Survey on
Device Fingerprinting

TABLE OF CONTENTS

1. INTRODUCTION

At present, the model that underlies most web services is based on providing a completely free service in exchange for the monetisation of the data gathered from users. In most cases, the information gathered from users is monetised through marketing services that run personalised advertising campaigns for clients looking to advertise their product or service. Therefore, in addition to identifying the user, tracking them and gathering the data, they need to profile those data with the aim of maximising the efficiency of the advertising on offer.

To identify users different tracking techniques are used, the best known of these is cookies, which are files stored on the user's computer created by the service provider's website and which are subsequently used to variouspurposes, such as improving the user experience with the web browser or studying the statistics of the user's website use. However, they are also used for other purposes including the profiling of users.

Article 22.2 of the Law on Information Society Services[1] states that service providers may use storage and data recovery devices on user's computers on the condition that they give their consent after being given clear and complete information on their use, in particular, for data processing purposes, in accordance with the applicable legislation. In the case we are concerned with here, in accordance with the provisions of Article 13 of Regulation (EU) 2016/679 of the European Parliament and of the Council, of 27 April 2016, relating to the protection of natural persons in relation to the processing of personal data and the free circulation of these data and repealing Directive 95/46/EC (GDPR) and Article 11 of Organic Law 3/2018 of 5 December, on the Protection of Personal Data and the Guarantee of Digital Rights (LOPDGDD).

Where technically possible and efficient to do so, the consent of the recipient to accept the processing of their data may be provided using the appropriate parameters of the browser or other application, whenever it appears during the installation or update via an express action to such effect.

Today, browsers can be configured, among other options, not to accept cookies or to accept only temporary cookies which are automatically deleted when the browser is closed. For their part, anti-virus systems have installed consistent protections to be able to schedule deletion of cookies and other files installed on the user's computer by web applications as well as anonymizers of the data of the terminals.

However, we find ourselves before a very dynamic market in which browsers and anti-viruses provided the tools to allow users to manage the exposure of their personal information, which implies a certain difficulty when it comes to accurately profiling potential clients. For this reason, the different stakeholders involved in the Internet market continue to research new ways of get around these restrictions to gather, and exploit, users' data.

[1] Law 34/2002, of 11 July, on information society services and eCommerce.

From a number of existing studies concerning the internet identification techniques, it is concluded that other, more advanced tracking techniques are being used and that they have overcomed cookies, based on the gathering of specific information from the browser and/or browsing device, the combination of which allows for a an identifier to single out and uniquely identify the user and the legitimacy of which remain unclear. This set of techniques is known as device fingerprinting, browser fingerprinting or simply fingerprinting. Through the text, these terms may be used interchangeably.

This study assesses an approximation of the digital fingerprint of the device; the techniques most used to obtain it; how they identify the device used by the user, some recommendations for users on how to protect their privacy through the uses of measures available to them and thus avoid the use of fingerprints for tracking and profiling purposes and, finally, recommendations for the industry.

2. DEVICE FINGERPRINTING

Device fingerprinting is the systematic gathering of information on a specific remote device with the aim of identifying, singling out and, thus being able to monitor its user's activity for the purpose of profiling.

The European Data Protection Board, in its document "Opinion 9/2014 on the application of Directive 2002/58/EC to device fingerprinting" assumes the definition of RFC6973[2] which identifies fingerprinting as *"a set of information elements that identifies a device or application instance".*[3]

In simpler terms, the digital fingerprint of a device is a data set extracted from the user's terminal device that allows that terminal device to be unequivocally uniquely identified. Given that people generally tend not to share terminals, whether it be a mobile phone, tablet, laptop or work computer, uniquely identifying the terminal means uniquely identifying the person using it. The entities that use digital fingerprinting mechanisms systematically compile information on all terminals that are connected to their servers with the aim of uniquely identifying them so as to monitor the user's browsing in order to build a profile.

Contrary to what some people may think, this profiling is not limited to compiling and analysing the user's browsing habits or the searches they make on the servers. More advanced techniques allow for the registration of the movements the user makes throughout the web page itself with their mouse, examining the parts of the screen they spend more time over[4]. On the other hand, the development of software for devices, for example JavaScript or Flash, facilitate the implementation of procedures to gather very specific information on the device, such as the

[2] *RFC 6973 Privacy Considerations for Internet Protocols*. Document that offers a guide with considerations on privacy to include in the development of internet protocol specifications. The aim is for designers, implementers and users of internet protocols to be conscious of design options relating to privacy.
[3] Opinion 9/2014 on the application of Directive 2002/58/EC to device fingerprinting.
[4] See, for example, the activity of www.hotjar.com or www.crazyegg.com, which allow for the recording of the user's mouse movements, clicks and page browsing, analysing the way users use web forms etc.

browser model, type and version of operating system, screen resolution, processor architecture, lists of text fonts, plugins or devices installed, IP addresses, etc. [5] The appropriate combination of all this information allows for the building of a type of unique device fingerprint which uniquely identifies it and, therefore, differentiates each internet user unequivocally.

Through these fingerprinting techniques, upon accessing a website, the browser executes on the user's device, and without their knowledge, a series of processes with the aim of gathering sufficiently detailed information to uniquely identify it and then transmits this to the server which stores it for subsequent use. This information is combined with other data the server receives from the user's browser, the purpose of which is initially technical (for example, to adapt the contents to the terminal device's screen) but which are reused for identification purposes.

It is widely known and accepted that a specific web service can track a user's browsing using cookies, with the guarantee that deleting the cookies will remove the link between the device and the personal information gathered. The reality is that the use of the device fingerprinting techniques allow for the linked information to be reassigned to the same user when identifying the deleted cookie, to prevent the loss of the traceability of the user's browser habits or indeed for such tracking to be carried out using only the digital fingerprint itself. In conclusion, if, when an identity cookie is generated, its device fingerprint is detected and stored, when the user deletes the cookies on their browser these can be restored using the digital fingerprint to re-identify the user, making deleting the cookies ineffective.

Digital fingerprinting techniques have been described in the specialised literature as "cookieless monsters" given that it is not necessary to install any type of cookie on the device to gather the information and if this happens in a manner that is fully transparent to the user, they can have no way of preventing it (N. Nikiforakis, 2013).

Among the different techniques that may be used to obtain digital fingerprinting of a device, there are a number of particularly advanced ones such as canvas fingerprint, canvas font fingerprint, webRTC fingerprint or audio fingerprint which allow for very precise profiles to be obtained.

The use of these techniques may have legitimate purposes such as, for example, forming part of multiple factor authentication mechanisms. However, they may also be used to monitor users during their web browsing and compile information on their habits and interests without the user being conscious of it.

With regard to the duty to inform, it is common to find privacy clauses on websites and applications that allow the user to consent to the use of cookies but it is not so common to find information for the user on the use of tracking techniques based on digital fingerprinting to build user profiles.

[5] https://amiunique.org/faq

3. FINGERPRINTING TECHNIQUES

There are numerous properties that can be gathered from a device through the browser and which allow for sufficient information to be gathered so that in certain situations the terminal device can be identified unequivocally. As shown previously, some of these characteristics are widely known because they are commonly used to present applications or websites adapted to the device that accesses them. However, others are much less known and may surprise some with their level of sophistication.

Annex I includes a number, but by no means all, of the terminal device characteristics that can be gathered through a browser and that could help obtain a digital fingerprint such as the type, version and personal configuration of the browser, the set of characteristics installed that give information on the applications installed, the language, the time zone, the screen configuration and the technical elements of the terminal, IP address, etc. It also includes information on other, more advanced techniques that in normal conditions might allow for more specific identification of the device such as canvas fingerprinting, canvas font fingerprinting, webRTC fingerprinting and audiocontext fingerprinting.

In addition to the aforementioned techniques there are many other properties that could be registered and in some cases they are used to form part of the digital signature of the device such as:

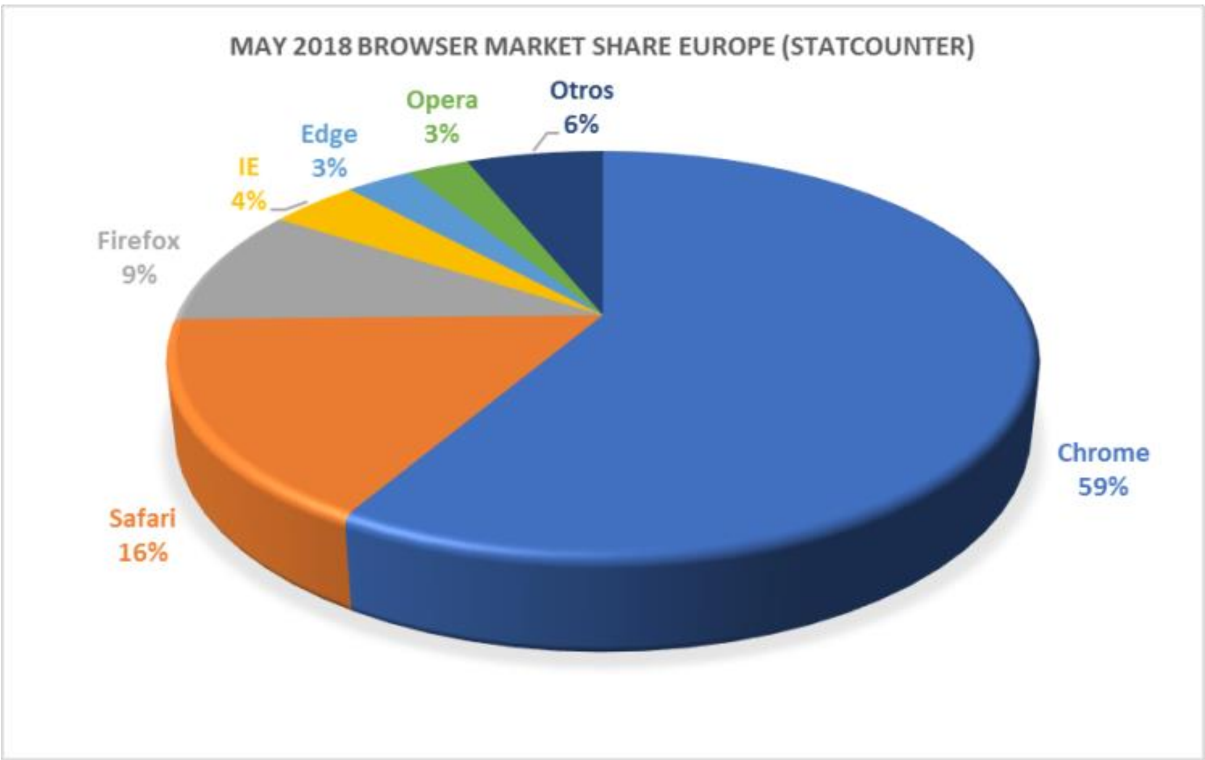
- Installation of ad-blockers on the browser
- Memory of the device
- Number of monitors connected to the device
- Device with accelerometer
- Presence of virtual keyboards
- List of supported actions in the case of multi-touch screens
- Available audio and video codecs
- Terminal device battery use profile
- List of applications installed

4. LEVEL OF IDENTIFICATION

Each of the digital fingerprint devices, on their own, would not allow for the unique identification of the device or the individual using it. However, when they combine a set of these techniques with the aim of individualising the device, the quantity of information generated offers a high probability that there are no cases between the identifiers assigned and different terminals. Without going into the detail, the quantity of information will be directly proportional to the number of properties analysed and inversely proportional to the probability that one of the characteristics is found to be present.

At present it is estimated that there are some 4 billion computers, smartphones and other terminal devices in the world. With a sufficient set of discriminating data it is possible to uniquely identify all of them and this is precisely what digital fingerprinting does. Moreover, it does so on a massive scale, as any terminal device that connects to a website that uses these techniques will be identified forever on this server and does so globally as the reach of the internet is worldwide.

