

## JRC TECHNICAL REPORTS

# Privacy safeguards and online anonymity

Pizzirani A., Di Gioia R., Chaudron S.,  
Draper Gil G., Sanchez I.

2017



This publication is a Technical report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication.

#### Contact information

Name: Alberto Pizzirani  
Address: Via Enrico Fermi 2749, I - 21027 Ispra (VA), Italia  
Email: [alberto.pizzirani@ec.europa.eu](mailto:alberto.pizzirani@ec.europa.eu)  
Tel.: 0332783663

#### JRC Science Hub

<https://ec.europa.eu/jrc>

JRC109792

EUR 28991 EN

---

PDF ISBN 978-92-79-77231-3 ISSN 1831-9424 doi:10.2760/30934

Ispra: European Commission, 2017

© European Union, 2017

The reuse of the document is authorised, provided the source is acknowledged and the original meaning or message of the texts are not distorted. The European Commission shall not be held liable for any consequences stemming from the reuse.

How to cite this report: Author(s), *Title*, EUR, doi

All images © European Union 2017, except:

- Frontpage: © EtiAmmos – Fotolia.com
- Page 7, Figure 1, 2017, Source: <http://www.amazon.com>
- Page 12, Figure 3, 2017, Source: Lightbeam for Mozilla Firefox <https://www.mozilla.org/it/lightbeam/>
- Page 17, Figure 5, 2017, Source: <http://www.zeit.de/datenschutz/malte-spitz-data-retention>
- Page 25, Figure 6, 2017, Source: Google Chrome, Microsoft Edge, Apple Safari and Mozilla Firefox
- Page 27, Figure 7, 2017, Source: <http://www.repubblica.it>
- Page 28, Figure 8, 2017, Source: <http://panopticlick.eff.org/>

# Contents

- 1 Introduction .....5
- 2 Motivations .....7
  - 2.1 Targeted advertising .....7
  - 2.2 Personalizing user’s experience .....8
  - 2.3 Malicious – Fraudulent – Illicit .....9
- 3 Behavioural Tracking..... 11
  - 3.1 Web tracking..... 11
    - 3.1.1 Cookies ..... 11
    - 3.1.2 Javascripts..... 14
    - 3.1.3 Etags ..... 14
  - 3.2 Social network tracking ..... 14
  - 3.3 Location tracking ..... 16
  - 3.4 Browser fingerprinting ..... 18
- 4 Profiling..... 19
  - 4.1 Definition..... 19
  - 4.2 Use Cases..... 20
    - 4.2.1 Targeted Advertising ..... 20
    - 4.2.2 Personalization of User’s Experience ..... 21
    - 4.2.3 Negative aspects of profiling ..... 21
- 5 Tools to prevent tracking and profiling ..... 23
  - 5.1 Technical means..... 23
    - 5.1.1 Privacy-Enhancing Technologies ..... 23
    - 5.1.2 Private browsing ..... 24
    - 5.1.3 Do Not Track ..... 24
    - 5.1.4 Anti-tracking browser plugins ..... 26
    - 5.1.5 TOR and VPNs ..... 29
- 6 User awareness and education ..... 31
  - 6.1 Privacy concerns..... 31
  - 6.2 Digital Competences ..... 31
  - 6.3 Privacy safeguards and online anonymity in the DigComp ..... 34
  - 6.4 Information Collection ..... 38
    - 6.4.1 Surveillance ..... 38
    - 6.4.2 Interrogation..... 39
  - 6.5 Information Processing ..... 42
    - 6.5.1 Aggregation ..... 42
    - 6.5.2 Identification..... 43

6.5.3	Insecurity .....	44
6.5.4	Secondary use.....	45
6.5.5	Exclusion .....	47
6.6	Information Dissemination .....	50
6.6.1	Breach of Confidentiality .....	50
6.6.2	Disclosure .....	50
6.6.3	Exposure .....	50
6.6.4	Increased Accessibility.....	52
6.6.5	Blackmail .....	52
6.6.6	Appropriation .....	53
6.6.7	Distortion .....	54
6.7	Invasion .....	59
6.7.1	Intrusion .....	59
6.7.2	Decisional Interference.....	59
7	Conclusions .....	63
	References .....	65
	List of figures.....	69
	List of tables .....	71

## **Abstract**

In a world that is increasingly more connected, digital citizens, actively or passively accept to transmit information, part of which are "personal data". This information is often collected and elaborated by third parties to infer further knowledge about users. The act of gathering the data is commonly called "tracking" and can be performed through several means. The act of analysing and processing those data and relate them to the individual is called "profiling".

The aim of this JRC Technical report is to be an instrument of support for the Digital Citizens to help them to protect and to manage their privacy during online activities.

After a brief introduction in Chapter 1, the following chapter is dedicated to the description of two legitimate use-cases to track and profile users on-line, namely target advertising and personalisation of the user experience. Chapter 3 and 4 identify and analyse the set of techniques currently used by online digital providers to track citizens and profile them based on their online behaviour. Chapter 5 deals with some of the available tools cited in chapter 6 that could be helpful to protect the privacy while browsing online. Chapter 6 aims to raise awareness among users and provide some guidelines to address specific issues related to privacy through a multidisciplinary approach. The report concludes highlighting the importance of raising awareness among digital users and empower them through education, technical and legal tools, including the General Data Protection Regulation (GDPR) to overcome possible privacy issues.

THIS PAGE IS INTENTIONALLY LEFT BLANK

# 1 Introduction

In a world that is increasingly more connected, digital citizens, actively or passively accept to transmit information. Some of these data, are “personal data” and they can be used to identify uniquely an individual, whereas others, even if they’re not strictly “personal” data are still related to the individual. There is a growing interest in using those data, both personal and not, and to gather further information from them.

In this scenario, the recent General Data Protection Regulation helps the citizen to better control his personal data. However, for those data that do not fall in the category of “personal data”, additional efforts have to be spent to improve the degree of control that the citizen has over them.

The act of collecting data is commonly called “tracking” and can be performed through several means. Sometimes those data are directly submitted by the citizen, whereas, in some other cases, third parties gather them through different mechanisms, such as observing the actions of the individuals.

Once these data have been collected, even if they are not personal data, they can still be processed to infer additional information of the individual. The act of analysing and processing those data and relate them to the individual is called profiling.

Whilst tracking and profiling have some legitimate use cases (for example to offer a more personalized experience to a specific visitor of a website), that is not always the case. Moreover, additional privacy issues could appear in the data flow where data are transferred, processed and used by third parties.

In Chapter 2, we present two of the typical use cases for user tracking and profiling. Chapter 3 describes possible technical means to track users online and Chapter 4 is dedicated to profiling.

Chapter 5 deals with some of the available tools that we have identified to be helpful to the data subject to protect the privacy while browsing online.

Chapter 6 is dedicated to user awareness and education. In this chapter, we started from the taxonomy of the privacy-related harms described by Solove (Solove, 2006) and then we used a multidisciplinary approach to analyse each harm. In this way, we created tables to regroup in a schematic way, for each specific harm highlighted by Solove, the possible legal means offered by the General Data Protection Regulation, the possible technical means and the digital competences involved both in the prevention and in the resolution of each privacy harm.

THIS PAGE IS INTENTIONALLY LEFT BLANK



## 2 Motivations

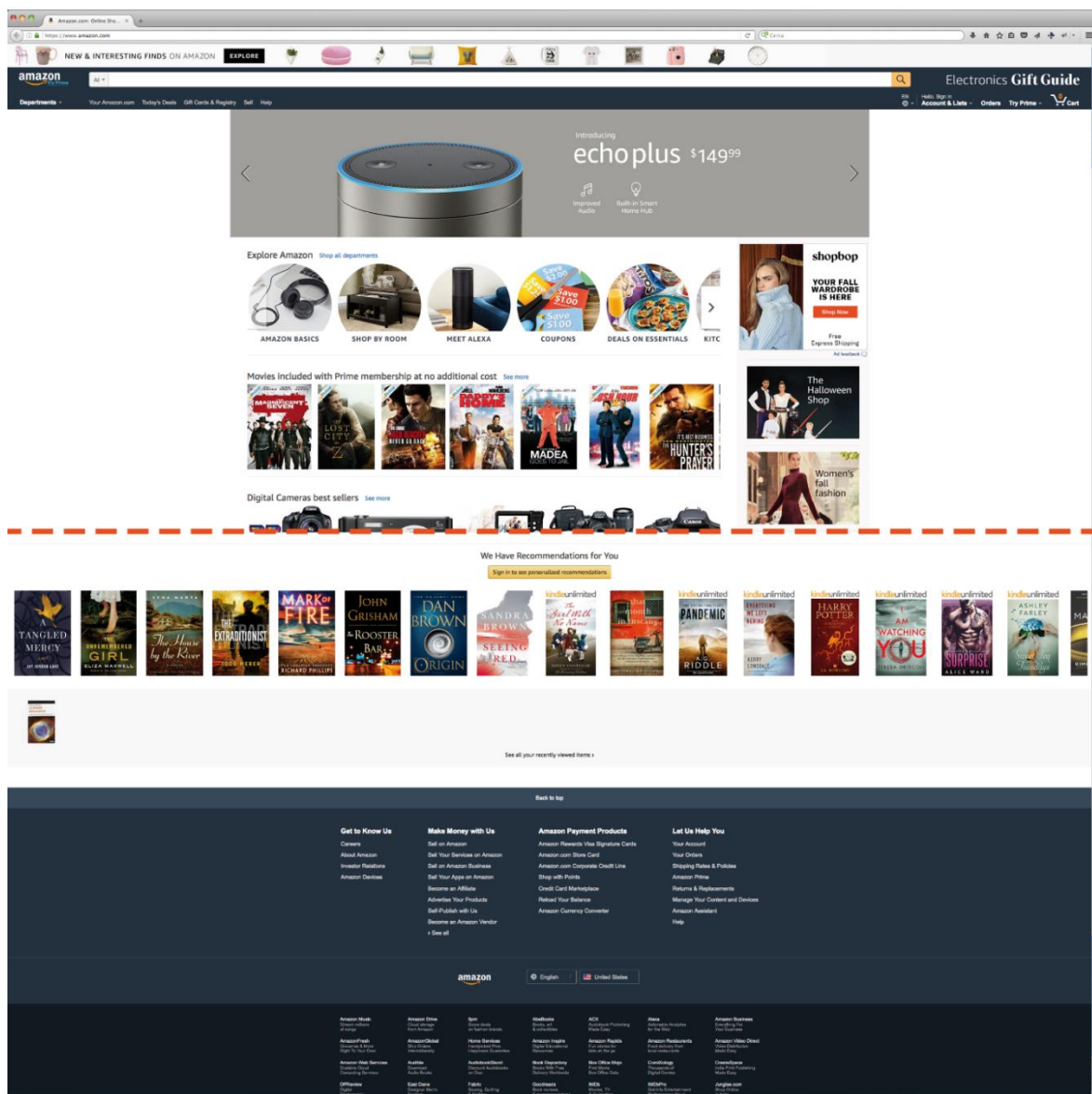
Collecting data is not harmful by itself; the problem could arise when those data are used into an unhallowed way or when some specific limit of the privacy of the individual is violated.

In this section, we give an overview of the possible scopes to process the data collected.

### 2.1 Targeted advertising

The term “targeted advertising” is commonly referred to a special form of advertising that is specifically aimed to a group of people or to a single person.

Figure 1: An example of targeted advertising where the recommended products are suggested through the analysis of the previous browsing history of the user.



This form of advertising, takes into account specific traits of the audience. The considered traits can be demographic (like race, economic status, sex, age, level of education, income

level and employment), or they can be psychographic (like attitudes, lifestyle interests, personality).

While this form of advertising can be performed also on traditional media, considering the typical audience of a specific transmission or journal, it is the one performed online that is the most refined displaying specific announces related to the website where the announce is going to appear, the content of the webpage, and the detailed information of the specific user.

Targeted advertising can be seen either as a good thing or as a malicious thing.

