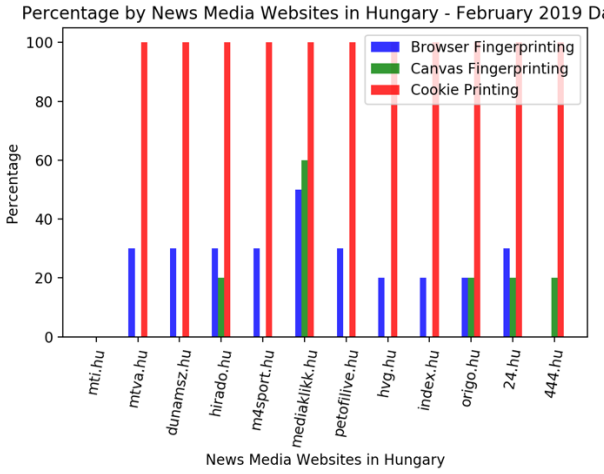

Percentage by News Media Websites in Estonia - February 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
delfi.ee
ohtuleht.ee
postimees.ee
rambler.ru
err.ee
News Media Websites in Estonia

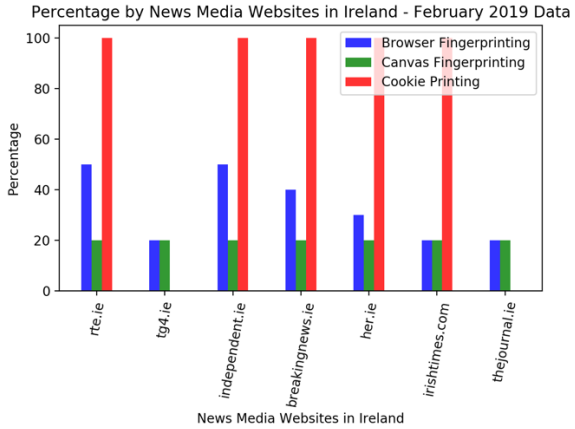

Percentage by News Media Websites in Finland - February 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
yle.fi
mtv.fi
hs.fi
iltalehti.fi
is.fi
katsomo.fi
kauppalehti.fi
aamulehti.fi
ksml.fi
suomenkuvalehti.fi
talouselama.fi
ts.fi
uusisuomi.fi
News Media Websites in Finland

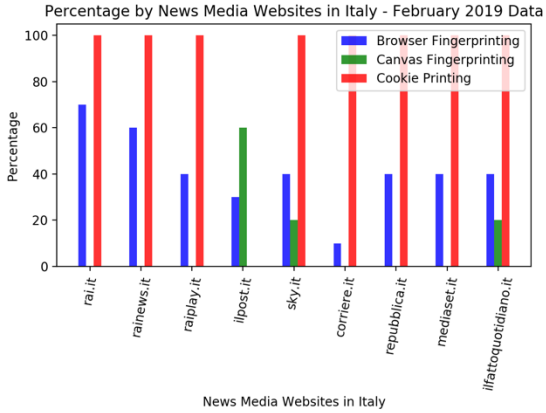

Percentage by News Media Websites in France - February 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
francemediasmonde.com
heute.de
radiofrance.fr
france.tv
mycanal.fr
europe1.fr
tf1.fr
ludo.fr
zouzous.fr
francetvinfo.fr
20minutes.fr
lefigaro.fr
lemonde.fr
bfmtv.com
huffingtonpost.fr
euronews.com
nicematin.com
News Media Websites in France

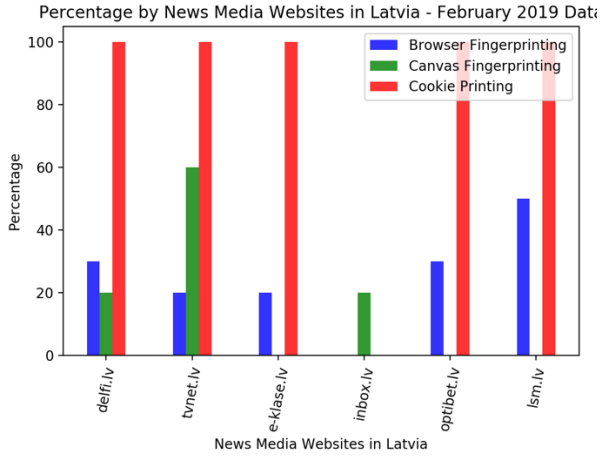

Percentage by News Media Websites in Germany - February 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
tagesschau.de
zdf.de
br.de
deutschlandfunk.de
deutschlandfunkkultur.de
deutschlandradio.de
hessenschau.de
hr-fernsehen.de
hr.de
hr1.de
hr2.de
hr3.de
kindernetz.de
mdr.de
ndr.de
radiobremen.de
sr.de
you-fm.de
dasding.de
deutschlandfunknova.de
planet-schule.de
rbb24.de
swr.de
wdr.de
dw.com
web.de
focus.de
spiegel.de
t-online.de
n-tv.de
News Media Websites in Germany

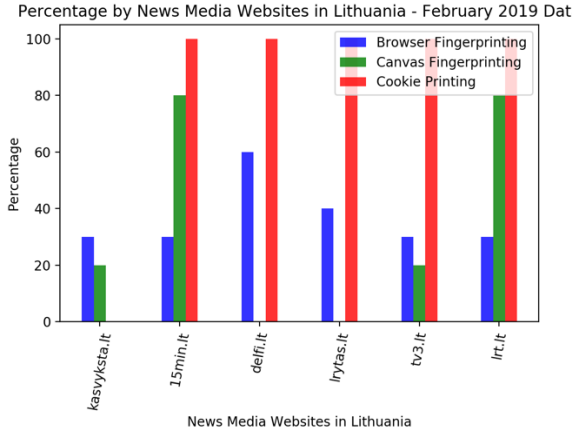

Percentage by News Media Websites in Greece - February 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
ert.gr
in.gr
newsbomb.gr
newsit.gr
skai.gr
yahoo.com
tvxs.gr
info-war.gr
News Media Websites in Greece

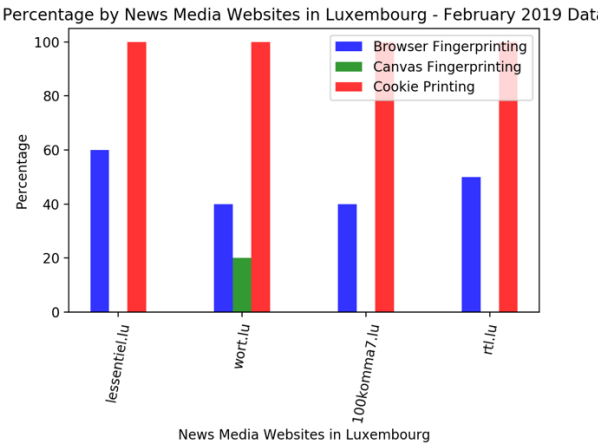

Percentage by News Media Websites in Hungary - February 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
mti.hu
mtva.hu
dunamsz.hu
hirado.hu
m4sport.hu
mediaklikk.hu
petofilive.hu
hvg.hu
index.hu
origo.hu
24.hu
444.hu
News Media Websites in Hungary

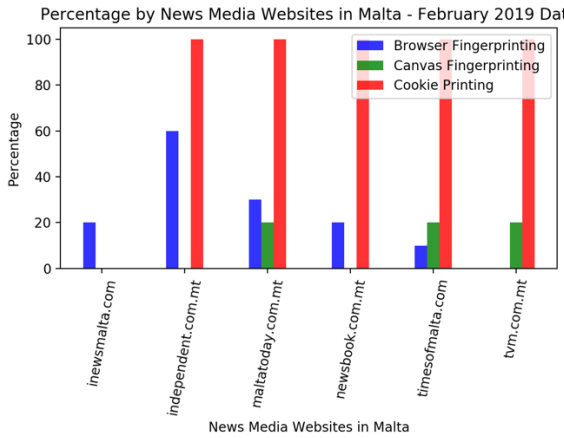

Percentage by News Media Websites in Ireland - February 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
rte.ie
tg4.ie
independent.ie
breakingnews.ie
her.ie
irishtimes.com
thejournal.ie
News Media Websites in Ireland

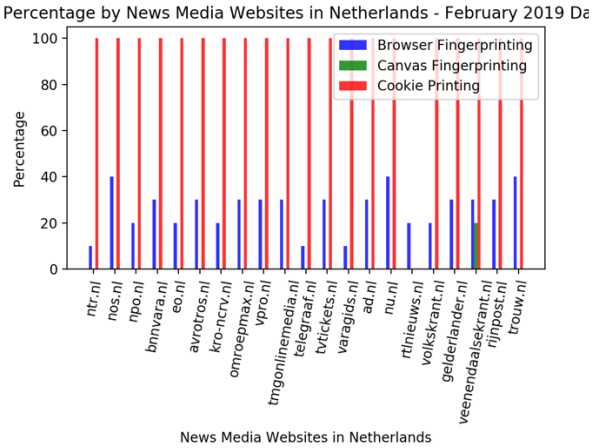

Percentage by News Media Websites in Italy - February 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
rai.it
rainews.it
raiplay.it
ilpost.it
sky.it
corriere.it
repubblica.it
mediaset.it
ilfattoquotidiano.it
News Media Websites in Italy

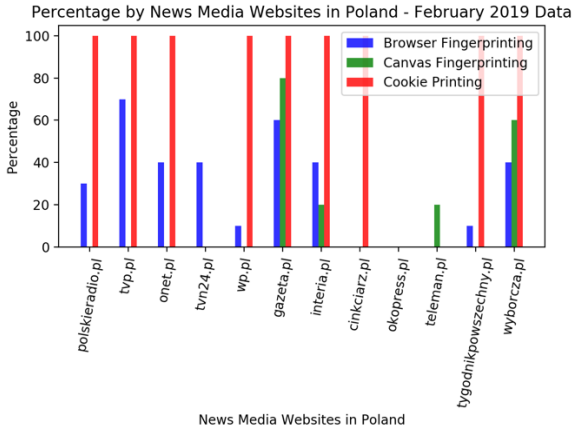

Percentage by News Media Websites in Latvia - February 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
delfi.lv
tvnet.lv
e-klase.lv
inbox.lv
optibet.lv
lsm.lv
News Media Websites in Latvia