In terms of web browsing, one might think that a service that attempts to obtain a digital fingerprint of a device simply detecting the model of the browser used. It seems obvious that a very efficient individual identification will not be achieved, given that looking at the statistics[6] around 59% of European users use Chrome, some 16% use Safari, 9% use Firefox and 4% use Internet Explorer with the remaining 12% using other browsers.



A web service that only detects the browser probably does so with the objective of adapting the content of the website to the user's browser. However, if in addition to the simple detection of the browser model, other characteristics are detected such as the language configured in the system, the time zone, the list of text fonts of the system, etc. and a combination of these characteristics is used to obtain a digital fingerprint of the device, the level of unique identification must be much greater and in certain circumstances it could be possible to uniquely identify a device from among all the users of a web service.

It would be useful to establish a metric that allows for the level of identification that could potentially be reached with each of the fingerprinting techniques studied and it is necessary to ask

[6] http://gs.statcounter.com/browser-market-share/all/europe/#monthly-201805-201805-bar

the following question: Is there any scientific way of quantifying the level of unique identification that could be reached?

The response is yes, and this quantification can be done by taking the concept of entropy as it is defined in the 'Theory of Information". Entropy measures in bits the degree of uncertainty in the result of any experiment or random event or, intuitively, the quantity of information provided by the occurrence of an event. For example, if an event has two equiprobable outcomes we say there is one bit of entropy, if there are four there are two bits of entropy and so on[7]. A six-sided dice would have six equiprobable results, which means the entropy or quantity of information a roll would provide is equal to $log_2 6 = 2{,}58\ bits$. If the dice is tampered with, not all the results would have the same probability and the quantity of information from each roll would be measured as $-\sum_1^6 P_n . log_2 P_n$, with P_n the probability of one result of the dice.

If we consider a world population of 7.5 billion people, the identity of an unknown person chosen at random would represent an entropy of just under 33 bits, given that 2^{33} is more than 8 billion.

As the characteristics of an individual are identified, a reduction of bits of entropy is accumulated in such a way that if the entropy is reduced 33 bits it can be said that the individual is definitively uniquely identified

For example, in the particular case of fingerprinting, it can be detected that they are using Internet Explorer, a reduction of entropy of four bits, while detecting that they are using Chrome would be a reduction in entropy of one bit[8].

In general terms, the closer a device is to generalities or to defect configuration, the fewer factors offered to make the unique identification possible. This statement is the main defence mechanism the user can exploit when it comes to minimising tracking actions used against them.

Even though, in general terms the factors detected can allow the reduction of the number of bits of entropy until full unique identification is reached, this reduction is not absolutely cumulative but depends on whether there is a correlation between the random variables. For example, knowing someone's date of birth is a reduction of entropy of 8.5 bits. Knowing another person's star sign of the zodiac is a reception of entropy of 3.6 bits. However, the knowledge of the date of birth and sign of the zodiac of the same individual does not represent a reduction of entropy of any more than 8.5 bits as the information providing a person's start sign is implicitly contained in their date of birth.

[7] (Eckersley, A Primer on Information Theory and Privacy, 2010)

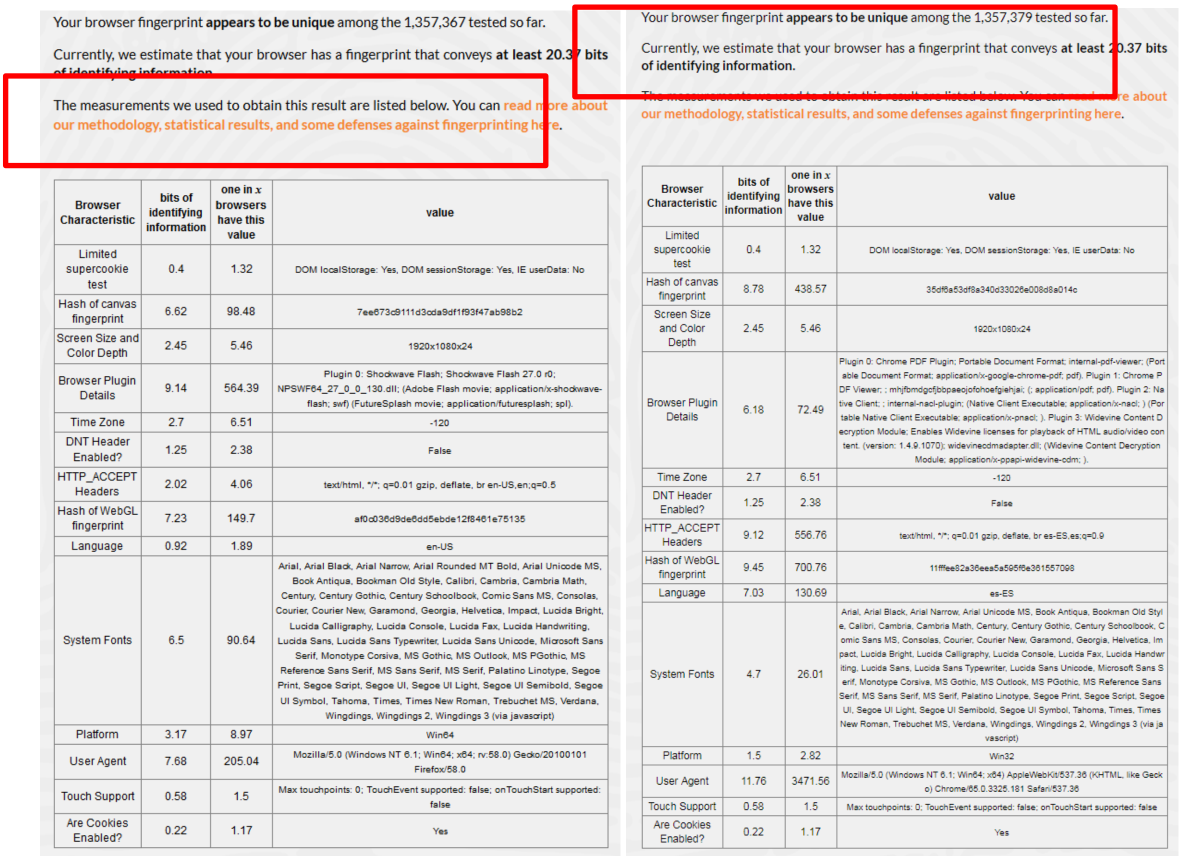
[8] As we have established, the quantity of information $\Delta S = -\log_2 P(X = x)$, where ΔS represents the reduction of average entropy through bits and $P(X = x)$ represents the probability that a specific case materializes. In more profane terms, we would say that the higher ΔS the more precise the unique identification of the user's device.

5. ASSESSMENT OF LEVEL OF IDENTIFICATION

There are various research projects that allow us to check if a browser/device is potentially identifiable through fingerprinting techniques.

- PANOPTICLICK[9].

This website performs a rapid test to check some of the techniques mentioned above. The figure below shows the result of a test performed by PANOPTICLICK with two different browsers. As can be read in the text highlighted in red, with both browsers the digital fingerprints of are unique among more than a million digital fingerprints generated in this protocol. This digital fingerprint represents a reduction of at least 20.37 bits of entropy in total. In addition, the bits of entropy estimated are indicated for each of the characteristics highlighted



Your browser fingerprint **appears to be unique** among the 1,357,367 tested so far.

Currently, we estimate that your browser has a fingerprint that conveys **at least 20.37 bits of identifying information.**

The measurements we used to obtain this result are listed below. You can read more about our methodology, statistical results, and some defenses against fingerprinting here.

Browser Characteristic	bits of identifying information	one in *x* browsers have this value	value
Limited supercookie test	0.4	1.32	DOM localStorage: Yes, DOM sessionStorage: Yes, IE userData: No
Hash of canvas fingerprint	6.62	98.48	7ee673c9111d3cda9df1f93f47ab98b2
Screen Size and Color Depth	2.45	5.46	1920x1080x24
Browser Plugin Details	9.14	564.39	Plugin 0: Shockwave Flash; Shockwave Flash 27.0 r0; NPSWF64_27_0_0_130.dll; (Adobe Flash movie; application/x-shockwave-flash; swf) (FutureSplash movie; application/futuresplash; spl).
Time Zone	2.7	6.51	-120
DNT Header Enabled?	1.25	2.38	False
HTTP_ACCEPT Headers	2.02	4.06	text/html, */*; q=0.01 gzip, deflate, br en-US,en;q=0.5
Hash of WebGL fingerprint	7.23	149.7	af0c036d9de6dd5ebde12f8461e75135
Language	0.92	1.89	en-US
System Fonts	6.5	90.64	Arial, Arial Black, Arial Narrow, Arial Rounded MT Bold, Arial Unicode MS, Book Antiqua, Bookman Old Style, Calibri, Cambria, Cambria Math, Century, Century Gothic, Century Schoolbook, Comic Sans MS, Consolas, Courier, Courier New, Garamond, Georgia, Helvetica, Impact, Lucida Bright, Lucida Calligraphy, Lucida Console, Lucida Fax, Lucida Handwriting, Lucida Sans, Lucida Sans Typewriter, Lucida Sans Unicode, Microsoft Sans Serif, Monotype Corsiva, MS Gothic, MS Outlook, MS PGothic, MS Reference Sans Serif, MS Sans Serif, MS Serif, Palatino Linotype, Segoe Print, Segoe Script, Segoe UI, Segoe UI Light, Segoe UI Semibold, Segoe UI Symbol, Tahoma, Times, Times New Roman, Trebuchet MS, Verdana, Wingdings, Wingdings 2, Wingdings 3 (via javascript)
Platform	3.17	8.97	Win64
User Agent	7.68	205.04	Mozilla/5.0 (Windows NT 6.1; Win64; x64; rv:58.0) Gecko/20100101 Firefox/58.0
Touch Support	0.58	1.5	Max touchpoints: 0; TouchEvent supported: false; onTouchStart supported: false
Are Cookies Enabled?	0.22	1.17	Yes

Your browser fingerprint **appears to be unique** among the 1,357,379 tested so far.

Currently, we estimate that your browser has a fingerprint that conveys **at least 20.37 bits of identifying information.**

The measurements we used to obtain this result are listed below. You can read more about our methodology, statistical results, and some defenses against fingerprinting here.

Browser Characteristic	bits of identifying information	one in *x* browsers have this value	value
Limited supercookie test	0.4	1.32	DOM localStorage: Yes, DOM sessionStorage: Yes, IE userData: No
Hash of canvas fingerprint	8.78	438.57	35df6a53df8a340d33026e008d8a014c
Screen Size and Color Depth	2.45	5.46	1920x1080x24
Browser Plugin Details	6.18	72.49	Plugin 0: Chrome PDF Plugin; Portable Document Format; 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; (Native Client Executable; application/x-nacl;) (Portable Native Client Executable; application/x-pnacl;). Plugin 3: Widevine Content Decryption Module; Enables Widevine licenses for playback of HTML audio/video content. (version: 1.4.9.1070); widevinecdmadapter.dll; (Widevine Content Decryption Module; application/x-ppapi-widevine-cdm;).
Time Zone	2.7	6.51	-120
DNT Header Enabled?	1.25	2.38	False
HTTP_ACCEPT Headers	9.12	556.76	text/html, */*; q=0.01 gzip, deflate, br es-ES,es;q=0.9
Hash of WebGL fingerprint	9.45	700.76	11fffee82a36eea5a595f6e361557098
Language	7.03	130.69	es-ES
System Fonts	4.7	26.01	Arial, Arial Black, Arial Narrow, Arial Unicode MS, Book Antiqua, Bookman Old Style, Calibri, Cambria, Cambria Math, Century, Century Gothic, Century Schoolbook, Comic Sans MS, Consolas, Courier, Courier New, Garamond, Georgia, Helvetica, Impact, Lucida Bright, Lucida Calligraphy, Lucida Console, Lucida Fax, Lucida Handwriting, Lucida Sans, Lucida Sans Typewriter, Lucida Sans Unicode, Microsoft Sans Serif, Monotype Corsiva, MS Gothic, MS Outlook, MS PGothic, MS Reference Sans Serif, MS Sans Serif, MS Serif, Palatino Linotype, Segoe Print, Segoe Script, Segoe UI, Segoe UI Light, Segoe UI Semibold, Segoe UI Symbol, Tahoma, Times, Times New Roman, Trebuchet MS, Verdana, Wingdings, Wingdings 2, Wingdings 3 (via javascript)
Platform	1.5	2.82	Win32
User Agent	11.76	3471.56	Mozilla/5.0 (Windows NT 6.1; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/65.0.3325.181 Safari/537.36
Touch Support	0.58	1.5	Max touchpoints: 0; TouchEvent supported: false; onTouchStart supported: false
Are Cookies Enabled?	0.22	1.17	Yes

- AmIUnique.org[10].