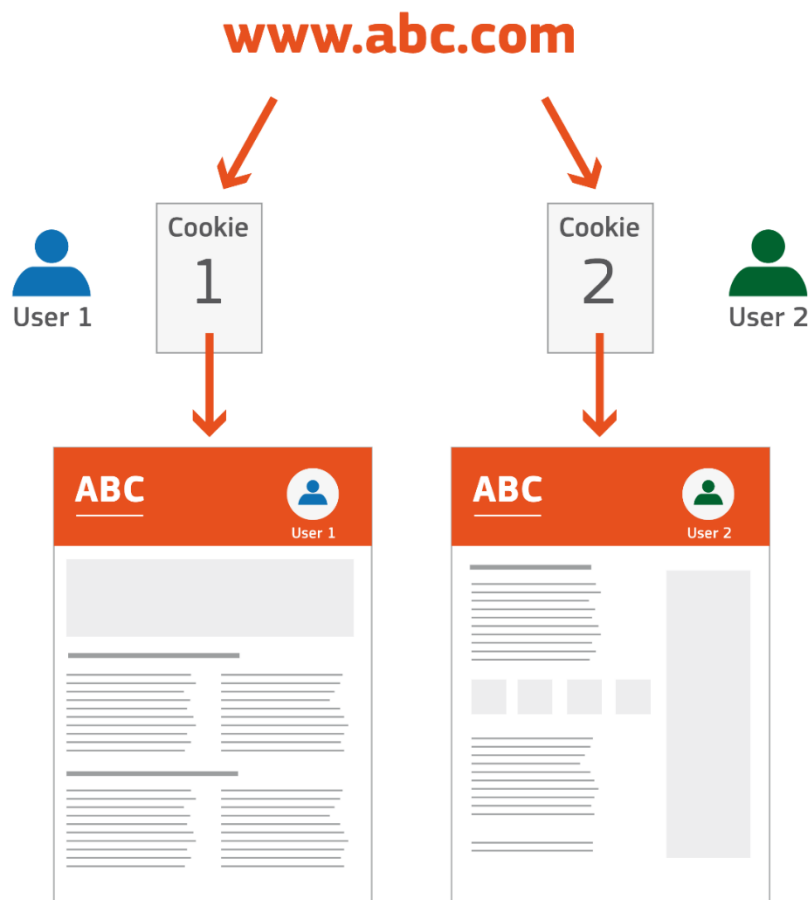
On one hand it is helpful since saves the user from specific advertisements in which he/she will (probably) never be interested, but on the other hand, it presses the right keys to push the user into buying items, and this, in some circumstances, could be dangerous when the audience of the advertisement is a easily influenced subject like a minor.

This form of advertising has evolved so much to the point of becoming a sub discipline whose main challenge is to find the best ad to show to a user when he/she is in a specific context.

## 2.2 Personalizing user's experience

Customizing user's experience is a highly requested feature by the users in several websites.

*Figure 2: The tracking of the users make it possible for a website to offer personalized contents for each one.*



The simplest usage of the cookies, to personalize the experience, is just to keep the user logged on the website, in such a way he/she do not need to login again every time he accesses the website from a trusted platform (his/her own pc, smartphone, laptop, etc...).

Another simple usage of the cookies is to store specific settings of the user, for example the layout of the page, the color scheme, the fonts and so on.

On blogs portals for example, the personalization of the page of the blog is a feature that is not only requested but in some way mandatory.

The change of appearance of the website, at higher levels, can show the user a completely different website taking into account the geographical data of the user itself, so for example a homepage in English, Italian, Japanese or Chinese according the location obtained from the device of the user.

In some way, also the targeted advertisements could be considered as a personalization of the user's experience since they just show aimed and very specific advertisements.

### **2.3 Malicious – Fraudulent – Illicit**

While the tracking and profiling of users can be used to sell specific products, it can also be used only to generate clicks for ads that follow the policy of pay per click.

Of course there is no specific threat for the user clicking on these announces, but the scope of the ad is not specifically to sell the product.

Going further, an entity that gathers information related to a subject can place itself in a position of advantage respect to the subject.

This could lead to threats related to the privacy not only online, but also in the real world (e.g. distortion, blackmail,...).

For example, the information could be gathered through malicious code executed on the machine of the data subject, or could be gathered through deceptive forms or questions, where answering, the user, could disclose information that otherwise he wouldn't.

Even if several tools are already available to protect the privacy of the users, the most powerful tool is always the education of the user.

THIS PAGE IS INTENTIONALLY LEFT BLANK

## 3 Behavioural Tracking

The phase where it is collected data related to a user that will be used to profile the user itself is often referred to as Behavioural Tracking.

With this term we refer to the collection of actions of the users in order to perform further analysis on the data collected.

Nowadays the technology offers several means to facilitate the tracking, in this chapter we offer an overview of the technical means that can be used online to perform Behavioural tracking of the users.

### 3.1 Web tracking

When we talk about "web tracking", we refer to the collection of data relative to a user, across different visits of a specific website, or across different websites.

This form of tracking is the main source of information used for profiling.

The tracking is mainly performed monitoring IP addresses and using well known techniques as the cookies or Javascripts and other less known techniques like the so called supercookies.

#### 3.1.1 Cookies

An HTTP cookie or Web cookie or simply cookie is a piece of data that is used from web pages as a token to identify the users.

The main reason to use cookies is because the HTTP protocol is stateless, so the cookies are useful to maintain stateful information.

The usage of the cookies has been disciplined both in the European legislation (European Commission, 2016) and in several national legislations.

Even if a cookie can be stored inside a text file, it is not mandatory that the container is a text file.

Usually the main fields inside a cookie are:

- Name/Value: it is a variable and it's a mandatory field
- Expiration date: this field is optional and define the date after which the cookies can be considered as expired.  
This value can be expressed as a date, number of days, "Now" (meaning that the cookies must be deleted immediately) or "Never".
- HttpOnly: it defines that the access to the cookie must be restricted strictly to HTTP.
- Secure: defines if the cookie must be sent through a secure connection (HTTPS).

"Domain" and "path" fields indicate that the cookie can be sent to the server only for the specific domain and path defined. If they're not specified, the value taken by default is the domain and path that originally requested it.

Cookies can be classified according the expiration date:

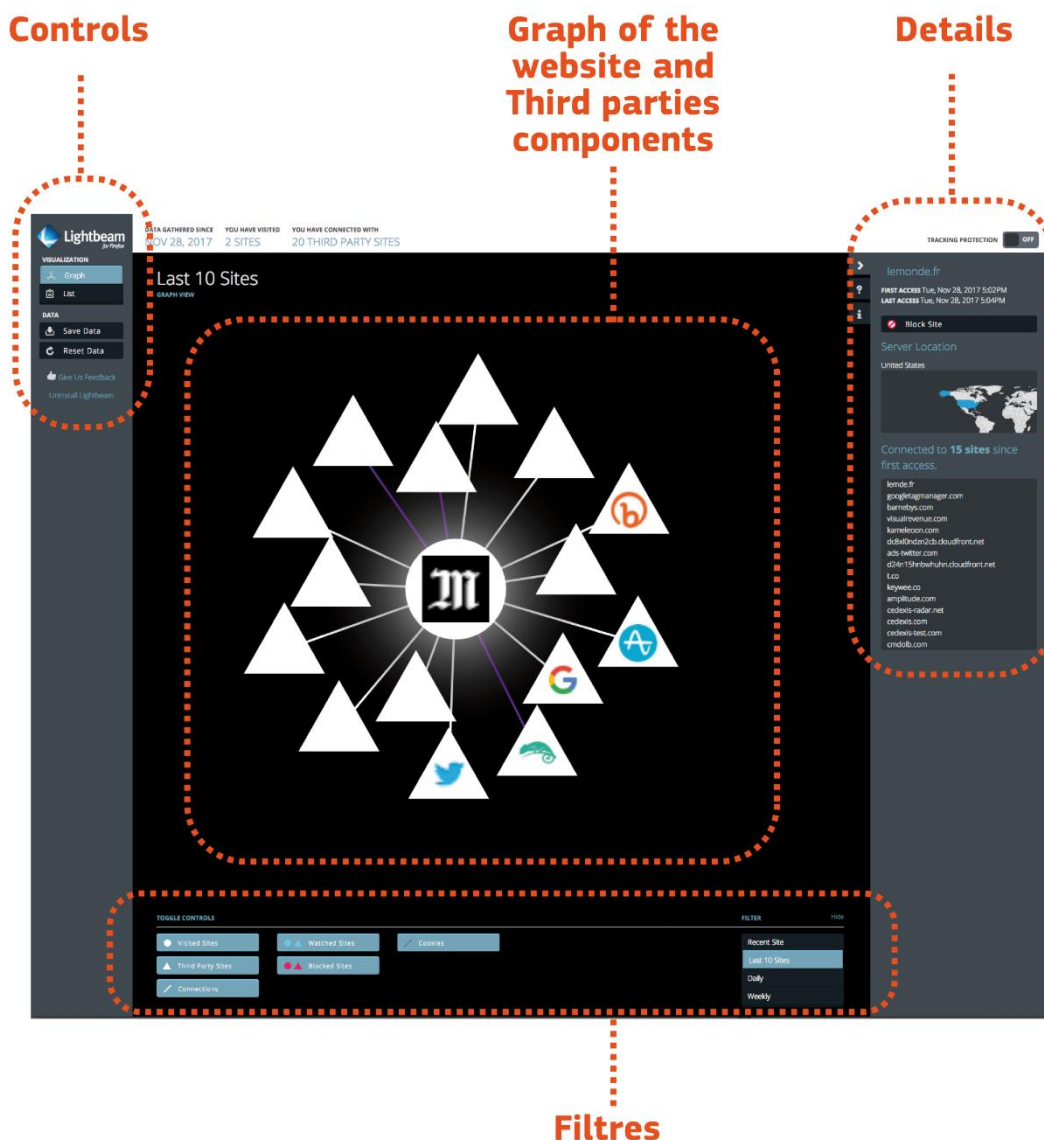
- Session cookie: these cookies will be deleted when the browser is closed. Session cookie do not have an expiration date in such way the browser can identify them.
- Persistent cookie: these cookies have an expiring date, that means that the cookie will be saved on the user's platform until that date (unless it is deleted earlier).

These cookies are useful to maintain information between separate visits to the same website (e.g. the "keep me logged" or "remember me" function on several website).

A further classification can be done according to the domain that creates them:

- First-party cookies: usually with this cookies the domain attribute will be the same of the website that the user is exploring.
- Third-party cookies: these cookies are usually saved when the web page on the website has embedded contents (like banners, web bugs, IFRAME, Javascripts, ...) coming from other domains.

Figure 3: Lightbeam is a plugin for several browser that make it possible for the user to know the third party components loaded while browsing a specific website and further information on each of them.



The use of the cookies have been regulated in the European Union, through The so-called E-Privacy Directive (Privacy and Electronic Communications Directive 2002/58/EC (European Parliament, 2002) and later amended with directive 2009/136/EC (European)).

While the directive itself recognizes the importance and usefulness of the cookies in the recital 25:

However, such devices, for instance so-called ‘cookies’, can be a legitimate and useful tool, for example, in analysing the effectiveness of website design and advertising, and in verifying the identity of users engaged in on-line transactions. 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. Users should have the opportunity to refuse to have a cookie or similar device stored on their terminal equipment. This is particularly important where users other than the original user have access to the terminal equipment and thereby to any data containing privacy-sensitive information stored on such equipment. Information and the right to refuse may be offered once for the use of various devices to be installed on the user’s terminal equipment during the same connection and also covering any further use that may be made of those devices during subsequent connections. The methods for giving information, offering a right to refuse or requesting consent should be made as user-friendly as possible. 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.

it regulates their usage with the article 5 paragraph 3:

Member States shall ensure that the use of electronic communications networks to store information or to gain access to information stored in the terminal equipment of a subscriber or user is only allowed on condition that the subscriber or user concerned is provided with clear and comprehensive information in accordance with Directive 95/46/EC, *inter alia* about the purposes of the processing, and is offered the right to refuse such processing by the data controller.

This shall not prevent any technical storage or access for the sole purpose of carrying out or facilitating the transmission of a communication over an electronic communications network, or as strictly necessary in order to provide an information society service explicitly requested by the subscriber or user.

The amendment 2009/136/EC to the directive, slightly change the article 5 paragraph 3, and instead of requiring an option for users to opt out of cookie storage, it requests a consent to the users for cookie storage.

Further recommendations and clarifications, related to cookies, are given, by the Article 29 Working Party in the “Working Document 02/2013 providing guidance on obtaining consent for cookies (Party, 2013)”.

The paper of the Article 29 explains in details what is the informed consent and how it should be requested to be valid under the european legislation, while for what concerns the recommendations given, the opinion is that there are some cases where it is not needed the explicit consent of the user for the storage of cookies that are not used for additional purposes or analytic cookies used by the originating website itself.

### **3.1.1.1 Super cookies**

Supercookies are a special kind of cookies that, differently from the common ones originated from a specific domain, are originated from a first level domain (e.g. “.com”) or from a public suffix.

Often this kind of cookies are blocked by default on the browsers, because otherwise they could be a potential security problem, since they can interfere with the requests from the users or from legitimate websites with the same extension of the domain.

A special case of Super cookies are the cookies saved through the Adobe Flash Player (AFP) application.

The AFP or more easily the Flash player, is an add on for almost all browser that usually is used to allow the reproduction of multimedia contents or to implement simple and portable application (like games) inside a web page.

The Flash cookies are downloaded or created when the flash code is executed by the flash plugin of the browser, but, unlikely conventional cookies, they're not under the control of the browser itself and the user has not direct control over them.

Furthermore there is no notification for the user when these cookies are set and, even more, they never expire.