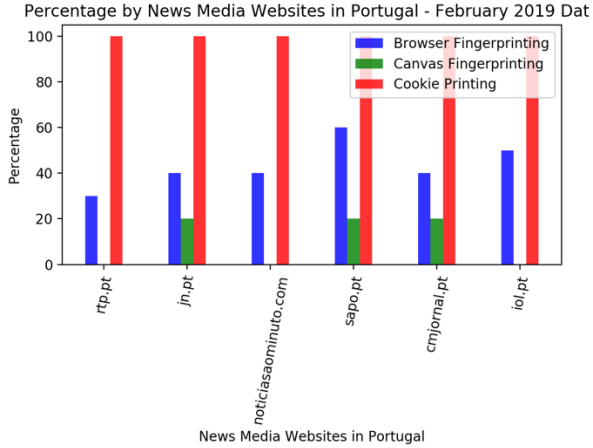

Percentage by News Media Websites in Lithuania - February 2019 Dat
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
kasyvyksta.lt
15min.lt
delfi.lt
lrytas.lt
tv3.lt
lrt.lt
News Media Websites in Lithuania

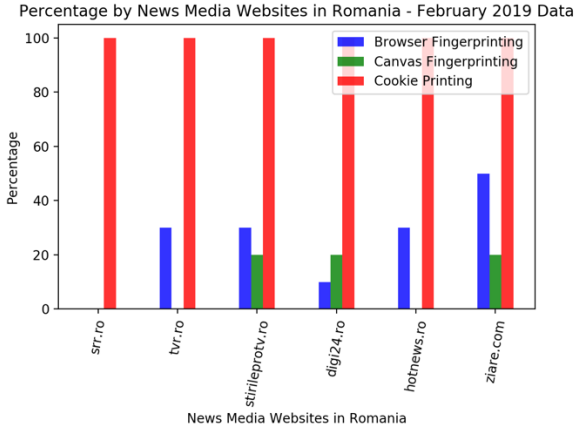

Percentage by News Media Websites in Luxembourg - February 2019 Dat
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
lessentiel.lu
wort.lu
100komma7.lu
rtl.lu
News Media Websites in Luxembourg

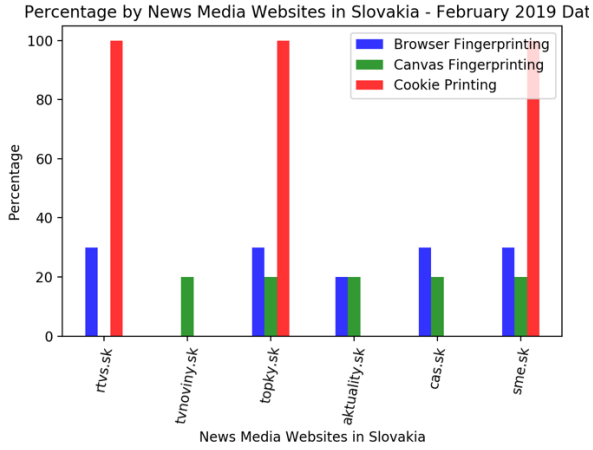

Percentage by News Media Websites in Malta - February 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
inewsmalta.com
independent.com.mt
maltatoday.com.mt
newsbook.com.mt
timesofmalta.com
tvm.com.mt
News Media Websites in Malta

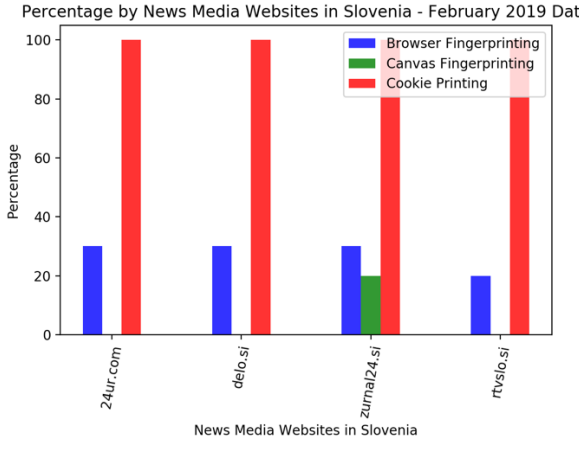

Percentage by News Media Websites in Netherlands - February 2019 Da
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
ntr.nl
nos.nl
npo.nl
bnnvara.nl
eo.nl
avrotros.nl
kro-ncrv.nl
omroepmax.nl
vpro.nl
tmgonlinemedia.nl
telegraaf.nl
tvtickets.nl
varagids.nl
ad.nl
nu.nl
rtlnieuws.nl
volkskrant.nl
gelderlander.nl
veenendaalsekrant.nl
rijnpost.nl
trouw.nl
News Media Websites in Netherlands

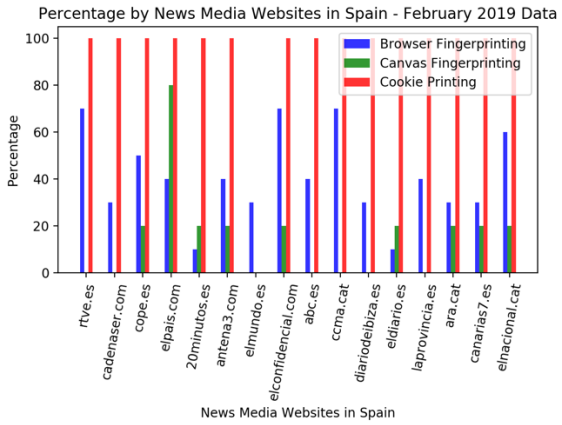

Percentage by News Media Websites in Poland - February 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
polskieradio.pl
tvp.pl
onet.pl
tvn24.pl
wp.pl
gazeta.pl
interia.pl
cinkciarz.pl
okopress.pl
teleman.pl
tygodnikpowszechny.pl
wyborcza.pl
News Media Websites in Poland

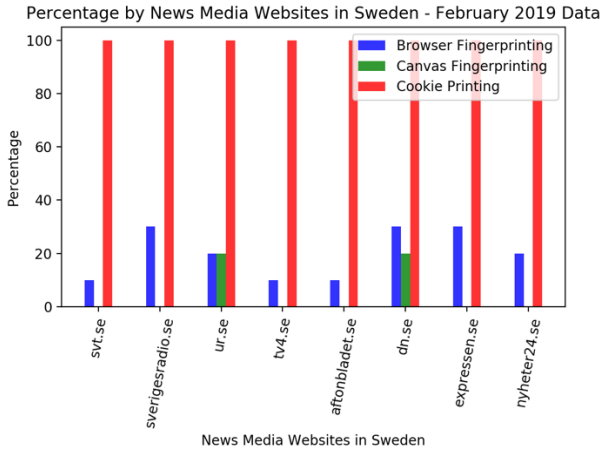

Percentage by News Media Websites in Portugal - February 2019 Dat
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
rtp.pt
jn.pt
noticiasaominuto.com
sapo.pt
cmjornal.pt
iol.pt
News Media Websites in Portugal

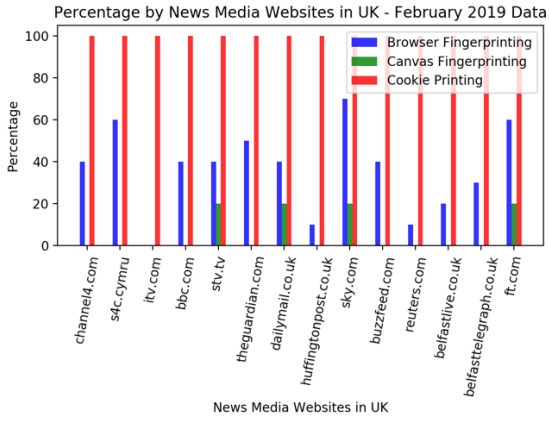

Percentage by News Media Websites in Romania - February 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
srr.ro
tvr.ro
stirileprotv.ro
digi24.ro
hotnews.ro
ziare.com
News Media Websites in Romania


Percentage by News Media Websites in Slovakia - February 2019 Dat
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
rtvs.sk
tvnoviny.sk
topky.sk
aktuality.sk
cas.sk
sme.sk
News Media Websites in Slovakia

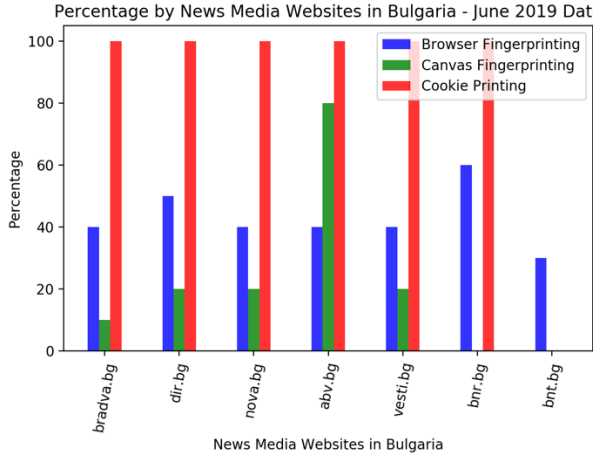

Percentage by News Media Websites in Slovenia - February 2019 Dat
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
24ur.com
delo.si
zurnal24.si
rtvslo.si
News Media Websites in Slovenia