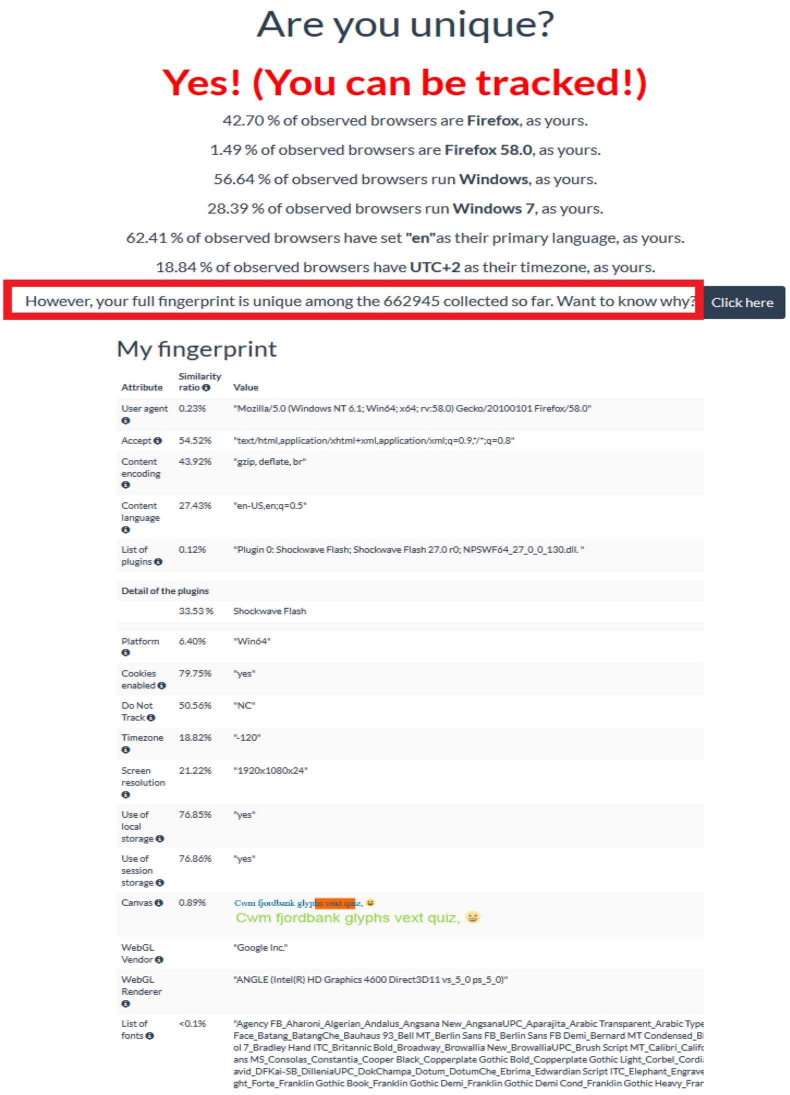
The aim of this page is investigate the use of digital fingerprints in web browsing, allowing for the user to be informed of certain details of their browser configuration and to what extent

[9] Electronic Frontier Foundation Research project. The Electronic Frontier Foundation is a non-profit organisation founded in 1990 with the aim of defending civil rights and liberties in the digital age.

[10] (Laperdrix, Rudametkin, & Baudry, 2016). This is a website created and maintained by a research group financed by the Project DIVERSIFY and the National Institute of Applied Sciences of Rennes.

they can consent to the tracking of same. Furthermore, it also intends to take advantage of information gathered on digital fingerprints to advise to configure their browser in a manner similar to other users thereby minimising the options for real tracking of browsing.

The figure below shows the result of a test performed on amiunique.org with two different browsers. As can be seen in the text highlighted in red, the complete digital fingerprintobtained is unique in more than half a million fingerprints collected so far.



Are you unique?

Yes! (You can be tracked!)

42.70 % of observed browsers are **Firefox**, as yours.

1.49 % of observed browsers are **Firefox 58.0**, as yours.

56.64 % of observed browsers run **Windows**, as yours.

28.39 % of observed browsers run **Windows 7**, as yours.

62.41 % of observed browsers have set **"en"**as their primary language, as yours.

18.84 % of observed browsers have **UTC+2** as their timezone, as yours.

However, your full fingerprint is unique among the 662945 collected so far. Want to know why?

My fingerprint

Attribute	Similarity ratio	Value
User agent	0.23%	"Mozilla/5.0 (Windows NT 6.1; Win64; x64; rv:58.0) Gecko/20100101 Firefox/58.0"
Accept	54.52%	"text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8"
Content encoding	43.92%	"gzip, deflate, br"
Content language	27.43%	"en-US,en;q=0.5"
List of plugins	0.12%	"Plugin 0: Shockwave Flash; Shockwave Flash 27.0 r0; NPSWF64_27_0_0_130.dll. "
Detail of the plugins		
	33.53 %	Shockwave Flash
Platform	6.40%	"Win64"
Cookies enabled	79.75%	"yes"
Do Not Track	50.56%	"NC"
Timezone	18.82%	"-120"
Screen resolution	21.22%	"1920x1080x24"
Use of local storage	76.85%	"yes"
Use of session storage	76.86%	"yes"
Canvas	0.89%	Cwm fjordbank glyphs vext quiz, 😃 Cwm fjordbank glyphs vext quiz, 😃
WebGL Vendor		"Google Inc."
WebGL Renderer		"ANGLE (Intel(R) HD Graphics 4600 Direct3D11 vs_5_0 ps_5_0)"
List of fonts	<0.1%	"Agency FB_Aharoni_Algerian_Andalus_Angsana New_AngsanaUPC_Aparajita_Arabic Transparent_Arabic Type Face_Batang_BatangChe_Bauhaus 93_Bell MT_Berlin Sans FB_Berlin Sans FB Demi_Bernard MT Condensed_B ol 7_Bradley Hand ITC_Britannic Bold_Broadway_Browallia New_BrowalliaUPC_Brush Script MT_Calibri_Califo ans MS_Consolas_Constantia_Cooper Black_Copperplate Gothic Bold_Copperplate Gothic Light_Corbel_Cordi avid_DFKai-SB_DilleniaUPC_DokChampa_Dotum_DotumChe_Ebrima_Edwardian Script ITC_Elephant_Engrave ght_Forte_Franklin Gothic Book_Franklin Gothic Demi_Franklin Gothic Demi Cond_Franklin Gothic Heavy_Fran

6. THE STUDY

In 2018, the Agency considered carrying out a study on the use of these techniques aimed at a Spanish audience and in Spain. To do so, the OpenWPM[11] tool was used, which is a complete

[11] OpenWPM (Web Privacy Measurement). This is the platform developed by the University of Princeton to carry out studies on Web privacy. Used in more than 20 studies carried out by different institutions OpenWPM is the free to use

framework with multiple tools to allow for the partial automation of the study and to compile browsing data on a large scale and a version of Firefox modified to register information on visits made automatically, in particular function calls executed on the device by the user and may be considered indicators of the use of fingerprinting techniques.

A function call which contains characteristics of the terminal does not necessarily mean that a website is using fingerprinting techniques, but if the call follows certain patterns or fulfils certain specific conditions, it can be deduced that the web service is potentially using these techniques.

It must also be taken into account that there are two distinct processes. On the one hand, there is the use of these functions on the user side to extract information and on the other hand there is the use or processing of this information to that might be carried out on the side of the web service provider (server).

In addition to OpenWPM, other tools have been used in the form of specific plugins for browsers and other browsers specifically developed to boost users' privacy with the aim of validating the detections made automatically with Open WPM.

On the other hand, at present, the majority of browsers allow for users' tracking preferences to be established, especially those widely used. Specifically, W3C[12] has proposed a mechanism that allows the user to express their privacy preference in a manner that a web service can deactivate their tracking techniques where the user makes a Do Not Track (DNT) request.

The DNT header field can accept two values: one in the event that the user does not want their browsing to be tracked and zero in the event that the user consents to tracking. There is also the possibility of not sending this header field in HTTP requests, therefore DNT will take the Null values which means that the user has not established a preference. The standard establishes by default that this header field is not sent unless the user activates it from the browser.

The tests carried out during the study are described below, along with the results obtained in each case.

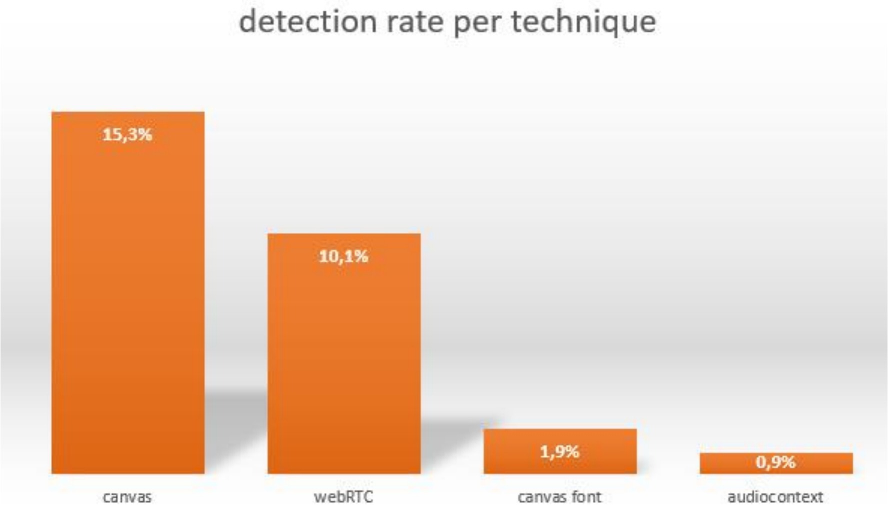
DETECTION OF FINGERPRINTING TECHNIQUES

In the first phase of the study, an analysis was carried out to detect the potential use of some advanced fingerprinting techniques, specifically canvas, webRTC, canvas font and audiocontext fingerprinting. Wirth this objective in mind, 5,006 URLs[13] corresponding to 2.503 domains were analysed, detecting the potential use of these techniques in **28.19%** of the requests from these URLs with the following details:

under GPLv3 license. Based on Python, OpenWPM allows use of Firefox to automate and simulate access to different websites compiling information such as the use of cookies, fingerprinting, tracking, etc.

[12] *World Wide Web Consortium* (W3C), is an international consortium that generates recommendations that and standards to ensure the long term growth of the World Wide Web.