### **3.1.1.2 Zombie cookies**

With the term "Zombie cookies", in a similar way to the fictional characters, we refer to a special kind of cookies that "comes back" after being eliminated.

This is possible because the information stored in the cookie is duplicated in multiple locations (e.g.: a flash cookie), and when the website detect that the cookie is missing can replicate the information from the secondary repository.

### **3.1.2 Javascripts**

Often websites use small Javascript files to perform several activities.

These small files are downloaded by the user, but have limited access to the user's data.

They perform computations and sometimes, since they're allowed to access the information stored in browser, they update first party cookies.

### **3.1.3 Etags**

With an increasing will to protect privacy and a better understanding of the problem by the non technical people, an higher number of persons block or delete cookies more frequently. But in this race, website with a will to track users moves towards alternative ways to track users. One of the newest methods used for user tracking is through the Etags.

The Etags are unique identifier assigned to resources (e.g. images inside a website), they are used to avoid to repeatedly download those resources if they are already cached. If the website detects that the current version of the resource is the same inside the cache of the browser of the user, it just tells the browser to use the one cached, otherwise the new resource, together with a new ETag is sent again to the browser.

The purpose of the ETags is to reduce the bandwidth consumption unusing resources already cached (if they are still valid) on the device of the user.

## **3.2 Social network tracking**

Outside the trivial aspects related to the information derived by the "social" actions performed while the user is logged on the network, like posting pictures (potentially of items like food or beverages that the user likes) or joining groups of interests or simply



putting some likes on other posts performed by friends, the relation between social networks and user tracking is a little deeper.

Some of the social networks offer as an additional service for the so-called "premium" users: the possibility to track other users.

This option is available for example:

- On job search portals to know the identity of other candidates for a position and to compare the curricula, or to see the profiles of the recruiters that visited your own profile,
- On academic social networks to know who downloaded your papers,
- On online dating websites to know further (personal) details of other users.

The possibility to know "more" about other users is usually, not only, accepted by the community, but often it is a feature requested by the users themselves, that implicitly accept to be tracked.

There is another hidden aspect of the social networks related to tracking.

Sometimes things that apparently are harmless have some hidden implications.

A typical example is the "like button" (a button typical of a huge social network displaying a blue hand with the thumb up) added by several websites to their pages to give the users a mean to show their appreciation.

The plugin that implements the button is a success from the marketing point of view with a increase in traffic that has been shown to exceed 200% on average, but according a study performed by Roosendaal (Roosendaal, 2012) the button revealed to be also an amazing tool to track users.

The button is a piece of HTML code that requests the image to the main server of the social network when the hosting website is loaded. In this way the button can be used to set third party cookies or to recognize them.

Third party cookies have always knowledge of the website hosting the plugin, since data related to the referrer is included in every HTTP request associated with the cookie.

Due to the high number of websites implementing this "feature", and the fact that the data related to multiple websites can be combined together, the cookie associated to the button can be used to build the browsing history of a web user.

Roosendaal underline that it is not needed that the user clicks on the button or not, the cookie that identify the user is sent to the server of the social network at the moment that the webpage is loaded.

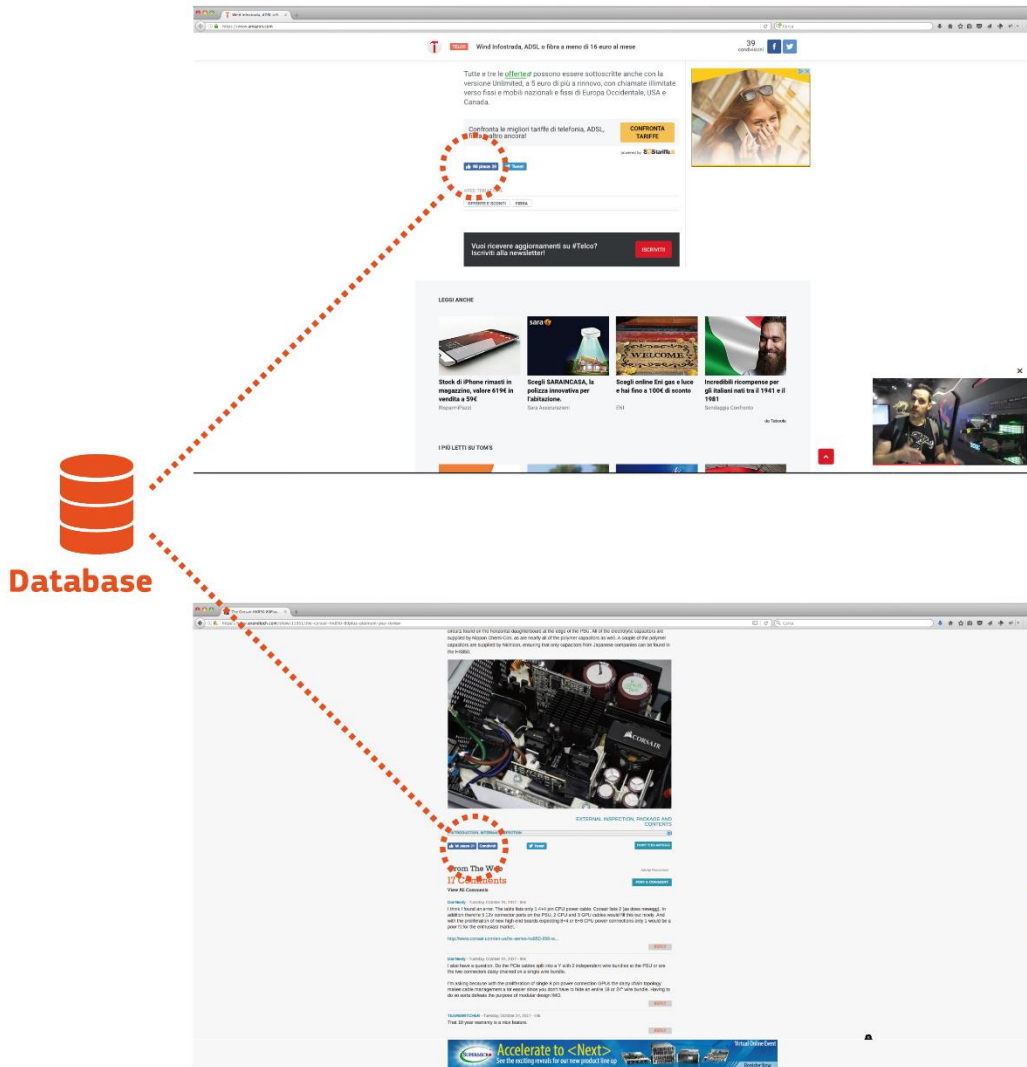
The paper of Roosendaal performs a detailed analysis of several scenarios, taking into account the differences of the cookie according the existence or not of an account on the social network of the user and if he/she is logged on the social network itself or not.

The results can be summarized in this way:

- If the user has an account, the cookie is created by the server of the social network when the account is created and is created again every time the user performs login from a new device. This cookie will be the one used by "the like button" component in the following visits to the websites implementing this feature.
- If the user does not have an account, it will be created a cookie anyway at the first visit to a website implementing a specific component of the social network. If the user later decides to create an account on the social network, the information stored in the cookie can be moved in the new cookie that identify the user.
- If the user deletes his/her account, he/she can still be tracked (through a previously stored cookie) and the browsing data can be connected to an individual data set. After deleting an account, to be considered really as a user that do not have an

account, every service connected to the social network, and every cookie related to them and the social network itself must be deleted.

Figure 4: Browsing different websites it may happen that the same resource is loaded from the same repository, making it possible for third parties to perform cross-site tracking.



### 3.3 Location tracking

In the latest years the availability of new technologies that can be moved easily in the physical world created a new kind of "location based services".

Like the name suggests, these services offer, or allow the user to perform specific actions based on their location.

Among the technologies on which these services are based, we can cite for example RFID that make it possible for the user the wireless payment of the highway toll, or the GSM, now integrated into any smartphone that can be used as a navigation system.

While these services are undeniably useful and nowadays considered essential, the market is evolving towards a different kind of location services with an increasing "social" aspect.

In the past years we have observed the appearance and an increase of interests into apps like:

- navigation systems with integrated social network to signal traffic jams or car accidents,
- apps that can tell the user the possible points of interests (restaurants, museums, attractions, etc. ...) in the surrounding based on the reviews of other users,
- apps to signal to "friends" your presence in a place or your will to go there to meet them.

As the time flows, the border between the location based services and the online social network becomes fuzzier and also the "traditional" online social networks have added options for users to obtain and make use of the location of the user.

While all of them are downloaded and installed by the user, there is an exchange of information related to the location that could be an issue for the privacy if not properly handled.

The possibility to be geolocated is not only linked to the usage of specific apps or location based services.

The use of the GSM itself is already enough to be geolocated.

In the past years, it become famous (Biermann, 2011) the case of a German politician that requested to his telecom operator six months of his phone data that later he made available for an online website.

Figure 5: ZEIT ONLINE newspaper showing the power of crossing information gathered from different sources.



This data then was combined with his public feeds, blog entries, interviews, all of them freely available online.

Then everything was put on a small application that reproduces like a movie his movements and actions on a map.

The result has a great visual impact and can give a good idea of what can be achieved when data coming from several sources, all related to a single individual are crossed together.

Another good example that illustrates the power of location tracking is the website "Pleaserobme" (<http://pleaserobme.com/>).

While the name sounds a little bit scaring, it is built with the specific scope of "Raising awareness about over-sharing". The site automatically scans Twitter feeds to find location check-ins that are being twitted out. Then it is posted a message on twitter like:

---

*Hi @NAME, did you know the whole world can see your location through Twitter? #pleaserobme.com*

---

Where "NAME" is the name of the account that revealed those information and "@NAME" is the way to alert the account owner that has been cited(tagged) into another post.

A similar work is performed by the website WeKnowYourHouse.com (linked to the omonymous Twitter account <https://twitter.com/weknowyourhouse>) that scans Twitter feeds for posts containing the word "Home" and posting a message to the owner of the feed pointing out that he's revealing the location where he lives.

Even more worrisome is the fact that a lot of people distribute information related to their exact position with anyone else (not only with their own telecom operator) and they do not even know that.

This happens, because people often distribute pictures taken by themselves, not knowing that they contain data of the location where the picture was taken.

The data related to the location are saved among the metadata of the Exif (Exchangeable Image File Format) and are taken from a gps connected with a normal camera or from the gps sensor of a smartphone.

While some of the social networks strip the images of the Exif data to protect the privacy of the users, not all of them do that.

This is because for some specific social networks, like the ones for photo enthusiast, the data relative to the location where the picture was taken are especially important, so the possibility to show them is a welcome feature.

### **3.4 Browser fingerprinting**

It is also possible to identify users, with a high level of accuracy, even if no information is saved on their machine. This is the so called "stateless tracking".

A study performed by Echersley (Eckersley, 2010) shown that online tracking tools are able to identify the browser of the user among a set of 286777 other browsers.

This is performed through the analisis of the information provided by the browser itself (User Agent, fonts, screen resolution, plugins, other flags set or not in the browser's settings ...) to the website when the user performs a connection.

## 4 Profiling

### 4.1 Definition.

Profiling is a process that transforms raw data into information that can be later used in a decision process. It applies to different fields, from psychology to law enforcement to computer science. To understand what is profiling, and to put into context in this report, we start with the definition in the Oxford English dictionary:

*"The recording and analysis of a person's psychological and behavioral characteristics, so as to assess or predict their capabilities in a certain sphere or to assist in identifying categories of people."*

It is a very general definition, but it introduces the basics of profiling as a process executed on data referring to an individual, with the objective to use the results in later decisions. In 2008 Hildebrandt et al. (Hildebrandt, 2008) elaborated a definition of profiling that extends its application to a generic subject, e.g. it can be applied to a person or a business. It also defines profiling as a process of discovery or a process of applying a profile. Profiles are defined as sets of correlated data.

*"The process of 'discovering' correlations between data in databases that can be used to identify and represent a human or nonhuman subject (individual or group) and/or the application of profiles (sets of correlated data) to individuate and represent a subject or to identify a subject as a member of a group or category."*

In 2014 the PROFILING project dedicated an entire report (Ferraris et al. (V. Ferraris, 2014)) to define the meaning of profiling, using as source the definition of Hildebrandt et al. (Hildebrandt, 2008) among others. This definition introduces the objective of making decisions as part of the profiling process, classifies the data into personal and non-personal, and it explicitly mentions that profiling is an automatic process. In their conclusions, Ferraris et al. (V. Ferraris, 2014) define profiling as:

*"Profiling is a technique to automatically process personal and non-personal data, aimed at developing predictive knowledge from the data in the form of constructing profiles that can subsequently be applied as a basis for decision-making. A profile is a set of correlated data that represents a (human or non-human, individual or group) subject. Constructing profiles is the process of discovering unexpected patterns between data in large data sets that can be used to create profiles. Applying profiles is the process of identifying and representing a specific subject or to identify a subject as a member of a specific group or category and taking some form of decision based on this identification and representation."*

In this chapter we will focus on the profiling definition of the GDPR, which constrains its previous definitions to an automated process of personal data. This definition is more precise, fits better in the context of online anonymity, the scope of this work. Article 4, (4) of the GDPR (Regulation (EU) 2016/679 of the European Parliament and of the Council, 2016) defines profiling as:

*"profiling means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements;"*

A typical classification of profiling describes it according to what a profile represent, how the profiles are built or depending on the relation of the members of a profile.