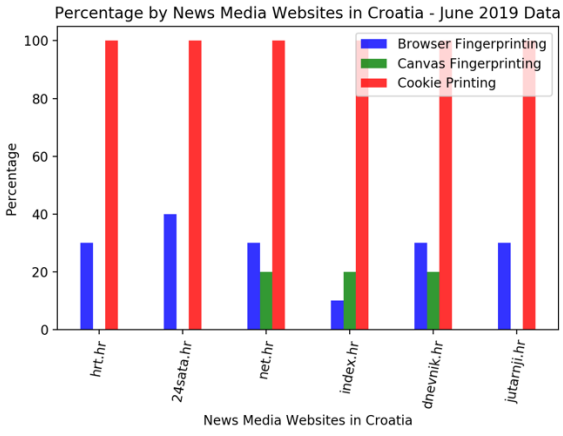


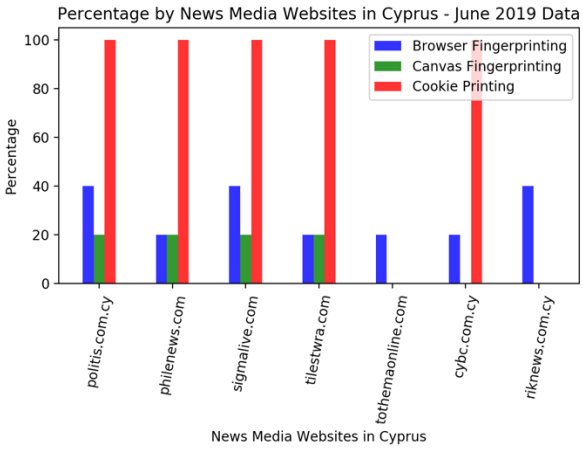


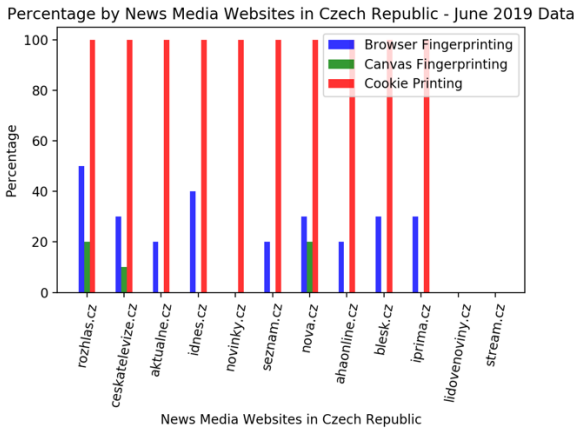


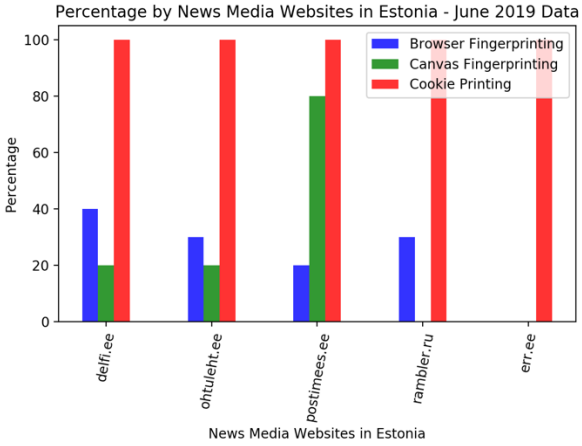
Results from June 2019

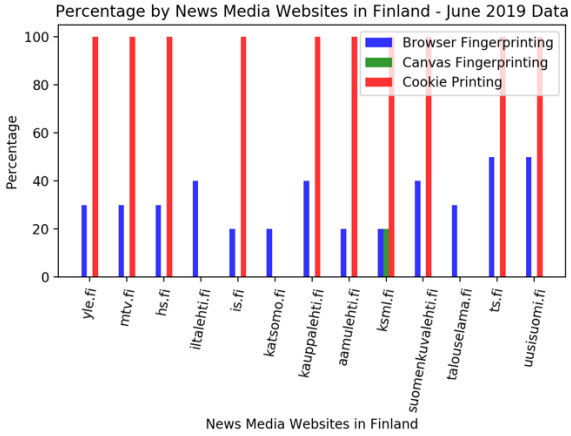


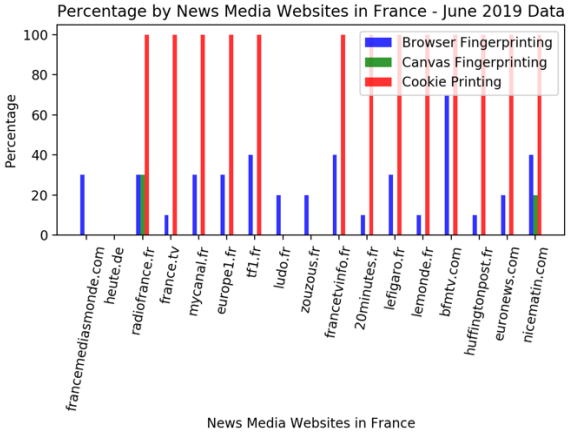


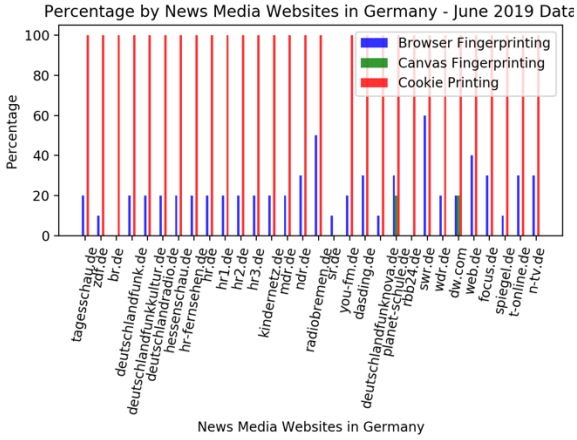

Percentage by News Media Websites in Cyprus - June 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
politis.com.cy
philenews.com
sigmalive.com
tilestwra.com
tothemaonline.com
cybc.com.cy
riknews.com.cy
News Media Websites in Cyprus

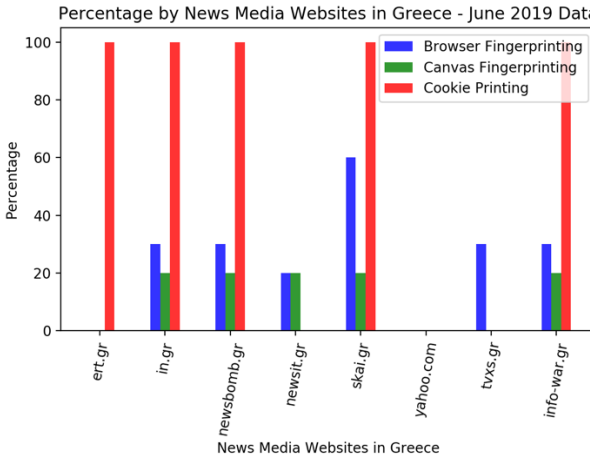

Percentage by News Media Websites in Czech Republic - June 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
rozhlas.cz
ceskatelevize.cz
aktualne.cz
idnes.cz
novinky.cz
seznam.cz
nova.cz
ahaonline.cz
blesk.cz
iprima.cz
lidovenoviny.cz
stream.cz
News Media Websites in Czech Republic

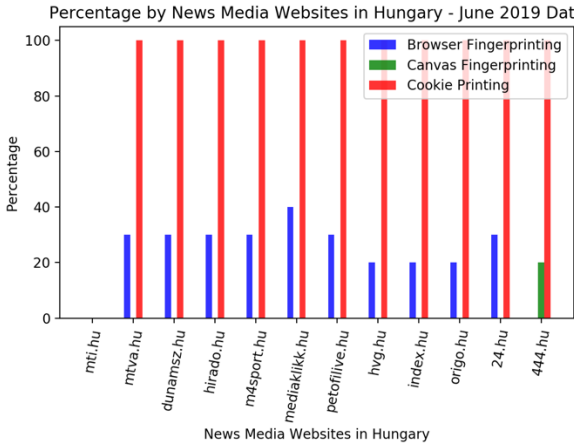

Percentage by News Media Websites in Estonia - June 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
delfi.ee
ohtuleht.ee
postimees.ee
rambler.ru
err.ee
News Media Websites in Estonia

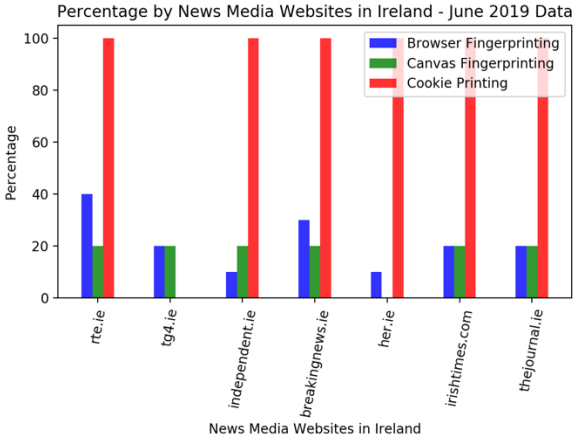

Percentage by News Media Websites in Finland - June 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
yle.fi
mtv.fi
hs.fi
iltalehti.fi
is.fi
katsomo.fi
kauppalehti.fi
aamulehti.fi
ksml.fi
suomenkuvalehti.fi
talouselama.fi
ts.fi
uusisuomi.fi
News Media Websites in Finland