[13] Web addresses



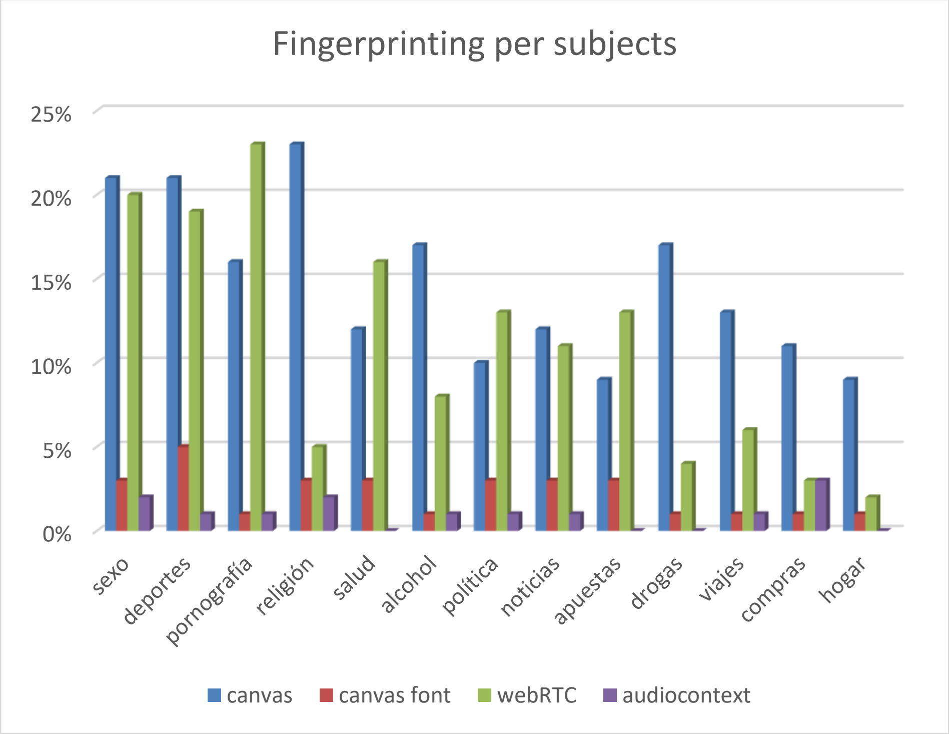
During the study, situations were detected in which several requests to the same URL showed variations in the use of fingerprinting techniques, observing the following behaviours:

- Websites that only use fingerprinting techniques when the device does not have certain cookies installed.
- Websites which initially used fingerprint detection techniques but which later, after a number of visits no longer used them.

More detailed information in the detection techniques used can be found in Annex II.

SUBJECT ANALYSIS

This phase saw analysis of the first one hundred search results on Google for the following terms: sex, drugs, alcohol, pornography, health, politics, news, sport, shopping, home, betting, travel and religion. For each search result, the homepage was also analysed.



Just as in the previous phase, the aim was to quantify the percentage of websites that potentially use advanced detection techniques such as canvas, webRTC, canvas font and audiocontext, The result is shown in the graphic above.

It can be observed that for some subjects percentages in excess of 20% were reached in the canvas and webRTC techniques. Particularly high percentages for search results for the terms sex, pornography, religion, health and politics attract attention. For the complete table with detailed percentages, see Annex III.

USE OF DO NOT TRACK REQUEST

On 5 April 2018, the Agency began a test of 14,442 websites aimed at Spanish users. The list of websites was obtained from combining all the listed URLs used in previous tests and removing any repetitions. The content of the list of URLs is very heterogeneous, including websites of all kinds: media, banking, online gaming, etc.

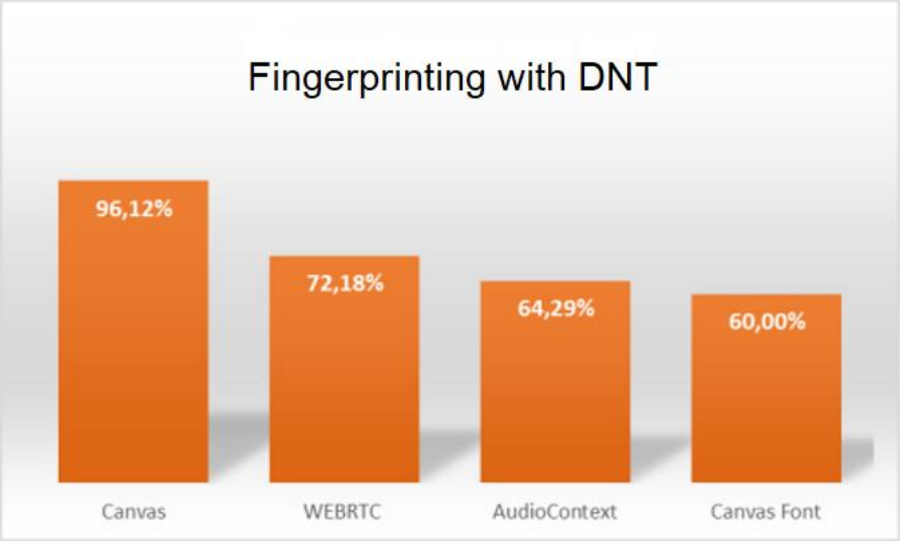
As in all tests carried out, the Flash function call was not included as it is technology that is blocked for security reasons and most of the latest browsers don’t include this technology or have it deactivated by default.

On this occasion, the focus of analysis was on verifying the use of the Do Not Track DNT) request by the websites visited. It was detected that 16.72% of sites check this parameter through javascript function calls to the user's device, which would be substantial for the data controller to detect It must also be highlighted that in 92% of cases in which the DNT parameter is used, this combination is made by third parties. These third party connections are pieces of code included on the website visited but which link to another website. For example, when visiting a website, a side panel containing ads appears. The information that appears does not originate in the website visited bot a third party page hosted on a different server.

Annex III included information on the detection of digital fingerprint techniques on the websites analysed and graphic representations with information on the main functions used.

This same sample was analysed to verify the level of compliance with the user's wishes as expressed using DNT, through an exploration of sites simulating a browser with DNT active, and discovering that fingerprint extraction programs make very diverse use of this option.

The table below shows the percentage of website visits which, despite having checked using javascript functions that the DNT request is activated by the user, continue to make suspicious function calls susceptible to use for compiling user fingerprints through the advanced techniques studied.



As can be seen, in most cases (Canvas Font) on 60% of the occasions on which the DNT option is activated, they continued to compile the fingerprint, ignoring the user's wishes.

As can be seen, in most cases (Canvas) on 96.12% of the occasions on which the DNT option is activated, they continued to compile the fingerprint, ignoring the user's wishes.,

Examining in greater detail, we were able to ascertain that it was that the case that the programme compiling the fingerprint was violating the user’s request expressed through the DNT option, but that these programs can even use the DNT request itself as an additional unique identification factor.

EFFICIENCY OF MITIGATION MEASURES

All the general use web browsers allow for extensions (also known as plugins or add-ons) to be installed which extend or modify the functionality of the browser. Among these there are some whose functionality is improved privacy for the user in the form of ad-blockers and, definitively, blockers of user tracking tools. Among the best known of these are uBlock Origin, Ghostery, Disconnect, Adguard, Adsafe and Adblock.

In the final phase of the study it is intended to assess if the installation of some of the extensions that promise to improve privacy can be of help and really efficiently deliver this functionality. Only open code options such as Disconnect, Ghostery and uBlock Origin were studied, as they are the most extensively used.

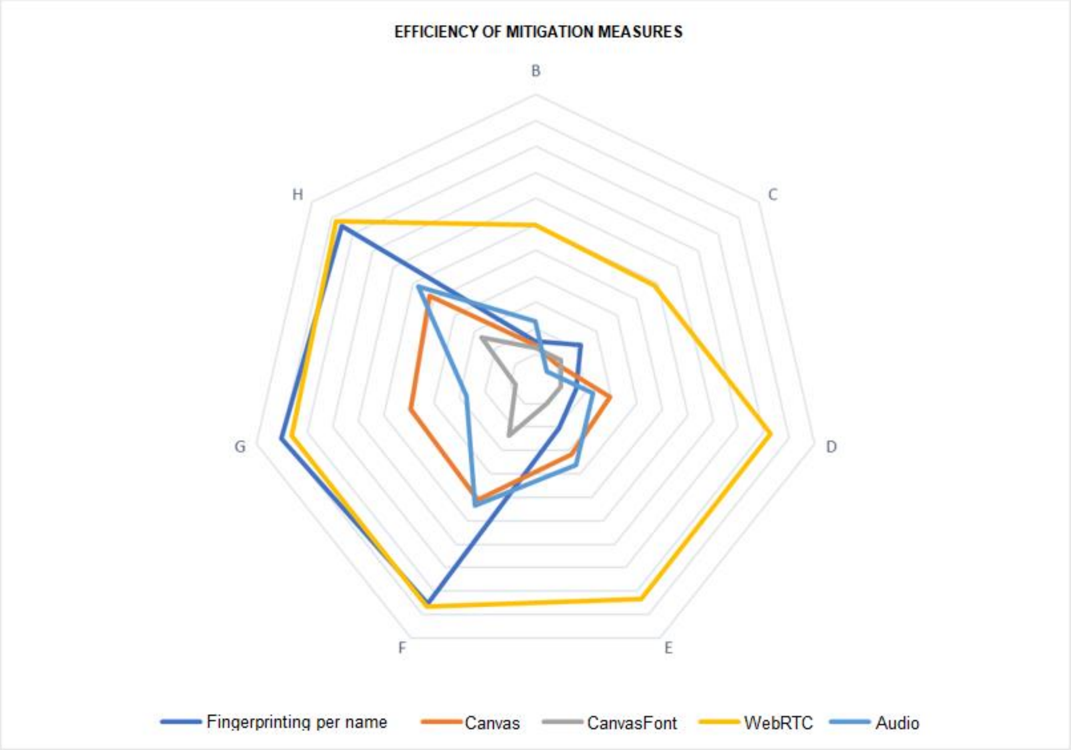
The tool OpenWPM allows for a study of this kind, carrying out automatic views with different extensions and configurations of privacy options on the browser. Thinking about minimising the study times, a list was generated containing only URLs on which fingerprinting techniques have been detected. The result is a list of approximately 1,400 websites on which to analyse the different extensions.

The following table shows the different tests that have been carried out the configuration of privacy settings and browser extensions on each of them.

	A	B	C	D	E	F	G	H
Accept 3rd party cookies	Always	Always	Never	Always	Always	Always	Always	Never
Flash disabled	No	Yes	Yes	No	Yes	Yes	Yes	Yes
Do Not Track request	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ghostery	No	No	No	No	No	No	Yes	No
Disconnect	No	No	No	Yes	Yes	No	No	No
ublock Origin	No	No	No	No	No	Yes	No	Yes

Test A is taken as the initial reference to measure the efficiency of the different configuration settings and for the rest of the tests the reduction of calls was quantified based on fingerprinting characteristics. The following table shows the results obtained during the different phases and below is the graphic representation of these data.

Technique detected	**B**	**C**	**D**	**E**	**F**	**G**	**H**
fingerprinting by function name	-5.0%	-12.0%	-5.9%	-10.5%	-85.3%	-90.1%	-85.3%
canvas	-3.6%	0.0%	-19.2%	-21.7%	-40.9%	-39.4%	-41.9%
canvas font	-2.5%	-2.5%	0.0%	0.0%	-13.8%	2.1%	-16.7%
webRTC	-49.7%	-48.5%	-82.8%	-83.4%	-86.6%	-86.0%	-87.9%
audiocontext	-12.7%	4.5%	-12.7%	-26.1%	-43.3%	-17.2%	-47.8%



As the table shows, the activation of privacy options that include browsers (tests B and C) and the disabling of Flash, send the request not to t rack and block third party cookies does not lead to significant reduction in the detection of fingerprinting techniques, except in the case of webRTC. However, the activation of ad-blockers does seem to offer some improvement against tracking by third parties through fingerprinting, producing a significant reduction of detections.

For more extensive information on the results obtained during this phase of the study, see Annex III.

7. MEASURES AVAILABLE TO THE USER

The user can protect their privacy by implementing measures available to them and thus prevent the use of fingerprinting for tracking and profiling purposes. Suggested below are a number of measures that, regrettably, are complex for the common user and make browsing difficult and have a limited effect:

- Use of the browser's Do Not Track (DNT) option

  The Do Not Track (DNT) option is the mechanism proposed by W3C[14] so that the user can express their preferences on tracking, in such a way that the web service can disable their

[14] *World Wide Web Consortium* (W3C), is an international consortium that generates recommendations that and standards to ensure the long term growth of the World Wide Web.

tracking techniques upon the request of the user. The user must visit their browser settings and enable this option, currently available for almost all browsers.