### **Individual vs Group profile**

Individual profiling is used to describe the behaviour, preferences, interests, etc. of a particular person, to the point of even being able to identify this person within a group. Group profiling describes the same traits of a group/category of individuals.

### **Explicit vs Predictive profiles**

In explicit profiling the profiles are built on data that is explicitly given by the user, e.g. when we fill in a list of topics we are interested or when we fill an online survey, whereas predictive profiling builds profiles on data obtained from observing the activity and behavior of the user, e.g. our browsing activity (types of webs we visit, queries, etc. ...). In practice, the best results are obtained using a hybrid approach, where the profiling process uses data from both sources.

### **Distributive vs Non-distributive**

In distributive profiling all the members of a profile share the same aspects, whereas in non-distributive profiling the members of a profile share only some aspects. As example, a group of parents with kids in a particular school would be distributive, everyone in the group has a kid in the same school. An example of a non-distributive group could be a list of clients with high risk in an insurance company. In this case, it is possible to end-up classified in that category without being a high risk driver, there are many variables involved.

## **4.2 Use Cases**

As use cases we present the two scenarios presented in Chapter 2: targeted advertising and personalization of user's experience. These two activities are strongly related to our daily online activity, and are two clear examples of how tracking and profiling work.

### **4.2.1 Targeted Advertising**

In the online advertising industry (G. Chen, 2016), user profiling is used to maximize the probability of an ad to be clicked. To do so, content providers rely on user profiling to decide which ad to show to a user. The set of data composed of our browsing activity, search queries, etc., allows the content providers to create a profile describing the habits, interests, etc. of an individual or group. Based on this profile, the content provider can decide which is the ad that is more likely to attract the user's interest.



In a typical architecture, we have the user, the publisher, the ad network and the advertiser. The user is the final user, the target of the ad. The publisher is the website (or application) the user is browsing. The ad network is the entity that connects publishers and advertisers. In a real scenario the architecture is more complex, we may have Ad exchangers, demand and supply platforms (Tuzhilin, 2005), etc.

Publishers sell space in their sites or applications to the ad networks, who use this space to send ads from their clients, the advertisers. When a user enters a website, within the content of the site visited he will receive one or more ads, but these ads will not be provided by the publisher, but by the ad network. The ad network again is the responsible to decide which ad to send, depending on the profile of the user that requested the ad.

A clear example that we are being profiled is when we start seeing ads related with our last Internet searches, or our last visits to online shops. Moreover, the ads can also be related to an interest that we may have, like sports or cooking, something that may not be related to a specific website but to a group of websites.

#### **4.2.2 Personalization of User's Experience**

Many online services offer the possibility of enhancing the user experience they provide, by processing our personal data. This data is usually acquired through one of the many existing tracking technologies, like the ones presented in Chapter 3, or directly from the user by answering surveys, rating items/services, etc.

Behind this personalization that suggests us articles to read, books to buy, music to listen to, etc. we usually have a so called recommender system. Recommender systems work by creating a profile (model) of the user and using this profile to offer suggestions. They are typically grouped in three categories (Tuzhilin, 2005):

*Content-based*: the recommendations are based on items/services that the user has previously liked, e.g., movies that the user has previously watched and/or reviewed positively.

*Collaborative*: The recommendations are based on items that similar users liked, e.g., music that other users with similar preferences have previously rated positively.

*Hybrid*: It is a recommender system that mixes both approaches, using data from the user itself and data from similar users.

As example, music stream services can generate music lists based on the music you listened and/or based on the music listened by other users, like the people who liked your music lists. Other examples could be services that suggest you to buy items that are similar to the ones you bought or similar to the ones that other people bought.

#### **4.2.3 Negative aspects of profiling**

There is no doubt that the use of profiling can benefit both users and advertisers in online advertising. On one hand the users can receive online ads on topics they are interested, and on the other hand, advertisers can communicate with their potential clients in a more effective way, their ads will only be sent to users who may be interested.

In the case of online services, profiling also has a positive impact in both sides. Users of online services can receive content tailored to their needs or even discover new services or content related to their interests. Providers of online services can use profiling as a tool to improve the experience of users, which can help them to obtain new users and/or keep the ones they already have.

Even though profiling has a clear positive impact, its application is not exempt of risks. In Gutwirth and Hildebrandt (2010) (Gutwirth S., 2010) the authors describe some of the concerns that the application of profiling presents. These concerns are related to privacy and data protection (data protection law applies only to personal data), dependency on decisions based on profiling, discrimination, auditability, knowledge asymmetries and transparency.

In a more practical approach, the limitations of profiling can also affect the user in a long term (e.g. If we only see content related to things we like, we will not be able to discover new topics of interest). These technical limitations have been addressed in different papers [(Požrl, 2017), (Verbert, 2016) and (Veijalainen, 2016)], where authors present a survey of previous work on recommender systems and propose research directions to improve the quality of the choices offered by recommender systems.



## 5 Tools to prevent tracking and profiling

### 5.1 Technical means

Some of the suggestions that can be given to reduce the risk of being tracked and to avoid (or at least reduce) privacy related issues are common good sense:

- Keep your operating system and programs updated to fix known security vulnerabilities,
- Keep the antivirus software updated with the latest virus signature definition,
- Be sure of what programs or apps you are going to install and if you can trust the source where you're going to download them,
- Do not disclose personal sensitive information with not trusted third parties

#### 5.1.1 Privacy-Enhancing Technologies

This term (or its acronym PETs) is commonly used to refer comprehensively to tools, methods or best practices to protect privacy and to improve the control of the data subject over his/her personal data, in accordance with the laws of data protection.

PETs are usually designed to address a single specific concern about privacy even if there are some of them that solve (or try to solve) more than one at once.

The main purpose of PETs is to improve the control of the data subject over his/her personal data or to minimize the amount of the data handled by data controllers.

There are already examples of existing PETs commonly used:

- Communication anonymizer: they aim to implement one of the most promising method to protect privacy: anonymity. There are already several services that offer methods to preserve privacy through complete anonymity or pseudonymity (i.e. anonymity that is reversible if needed). These tools operate at different levels, and among them, as examples we can cite disposable/one-time email addresses, pseudonyms for online payment, anonymizers for web browsing. Mainly these services have been proposed as "countermeasures to surveillance".
- Shared bogus online accounts: the main purpose of these accounts is to have an account for a service not related with the real identity of the user. Those accounts are created with bogus data and then the user-ID and the associated password are shared online. In such way, every user that needs an account for that service, and doesn't want to register with his/her data, can use the service anyway.
- Access to personal data: these kind of PETs are usually provided by Data Controllers in order to grant the users the right to handle their personal data easily, to fulfill the requirements of the General Data Protection Regulation.
- Enhanced privacy ID (EPID): is an algorithm for attestation of a trusted system while preserving privacy. The algorithm is compliant with the international standards ISO 20008 (ISO/IEC JTC 1/SC 27, 2013), ISO 20009 (ISO/IEC JTC 1/SC 27, 2013) and the Trusted Platform Module 2.0 (ISO/IEC JTC 1, 2015) of the Trusted Computing Group. It has been heavily supported by Intel and has been incorporated into intel chipsets since 2008 and in the processors since 2011. The idea is that each entity has a unique public verification key associated with multiple private signature keys. In this way, a device supporting EPID, could prove to an external party what kind of device it is without revealing all the informations.

The list of PETs is by no means exhaustive and it is growing more as the time passes.

In the following we will give some examples of PETs and we will describe some applications or technologies that can be helpful for the citizens to preserve privacy while performing their online experiences.

### **5.1.2 Private browsing**

Several browser are now implementing the feature of "Private Browsing" (some browsers call them Incognito Mode or InPrivate Browsing instead of Private Browsing).

While this mode is enabled, data like browsing history and search history, cookies and temporary files are only saved until the end of the private session, while downloaded files and bookmarks are saved and if needed must be deleted manually.

Some browsers perform a further step towards privacy and offer the option to log into a VPN (see later 5.1.5) while performing private browsing.

While the private browsing is not the perfect solution to the privacy problems, since it works only locally, it could be helpful for tracking problems even if with some drawbacks: if all cookies of the session are deleted, also the ones saving your settings or those that keep the user logged will be lost and the user has to re-log manually every time.

### **5.1.3 Do Not Track**

The Do Not Track (DNT) is a proposed HTTP header field that is used to request to web sites and applications to disable tracking or cross site user tracking.

The field has been proposed in 2009 by researchers Christopher, Sid Stamm, and Dan Kaminsky (Soghoian, 2011) and has been added as extension to the HTTP protocol in a document (W3C, 2015) by World Wide Web Consortium (W3C).

The main purpose of the Do Not Track header is to express the preference of a user related to the tracking behaviour of websites.

The header field can assume 3 different values: 1 if the user do not want to be tracked (opt out), 0 if the user consents to being tracked, or *null* if the user hasn't expressed yet a preference.

The extension to the HTTP protocol to include the DNT header, not only defines a way for the user to express his/her preferences regarding the tracking, but it enables also servers to communicate in a machine-readable form, their own settings for what concerns the tracking behaviour.

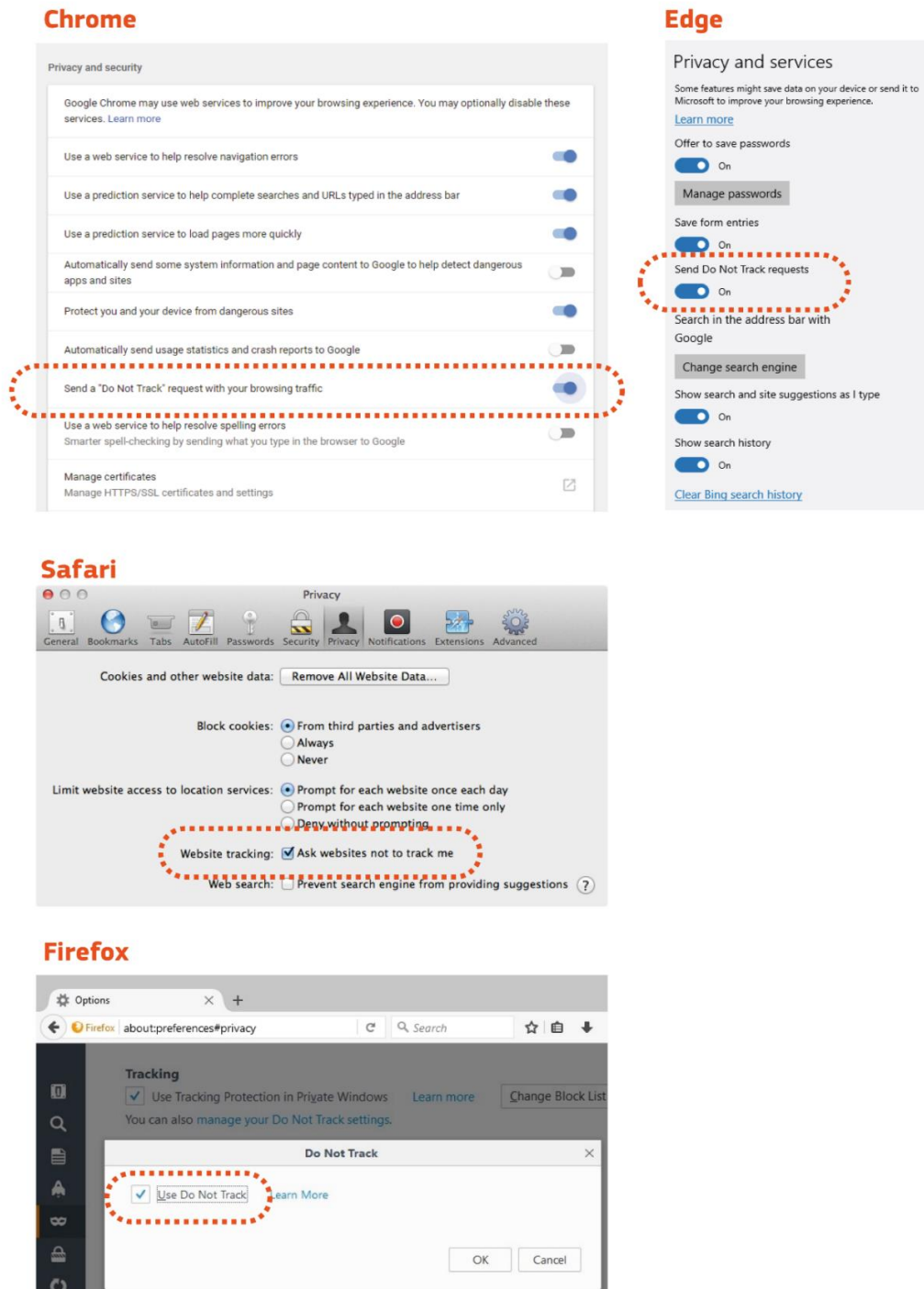
It offers means to declare:

- The identity of the site's owner (or like it's called in the GDPR the Data Controller),
- Its tracking policy (purposes for which the collected personal data is used),
- The compliance regimes it operates under,
- The other host domains it controls,
- How consent can be given or revoked,
- How its tracking behaviour has been modified in the light of a specific tracking preference.