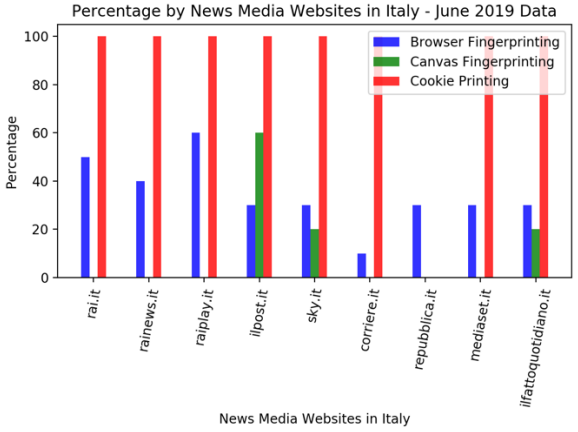

Percentage by News Media Websites in France - June 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
francemediasmonde.com
heute.de
radiofrance.fr
france.tv
mycanal.fr
europe1.fr
tf1.fr
ludo.fr
zouzous.fr
francetvinfo.fr
20minutes.fr
lefigaro.fr
lemonde.fr
bfmtv.com
huffingtonpost.fr
euronews.com
nicematin.com
News Media Websites in France

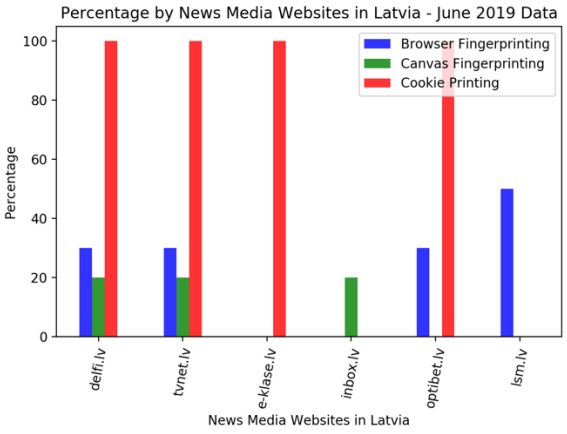

Percentage by News Media Websites in Germany - June 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
tagesschau.de
zdf.de
br.de
deutschlandfunk.de
deutschlandfunkkultur.de
deutschlandradio.de
hessenschau.de
hr-fernsehen.de
hr.de
hr1.de
hr2.de
hr3.de
kindernetz.de
mdr.de
ndr.de
radiobremen.de
sr.de
you-fm.de
dasding.de
deutschlandfunknova.de
planet-schule.de
rbb24.de
swr.de
wdr.de
dw.com
web.de
focus.de
spiegel.de
t-online.de
n-tv.de
News Media Websites in Germany

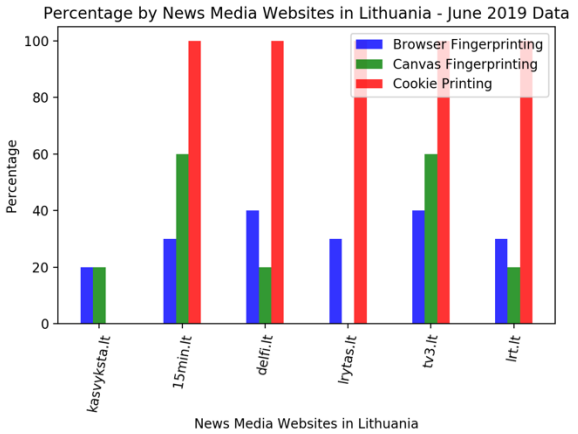

Percentage by News Media Websites in Greece - June 2019 Dat
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
ert.gr
in.gr
newsbomb.gr
newsit.gr
skai.gr
yahoo.com
tvxs.gr
info-war.gr
News Media Websites in Greece

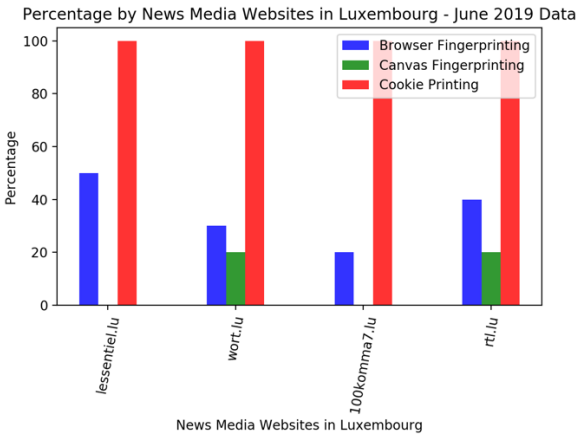

Percentage by News Media Websites in Hungary - June 2019 Dat
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
mti.hu
mtva.hu
dunamsz.hu
hirado.hu
m4sport.hu
mediaklikk.hu
petofilive.hu
hvg.hu
index.hu
origo.hu
24.hu
444.hu
News Media Websites in Hungary

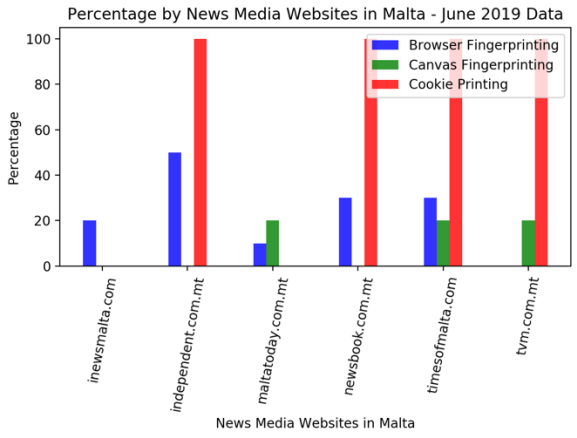

Percentage by News Media Websites in Ireland - June 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
rte.ie
tg4.ie
independent.ie
breakingnews.ie
her.ie
irishtimes.com
thejournal.ie
News Media Websites in Ireland

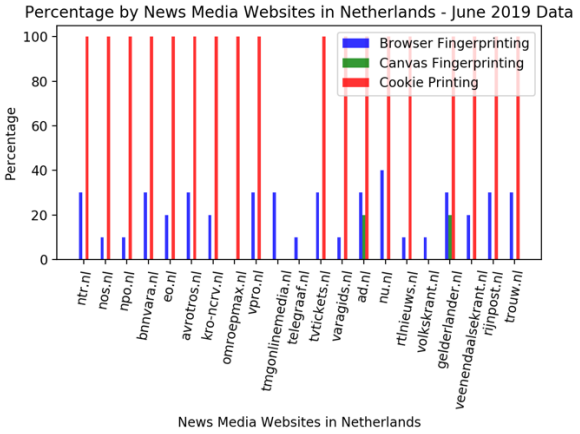

Percentage by News Media Websites in Italy - June 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
rai.it
rainews.it
raiplay.it
ilpost.it
sky.it
corriere.it
repubblica.it
mediaset.it
ilfattoquotidiano.it
News Media Websites in Italy

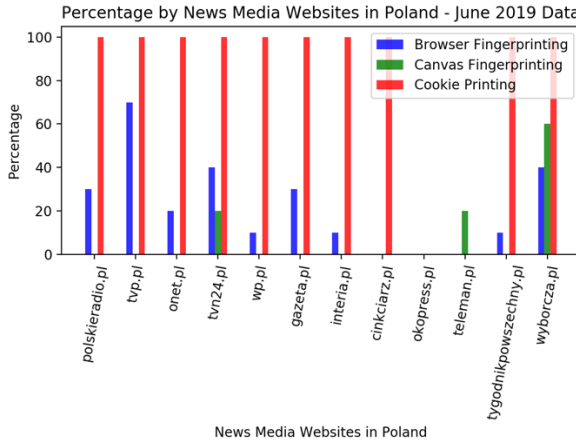

Percentage by News Media Websites in Latvia - June 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
delfi.lv
tvnet.lv
e-klase.lv
inbox.lv
optibet.lv
lsm.lv
News Media Websites in Latvia

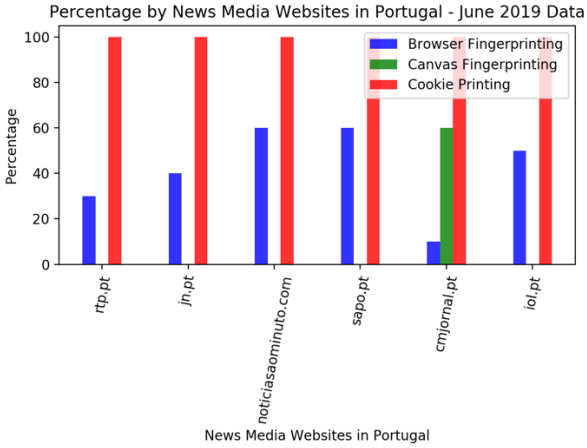

Percentage by News Media Websites in Lithuania - June 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
kasvyksta.lt
15min.lt
delfi.lt
lrytas.lt
tv3.lt
lrt.lt
News Media Websites in Lithuania

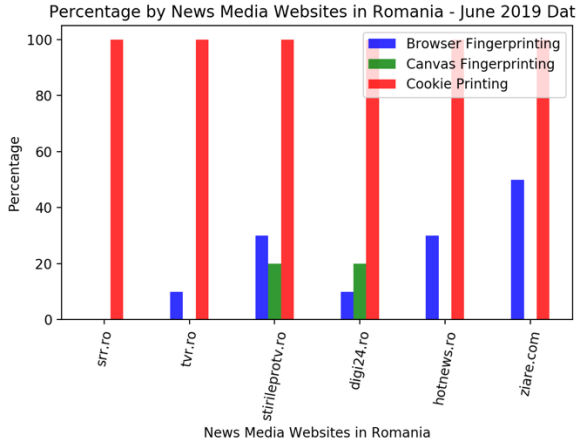

Percentage by News Media Websites in Luxembourg - June 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
lessentiel.lu
wort.lu
100komma7.lu
rtl.lu
News Media Websites in Luxembourg