Regrettably, not all web services comply with or respect the user's DNT request, primarily because there is still no firm recommendation on the part of W3C and because the compliance requirement[15] are still only in the draft stage. In fact, some web services use this information as another factor in the digital fingerprint of the user.

- Installation of blockers

  Browser extensions (also known as plugins or add-ons) that extend or modify their functionality have become popular. Blockers are one type of such extensions, allowing the user to elude advertising and user tracking.

  Tests was carried out testing one browser with different configurations on websites on which fingerprinting activity had previously been detected. The tests compared the use of browser privacy options. As a result of these analyses the following conclusions were reached:

  ✓ Activation of privacy options included on browsers, such as disabling Flash, sending a Do Not Track request and blocking third part cookies; these measures can be effective against other techniques but did not lead to a significant reduction in detections of fingerprinting techniques.

  ✓ However, the activation of ad-blockers does seem to offer some improvement against tracking by third parties, producing a significant reduction of detections.

  Based on our observations, the blockers *Ghostery* y *uBlock Origin* stood out as the most efficient.

- Disabling use of Javascript

  The disabling of Javascript prevents the capture of data from the terminal device, although not in all cases and may prevent effective browsing on many websites.

- Alternating browser

  Using different browsers on the same device does not remove the use of the fingerprinting but it will ensure that all their information cannot be consolidated and associated with the same identifier.

  On the other hand, the use of the TOR browser masks the fingerprint of the terminal device when accessing the internet.

- Execution of access to internet in virtual machines.

  This is an option within reach for more advanced users and consists of the execution of applications that simulate devices that use different operating systems and browser configurations. This allows for internet access in a controlled environment without providing

[15] https://www.w3.org/2011/tracking-protection/drafts/tracking-compliance.html

any access to the terminal device, even if the filtering of certain information, such as the IP address, cannot be prevented.

With respect to tracking prevention measures, below we provide some notes on two of them that are not effective:

- Private browsing: Many browsers have the option of private or incognito browsing. With this option, users get the impression that their browsing is secure and is not trackable. With this option the browser does not save information on websites or browsing history, web caché, passwords, information forms, cookies or other website data, and upon closing the tab deletes all the information from the user's device. It may give the sense that browsing allows the user to be protected against the use of fingerprinting but is a false sense of security. Private browsing is transparent for the techniques used in digital fingerprinting, as the characteristics checked by the fingerprint are the same, with or without private browsing and the user will be just as uniquely identifiable. In this sense, private browsing is not effective.
- Use of anonymization networks or VPNs. Even though they prevent the disclosure of IP addresses to the destination server, they do not filter the collection data on the characteristics of the terminal. Moreover, it is necessary to be conscious of the fact that behind a free service there is a money-making strategy.

Finally, in terms of conclusions, we must add that one of the main recommendations to increase the privacy of the browser is, insofar as possible, to reduce the installation of other types of extensions on the browser.

1. One of the factors of identification by fingerprinting consists of obtaining a list of extensions or browser plugins. The more extensions installed and the further the browser configuration from the defect settings, the greater the capacity of the list of extensions to uniquely identify the user.
2. Installing extensions constitutes adding a piece of software in our browser developed by a third party removed from the development of the browser, with all the implications that involves.

8. INDUSTRY RECOMMENDATIONS

The study carried out arrived at the following recommendations for both developers of products and services for accessing the internet and for those entities that exploit the data obtained from the device fingerprint:

RECOMMENDATIONS FOR MANUFACTURERS AND/OR DEVELOPERS

Just as there are browsers that include the DNT option and various settings for accepting cookies, manufacturers and developers of devices susceptible to fingerprinting technologies should include in their products the settings necessary so that the user can avail of capacities to deny or accept, in full or in part, use of these technologies.

Moreover, they should provide the consumers said devices with the maximum security setting configured by default and the user may modify these options if they so wish. As best practice, browsers may have the DNT option activated by default.

RECOMMENDATIONS FOR ENTITIES THAT WANT TO USE FINGERPRINTING

The digital fingerprinting procedure should follow, in the terms provided for in Article 22.2 of the LSSI, which transposes Directive 2002/58/EC into Spanish Law, the requirements concerning information and obtaining consent. Requirements for the application of detailed criteria in the AEPD’s “Guide to Use of Cookies”.

Where the user has not consented to processing, the data controller must refrain from compiling and processing the fingerprint and any other data associated with same. Moreover, all applications of fingerprinting should check the status of the DNT option. If the user has activated said option, it must be interpreted as a clear negative, acting appropriately. As best practice, the providers of the service must consider activating the DNT where no preference has been established by the user.

Even where the DNT option is disabled, it should nonetheless offer the opportunity of giving prior consent for processing of fingerprinting for purposes beyond the strict provision of the service and the possibility of subsequently withdrawing that consent.

In any case insofar as device fingerprinting techniques gather personal data in accordance with the provisions of Article 4.1 and Recital 26 of the GDPR, the processing regime is subject to the provisions of same, in particular in relation to the exercise of rights.

The company must compile a register of processing activities, including processes that use fingerprinting.

They must also evaluate whether they comply with the criteria for a Data Protection Officer and contract one in accordance with the criterial set by the GDPR. The advice of this figure will be important in adapting to the GDPR.

Similarly, a data protection risk analysis must also be carried out on the rights and freedoms of those affected. If said analysis shows a high levels of risk, they will then be obliged to complete a Data Protection Impact Assessment (DPIA) to establish the necessary measures to guarantee the protection of users’ rights.

This impact assessment must consider, at least, the following risks:

- The impact of the filtration of profiling information contained in the database.
- In relation to the above, access to said information by governmental or political organisations.
- The use of social, cultural or racial bias leading to automatic decisions.

- Access by employees or third parties to specific users' data.
- The use of the data to social, political or general harassment.
- The excessive collection of data and their retention for excessive periods.
- The impact on the perception of the freedom of use of profiling information.
- The manipulation of user's wishes, beliefs and emotional state.
- In relation to the above, the risk of re-identification.

As a result of the DPIA, the Privacy requirements must be established by default, applying the Privacy measures from the Design and defining the specific requirements in relation to confidentiality, availability, integrity, authentication and traceability that, from the data protection perspective are used in the security risk management for the information systems that process said data.

9. CONCLUSIONS

From the analysis of the information in the above sections, the problems arising from digital fingerprinting are the following:

- Fingerprinting techniques collect information on the terminal devices of the user, generally without their knowledge or consent, processing information on the characteristics of the device, in some cases for purposes other than the initial technical purpose envisaged, through the execution on the terminal of applications that capture and transmit data to the data controller's server.
- The set of data collected can be so extensive, or enriched to such an extent that it can unequivocally identify the user and, among other things, find some data defined as special categories according to the GDPR. This set is, a priori, unknown. There are serious doubts as to the application of the minimisation of data principle and the period for which they can be stored.
- There are not sufficient tools provided to evade the data collection as once the accessing of a website is initiated and before the user has even been able to view it, the server already has their fingerprint information.
- Cases have even been detected where, in general, the DNT configuration established by the user to disable the collection of fingerprint information on internet services is not observed.
- The obligation to obtain informed consent and, in particular with respect to the purpose for which the data are collected, as said techniques are generally used for user profiling (including making decisions with consequences for the service) and analysing internet activity, is not adhered to
- The user has no means of exercising the rights established in the GDPR where there collect or are associated with personal data.

The impact of the use of these techniques on the rights and freedoms of users has never been analysed by the data controllers of device fingerprinting models, nor have they provided information on the measures established to minimise the risk and to prevent any breach in security.

The processing of data using device fingerprinting techniques must follow the criteria contained in the Spanish Data Protection Agency's “Guide to Use of Cookies” and the provisions of the General Data Protection Regulation on the processing of personal data.

References

Eckersley, P. (2010). *A Primer on Information Theory and Privacy*. Retrieved from Electronic Frontier Foundation: https://www.eff.org/deeplinks/2010/01/primer-information-theory-and-privacy

Eckersley, P. (2010). How Unique Is Your Web Browser? *Privacy Enhancing Technologies*.

Englehardt, S., & Narayanan, A. (2016). Online Tracking: A 1-million-site Measurement and Analysis. *Proceedings of ACM CCS 2016*.

Laperdrix, P., Rudametkin, W., & Baudry, B. (2016). Beauty and the Beast: Diverting modern web browsers to build unique browser fingerprints. *37th IEEE Symposium on Security and Privacy*.

Mowery, K., & Shacham, H. (2012). Pixel Perfect: Fingerprinting Canvas in HTML5.

N. Nikiforakis, A. K. (2013). Cookieless Monster: Exploring the Ecosystem of Web-Based Device Fingerprinting,. *2013 IEEE Symposium on Security and Privacy*, 541-555.

ANNEX I

Identifying characteristics:

Some of the characteristics that may be detected through the web browser and which might contribute to obtaining the digital fingerprint of a device are:

- User Agent: This is a string that the browser sends to the server in the HTTP request headers. This text chain contains information on the browser being used and the operating system of the device. It also contains information on the versions of the browser and operating system.
- HTTP Accept Header: HTTP Accept Header sent to the server in the HTTP requests to indicate the type of content the browser will accept in the server responses.
- HTTP Accept-Charset: HTTP Accept-Charset Header sent to the server in HTTP requests to indicate the set of characters accepted in HTTP requests, for example 'utf-8'.
- HTTP Accept-Encoding: HTTP Accept-Encoding Header sent to the server in HTTP requests to indicate the set of characters accepted in requests, for example 'gzip, deflate'.
- HTTP Accept-Language: HTTP Accept--Encoding Header sent to the server in HTTP requests to indicate the set of characters accepted in requests, for example 'en- US'.
- List of plugins activated in the browser: Through javascript, the list of plugins activated in the browser can be obtained.
- Platform on which to execute the browser: The platform on which the instance of the browser is executing, for example 'Win32', can be obtained through javascript.
- Cookies enabled: Through javascript it can detected in the browser has cookies enabled or not.
- HTTP Do not track Header: Most current browser are able to inform the websites they visit, their advertisers and their content providers that the user does not want their browsing to be tracked. This is done through a HTTP header request. Using javascript, once can detect if this characteristic has been activated or not.
- Time zone of the browser: Can be obtained using javascript.
- Resolution of the screen: Can be obtained using javascript.
- Use of local storage Javascript is used to check if the local storage can be used providing HTML5.
- Use of session storage: Javascript is used to check if the local storage can be used providing HTML5.
- WebGL Vendor: Some browsers provide complete identification for the graphic card installed on the system.
- WebGL Renderer: Some browsers provide the complete name of the graphic driver installed on the system.
- List of text fonts: Can be detected through javascript.

- ➢ Use of ad-blockers: Various checks are performed to determine the use of ad-blockers in the browser.
- ➢ Touchscreen device: It is detected if the device has a touch screen.
- ➢ Public IP with which the device connects to the internet.

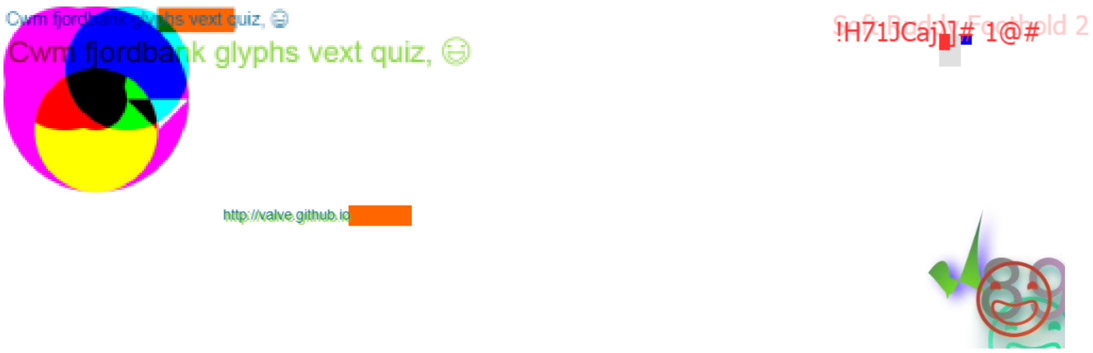
Advanced digital techniques on which this study focuses:

a) CANVAS:

Use is made of the canvas element of HTML5 to calculate a certain image through javascript. This image is rendered subtly different due to the differences of hardware/software of each device. These subtle differences can be detected for the purpose of identifying the devices.