The W3C has several guides for the correct application of the DNT as well guides for the formal verification of its correct implementation. This is useful to meet the requirements of the General Data Protection Regulation and ePrivacy directive, in particular the right to object to processing of personal data or the requirement for prior consent.

The guide to the formal check of the compliance to the DNT has been proposed on 07/03/2016 by W3C but, while it could help in the practical implementation of a useful tool for the users, there is no legal obligation for the web servers to consider the DNT header field. In this way, the servers can either honour the request of a user that do not want to be tracked (DNT field=1) and respect his/her will or they can just discard the field while processing the HTTP request.

Figure 6: How to enable Do Not Track in Chrome, Edge, Safari and Firefox.



To grant the user the ability to express their consent, even if they have the DNT field set to 1, several Java-Script Consent API (also known as the Tracking Exception API) have

been developed, allowing websites to register the user's consent. At the time of writing this report, not all browsers support the Consent API, and some of them implemented it in a proprietary way.

To help the user, some browser have their own means to handle a white list of website for which the user wants the DNT flag set to 0, for example to keep the user logged on the website or to maintain the so-called "user experience" with the customization of the website.

#### **5.1.4 Anti-tracking browser plugins**

To better protect the privacy of the users, several tools have been developed as plug in of the web browsers.

In the following we give a description of some of them that act from different perspectives.

##### **5.1.4.1 EFF plugins**

Electronic Frontier Foundation (<https://www.eff.org>) is a non profit organization whose aim is to defend civil liberties in the digital world.

The organization offers a full suite of stand alone programs and plugins for several browser to protect the privacy of the users.

The projects are released under free or open source licenses (e.g. GNU or Creative Commons) and are made publicly available.

The tools are:

- **Privacy Badger:** puts you back in control by spotting and then blocking third-party domains that seem to be tracking your browsing habits. The tool, on the other hand, allows contents from domains that respect the "Do Not Track" policy. In this way, they promote the respect for the users that want to maintain their anonymity.
- **Panoptlick:** is a tool that performs a detailed analysis of your browser fingerprint and attributes a "uniqueness score". This score should give to the users of the tool an idea of how easily identifiable they are through their browser.
- **HTTPS://EVERYWHERE:** this tool, offered as extension to several browsers, has been developed together with the Tor Project (<https://www.torproject.org/>). The purpose of this tool is to encrypt the communication whenever possible while browsing. Some websites, even if they support HTTPS, do not use it in a consistent way, or use HTTP as default or even more there are links from HTTPS pages to HTTP pages. HTTPS://EVERYWHERE by default rewrites all requests to websites to HTTPS activating encryption and HTTPS protection by default.
- **Certbot:** while the previous tool offers the users a mean to use HTTPS on the websites, this tool offers to the website administrators a simple and effective way to setup HTTPS. Certbot is a client for the Let's Encrypt (<https://letsencrypt.org/>) certification authority operated by the Internet Security Research Group (<https://letsencrypt.org/isrg/>). The tool will help the website administrator to deploy Let's Encrypt certificates with easy to follow, interactive instructions based on webserver and operating system.
- **Surveillance Self-Defence:** is a guide to introduce the user to specific topic related to the privacy and the security like threat modeling, the importance of strong passwords and protecting metadata. Together with this "educational" side, this guide helps the user to install and setup security friendly software.

### 5.1.4.2 Adblock

AdBlock is a free plugin available for all the most important web browsers on the market.

The purpose of the tool is to prevent the display of annoying pop ups that could open while browsing and to block the tracking cookies.

The program has a good impact for what concerns the protection of the privacy of the users, but, since a lot of personal blogs and news websites earn money showing advertisement, the web administrators are implementing countermeasures for people using AdBlock (e.g. showing only a small part of the contents of the page or not showing them at all).

It's usage is really simple, with a "deny by default" policy, but, if needed, every website can be added to a white list.

Figure 7: Example of website before and after the activation of AdBlock with the removed ads highlighted.





### 5.1.4.3 Anti-browser-fingerprint

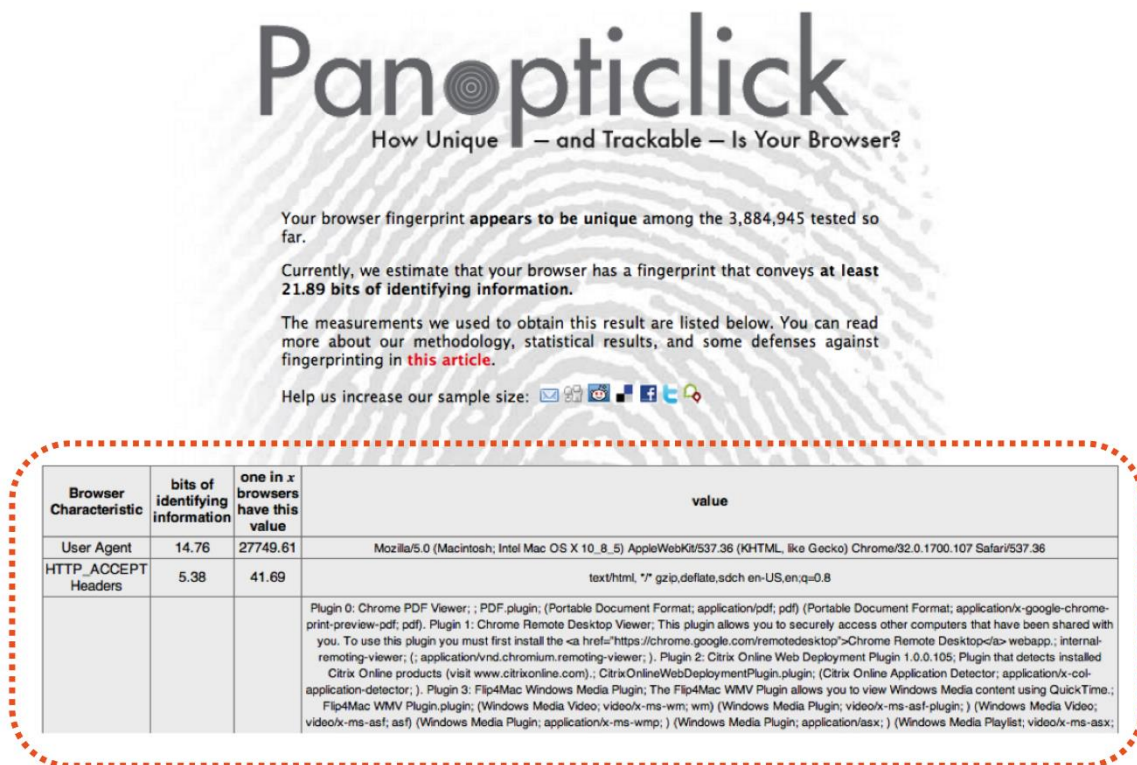
As said earlier (see paragraph 3.4) there are means to track users that are not based on data stored on the personal computer or smartphone of the user.

To have a simple idea of how easily identifiable the user is through his/her own browser it is enough to use the Panoptick tool provided by EFF, that gives an idea of how easily the user could be identified according only to his/her browser fingerprint.

Unless the user do not want to change every time the parametes of the operating system and the browser itself, to generate a different fingerprint, there are plugins for several browsers or browsers built ad-hoc that prevent the fingerprint or confuse the methods that create the fingerprint.

The idea behind this plugins is to sent false data related for example to the browser or the operating system or the language settings of the user.

Figure 8: Analysis of the browser performed by Panoptick with highlighted the specific settings of the browser taken into account for browser fingerprinting.



### 5.1.5 TOR and VPNs

There are some cases where avoiding fingerprinting or just deleting cookies is not enough to protect the privacy of the individual.

Some users need a further layer of protection in cases where the traffic of network that they use is under surveillance or there are serious threats for the freedom or the life of the individual.

In recent years, for example, some government institutions or regimes performed mass surveillance and data collection, but without going so far, a user would just like to protect his own confidential business activities and relationship or even just the privacy while being connected to a public wi-fi hot spot.

The use of a Virtual Private Network (VPN) could help in doing so.

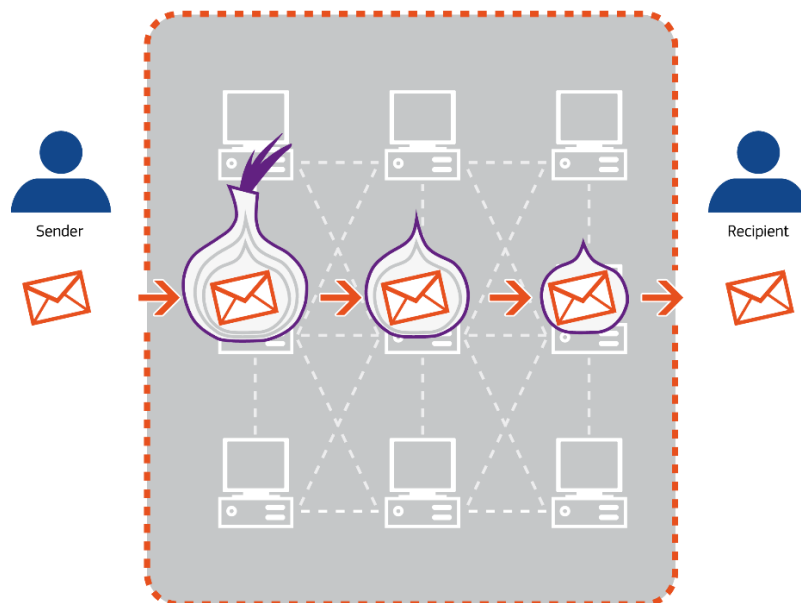
A VPN extends the concept of private network across a public network. This is achieved creating a (virtual) point to point connection through wide area networks.

Typically the VPNs only allow authenticated remote access using tunneling protocols and encryption of the data, assuring in this way:

- **Confidentiality:** if the traffic is intercepted at the packet level, the eavesdropper will only collect encrypted data.
- Sender **authentication:** because the access to the VPN is prevented to unauthorized users.
- Message **integrity:** if the VPN supports also IPsec, an hash of the message will be used to verify the message itself upon reception.

A step ahead towards privacy protection is the use of TOR. TOR is the acronym of "The Onion Router", and the idea behind this project is to route the traffic coming from a user through an overlay network. The term "onion" comes from the fact that the messages through the network are encapsulated into several layers of encryption, just like the layers of an onion. When the packets pass through the mesh of the network, at every hop, a layer of encryption is removed, with the latest layer removed when the message is received by the recipient.

*Figure 9: The flow of a message inside the TOR network from a sender to a recipient.*



This kind of onion encryption assures that, at every hop, the current node knows only the next hop and the previous hop, but not the full virtual circuit and neither the content of

the message, while a potential eavesdropper will only collect encrypted data. Furthermore the virtual circuit through the nodes that messages follow is recreated periodically by the TOR network.



## 6 User awareness and education

Nowadays more and more users share huge amounts of data and personal information, sometimes without being fully aware of what and with whom they are sharing those data and information.

Even the term “privacy” has not a worldwide-recognised definition and it is opened to discussion and to cultural influences. As general attitude, privacy concerns are addressed by using the construct of perceived risk and trust.

As proposed by first privacy theorists (Warren & Brandeis, 1890) the “right to privacy” is acknowledged by European Union promoting a rights-based capacity building model, in line with new General Data Protection Regulation and the Digital4Development approach.

### 6.1 Privacy concerns

To assess EU citizens’ perception on privacy and data protection some initiatives have been put in place. In the EU, according to the Euro barometer on data protection, the protection of personal data is seen as an important concern for citizens. The central finding of the survey shows that trust in digital environments remains low. Two-thirds of the respondents (67%) say that they are worried about having no control over the information they provide online, while only 15% feel they have complete control. This outcome confirms the need to finalise the data protection reform.

In the frame of FP7-Security Program funding the PACT ( (European Commission, 2016), (PACT project, 2014)) has been run from 2012 to 2015 with the aim to assess existing knowledge about the relation between security, privacy, trust and concern.

The PACT project has contributed to the further understanding of the sensitivity of the privacy-security relations. PACT provided insights on the tensions and arrangements raised by individuals in different fields, however further research is needed when it comes to aspects of everyday life in technology-dense societies.