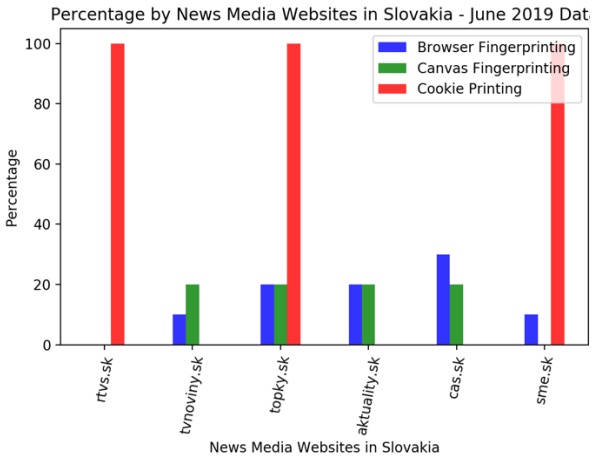

Percentage by News Media Websites in Malta - June 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
inewsmalta.com
independent.com.mt
maltatoday.com.mt
newsbook.com.mt
timesofmalta.com
tvm.com.mt
News Media Websites in Malta

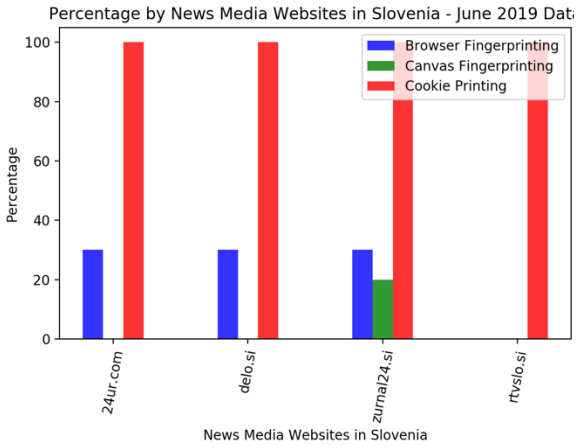

Percentage by News Media Websites in Netherlands - June 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
ntr.nl
nos.nl
npo.nl
bnnvara.nl
eo.nl
avrotros.nl
kro-ncrv.nl
omroepmax.nl
vpro.nl
tmgonlinemedia.nl
telegraaf.nl
tvtickets.nl
varagids.nl
ad.nl
nu.nl
rtlnieuws.nl
volkskrant.nl
gelderlander.nl
veenendaalsekrant.nl
rijnpost.nl
trouw.nl
News Media Websites in Netherlands

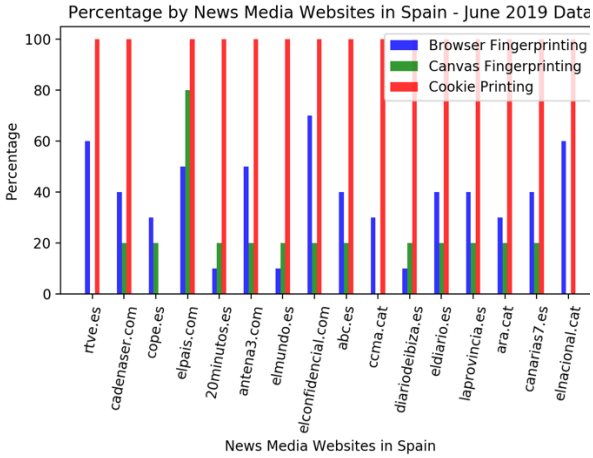

Percentage by News Media Websites in Poland - June 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
polskieradio.pl
tvp.pl
onet.pl
tvn24.pl
wp.pl
gazeta.pl
interia.pl
cinkciarz.pl
okopress.pl
teleman.pl
tygodnikpowszechny.pl
wyborcza.pl
News Media Websites in Poland

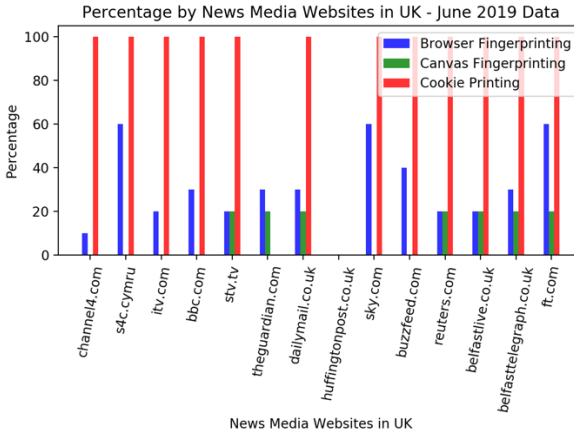

Percentage by News Media Websites in Portugal - June 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
rtp.pt
jn.pt
noticiasaominuto.com
sapo.pt
cmjornal.pt
iol.pt
News Media Websites in Portugal


Percentage by News Media Websites in Romania - June 2019 Dat
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
srr.ro
tvr.ro
stirileprotv.ro
digi24.ro
hotnews.ro
ziare.com
News Media Websites in Romania

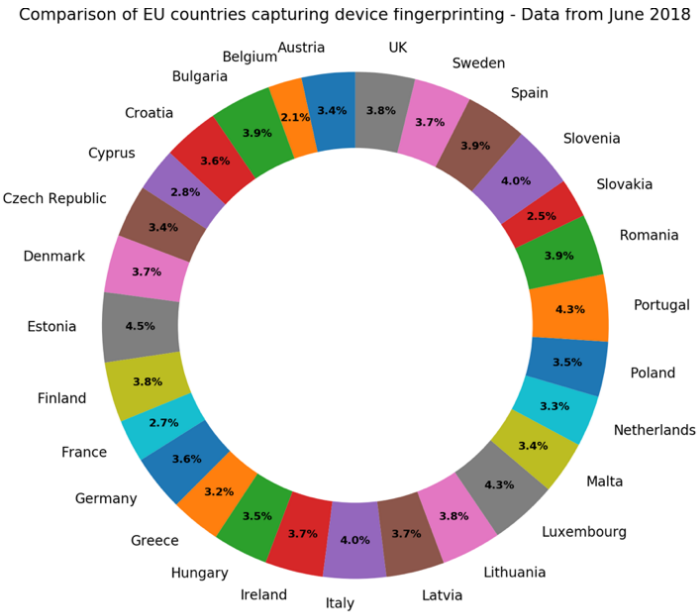

Percentage by News Media Websites in Slovakia - June 2019 Dat
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
rtvs.sk
tvnoviny.sk
topky.sk
aktuality.sk
cas.sk
sme.sk
News Media Websites in Slovakia

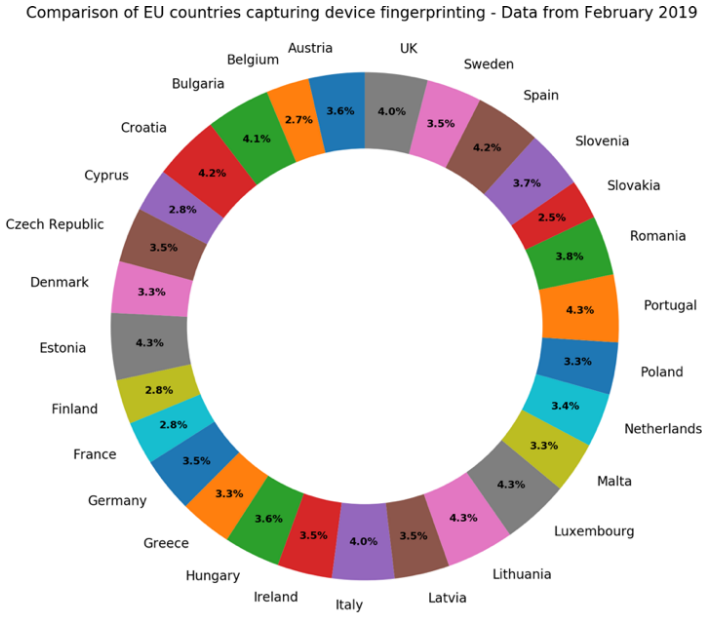

Percentage by News Media Websites in Slovenia - June 2019 Dat
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
24ur.com
delo.si
zurnal24.si
rtvslo.si
News Media Websites in Slovenia

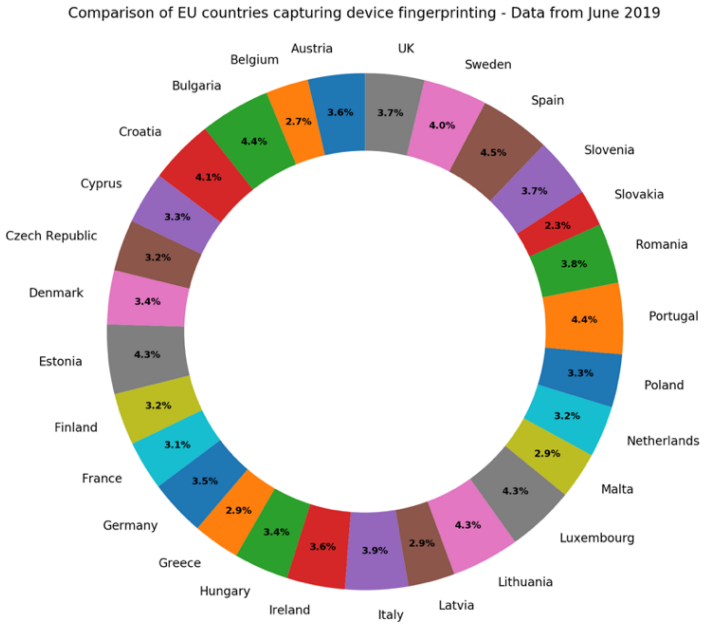

Percentage by News Media Websites in Spain - June 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
rtve.es
cadenaser.com
cope.es
elpais.com
20minutos.es
antena3.com
elmundo.es
elconfidencial.com
abc.es
ccma.cat
diariodelbiza.es
eldiario.es
laprovincia.es
ara.cat
canarias7.es
elnacional.cat
News Media Websites in Spain