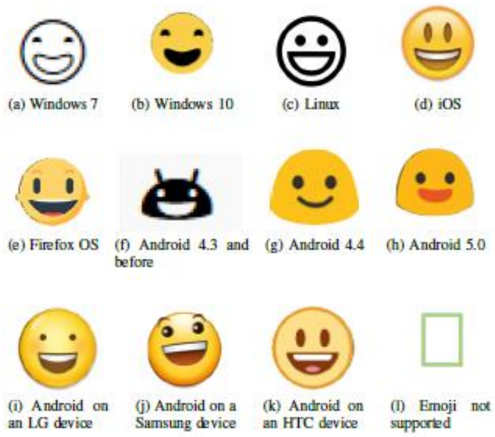
HTML Canvas is an element used to draw websites in real time using JavaScript code that is executed in the user's browser. The Canvas element is only a container for graphics, it needs to use languages like JavaScript to draw on the container. HTML Canvas has different methods that allow you to draw lines, rectangles, arches, text and to add images. This element, HTML Canvas, can be used for *fingerprinting* users (Mowery & Shacham, 2012). Differences in the rendering of fonts, smoother, anti-aliasing and other characteristics that make each device draw the image in a subtly different manner, which allows for the user's digital fingerprint to be obtained. The factor that forms part of the digital fingerprint of the device is a hash of the particular image rendered by the device.

Some examples of images generated through JavaScript on websites to identify users are shown in Figure 1. These images are rendered in the browser but are not shown to the user.


Cwm fjordbank glyphs vext quiz,
Cwm fjordbank glyphs vext quiz,
http://valve.github.io
Soft Ruddy Foothold 2
!H71JCaj)]# 1@#

➢ **Figure 1: Examples of images used in Canvas Fingerprinting**

Most strings include special Unicode characters, in which the differences in rendering are made more evident if possible. Such is the case of UNICODE character U+1F603, which represented the smiley face' emoji. Figure 2 shows the rendering differences of this character on different devices (Laperdrix, Rudametkin, & Baudry, 2016).



(a) Windows 7
(b) Windows 10
(c) Linux
(d) iOS
(e) Firefox OS
(f) Android 4.3 and before
(g) Android 4.4
(h) Android 5.0
(i) Android on an LG device
(j) Android on a Samsung device
(k) Android on an HTC device
(l) Emoji not supported

➢

➢ Figure 2: Different rendering of the same UNICODE character.

CANVAS FONT:

The list of text fonts of a device or browser is a characteristic that, together with others, can be used to obtain a unique identifier for each user (Eckersley, How Unique Is Your Web Browser?, 2010).

Canvas Font Fingerprinting is considered a variation of Canvas Fingerprinting, in which a list of text fonts are used to generate images of the same string multiple times (generally several dozen times). The variety of fonts together with the subtle rendering differences allow for metrics to be extracted from the text generated in the images which serve to generate a unique identifier of the browser.

When the browser does not facilitate the list of sources through Flash or Javascript, once can use the same technique to detect the presence of certain fonts in the system. A first rendering is carried out with a non-existent fonts that causes the browser to render with the defect source. For comparison, obtaining the metrics of the font by default allows for a list of fonts present in the system to be prepared base on rendering the same text chain with a predetermined list.

WEBRTC

This technique consists of the user of the HTML5 API WebRTC to obtain the local IP (IP behind a NAT) of a device. The local IP, combined with the Public IP, constitutes a very consistent identification factor of the device.

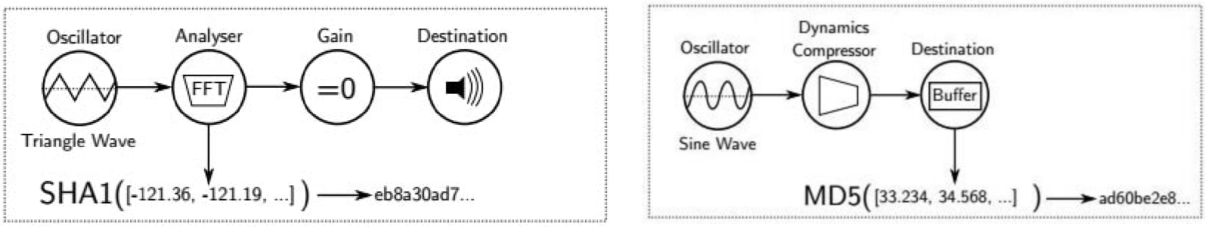
WebRTC is free and open code framework that provides browsers and mobile applications with p2p (peer to peer) Real-Time Communications (RTC) between devices To determine the best route in the network between the devices, each of them collects information on the addresses of the other, including IP addresses for local networks (Ethernet or WiFi) and directions from the public side of the NAT, making them available to the web application without the explicit consent of the user [Section 6.3 de (Englehardt & Narayanan, 2016)]. Web applications can access users' local IP addresses behind an NAT (Network Address Translation), and this information is very useful for tracking purposes. We use the example of a local network with 20 devices connected to the network and the internet through a router. A web application could access the public IP of the network/router and the private IP of each of the devices, uniquely identifying each device perfectly despite being behind a router.

AUDIOCONTEXT:

HTML5 API Audiocontext is used for a series of processes on a fixed audio signal. The slight differences in the result of processing according to the hardware/software of the specific system allow for the device to be particularized.

This type of technique works very similarly to Canvas Fingerprinting but using audio rather than images. Through the use of the AudioContext library available in most more recent browsers, the subtle differences in rendering a specific audio signal, for example a sinusoidal or triangular signal, to generate a digital fingerprint. It must be highlighted that this technique does not involve collecting audio signals played or recorded on the device, but a property of the device's audio process stack.

The two methods most used for AudioContext fingerprinting are shown in Figure 2 (Englehardt & Narayanan, 2016). Both methods process a signal generated through an *OscillatorNode* to then read the resulting signal and generate a hash that constitutes the digital fingerprint generated. The same audio signal processed by different devices or browsers will have subtle differences due to the differences of hardware/software between devices.



➢ **Figure 3: AudioContext Fingerprinting Methods**

ANNEX II

Fingerprinting detection patterns

The detection of fingerprinting techniques has been performed through the identification of certain patterns of data registered during automatic navigations with the OpenWPM tool. These patterns are based on those described in (Englehardt & Narayanan, 2016).

a) DETECTION OF FINGERPRINTING TECHNIQUES IN GENERAL

This pattern of detection is the simplest of all, as it consists of identifying within the code of the websites analysed the calls to javascript functions with names that might allude to such techniques. That is, function calls with names like *'getCanvasFingerprint'*, '*getFP*', '*getFingerprint'* are identified. This detection pattern has the advantage of permitting detection of all kinds of digital fingerprint techniques and the disadvantage of not detecting the use of digital fingerprint techniques when the names of the functions do not allude to these kinds of techniques.

Table 1 shows examples of real detections identified using the OpenWPM tool-

visit_id	func_name	short_script_url
3087	e.prototype.**getCanvasFp**	crm.clubenvero.es/mtc.js
3857	window.SN</</**FingerPrint**</t</e.prototype.getHasLiedBrowser	d1af033869koo7.cloudfront.net/psp/platform/247px.js
441	q.**getFingerPrint**	mc.yandex.ru/metrika/watch.js
215	a.prototype.**getCanvasFingerprint**	prod-js.aws.y-track.com/v5/profile-hub.min.js
216	a.prototype.**getCanvasFingerprint**	prod-js.aws.y-track.com/v5/profile-hub.min.js
3035	Fingerprint.prototype.**getCanvasFingerprint**	s3.amazonaws.com/dmp-pr-production/JScript/fingerprintjs/fingerprint.js
1371	hj.**fingerprinter**.prototype.getHasLiedBrowser	script.hotjar.com/modules-b4b50aa474eaa7a39e3ccc9eed6884eb.js
3155	b.prototype.**getCanvasFp**	static.brandcrumb.com/bbva.js
3155	b.prototype.**getCanvasFp**	static.brandcrumb.com/bc.js
2053	[1]</a.prototype.**getCanvasFp**	www.edreams.es/drmsdstl.js
2221	e.**Fingerprint**2</t.prototype.getHasLiedBrowser	www.thehotelsnetwork.com/js/hotel_price_widget.js
4075	**Fingerprint**2.prototype.getHasLiedBrowser	www.thehotelsnetwork.com/js/hotel_price_widget.js

Table 1: Javascript functions with names alluding to the use of fingerprinting techniques.

DETECTION OF FINGERPRINTING CANVAS:

To detect this technique, two patterns of identification have been used:

- Pattern C1: Calls from the same JavaScript function to canvas.toDataURL to obtain a hash of an image and canvas.fillText to write strings in an image. Specifically, those functions that call twice to canvas.fillText and once to canvas.toDataURL are particularly suspicious, even though other combinations cannot be discounted either. This pattern of identification is typical of use of canvas fingerprint and some examples of detection are shown in table 2.

visit_id	func_name	toDataURL Count	fillText Count	short_script_url
215	a.prototype.getCanvasFingerprint	1	2	prod-js.aws.y-track.com/v5/profile-hub.min.js
216	a.prototype.getCanvasFingerprint	1	2	prod-js.aws.y-track.com/v5/profile-hub.min.js
683	cv/</</dF</N[112]</d	1	2	mmesbkildq-a.akamaihd.net/FLE5J21L2U.js
861	p.prototype.getCanvasPrint	1	2	cdn3.streamlike.com/secure/player/js/clientjs.js
917	l	1	2	www.logistics.dhl/akam/10/4731bef2
1023	p.prototype.getCanvasPrint	1	2	cdn3.streamlike.com/secure/player/js/clientjs.js
1285	p.prototype.getCanvasPrint	1	2	cdn3.streamlike.com/secure/player/js/clientjs.js
2221	e.Fingerprint2</t.prototype.getCanvasFp	1	2	www.thehotelsnetwork.com/js/hotel_price_widget.js
3035	Fingerprint.prototype.getCanvasFingerprint	1	2	s3.amazonaws.com/dmp-pr-production/JScript/fingerprintjs/fingerprint.js
3155	b.prototype.getCanvasFp	1	2	static.brandcrumb.com/bc.js
3723	f	1	2	cdn.doubleverify.com/dvtp_src_internal121.js
3857	window.SN</</FingerPrint</t</e.prototype.getCanvasFp	1	2	d1af033869koo7.cloudfront.net/psp/platform/247px.js
4075	Fingerprint2.prototype.getCanvasFp	1	2	www.thehotelsnetwork.com/js/hotel_price_widget.js
4443	StripeM</t.default<	1	2	m.stripe.network/inner.html

Table 2: Detections through C1 pattern: javascript functions with 2 calls to fillText and 1 to toDataURL.

- Pattern C2: Calls from the same javascript function to canvas.fillText to render specific strings that have previously been identified as tell-tale signs of the use of canvas fingerprinting techniques. In many cases these are open code functions very accessible for any developer. This indicator produces results with practically no false positives, but in return can entail a high number of false negatives due to the use of non-identified strings.