Finally, a practical lesson learned concerns the importance of dissemination and popularisation.

### 6.2 Digital Competences

Education and user awareness are fundamental dimensions of an effective privacy safeguards strategy that also relies on the skills of the people concerned.

As also set out in the last Joint Communication to the European Parliament and the Council (European Commission, 2017) to better place EU to face cybersecurity and privacy threats, EU needs to affirm a resilient and complete strategy to boost citizen’s skills in term of technology, awareness and education.

To respond to this need, already in 2013 the Institute of Prospective Technological Studies of the Directorate General JRC, on behalf of DG Education and Culture and after on behalf of DG Employment, has developed and published a detailed **Digital Competence framework** (Brecko Barbara, 2017). This framework, developed with intensive consultation of stakeholders, is tied to needs that every citizen faces while interacting with digital devices and environments and it has become a general reference model for all EU member States for many digital competence initiatives with the aim to create a common language on the development of digital competences.

This first paragraph aims at illustrate the possible tools conceived at European level to boost citizens’ digital competences. Dedicated frameworks are available for enterprises, teachers, consumers and organisations.

The framework foresees 21 competences (with three proficiency levels), divided in 5 areas, which can be summarised as below:

- 1) **Information and data literacy:** To articulate information needs, to locate and retrieve digital data, information and content. To judge the relevance of the source and its content. To store, manage, and organise digital data, information and content.
- 2) **Communication and collaboration:** To interact, communicate and collaborate through digital technologies while being aware of cultural and generational diversity. To participate in society through public and private digital services and participatory citizenship. To manage one's digital identity and reputation.
- 3) **Digital content creation:** To create and edit digital content to improve and integrate information and content into an existing body of knowledge while understanding how copyright and licences are to be applied. To know how to give understandable instructions for a computer system.
- 4) **Safety:** To protect devices, content, personal data and privacy in digital environments. To protect physical and psychological health, and to be aware of digital technologies for social well-being and social inclusion. To be aware of the environmental impact of digital technologies and their use.
- 5) **Problem solving:** To identify needs and problems, and to resolve conceptual problems and problem situations in digital environments. To use digital tools to innovate processes and products. To keep up-to-date with the digital evolution.

In a prospective of a life-long term learning approach, DigComp is experiencing different phases. Result of phase 1 is an update of the framework, named **DigComp 2.0** (Carretero Stephanie, 2017) with a focus on the conceptual reference model, new vocabulary and streamlined descriptors. In comparison with first version for example, new focuses are on **mobile devices, new environments, data literacy, privacy legislation and social inclusion** (Vourikari Riina, 2017).

Today, a new version is available. The current version is labelled **DigComp 2.1** (Carretero Gomez Stephanie, 2017) and it focuses on expanding the initial three proficiency levels to a more fine-grained eight level description as well as providing examples of use for these eight levels. Its aim is to support stakeholders with the further implementation of DigComp.

Other related JRC works enhancing the development of digital competence have as results the following frameworks:

- **DigCompConsumers** (Brečko, 2017)
- **DigCompOrg** (Kampylis Panagiotis, European Framework for Digitally Competent Educational Organisations, 2016)
- **DigCompEdu** (Punie Yves, 2017)

A framework for opening-up Higher Education Institutions (OpenEdu) (Inamorato dos Santos Andreia, 2016) was also published in 2016, as well as a competence framework for entrepreneurship (EntreComp) (Bacigalupo Margherita, 2017). Some of these frameworks are accompanied by (self-)assessment instruments. Additional research has been undertaken on computational thinking (CompuThink) (Kampylis Panagiotis, The Computational Thinking Study, 2016), Learning Analytics, MOOC learners (MOOCknowledge) (FERGUSON Rebecca, 2017) and MOOCs and free digital learning opportunities for migrants and refugees (MOOCs4inclusion) (Charalambos Vrasidas).

Table 1: DigComp 2.1

DigComp 2.0 (year 2016)		DigComp 2.1 (year 2017)	
Competence areas (dimension 1)	Competences (dimension 2)	Proficiency levels (dimension 3)	Examples of use (dimension 5)
1. Information and data literacy	1.1 Browsing, searching and filtering data, information and digital content 1.2 Evaluating data, information and digital content 1.3 Managing data, information and digital content	Eight proficiency levels for each of the 21 competences	Examples of use of the eight proficiency levels applied to learning and employment scenario in the 21 competences
2. Communication and collaboration	2.1 Interacting through digital technologies 2.2 Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies 2.5 Netiquette 2.6 Managing digital identity		
3. Digital content creation	3.1 Developing digital content 3.2 Integrating and re-elaborating digital content 3.3 Copyright and licences 3.4 Programming		
4. Safety	4.1 Protecting devices 4.2 Protecting personal data and privacy 4.3 Protecting health and well-being 4.4 Protecting the environment		
5. Problem solving	5.1 Solving technical problems 5.2 Identifying needs and technological responses 5.3 Creatively using digital technologies 5.4 Identifying digital competence gaps		

### 6.3 Privacy safeguards and online anonymity in the DigComp

The digital transformation enhances new requirements for digital competences and new vocabulary updates for such competences. Already in its update of 2016 and in its current version the DigComp Area 1 has been updated from "Information" only to "Information and **data literacy**". This to emphasise both the importance of data *per se* and the skills needed to critically evaluate and manage data in a safe and awareness-based way.

**Area 4. Safety**, section 4.2 has also been moved on from "Protection personal data" to "Protection personal data **and privacy**". This update aims at raising awareness about data privacy as a concept, meaning that data and technology are related to public and legal (Regulation (EU) 2016/679 of the European Parliament and of the Council, 2016) expectations of privacy. According to DigComp, by acquiring digital skills on safety, user is able:

- To protect personal data and privacy in digital environments
- To understand how to use and share personally identifiable information while being able to protect oneself and others from damages
- To understand that digital services use a "Privacy policy" to inform how personal data is used

**Privacy and Data Protection, Profiling and targeting, behavioural tracking** are extensively analysed in the **DigComp for Consumers**.

The most detailed taxonomy of possible problems related to privacy has been published by Solove (Solove, 2006). Solove groups the possible harms related to the privacy into four categories:

- **Information Collection:**
  - Harms: *Surveillance, Interrogation*
- **Information Processing:**
  - Harms: *Aggregation, Identification, Insecurity, Secondary Use, Exclusion*
- **Information Dissemination:**
  - Harms: *Breach of Confidentiality, Disclosure, Exposure, Increased Accessibility, Blackmail, Appropriation, Distortion*
- **Invasion:**
  - Harms: *Intrusion, Decisional Interference.*

While some of the harms are recognized crimes (e.g. blackmail, appropriation and distortion) that already had (before the "digital world") specific laws to prevent and punish them, others, instead, become "problems" only after a threshold (e.g. surveillance, interrogation) otherwise they are perfectly legitimate if performed by law enforcement agencies following the law and respecting human rights.

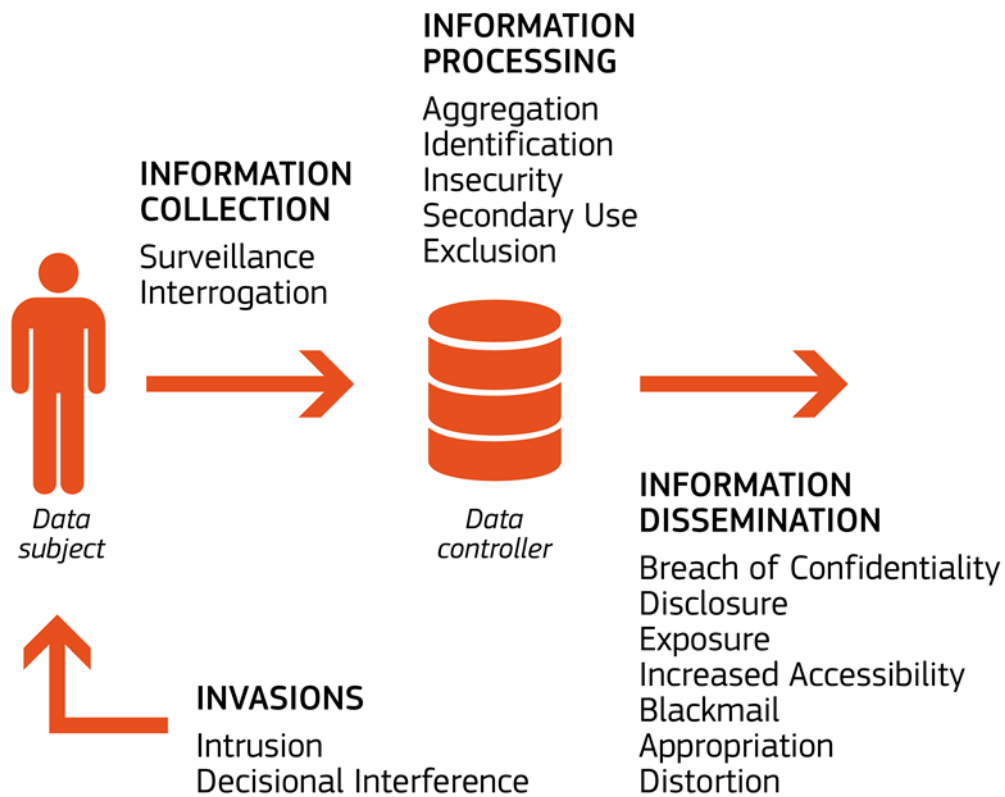
Table 2: Solove's taxonomy

A Taxonomy of Privacy Harms (compiled from (Solove, 2006))		
Domain	Privacy breach	Description
<b>Information Collection</b>	Surveillance	Watching, listening to, or recording of an individual's activities
	Interrogation	Various forms of questioning or probing for information
<b>Information Processing</b>	Aggregation	The combination of various pieces of data about a person
	Identification	Linking information to particular individuals
	Insecurity	Carelessness in protecting stored information from leaks and improper access
	Secondary Use	Use of information collected for one purpose for a different purpose without the data subject's consent
<b>Information Dissemination</b>	Exclusion	Failure to allow the data subject to know about the data that others have about her and participate in its handling and use, including being barred from being able to access and correct errors
	Breach of Confidentiality	Breaking a promise to keep a person's information confidential
	Disclosure	Revelation of information about a person that impacts the way others judge her character
	Exposure	Revealing another's nudity, grief, or bodily functions
	Increased Accessibility	Amplifying the accessibility of information
	Blackmail	Threat to disclose personal information
	Appropriation	The use of the data subject's identity to serve the aims and interests of another
<b>Invasion</b>	Distortion	Dissemination of false or misleading information about individuals
	Intrusion	Invasive acts that disturb one's tranquillity or solitude
	Decisional Interference	Incursion into the data subject's decisions regarding her private affairs

In this report, with reference to Solove's taxonomy in the digital domain, we deemed as worth to provide possible practical scenarios where digital competences are identified to complement technical and legal resources seen as strategies to deal with privacy harms.

In the following, we analyzed the privacy harms in four different sections (one for each category identified by Solove). At the end of each section corresponding to a specific domain, there is a table that summarize the findings.

Figure 10: Privacy harms as grouped by Solove.



In the table, the first two columns refer to the Solove's taxonomy:

1. The privacy harm itself;
2. The description of the specific privacy harm (from (Robin Mansell, 2015)).

The second group of columns refers to the tools that could be helpful to prevent or resolve the specific harm:

3. Legal resources<sup>1</sup>;
4. Technical resources

The third group of columns refers to the Digital Competence Framework (DigComp 2.1):

5. Competence Area (Dimension 1);
6. Competences (Dimension 2)

While the DigComp 2.1 together with the previously named dimensions foresees the proficiency levels, we have decided to not include them. In such a way, our indications should not be considered as a hinder for users not having a minimum level of proficiency.

<sup>1</sup> The content of this column do not intend at all to be an exhaustive list of tools offered by the legislation. For what concerns the articles suggested, they could apply only in specific circumstances. Furthermore, some of the problems related to the privacy identified by Solove were already addressed by the legislation before the existence of the "digital world".

Privacy is a hot topic of discussion, and in particular for the online life of the individuals. All privacy problems involve personal data or sensitive information. In the digital competence framework, we have identified some competences that apply in the prevention and/or in the solution of almost all of the privacy harms identified by Solove:

**1.3 Managing data, information and digital content:** all the harms related to privacy involve data management, thus it is needed a careful handling of data and information by user;

**2.2 Sharing through digital technologies:** the user should pay attention to what he/she is going to share online and he/she should be careful with whom he/she is going to share that information;

**2.3 Engaging in citizenship through digital technologies:** the user should participate in society through the use of public and private digital services and should seek for opportunities for self-empowerment. Furthermore, the user should participate actively as a citizen of the digital world through appropriate technologies.

**2.4 Collaborating through digital technologies:** the user could collaborate, through digital tools, with other persons and law enforcement to build knowledge related to privacy harms (e.g. reporting privacy harms in which he/she was involved) and to raise awareness.