Percentage by News Media Websites in UK - June 2019 Data
Browser Fingerprinting
Canvas Fingerprinting
Cookie Printing
Percentage
0
20
40
60
80
100
channel4.com
s4c.cymru
itv.com
bbc.com
stv.tv
theguardian.com
dailymail.co.uk
huffingtonpost.co.uk
sky.com
buzzfeed.com
reuters.com
belfastlive.co.uk
belfasttelegraph.co.uk
ft.com
News Media Websites in UK

Appendix C: Comparative Result of the Countries in Percentages







Appendix D: List of countries and websites used for current research

No	Id	Country	Europe	Europe2	PublicPrivate	SiteCategory	URLtype	TopLevelDomainLookUp
1	29	Austria	EU	EU/EEA	Public	News	PSM possible	orf.at
2	50	Austria	EU	EU/EEA	Private	News	PrivateMedia EU	derstandard.at
3	51	Austria	EU	EU/EEA	Private	News	PrivateMedia EU	kurier.at
4	52	Austria	EU	EU/EEA	Private	News	PrivateMedia EU	heute.at
5	53	Austria	EU	EU/EEA	Private	News	PrivateMedia EU	events.at
6	54	Austria	EU	EU/EEA	Private	News	PrivateMedia EU	gmx.at
7	55	Austria	EU	EU/EEA	Private	News	PrivateMedia EU	krone.at
8	81	Belgium	EU	EU/EEA	Public	News	PSM possible	deredactie.be
9	82	Belgium	EU	EU/EEA	Public	News	PSM possible	sporza.be
10	83	Belgium	EU	EU/EEA	Public	News	PSM possible	een.be
11	84	Belgium	EU	EU/EEA	Public	News	PSM possible	rtbf.be
12	103	Belgium	EU	EU/EEA	Private	News	PrivateMedia EU	vtmkids.be
13	104	Belgium	EU	EU/EEA	Private	News	PrivateMedia EU	vtm.be
14	105	Belgium	EU	EU/EEA	Public	News	PSM notSeen	vrt.be
15	106	Belgium	EU	EU/EEA	Private	News	PrivateMedia EU	hln.be
16	107	Belgium	EU	EU/EEA	Private	News	PrivateMedia EU	nieuwsblad.be
17	108	Belgium	EU	EU/EEA	Private	News	PrivateMedia EU	standaard.be
18	109	Belgium	EU	EU/EEA	Private	News	PrivateMedia EU	gva.be
19	110	Belgium	EU	EU/EEA	Private	News	PrivateMedia EU	dhnet.be
20	111	Belgium	EU	EU/EEA	Private	News	PrivateMedia EU	lavenir.net
21	112	Belgium	EU	EU/EEA	Private	News	PrivateMedia EU	rtl.be
22	113	Belgium	EU	EU/EEA	Private	News	PrivateMedia EU	7sur7.be
23	114	Belgium	EU	EU/EEA	Private	News	PrivateMedia EU	lesoir.be
24	115	Belgium	EU	EU/EEA	Private	News	PrivateMedia EU	neweurope.eu
25	139	Bulgaria	EU	EU/EEA	Private	News	PrivateMedia EU	bradva.bg
26	140	Bulgaria	EU	EU/EEA	Private	News	PrivateMedia EU	dir.bg
27	141	Bulgaria	EU	EU/EEA	Private	News	PrivateMedia EU	nova.bg
28	142	Bulgaria	EU	EU/EEA	Private	News	PrivateMedia EU	abv.bg
29	143	Bulgaria	EU	EU/EEA	Private	News	PrivateMedia EU	vesti.bg
30	146	Bulgaria	EU	EU/EEA	Public	News	PSM possible	bnr.bg
31	147	Bulgaria	EU	EU/EEA	Public	News	PSM possible	bnt.bg
32	172	Croatia	EU	EU/EEA	Public	News	PSM possible	hrt.hr
33	194	Croatia	EU	EU/EEA	Private	News	PrivateMedia	24sata.hr
34	195	Croatia	EU	EU/EEA	Private	News	PrivateMedia EU	net.hr
35	196	Croatia	EU	EU/EEA	Private	News	PrivateMedia EU	index.hr

36	**197**	Croatia	EU	EU/EEA	Private	News	PrivateMedia EU	dnevnik.hr
37	**198**	Croatia	EU	EU/EEA	Private	News	PrivateMedia EU	jutarnji.hr
38	**199**	Cyprus	EU	EU/EEA	Private	News	PrivateMedia EU	politis.com.cy
39	**200**	Cyprus	EU	EU/EEA	Private	News	PrivateMedia EU	philenews.com
40	**201**	Cyprus	EU	EU/EEA	Private	News	PrivateMedia EU	sigmalive.com
41	**202**	Cyprus	EU	EU/EEA	Private	News	PrivateMedia EU	tilestwra.com
42	**203**	Cyprus	EU	EU/EEA	Private	News	PrivateMedia EU	tothemaonline.com
43	**204**	Cyprus	EU	EU/EEA	Public	News	PSM notSeen	cybc.com.cy
44	**224**	Cyprus	EU	EU/EEA	Private	News	PrivateMedia EU	riknews.com.cy
45	**225**	Czech Republic	EU	EU/EEA	Public	News	PSM forbidden	rozhlas.cz
46	**226**	Czech Republic	EU	EU/EEA	Public	News	PSM possible	ceskatelevize.cz
47	**247**	Czech Republic	EU	EU/EEA	Private	News	PrivateMedia EU	aktualne.cz
48	**248**	Czech Republic	EU	EU/EEA	Private	News	PrivateMedia EU	idnes.cz
49	**249**	Czech Republic	EU	EU/EEA	Private	News	PrivateMedia EU	novinky.cz
50	**250**	Czech Republic	EU	EU/EEA	Private	News	PrivateMedia EU	seznam.cz
51	**251**	Czech Republic	EU	EU/EEA	Private	News	PrivateMedia EU	nova.cz
52	**258**	Czech Republic	EU	EU/EEA	Private	News	PrivateMedia EU	ahaonline.cz
53	**259**	Czech Republic	EU	EU/EEA	Private	News	PrivateMedia EU	blesk.cz
54	**260**	Czech Republic	EU	EU/EEA	Private	News	PrivateMedia EU	iprima.cz
55	**261**	Czech Republic	EU	EU/EEA	Private	News	PrivateMedia EU	lidovenoviny.cz
56	**262**	Czech Republic	EU	EU/EEA	Private	News	PrivateMedia EU	stream.cz
57	**277**	Denmark	EU	EU/EEA	Public	News	PSM forbidden	dr.dk
58	**278**	Denmark	EU	EU/EEA	Public	News	PSM possible	tv2.dk
59	**310**	Denmark	EU	EU/EEA	Private	News	PrivateMedia	ekstrabladetcasino.dk
60	**311**	Denmark	EU	EU/EEA	Private	News	PrivateMedia EU	ekstrabladet.dk
61	**312**	Denmark	EU	EU/EEA	Private	News	PrivateMedia EU	politiken.dk
62	**313**	Denmark	EU	EU/EEA	Private	News	PrivateMedia EU	bt.dk
63	**314**	Denmark	EU	EU/EEA	Private	News	PrivateMedia EU	b.dk
64	**331**	Denmark	EU	EU/EEA	Private	News	PrivateMedia EU	borsen.dk
65	**332**	Denmark	EU	EU/EEA	Private	News	PrivateMedia EU	lokalavisen.dk
66	**333**	Denmark	EU	EU/EEA	Private	News	PrivateMedia EU	information.dk
67	**348**	Estonia	EU	EU/EEA	Private	News	PrivateMedia EU	delfi.ee
68	**349**	Estonia	EU	EU/EEA	Private	News	PrivateMedia EU	ohtuleht.ee
69	**350**	Estonia	EU	EU/EEA	Private	News	PrivateMedia EU	postimees.ee
70	**351**	Estonia	EU	EU/EEA	Private	News	PrivateMedia EU	rambler.ru
71	**352**	Estonia	EU	EU/EEA	Public	News	PSM forbidden	err.ee
72	**374**	Finland	EU	EU/EEA	Public	News	PSM forbidden	yle.fi
73	**375**	Finland	EU	EU/EEA	Public	News	PSM possible	mtv.fi
74	**397**	Finland	EU	EU/EEA	Private	News	PrivateMedia EU	hs.fi
75	**398**	Finland	EU	EU/EEA	Private	News	PrivateMedia EU	iltalehti.fi