  Examples of detections using this pattern can be seen in Table 3 with the identified signatures in bold.

visit_id	arguments	KnowText Count	short_script_url
92	{"0":"**Hel$&?6%){mZ+#@**","1":2,"2":2}	1	www.iberia.com/ibcomv3/rbrand/scripts/libs/iberialib.js
175	{"0":"**!H71JCaj)]# 1@#**","1":4,"2":8}	2	www.vueling.com/akam/10/392f2669
215	{"0":"**http://valve.github.io**","1":4,"2":17}	2	prod-js.aws.y-track.com/v5/profile-hub.min.js
683	{"0":"**al;kscja;lkdfjkAKJKJX**","1":4,"2":45}	2	mmesbkildq-a.akamaihd.net/FLE5J21L2U.js
861	{"0":"**ClientJS,org <canvas> 1.0**","1":4,"2":17}	2	cdn3.streamlike.com/secure/player/js/clientjs.js
917	{"0":"**!H71JCaj)]# 1@#**","1":4,"2":8}	2	www.logistics.dhl/akam/10/4731bef2
927	{"0":"**<@nv45. F1n63r,Pr1n71n6!**","1":10,"2":40}	1	www.adidas.es/_bm/async.js
1897	{"0":"**Cwm fjordbank glyphs vext quiz,** ","1":2,"2":15}	1	fba.omniretailgroup.net/main-bru-built.js
1897	{"0":"**<@nv45. F1n63r,Pr1n71n6!**","1":10,"2":40}	1	www.toysrus.com/_bm/async.js
3723	{"0":"**!image!**","1":4,"2":17}	2	cdn.doubleverify.com/dvtp_src_internal121.js
3857	{"0":"**Cwm fjordbank glyphs vext quiz,** ","1":4,"2":45}	2	d1af033869koo7.cloudfront.net/psp/platform/247px.js

Table 3: Scripts that use text chains typical of libraries for canvas fingerprinting

DETECTION OF CANVAS FONT FINGERPRINTING:

To detect this technique, two patterns of identification have been used:

- Pattern CF1: Calls to canvas.measure Text from the same javascript function on multiple occasions (more than 30 for example). This form of detection can be refined if all calls use the same strings and changing the text font used on each occasion.

Table 4 shows examples of detections of this kind, as the "measureTextCount" column shows the number of calls to measure Text.

visit_id	func_name	measureText Count	short_script_url
211	nds.common.bi.getFontMetrics	68	api-ob.nd.nudatasecurity.com/2.2/w/w-766580/sync/js/
212	nds.common.bi.getFontMetrics	68	api-ob.nd.nudatasecurity.com/2.2/w/w-766580/sync/js/
395	td_2C	174	regstat.betfair.com/fp/check.js
1213	td_0s	174	cdn1.f-cdn.com/fp/check.js
1227	cp<	82	www.iberia.com/ibcomv3/rbrand/scripts/libs/iberialib.js
1605	td_1T	174	datawi.pokerstars.com/fp/check.js
2599	cp<	82	www.westernunion.com/etc/clientlibs/westernunion/wu_common.js
3023	d	497	mathid.mathtag.com/d/i.js
3481	r	58	m.stripe.network/inner.html

Table 4: Scripts with functions that call to the measureText function on more than 30 occasions.

- Pattern CF2: Calls from canvas.measureText with typical strings of functions developed to perform Canvas Font Fingerprinting.
  Table 5 shows examples of real detections that comply with these parameters. Just as in the previous case, this method provides practically no false positive in exchange for potentially many false negatives for use of unidentified text chains.

visit_id	arguments	KnowFontText Count	short_script_url
211	{"0":"**mmmmmmmmmmlli**"}	68	api-ob.nd.nudatasecurity.com/2.2/w/w-766580/sync/js/
212	{"0":"**mmmmmmmmmmlli**"}	68	api-ob.nd.nudatasecurity.com/2.2/w/w-766580/sync/js/
395	{"0":"**gMcdefghijklmnopqrstuvwxyz0123456789**"}	174	regstat.betfair.com/fp/check.js
1213	{"0":"**gMcdefghijklmnopqrstuvwxyz0123456789**"}	174	cdn1.f-cdn.com/fp/check.js
1227	{"0":"**0-_{w.**"}	82	www.iberia.com/ibcomv3/rbrand/scripts/libs/iberialib.js
1415	{"0":"**mmmmmmmmmmlli**"}	58	m.stripe.network/inner.html
1435	{"0":"**mmmmmmmmmmlli**"}	58	m.stripe.network/inner.html
1605	{"0":"**gMcdefghijklmnopqrstuvwxyz0123456789**"}	174	datawi.pokerstars.com/fp/check.js
2599	{"0":"**0-_{w.**"}	82	www.westernunion.com/etc/clientlibs/westernunion/wu_common.js

Table 5: Scripts that use text chains typical of canvas font fingerprinting

DETECTION OF WEBRTC AND AUDIOCONTEXT FINGERPRINTING:

The detection of WebRTC and AudioContext fingerprinting is based on the identification of functions that make use of certain functions typical of these identification techniques.

ANNEX III

a) TABLE RESULTS OF ANALYSIS BY SUBJECT:

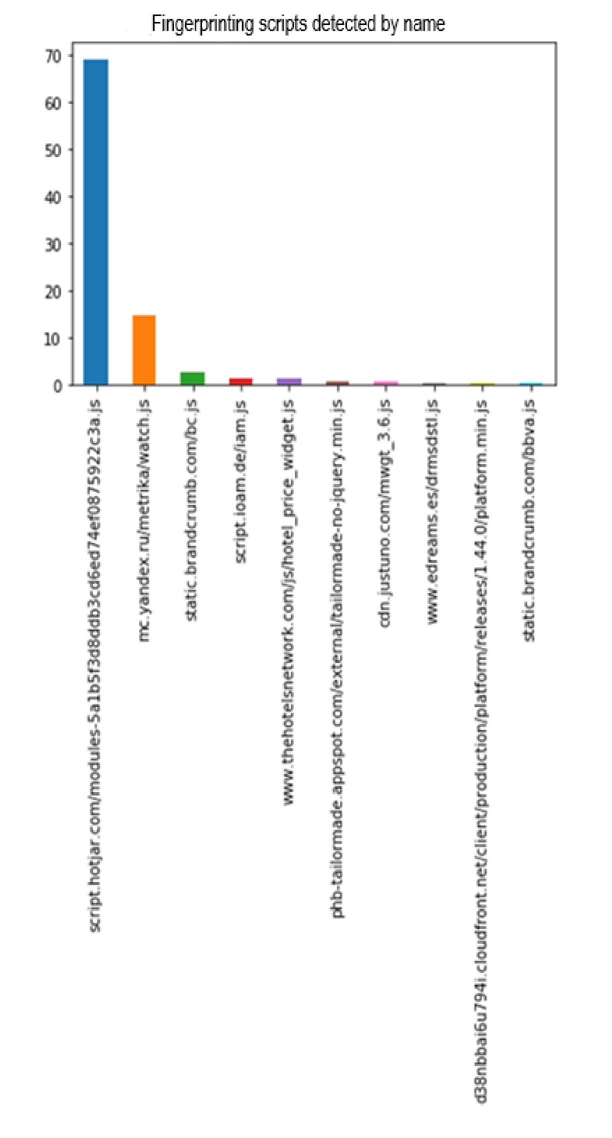
	canvas	canvas font	webRTC	audiocontext	Total
sex	21%	3%	20%	2%	46%
sport	21%	5%	19%	1%	46%
pornography	16%	1%	23%	1%	41%
religion	23%	3%	5%	2%	33%
health	12%	3%	16%	0%	31%
alcohol	17%	1%	8%	1%	27%
politics	10%	3%	13%	1%	27%
news	12%	3%	11%	1%	27%
betting	9%	3%	13%	0%	25%
drugs	17%	1%	4%	0%	22%
travel	13%	1%	6%	1%	21%
shopping	11%	1%	3%	3%	18%
home	9%	1%	2%	0%	12%

FINGERPRINTING TECHNIQUES DETECTED IN TEST OF 14,442 WEBSITES

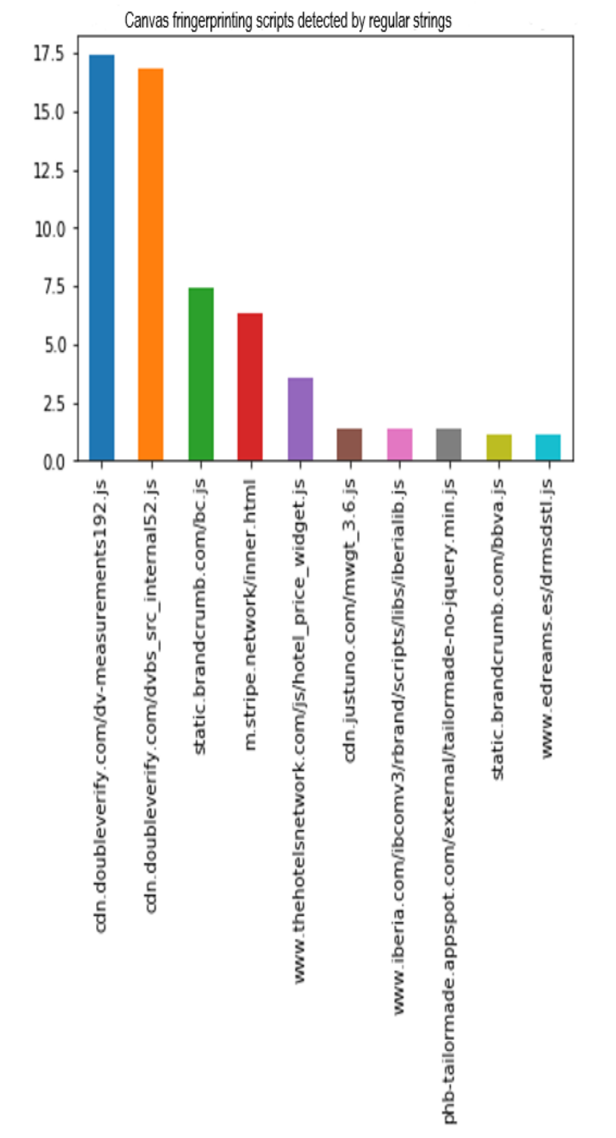
- Websites that call functions with names alluding to fingerprinting on 1,107 visits, or 7.7% of the website visits.
- Websites with functions that use calls to toDataURL and fillText (C1 identification patter) typically used for Canvas Fingerprint, on 402 visits, or 2.8% of the total of the website visits.
- Websites with functions that use strings typically used for Canvas Fingerprint (identification pattern C2), on 369 visits, or 2.6% of the total of the website visits.
- Websites with functions that perform more than 30 calls to the measureText function (identification pattern CF1), typical in Canvas Font Fingerprinting, on 38 visits, or 0.26% of the website visits.
- Websites with functions that use strings typical of for Canvas Font Fingerprint (identification pattern C2), on 43 visits, or 0.3% of the total of the websites visited.
- Websites with calls to the onicecandidate, typically used for WebRTC fingerprinting techniques to obtain the IP address of the device on 221 visits, or 1.5% of the visits.
- Websites with function calls typical of AudioContext Fingerprinting on 29 visits, or 0.2% of the website visits.