**2.6 Managing digital identity:** the user should be able to manage and protect his/her own online data, information and reputation that build his/her own digital identity (or multiple identities).

**4.2 Protecting personal data and privacy:** the user should protect personal data and privacy in digital environments. He/she should understand how to use and share personally identifiable information in a safe and responsible way. He/she should also be able to protect oneself and others from damages. Further, the user should understand that digital services use a "Privacy policy" to inform how personal data is used.

**4.3 Protecting health and well-being:** the user should be aware of possible risks or threats both physical and psychological arising from privacy harms. In addition, the user should be able to protect oneself and others from those dangers and should use, if possible, digital technologies for social wellbeing and social inclusion helping to fight those harms.

**5.4 Identifying digital competence gaps:** the user that operates in the digital world should improve his/her digital competences to be able to identify possible risks and try to solve them, both by his own, or asking for help to a proficient user. On the other hand, a proficient user should provide support to users' not having a sufficient proficiency level.

In the following section, we are going to analyse each single privacy harm identified by Solove placing it in a digital context. We will give the description of the specific harm and an example to allow the reader to better understand the harm itself. **The examples provided are inspired by real cases of recent past not necessarily happened in the European Union.**

## 6.4 Information Collection

### 6.4.1 Surveillance

**Description:** *"The watching, listening to, or recording of an individual's activities"* (Robin Mansell, 2015)

*"What is the harm if people or the government watch or listen to us? Certainly, we all watch or listen, even when others may not want us to, and we often do not view this as problematic. However, when done in a certain manner—such as continuous monitoring—surveillance has problematic effects. For example, people expect to be looked at when they ride the bus or subway, but persistent gawking can create feelings of anxiety and discomfort.*

*Not only can direct awareness of surveillance make a person feel extremely uncomfortable, but it can also cause that person to alter her behaviour. Surveillance can lead to self-censorship and inhibition. Because of its inhibitory effects, surveillance is a tool of social control, enhancing the power of social norms, which work more effectively when people are being observed by others in the community. [...] This aspect of surveillance does not automatically make it harmful, though, since social control can be beneficial and every society must exercise a sizeable degree of social control. [...] Too much social control, however, can adversely impact freedom, creativity, and self-development."* (Solove, 2006)

As a general rule surveillance for a long time has been considered as troubling, in particular if it is set up with the purpose of spying upon or invading the privacy of the persons spied upon. Video surveillance and Audio surveillance enable the same perceptions and can create feelings of anxiety and discomfort. Another point is that surveillance can have a chilling effect on behavior, when one is aware of the possibility of surveillance. This conformity to the rules effect is known as Panoptic, based on Jeremy Bentham's architectural design for a prison Panoptic. Nevertheless, surveillance is applied as a deterrent to crime and some people desire the discipline and control that surveillance can bring. In this sense, surveillance can be perceived as a guardian with a "friendly eye on their lives".

**Example:** Let's assume that a European citizen travels into a country where the regime reduces the freedom of expression, analysing the internet traffic passing through the national gateways and performing censorship.

#### **Possible means for prevention and-or resolution:**

In those cases, there are several solutions like using of anonymizers to perform the most common activities online (disposable e-mail accounts, pseudo anonymizers for electronic payments, etc ...). Furthermore, passing through virtual private networks (VPNs) or *The Onion Router* (TOR) network, the traffic generated by the user is encrypted. In such a way the traffic is harder to be analyzed by any eavesdropper.

#### **Digital Competences involved:**

The use of those tools requires that the user has a sufficient level of skills to set up them. The digital competences involved in the management of this problem include 1.3, 2.2, 2.3, 2.4, 2.6, 4.2, 4.3, 5.4 (see paragraph 6.1). Furthermore, outside the above mentioned the following ones are needed:

- "4.1 Protecting devices" (unprotected devices can be easily kept under surveillance by malicious intent),
- "5.1 Solving technical problems",
- "5.2 Identifying needs and technological responses".



## 6.4.2 Interrogation

**Description:** "Various forms of questioning or probing for information" (Robin Mansell, 2015).

*"Interrogation is the pressuring of individuals to divulge information. Interrogation has many benefits; it is useful for ferreting out information that others want to know.*

*However, interrogation can create harm. Part of this harm arises from the degree of coerciveness involved. [...] However, for interrogation generally, the compulsion need not be direct; nor must it rise to the level of outright coercion. Compulsion can consist of the fear of not getting a job or of social opprobrium. People take offense when others ask an unduly probing question—even if there is no compulsion to answer. One explanation may be that people still feel some degree of compulsion because not answering might create the impression that they have something to hide. This is why, I believe, there are social norms against asking excessively probing or prying questions: they make the person being questioned feel uncomfortable. Interrogation forces people to be concerned about how they will explain themselves or how their refusal to answer will appear to others.*

*Interrogation resembles intrusion in its invasiveness, for interrogation is a probing, a form of searching. Like disclosure, interrogation often involves the divulging of concealed information; unlike disclosure, interrogation can create discomfort even if the information is barely disseminated. To some degree, surveillance resembles interrogation, for both involve the involuntary gathering of information. Interrogation, however, occurs with the conscious awareness of the subject; surveillance can be clandestine." (Solove, 2006)*

Interrogation can have some benefits and it can be useful during some activities, as for examples the interrogation of suspects in criminal investigations. However, depending on the grade of coerciveness involved, interrogation can create uncomfortable feelings. Moreover, a skilled interrogator can orchestrate a dialogue to elicit specific responses. Communication and persuasion methods are also applied in the Information and Communication Technologies in order to create individuals' disclosure.

**Example:** Although the *Interrogation* issue can be put in place outside the digital context, we may assume that third parties can deliberately trying to extort information by internet users. The issue could take place through digital technologies (e.g. skype, WhatsApp, text chat or VoIP, through which the user might share any kind of information). A user registered on a social network could receive a request of "friendship" by another user that he does not know. The other user could start asking to the first one for personal information.

### **Possible means for prevention and-or resolution:**

The users should be able to understand the possible risks of answering to detailed questions and evaluate if the other party can be trusted.

The only countermeasure against interrogation, if there is no coercion, is good sense:

- If you receive a message from somebody pretending to be a data controller that asks for your user name or a password, probably (almost for sure) it is a scam e-mail.
- If somebody that you do not know (or you know really few about him/her), starts to question you about personal data, it is not a good idea to reveal too much.

Since the hardest part related to the prevention of the interrogation performed online is the evaluation of the trust, the interrogation is a problem in particular for the minors. For this reason, it is important to perform awareness raising campaigns.

**Digital Competences involved:**

The digital competences involved in the management of this problem include 1.3, 2.2, 2.3, 2.4, 2.6, 4.2, 4.3, 5.4 (see paragraph 6.1).

To manage this issue, following digital competence is desirables:

- "1.2 *Evaluating data, information and digital content*" the user should be able to critically evaluate to whom he/she is answering and if can be trusted,

Table 3: Information Collection

Information Collection					
Privacy harm	Description	Possible means for prevention and-or resolution		Digital competencies involved	
		Legal resources	Technical resources	Area (dimension 1)	Competence (dimension 2)
<b>Surveillance</b>	<i>"The watching, listening to, or recording of an individual's activities"</i>	General Data Protection Regulation (e.g. Art. 5, Art. 6, Art. 7, Art. 9, Art. 10, Art. 12, Art. 13, Art. 14, Art. 23). -	Communication anonymizer (e.g. TOR, VPNs, disposable email addresses, pseudonyms for online payment)	1. Information and data literacy	1.3. Managing data, information and digital content
				2. Communication and collaboration	2.2. Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies 2.6 Managing digital identity
				3. Digital content creation	-
				4. Safety	4.1. Protecting devices 4.2. Protecting personal data and privacy 4.3 Protecting health and well-being
				5. Problem solving	5.1. Solving technical problems 5.2. Identifying needs and technological responses 5.4. Identifying digital competence gaps
<b>Interrogation</b>	<i>"Various forms of questioning or probing for information" (possibly also with coercion)</i>	See end of page 34		1. Information and data literacy	1.2 Evaluating data, information and digital content 1.3. Managing data, information and digital content
				2. Communication and collaboration	2.2. Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies 2.6. Managing digital identity
				3. Digital content creation	-
				4. Safety	4.2. Protecting personal data and privacy 4.3 Protecting health and well-being
				5. Problem solving	5.4 Identifying digital competence gaps

## 6.5 Information Processing

### 6.5.1 Aggregation

**Description:** *"The combination of various pieces of data about a person"* (Robin Mansell, 2015).

*"Aggregation is the gathering together of information about a person. A piece of information here or there is not very telling. But when combined together, bits and pieces of data begin to form a portrait of a person. The whole becomes greater than the parts. This occurs because combining information creates synergies. When analyzed, aggregated information can reveal new facts about a person that she did not expect would be known about her when the original, isolated data was collected.*

*Aggregating information is certainly not a new activity. It was always possible to combine various pieces of personal information, to put two and two together to learn something new about a person. But aggregation's power and scope are different in the Information Age; the data gathered about people is significantly more extensive, the process of combining it is much easier, and the computer technologies to analyze it are more sophisticated and powerful.*

*Combining data and analyzing it certainly can be put to beneficial uses. Amazon.com, for example, uses aggregated data about a person's book-buying history to recommend other books that the person might find of interest. [...] These developments make sense in a world where there are billions of people and word-of-mouth is insufficient to assess reputation.*

*Alongside these benefits, however, aggregation can cause dignitary harms because of how it unsettles expectations. People expect certain limits on what is known about them and on what others will find out. Aggregation upsets these expectations, because it involves the combination of data in new, potentially unanticipated ways to reveal facts about a person that are not readily known."* (Solove, 2006)

Everyone spreads selectively pieces of information in the daily activities, offline and online, and the boundaries of such disclosure is decided by the data subject that probably reveals very little about him/herself. The perception of this disclosure can indeed change in case of aggregation of data as it can reveal much more than what one expected when gave out data separately.

**Example:** Several online market places make use of the information provided by the users both directly and/or by their browsing history to create a profile of the user. These websites might identify the user because he/she is not logged into the service according to public records. Nevertheless, by aggregating and merging data they can make an idea of user's preferences and make some inferences. Those inferences are then used to propose to the users, products to which they could be interested, to increase the chances of a purchase.

#### **Possible means for prevention and-or resolution:**

The aggregated data could be useful for the data subject in some circumstances like online shopping. In the web-shop, analysing the data subject's behaviour and his/her browsing activity, are proposed specific products that users with a similar history found interesting.

Sometimes the data collection and aggregation goes too further, gathering data also when it is not strictly needed or when the user doesn't want it.

As first measure to prevent the aggregation, the user should be aware of the possible risks of his/her actions online and how he/she can be tracked and profiled.

The user could reduce the risk of providing too much data while performing his activities online using:

- Private browsing sessions,

- Do Not Track,
- Anti-tracking plugins (available for the most widespread browsers),
- Anonymizers for payments and disposable email addresses.

The user could further reduce the information provided, disabling the GPS on the mobile device when it's not needed and when taking pictures (or disabling the function to geotag the pictures).

Moreover, in the General Data Protection Regulation (GDPR), there are the following articles related to the aggregation:

- Article 13: Information to be provided where personal data are collected from the data subject
- Article 21: Right to object
- Article 22: Automated individual decision-making, including profiling.

### **Digital Competences involved:**

The digital competences involved in the management of this problem include 1.3, 2.2, 2.3, 2.4, 2.6, 4.2, 4.3, 5.4 (see paragraph 6.1).

Furthermore:

- "1.2 *Evaluating data, information and digital content*" the user should be able to critically evaluate if the third party processing his/her personal data can be trusted,
- "5.2 *Identifying needs and technological responses*" the user should be able to evaluate if he/she needs further protection while browsing online and apply countermeasures if needed.

### **6.5.2 Identification**

**Description:** "*The linking of information to a particular individuals*" (Robin Mansell, 2015) (*It can be with high or low aggregation*).

*"Identification" is connecting information to individuals. [...] Identification is related to disclosure in that both involve the revelation of true information. Identification involves a particular form of true information (one's identity), which enables databases of information to be linked to people. Identification is similar to aggregation as both involve the combination of different pieces of information, one being the identity of a person. However, identification differs from aggregation in that it entails a link to the person in the flesh. For example, there can be extensive aggregations of data about a person in many databases, but these aggregations might be rarely connected to that person as she goes through her day-to-day activities. This is a situation involving high aggregation and low identification. On the flip side, one can have high identification and low aggregation, such as in a world of checkpoints, where people constantly have to show identification but where there are few linkages to larger repositories of data about people.*

*Identification has many benefits. In order to access various accounts, people's identity must be verified, a step that can reduce fraud and enhance accountability.[...] Although identification of people or sources of particular messages can be beneficial, it also creates problems.[...] Identification goes a step further--it links the digital person directly to a person in real space.[...] Identification can inhibit one's ability to be anonymous or pseudonymous. Anonymity and pseudonymity protect people from bias based on their identities and enable people to vote, speak, and associate more freely by protecting them from the danger of reprisal."* (Solove, 2006)

Identification can be perceived as demeaning to dignity as some argue that it reduces people to a number of bodily characteristics, or data. According to Solove "[...] *identification is a means to link people to data, not necessarily an indication that people are the equivalent of their identifying characteristics.*" Identification markers like scarlet letters, tattoos, branding have been used in different social contexts to recognise people belonging to specific social categories and often they bear particular stigmas. This can bring to people's inhibition ability to change and disclosure. On the contrary, non-expressive means of identification such as fingerprints, identify people without signalling anything to the public. Anonymity can enhance prejudice reduction and increase online safety.