76	**399**	Finland	EU	EU/EEA	Private	News	PrivateMedia EU	is.fi
77	**400**	Finland	EU	EU/EEA	Private	News	PrivateMedia EU	katsomo.fi
78	**401**	Finland	EU	EU/EEA	Private	News	PrivateMedia EU	kauppalehti.fi
79	**422**	Finland	EU	EU/EEA	Private	News	PrivateMedia EU	aamulehti.fi
80	**423**	Finland	EU	EU/EEA	Private	News	PrivateMedia EU	ksml.fi
81	**424**	Finland	EU	EU/EEA	Private	News	PrivateMedia EU	suomenkuvalehti.fi
82	**425**	Finland	EU	EU/EEA	Private	News	PrivateMedia EU	talouselama.fi
83	**426**	Finland	EU	EU/EEA	Private	News	PrivateMedia EU	ts.fi
84	**427**	Finland	EU	EU/EEA	Private	News	PrivateMedia EU	uusisuomi.fi
85	**475**	France	EU	EU/EEA	Public	News	PSM notSeen	francemediasmonde.com
86	**476**	France	EU	EU/EEA	Public	News	PSM forbidden	heute.de
87	**477**	France	EU	EU/EEA	Public	News	PSM possible	radiofrance.fr
88	**478**	France	EU	EU/EEA	Public	News	PSM possible	france.tv
89	**479**	France	EU	EU/EEA	Public	News	PSM possible	mycanal.fr
90	**480**	France	EU	EU/EEA	Public	News	PSM possible	europe1.fr
91	**481**	France	EU	EU/EEA	Public	News	PSM possible	tf1.fr
92	**502**	France	EU	EU/EEA	Public	News	PSM notSeen	ludo.fr
93	**503**	France	EU	EU/EEA	Public	News	PSM notSeen	zouzous.fr
94	**504**	France	EU	EU/EEA	Public	News	PSM possible	francetvinfo.fr
95	**505**	France	EU	EU/EEA	Private	News	PrivateMedia EU	20minutes.fr
96	**506**	France	EU	EU/EEA	Private	News	PrivateMedia EU	lefigaro.fr
97	**507**	France	EU	EU/EEA	Private	News	PrivateMedia EU	lemonde.fr
98	**508**	France	EU	EU/EEA	Private	News	PrivateMedia EU	bfmtv.com
99	**509**	France	EU	EU/EEA	Private	News	PrivateMedia EU	huffingtonpost.fr
100	**512**	France	EU	EU/EEA	Private	News	PrivateMedia EU	euronews.com
101	**513**	France	EU	EU/EEA	Private	News	PrivateMedia EU	nicematin.com
102	**519**	Germany	EU	EU/EEA	Public	News	PSM forbidden	tagesschau.de
103	**520**	Germany	EU	EU/EEA	Public	News	PSM forbidden	zdf.de
104	**521**	Germany	EU	EU/EEA	Public	News	PSM forbidden	br.de
105	**522**	Germany	EU	EU/EEA	Public	News	PSM forbidden	deutschlandfunk.de
106	**523**	Germany	EU	EU/EEA	Public	News	PSM forbidden	deutschlandfunkkultur.de
107	**524**	Germany	EU	EU/EEA	Public	News	PSM forbidden	deutschlandradio.de
108	**525**	Germany	EU	EU/EEA	Public	News	PSM forbidden	hessenschau.de
109	**526**	Germany	EU	EU/EEA	Public	News	PSM forbidden	hr-fernsehen.de
110	**527**	Germany	EU	EU/EEA	Public	News	PSM forbidden	hr.de
111	**528**	Germany	EU	EU/EEA	Public	News	PSM forbidden	hr1.de
112	**529**	Germany	EU	EU/EEA	Public	News	PSM forbidden	hr2.de
113	**530**	Germany	EU	EU/EEA	Public	News	PSM forbidden	hr3.de
114	**531**	Germany	EU	EU/EEA	Public	News	PSM forbidden	kindernetz.de
115	**532**	Germany	EU	EU/EEA	Public	News	PSM forbidden	mdr.de

116	**533**	Germany	EU	EU/EEA	Public	News	PSM forbidden	ndr.de
116	**534**	Germany	EU	EU/EEA	Public	News	PSM forbidden	radiobremen.de
118	**535**	Germany	EU	EU/EEA	Public	News	PSM forbidden	sr.de
119	**536**	Germany	EU	EU/EEA	Public	News	PSM forbidden	you-fm.de
120	**537**	Germany	EU	EU/EEA	Public	News	PSM forbidden	dasding.de
121	**538**	Germany	EU	EU/EEA	Public	News	PSM forbidden	deutschlandfunknova.de
122	**539**	Germany	EU	EU/EEA	Public	News	PSM forbidden	planet-schule.de
123	**540**	Germany	EU	EU/EEA	Public	News	PSM forbidden	rbb24.de
124	**541**	Germany	EU	EU/EEA	Public	News	PSM forbidden	swr.de
125	**542**	Germany	EU	EU/EEA	Public	News	PSM forbidden	wdr.de
126	**543**	Germany	EU	EU/EEA	Public	News	PSM possible	dw.com
127	**567**	Germany	EU	EU/EEA	Private	News	PrivateMedia EU	web.de
128	**568**	Germany	EU	EU/EEA	Private	News	PrivateMedia EU	focus.de
129	**569**	Germany	EU	EU/EEA	Private	News	PrivateMedia EU	spiegel.de
130	**570**	Germany	EU	EU/EEA	Private	News	PrivateMedia EU	t-online.de
131	**571**	Germany	EU	EU/EEA	Private	News	PrivateMedia EU	n-tv.de
132	**580**	Greece	EU	EU/EEA	Public	News	PSM notSeen	ert.gr
133	**601**	Greece	EU	EU/EEA	Private	News	PrivateMedia EU	in.gr
134	**602**	Greece	EU	EU/EEA	Private	News	PrivateMedia EU	newsbomb.gr
135	**603**	Greece	EU	EU/EEA	Private	News	PrivateMedia EU	newsit.gr
136	**604**	Greece	EU	EU/EEA	Private	News	PrivateMedia EU	skai.gr
137	**605**	Greece	EU	EU/EEA	Private	News	PrivateMedia EU	yahoo.com
138	**608**	Greece	EU	EU/EEA	Private	News	PrivateMedia EU	tvxs.gr
139	**609**	Greece	EU	EU/EEA	Private	News	PrivateMedia EU	info-war.gr
140	**611**	Hungary	EU	EU/EEA	Public	News	PSM possible	mti.hu
141	**612**	Hungary	EU	EU/EEA	Public	News	PSM possible	mtva.hu
142	**613**	Hungary	EU	EU/EEA	Public	News	PSM possible	dunamsz.hu
143	**614**	Hungary	EU	EU/EEA	Public	News	PSM possible	hirado.hu
144	**615**	Hungary	EU	EU/EEA	Public	News	PSM possible	m4sport.hu
145	**616**	Hungary	EU	EU/EEA	Public	News	PSM possible	mediaklikk.hu
146	**617**	Hungary	EU	EU/EEA	Public	News	PSM possible	petofilive.hu
147	**636**	Hungary	EU	EU/EEA	Private	News	PrivateMedia EU	hvg.hu
148	**637**	Hungary	EU	EU/EEA	Private	News	PrivateMedia EU	index.hu
149	**638**	Hungary	EU	EU/EEA	Private	News	PrivateMedia EU	origo.hu
150	**639**	Hungary	EU	EU/EEA	Private	News	PrivateMedia EU	24.hu
151	**640**	Hungary	EU	EU/EEA	Private	News	PrivateMedia EU	444.hu
152	**666**	Ireland	EU	EU/EEA	Public	News	PSM possible	rte.ie
153	**667**	Ireland	EU	EU/EEA	Public	News	PSM possible	tg4.ie
154	**689**	Ireland	EU	EU/EEA	Private	News	PrivateMedia	independent.ie
155	**690**	Ireland	EU	EU/EEA	Private	News	PrivateMedia	breakingnews.ie

156	**691**	Ireland	EU	EU/EEA	Private	News	PrivateMedia	her.ie
157	**692**	Ireland	EU	EU/EEA	Private	News	PrivateMedia EU	irishtimes.com
158	**693**	Ireland	EU	EU/EEA	Private	News	PrivateMedia EU	thejournal.ie
159	**695**	Italy	EU	EU/EEA	Public	News	PSM possible	rai.it
160	**696**	Italy	EU	EU/EEA	Public	News	PSM possible	rainews.it
161	**697**	Italy	EU	EU/EEA	Public	News	PSM possible	raiplay.it
162	**708**	Italy	EU	EU/EEA	Private	News	PrivateMedia EU	ilpost.it
163	**719**	Italy	EU	EU/EEA	Private	News	PrivateMedia EU	sky.it
164	**720**	Italy	EU	EU/EEA	Private	News	PrivateMedia EU	corriere.it
165	**721**	Italy	EU	EU/EEA	Private	News	PrivateMedia EU	repubblica.it
166	**722**	Italy	EU	EU/EEA	Private	News	PrivateMedia EU	mediaset.it
167	**723**	Italy	EU	EU/EEA	Private	News	PrivateMedia EU	ilfattoquotidiano.it
168	**729**	Latvia	EU	EU/EEA	Private	News	PrivateMedia EU	delfi.lv
169	**730**	Latvia	EU	EU/EEA	Private	News	PrivateMedia EU	tvnet.lv
170	**731**	Latvia	EU	EU/EEA	Private	News	PrivateMedia EU	e-klase.lv
171	**732**	Latvia	EU	EU/EEA	Private	News	PrivateMedia EU	inbox.lv
172	**733**	Latvia	EU	EU/EEA	Private	News	PrivateMedia EU	optibet.lv
173	**735**	Latvia	EU	EU/EEA	Public	News	PSM possible	lsm.lv
174	**760**	Lithuania	EU	EU/EEA	Private	News	PrivateMedia EU	kasvyksta.lt
175	**761**	Lithuania	EU	EU/EEA	Private	News	PrivateMedia EU	15min.lt
176	**762**	Lithuania	EU	EU/EEA	Private	News	PrivateMedia EU	delfi.lt
177	**763**	Lithuania	EU	EU/EEA	Private	News	PrivateMedia EU	lrytas.lt
178	**764**	Lithuania	EU	EU/EEA	Private	News	PrivateMedia EU	tv3.lt
179	**767**	Lithuania	EU	EU/EEA	Public	News	PSM possible	lrt.lt
180	**786**	Luxembourg	EU	EU/EEA	Private	News	PrivateMedia EU	lessentiel.lu
181	**787**	Luxembourg	EU	EU/EEA	Private	News	PrivateMedia EU	wort.lu
182	**790**	Luxembourg	EU	EU/EEA	Public	News	PSM notSeen	100komma7.lu
183	**791**	Luxembourg	EU	EU/EEA	Public	News	PSM possible	rtl.lu
184	**804**	Malta	EU	EU/EEA	Private	News	PrivateMedia EU	inewsmalta.com
185	**805**	Malta	EU	EU/EEA	Private	News	PrivateMedia EU	independent.com.mt
186	**806**	Malta	EU	EU/EEA	Private	News	PrivateMedia EU	maltatoday.com.mt
187	**807**	Malta	EU	EU/EEA	Private	News	PrivateMedia EU	newsbook.com.mt
188	**808**	Malta	EU	EU/EEA	Private	News	PrivateMedia EU	timesofmalta.com
189	**809**	Malta	EU	EU/EEA	Public	News	PSM possible	tvm.com.mt
190	**869**	Netherlands	EU	EU/EEA	Public	News	PSM notSeen	ntr.nl
191	**870**	Netherlands	EU	EU/EEA	Public	News	PSM possible	nos.nl
192	**871**	Netherlands	EU	EU/EEA	Public	News	PSM possible	npo.nl
193	**872**	Netherlands	EU	EU/EEA	Public	News	PSM possible	bnnvara.nl
194	**873**	Netherlands	EU	EU/EEA	Public	News	PSM possible	eo.nl
195	**874**	Netherlands	EU	EU/EEA	Public	News	PSM possible	avrotros.nl