ANALYSIS OF THE MOST POPULARLY USED SCRIPTS

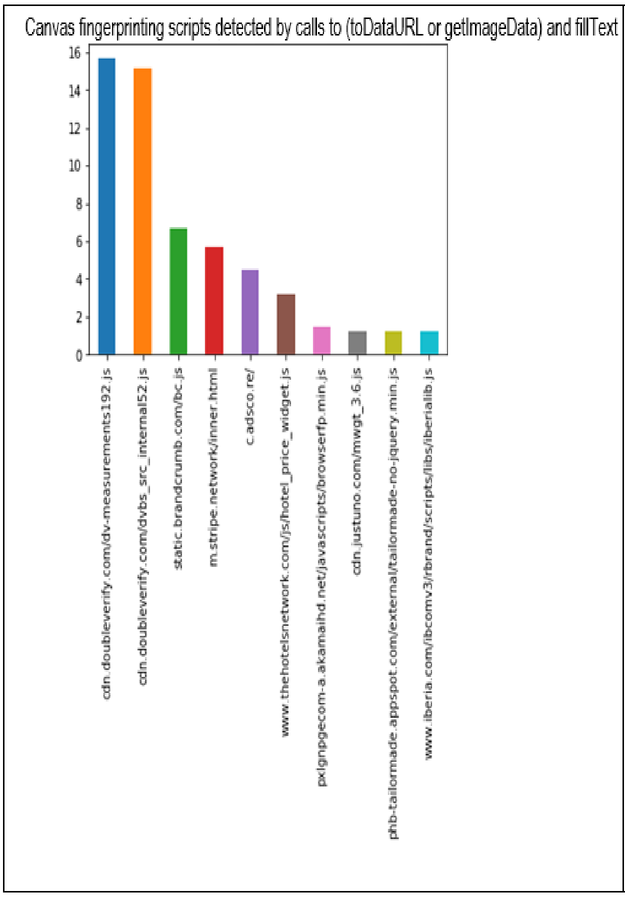
Percentage of use of some scripts where fingerprinting techniques have been detected.

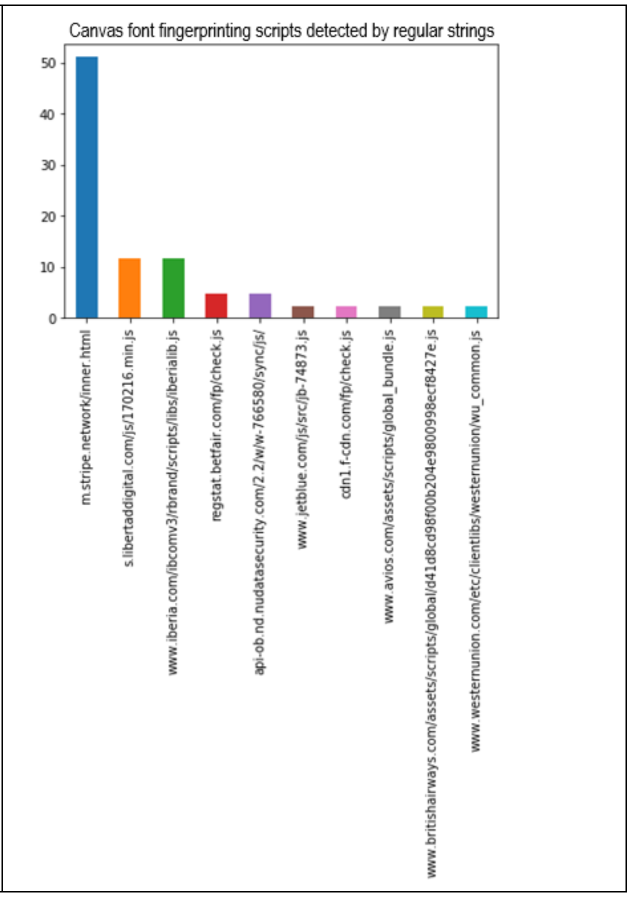


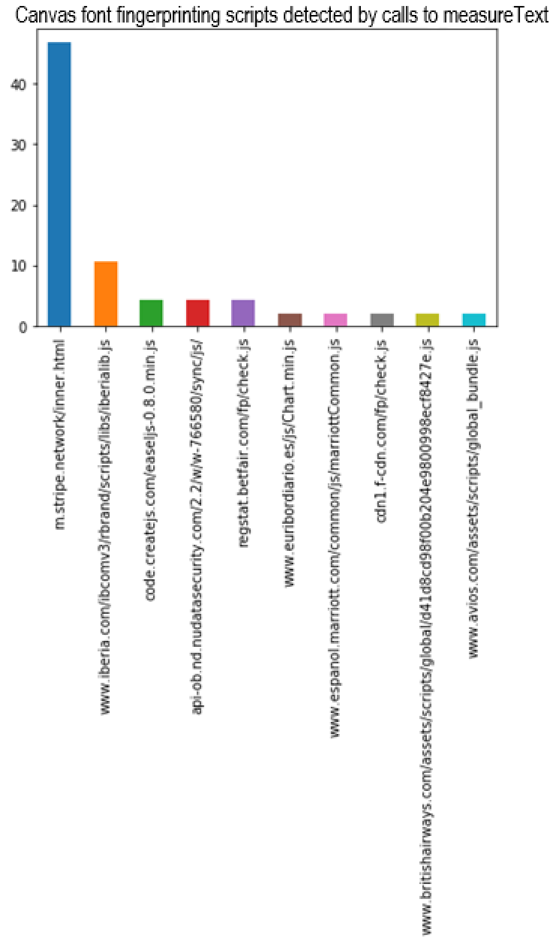
Fingerprinting scripts detected by name
70
60
50
40
30
20
10
0
script.hotjar.com/modules-5a1b5f3d8ddb3cd6ed74ef0875922c3a.js
mc.yandex.ru/metrika/watch.js
static.brandcrumb.com/bc.js
script.ioam.de/iam.js
www.thehotelsnetwork.com/js/hotel_price_widget.js
phb-tailormade.appspot.com/external/tailormade-no-jquery.min.js
cdn.justuno.com/mwgt_3.6.js
www.edreams.es/drmsdstl.js
d38nbbai6u794i.cloudfront.net/client/production/platform/releases/1.44.0/platform.min.js
static.brandcrumb.com/bbva.js



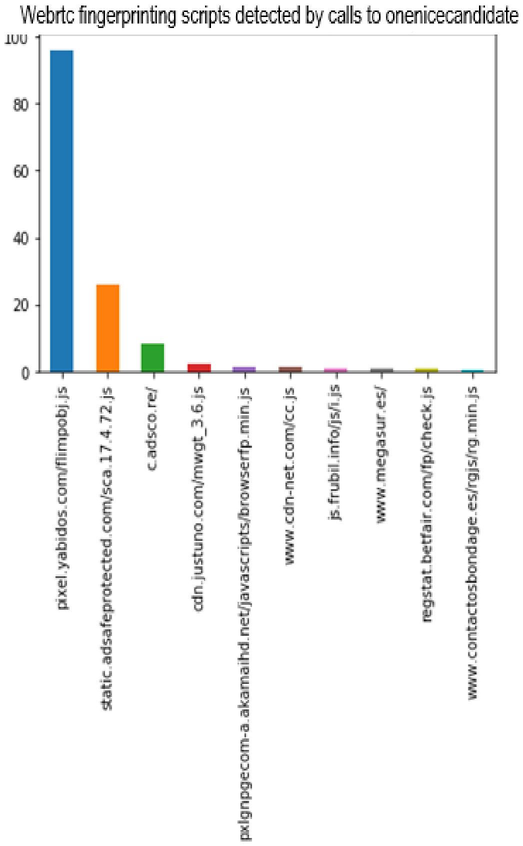
Canvas fringerprinting scripts detected by regular strings
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
cdn.doubleverify.com/dv-measurements192.js
cdn.doubleverify.com/dvbs_src_internal52.js
static.brandcrumb.com/bc.js
m.stripe.network/inner.html
www.thehotelsnetwork.com/js/hotel_price_widget.js
cdn.justuno.com/mwgt_3.6.js
www.iberia.com/ibcomv3/rbrand/scripts/libs/iberialib.js
phb-tailormade.appspot.com/external/tailormade-no-jquery.min.js
static.brandcrumb.com/bbva.js
www.edreams.es/drmsdstl.js


Canvas fingerprinting scripts detected by calls to (toDataURL or getImageData) and fillText
16
14
12
10
8
6
4
2
0
cdn.doubleverify.com/dv-measurements192.js
cdn.doubleverify.com/dvbs_src_internal52.js
static.brandcrumb.com/bc.js
m.stripe.network/inner.html
c.adsco.re/
www.thehotelsnetwork.com/js/hotel_price_widget.js
pxlgnpgecom-a.akamaihd.net/javascripts/browserfp.min.js
cdn.justuno.com/mwgt_3.6.js
phb-tailormade.appspot.com/external/tailormade-no-jquery.min.js
www.iberia.com/ibcomv3/rbrand/scripts/libs/iberialib.js

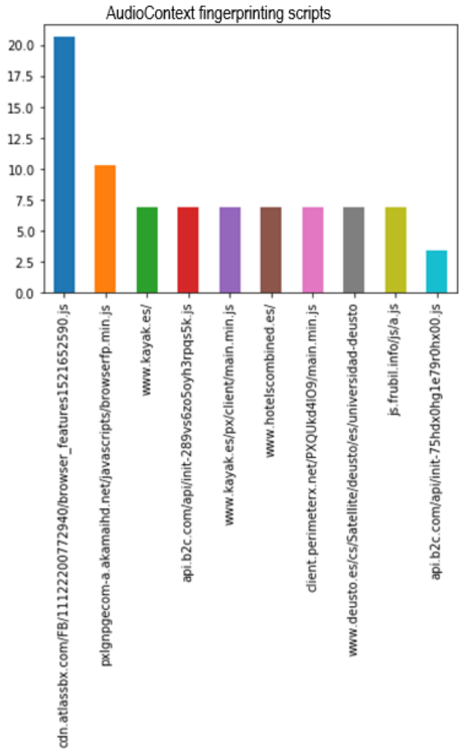

Canvas font fingerprinting scripts detected by regular strings
50
40
30
20
10
0
m.stripe.network/inner.html
s.libertaddigital.com/js/170216.min.js
www.iberia.com/ibcomv3/rbrand/scripts/libs/iberialib.js
regstat.betfair.com/fp/check.js
api-ob.nd.nudatasecurity.com/2.2/w/w-766580/sync/js/
www.jetblue.com/js/src/jb-74873.js
cdn1.f-cdn.com/fp/check.js
www.avios.com/assets/scripts/global_bundle.js
www.britishairways.com/assets/scripts/global/d41d8cd98f00b204e9800998ecf8427e.js
www.westernunion.com/etc/clientlibs/westernunion/wu_common.js



Canvas font fingerprinting scripts detected by calls to measureText



Webrtc fingerprinting scripts detected by calls to onenicecandidate



AudioContext fingerprinting scripts

USE OF DO NOT TRACK REQUEST

Fingerprinting technique	% fingerprinting + DNT[16]
canvas: detection pattern C2[17]	96.12%
canvas: detection pattern C1	93.97%
WebRTC: suspicious FP functions	72.18%
audiocontext: suspicious FP functions	64.29%
canvas: detection pattern CF2	60.00%
canvas font: detection pattern CF1	58.14%

EFFICIENCY OF MITIGATION MEASURES

Fingerprinting technique	B	C	D	E	F	G	H
fingerprinting by function name	**-5.0%**	**-12.0%**	**-5.9%**	**-10.5%**	**-85.3%**	**-90.1%**	**-85.3%**
canvas: pattern C2	**-3.6%**	**0.0%**	-13.5%	-15.1%	-35.4%	-34.4%	-36.5%
canvas: pattern C1	-2.5%	1.0%	**-19.2%**	**-21.7%**	**-40.9%**	**-39.4%**	**-41.9%**
Canvas font: pattern CF2	**-2.5%**	**-2.5%**	**0.0%**	**0.0%**	-12.5%	2.1%	-14.6%
canvasfont: pattern CF1	-2.4%	-2.4%	0.0%	0.0%	**-13.8%**	**2.8%**	**-16.7%**
webRTC	**-49.7%**	**-48.5%**	**-82.8%**	**-83.4%**	**-86.6%**	**-86.0%**	**-87.9%**
audiocontext	**-12.7%**	**4.5%**	**-12.7%**	**-26.1%**	**-43.3%**	**-17.2%**	**-47.8%**

[16] % of websites where fingerprinting techniques continue to be detected despite having previously checked that the user has activated the DNT request.

[17] Information on some detection patterns Annex II.