196	**875**	Netherlands	EU	EU/EEA	Public	News	PSM possible	kro-ncrv.nl
197	**876**	Netherlands	EU	EU/EEA	Public	News	PSM possible	omroepmax.nl
198	**877**	Netherlands	EU	EU/EEA	Public	News	PSM possible	vpro.nl
199	**897**	Netherlands	EU	EU/EEA	Private	News	PrivateMedia EU	tmgonlinemedia.nl
200	**898**	Netherlands	EU	EU/EEA	Private	News	PrivateMedia EU	telegraaf.nl
201	**899**	Netherlands	EU	EU/EEA	Public	News	PSM possible	tvtickets.nl
202	**900**	Netherlands	EU	EU/EEA	Public	News	PSM possible	varagids.nl
203	**901**	Netherlands	EU	EU/EEA	Private	News	PrivateMedia	ad.nl
204	**902**	Netherlands	EU	EU/EEA	Private	News	PrivateMedia EU	nu.nl
205	**903**	Netherlands	EU	EU/EEA	Private	News	PrivateMedia EU	rtlnieuws.nl
206	**904**	Netherlands	EU	EU/EEA	Private	News	PrivateMedia EU	volkskrant.nl
207	**911**	Netherlands	EU	EU/EEA	Private	News	NewsUser	gelderlander.nl
208	**912**	Netherlands	EU	EU/EEA	Private	News	PrivateMedia EU	veenendaalsekrant.nl
209	**913**	Netherlands	EU	EU/EEA	Private	News	PrivateMedia EU	rijnpost.nl
210	**914**	Netherlands	EU	EU/EEA	Private	News	PrivateMedia EU	trouw.nl
211	**957**	Poland	EU	EU/EEA	Public	News	PSM possible	polskieradio.pl
212	**958**	Poland	EU	EU/EEA	Public	News	PSM possible	tvp.pl
213	**977**	Poland	EU	EU/EEA	Private	News	PrivateMedia EU	onet.pl
214	**978**	Poland	EU	EU/EEA	Private	News	PrivateMedia EU	tvn24.pl
215	**979**	Poland	EU	EU/EEA	Private	News	PrivateMedia EU	wp.pl
216	**980**	Poland	EU	EU/EEA	Private	News	PrivateMedia EU	gazeta.pl
217	**981**	Poland	EU	EU/EEA	Private	News	PrivateMedia EU	interia.pl
218	**982**	Poland	EU	EU/EEA	Private	News	PrivateMedia EU	cinkciarz.pl
219	**983**	Poland	EU	EU/EEA	Private	News	PrivateMedia EU	okopress.pl
220	**984**	Poland	EU	EU/EEA	Private	News	PrivateMedia EU	teleman.pl
221	**985**	Poland	EU	EU/EEA	Private	News	PrivateMedia EU	tygodnikpowszechny.pl
222	**986**	Poland	EU	EU/EEA	Private	News	PrivateMedia EU	wyborcza.pl
223	**992**	Portugal	EU	EU/EEA	Public	News	PSM possible	rtp.pt
224	**1013**	Portugal	EU	EU/EEA	Private	News	PrivateMedia EU	jn.pt
225	**1014**	Portugal	EU	EU/EEA	Private	News	PrivateMedia EU	noticiasaominuto.com
226	**1015**	Portugal	EU	EU/EEA	Private	News	PrivateMedia EU	sapo.pt
227	**1016**	Portugal	EU	EU/EEA	Private	News	PrivateMedia EU	cmjornal.pt
228	**1017**	Portugal	EU	EU/EEA	Private	News	PrivateMedia EU	iol.pt
229	**1019**	Romania	EU	EU/EEA	Public	News	PSM notSeen	srr.ro
230	**1020**	Romania	EU	EU/EEA	Public	News	PSM possible	tvr.ro
231	**1041**	Romania	EU	EU/EEA	Private	News	PrivateMedia EU	stirileprotv.ro
232	**1042**	Romania	EU	EU/EEA	Private	News	PrivateMedia EU	digi24.ro
233	**1043**	Romania	EU	EU/EEA	Private	News	PrivateMedia EU	hotnews.ro
234	**1044**	Romania	EU	EU/EEA	Private	News	PrivateMedia EU	ziare.com
235	**1073**	Slovakia	EU	EU/EEA	Public	News	PSM forbidden	rtvs.sk

236	**1093**	Slovakia	EU	EU/EEA	Private	News	PrivateMedia EU	tvnoviny.sk
237	**1094**	Slovakia	EU	EU/EEA	Private	News	PrivateMedia EU	topky.sk
238	**1095**	Slovakia	EU	EU/EEA	Private	News	PrivateMedia EU	aktuality.sk
239	**1096**	Slovakia	EU	EU/EEA	Private	News	PrivateMedia EU	cas.sk
240	**1097**	Slovakia	EU	EU/EEA	Private	News	PrivateMedia EU	sme.sk
241	**1098**	Slovenia	EU	EU/EEA	Private	News	PrivateMedia EU	24ur.com
242	**1099**	Slovenia	EU	EU/EEA	Private	News	PrivateMedia EU	delo.si
243	**1100**	Slovenia	EU	EU/EEA	Private	News	PrivateMedia EU	zurnal24.si
244	**1101**	Slovenia	EU	EU/EEA	Public	News	PSM possible	rtvslo.si
245	**1102**	Spain	EU	EU/EEA	Public	News	PSM forbidden	rtve.es
246	**1103**	Spain	EU	EU/EEA	Public	News	PSM possible	cadenaser.com
247	**1104**	Spain	EU	EU/EEA	Public	News	PSM possible	cope.es
248	**1124**	Spain	EU	EU/EEA	Private	News	PrivateMedia EU	elpais.com
249	**1125**	Spain	EU	EU/EEA	Private	News	PrivateMedia EU	20minutos.es
250	**1126**	Spain	EU	EU/EEA	Private	News	PrivateMedia EU	antena3.com
251	**1127**	Spain	EU	EU/EEA	Private	News	PrivateMedia EU	elmundo.es
252	**1128**	Spain	EU	EU/EEA	Private	News	PrivateMedia EU	elconfidencial.com
253	**1132**	Spain	EU	EU/EEA	Private	News	PrivateMedia EU	abc.es
254	**1133**	Spain	EU	EU/EEA	Private	News	PrivateMedia EU	ccma.cat
255	**1134**	Spain	EU	EU/EEA	Private	News	PrivateMedia EU	diariodeibiza.es
256	**1135**	Spain	EU	EU/EEA	Private	News	PrivateMedia EU	eldiario.es
257	**1136**	Spain	EU	EU/EEA	Private	News	PrivateMedia EU	laprovincia.es
258	**1137**	Spain	EU	EU/EEA	Private	News	PrivateMedia EU	ara.cat
259	**1138**	Spain	EU	EU/EEA	Private	News	PrivateMedia EU	canarias7.es
260	**1139**	Spain	EU	EU/EEA	Private	News	PrivateMedia EU	elnacional.cat
261	**1153**	Sweden	EU	EU/EEA	Public	News	PSM forbidden	svt.se
262	**1154**	Sweden	EU	EU/EEA	Public	News	PSM forbidden	sverigesradio.se
263	**1155**	Sweden	EU	EU/EEA	Public	News	PSM forbidden	ur.se
264	**1156**	Sweden	EU	EU/EEA	Public	News	PSM possible	tv4.se
265	**1172**	Sweden	EU	EU/EEA	Private	News	PrivateMedia EU	aftonbladet.se
266	**1173**	Sweden	EU	EU/EEA	Private	News	PrivateMedia EU	dn.se
267	**1174**	Sweden	EU	EU/EEA	Private	News	PrivateMedia EU	expressen.se
268	**1175**	Sweden	EU	EU/EEA	Private	News	PrivateMedia EU	nyheter24.se
269	**1206**	UK	EU	EU/EEA	Public	News	PSM notSeen	channel4.com
270	**1207**	UK	EU	EU/EEA	Public	News	PSM notSeen	s4c.cymru
271	**1208**	UK	EU	EU/EEA	Public	News	PSM notSeen	itv.com
272	**1209**	UK	EU	EU/EEA	Public	News	PSM possible	bbc.com
273	**1210**	UK	EU	EU/EEA	Public	News	PSM possible	stv.tv
274	**1227**	UK	EU	EU/EEA	Private	News	PrivateMedia EU	theguardian.com
275	**1228**	UK	EU	EU/EEA	Private	News	PrivateMedia EU	dailymail.co.uk

276	**1229**	UK	EU	EU/EEA	Private	News	PrivateMedia EU	huffingtonpost.co.uk
277	**1230**	UK	EU	EU/EEA	Private	News	PrivateMedia EU	sky.com
278	**1231**	UK	EU	EU/EEA	Private	News	PrivateMedia EU	buzzfeed.com
279	**1239**	UK	EU	EU/EEA	Private	News	PrivateMedia EU	reuters.com
280	**1240**	UK	EU	EU/EEA	Private	News	PrivateMedia EU	belfastlive.co.uk
281	**1241**	UK	EU	EU/EEA	Private	News	PrivateMedia EU	belfasttelegraph.co.uk
282	**1242**	UK	EU	EU/EEA	Private	News	PrivateMedia EU	ft.com