**Example:** Services like Pleaserobme or WeKnowYourHouse are good examples to illustrate how information shared online can be used (or misused) when it is linked to persons outside the digital world.

While the names of those Twitter accounts are a little bit alarmist, as explained also in 3.3, they are used only to raise awareness in the users, with the aim to teach them to not over-sharing.

### **Possible means for prevention and-or resolution:**

The first tool to prevent problems that derive by the identification is education. The user should be able to discriminate trusted person and/or sources and be able to understand which information he/she should disclose or not.

The technical tools that the user should use are the same suggested for the aggregation, keeping in mind that those tools are not helpful at all if the user is inclined to share information with non-trusted parties.

### **Digital Competences involved:**

The digital competences involved in the management of this problem include 1.3, 2.2, 2.3, 2.4, 2.6, 4.2, 4.3, 5.4 (see paragraph 6.1).

Furthermore:

- "1.2 *Evaluating data, information and digital content*" the user should be able to critically evaluate if the third party processing his/her personal data can be trusted,
- "5.2 *Identifying needs and technological responses*" the user should be able to evaluate if he/she needs further protection while browsing online and apply countermeasures if needed.

### **6.5.3 Insecurity**

**Description:** "*Carelessness in protecting stored information from leaks and improper access*" (Robin Mansell, 2015).

*"Insecurity, in short, is a problem caused by the way our information is handled and protected. [...] Insecurity exposes people to potential future harm. [...] Many privacy statutes require that information be kept secure."* (Solove, 2006)

In this issue two entities might be responsible of the insecurity risks: the data subject and the data controller.

**Example:** Insecurity by itself is not a problem for the privacy, but it can be the cause of several other problems related to the privacy. Let's suppose that a shop online that do not use the minimum cryptographic standard to protect the data (in particular credit card

numbers) of their customers. In case of data breach, those data can be read easily by whoever.

Insecurity can be the cause of several others harms identified by Solove:

- Surveillance, intrusion: insecurity in a domestic network could allow unauthorized access from other users.
- Breach of confidentiality, disclosure, exposure, identification, secondary use, blackmail, appropriation: insecurity in the network of a data controller can cause data breaches that could lead to a loss of trust in the data controller from the data subject and various harms, according to the kind of information that has been leaked.

### **Possible means for prevention and-or resolution:**

The user should be able to understand his level of proficiency to evaluate if he/she needs help to fix security holes. A common best practice is to update the operating system, the applications used and the antivirus as soon as a new patch is available to avoid that a malicious users takes advantage of possible vulnerabilities left unresolved.

The user should be also able to evaluate which information everybody can see and which one can be seen only by his friends or trusted persons/websites, and acts consequently to restrict the access to the most sensible ones.

As for the data controller, the GDPR establishes some specific articles (and heavy fines) that should be sufficient to put pressure on the data controller to invest in the security of their infrastructure.

### **Digital Competences involved:**

The digital competences involved in the management of this problem include 1.3, 2.2, 2.3, 2.4, 2.6, 4.2, 4.3, 5.4 (see paragraph 6.1).

Furthermore:

- "1.2 Evaluating data, information and digital content" the user should be able to critically evaluate if the third party processing his/her personal data can be trusted,
- "4.1 Protecting personal data and privacy",
- "5.1 Solving technical problems" if the user detects a problem related to his/her devices, he/she should be able to solve it or ask somebody (trusted) with a sufficient proficiency level to solve it,
- "5.2 Identifying needs and technological responses" the user should be able to evaluate if he/she needs further protection while browsing online and apply countermeasures if needed.

### **6.5.4 Secondary use**

**Description:** "Secondary use of information collected for one purpose without the data subject's consent" (Robin Mansell, 2015).

*"Secondary use" is the use of data for purposes unrelated to the purposes for which the data was initially collected without the data subject's consent. There are certainly many desirable instances of secondary use. Information might be used to stop a crime or to save a life. The variety of possible secondary uses of data is virtually infinite, and they range from benign to malignant.[...] Secondary use can cause problems.[...] Secondary uses thwart people's expectations about how the data they give out will be used. People might not give out data if they know about a potential secondary use, such as for telemarketing, spam, or other forms of intrusive advertising.[...] The potential for secondary use generates*



*fear and uncertainty over how one's information will be used in the future, creating a sense of powerlessness and vulnerability. In this respect, secondary use resembles the harm created by insecurity. [...] Secondary use also creates architectural problems. The secondary use of information can create problems because the information may not fit as well with the new use. When removed from the original context in which it was collected, data can more readily be misunderstood.” (Solove, 2006)*

Even though, there are benign secondary uses of data (e.g. saving a life, stopping crime), this harm can be perceived as a breach of confidentiality (see 6.6.1) as it involves using information in ways to which a person does not consent and might not find desirable. Even for those privacy policies stating that information might be used in secondary ways, people often do not read and/or do not (completely) understand these policies and have little idea about the range of potential secondary uses.

**Example:** Let's assume that an online platform that hosts online petitions ask consent to voters for the petition to treat their data only for the petitions subscribed. If the online platform starts to sell those data for marketing purposes, from the subscribed petitions some of the specific interests of the user can be inferred.

Of course this is a secondary use for which the subscriber of the petitions didn't provide their consent.

### **Possible means for prevention and-or resolution:**

The first tool to protect from the secondary use is to carefully read the terms and conditions of a service upon registration. When the user is going to give consent for specific services, he/she should be also aware that (unless special conditions hold) the data controller is not allowed to perform actions for which it hasn't the consent.

The GDPR is clear for what concerns the meaning of "consent", how it should be requested and the conditions that make it valid. Several articles give details about the special processing of personal data for which it is not needed an explicit consent from the data subject and at the same time, they give restriction to the unauthorized processing of those data.

### **Digital Competences involved:**

The digital competences involved in the management of this problem include 1.3, 2.2, 2.3, 2.4, 2.6, 4.2, 4.3, 5.4 (see paragraph 6.1).

Furthermore:

- "1.2 *Evaluating data, information and digital content*" the user should be able to critically evaluate if the third party processing his/her personal data can be trusted,
- "4.1 *Protecting personal data and privacy*",
- "5.1 *Solving technical problems*" if the user detects a problem related to his/her devices, he/she should be able to solve it or ask somebody (trusted) with a sufficient proficiency level to solve it,
- "5.2 *Identifying needs and technological responses*" the user should be able to evaluate if he/she needs further protection while browsing online and apply countermeasures if needed.



### 6.5.5 Exclusion

**Description:** *"The failure to allow the data subject to know about the data that others have about his/her and participate in its handling and use"* (Robin Mansell, 2015).

*"I refer to the failure to provide individuals with notice and input about their records as exclusion. [...] it is a harm created by being shut out from participating in the use of one's personal data, by not being informed about how that data is used, and by not being able to do anything to affect how it is used. [...] As with secondary use and insecurity, exclusion creates a sense of vulnerability and uncertainty in individuals. An inability to participate in the maintenance and use of one's information can lead to feelings of powerlessness and frustration."* (Solove, 2006)

Inability to participate in the managing and use of one's personal information can lead to feelings of powerlessness and frustration, especially when one party stands in a special position of power over another person.

**Example:** Let's assume that a user registers on a shop online and gives his consent to the shop to manage the data provided. Now the shop online is the data controller of the information related to the user (the data subject). Let's suppose that the same user, after a while, do not use anymore the services provided by the website and wants to delete his/her personal information stored by the data controller. Of course, if there's no special reason for the data controller to preserve the data related to the user, the user must be able to modify and/or delete those data.

#### **Possible means for prevention and-or resolution:**

The user should be aware that he/she has the right to access his/her own personal data collected by the data controller explicitly stated in the GDPR in the article 15 (Right of access). Furthermore, the GDPR grants the user other rights to manage his/her own personal data like:

- The right to rectification,
- The right to data portability,
- The right of erasure.

#### **Digital Competences involved:**

The digital competences involved in the management of this problem include 1.3, 2.2, 2.3, 2.4, 2.6, 4.2, 4.3, 5.4 (see paragraph 6.1), furthermore:

- *"1.2 Evaluating data, information and digital content"* the user should be able to critically evaluate if the third party processing his/her personal data can be trusted.

Table 4: Information Processing

Information Processing					
Privacy harm	Description	Possible means for prevention and-or resolution		Digital competencies involved	
		Legal resources	Technical resources	Area (dimension 1)	Competence (dimension 2)
<b>Aggregation</b>	<i>"The combination of various pieces of data about a person"</i>	General Data Protection Regulation (e.g. Art. 13, Art. 21, Art. 22).	Countermeasures to tracking.	1. Information and data literacy	1.2 Evaluating data, information and digital content 1.3. Managing data, information and digital content
				2. Communication and collaboration	2.2. Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies 2.6. Managing digital identity
				3. Digital content creation	-
				4. Safety	4.2. Protecting personal data and privacy 4.3 Protecting health and well-being
				5. Problem solving	5.2 Identifying needs and technological responses 5.4 Identifying digital competence gaps
<b>Identification</b>	<i>"The linking of information to a particular individuals" (it can be with high or low aggregation)</i>	General Data Protection Regulation (e.g. Art. 13, Art. 21, Art. 22).	Countermeasures to profiling	1. Information and data literacy	1.2 Evaluating data, information and digital content 1.3. Managing data, information and digital content
				2. Communication and collaboration	2.2. Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies 2.6. Managing digital identity
				3. Digital content creation	-
				4. Safety	4.2. Protecting personal data and privacy 4.3 Protecting health and well-being
				5. Problem solving	5.2 Identifying needs and technological responses 5.4 Identifying digital competence gaps
<b>Insecurity</b>	<i>"Carelessness in protecting stored information from leaks and improper access"</i>	General Data Protection Regulation (e.g. Art. 24, Art. 25, Art. 32, Art. 33, Art. 34).	User awareness and education to the privacy management. -	1. Information and data literacy	1.2 Evaluating data, information and digital content 1.3. Managing data, information and digital content
				2. Communication and collaboration	2.2. Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies

					2.6. Managing digital identity
				3. Digital content creation	-
				4. Safety	4.2. Protecting personal data and privacy 4.3 Protecting health and well-being
				5. Problem solving	5.1 Solving technical problems 5.2 Identifying needs and technological responses 5.4 Identifying digital competence gaps
<b>Secondary Use</b>	<i>"Secondary use of information collected for one purpose without the data subject's consent"</i>	General Data Protection Regulation (e.g. Art. 5, Art. 6, Art. 13 par. 3).		1. Information and data literacy	1.2 Evaluating data, information and digital content 1.3. Managing data, information and digital content
				2. Communication and collaboration	2.2. Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies 2.6. Managing digital identity
				3. Digital content creation	-
				4. Safety	4.2. Protecting personal data and privacy 4.3 Protecting health and well-being
				5. Problem solving	5.1 Solving technical problems 5.2 Identifying needs and technological responses 5.4 Identifying digital competence gaps
<b>Exclusion</b>	<i>"The failure to allow the data subject to know about the data that others have about his/her and participate in its handling and use"</i>	-General Data Protection Regulation (e.g. Art. 15).		1. Information and data literacy	1.2 Evaluating data, information and digital content 1.3. Managing data, information and digital content
				2. Communication and collaboration	2.2. Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies 2.6. Managing digital identity
				3. Digital content creation	-
				4. Safety	4.2. Protecting personal data and privacy 4.3 Protecting health and well-being
				5. Problem solving	5.4 Identifying digital competence gaps

## 6.6 Information Dissemination

### 6.6.1 Breach of Confidentiality

**Description:** *"The breaking of a promise to keep a person's information confidential"* (Robin Mansell, 2015).

*"The harm from a breach of confidence, then, is not simply that information has been disclosed, but that the victim has been betrayed. [...] Breach of confidentiality requires only a betrayal of trust, regardless of the nature of the data revealed. [...] When people establish a relationship with banks, Internet service providers, phone companies, and other businesses, they are not disclosing their information to the world. They are giving it to a party with implicit (and often explicit) promises that the information will not be disseminated."* (Solove, 2006)

Breach of confidentiality violate the trust in a specific relationship and can create feeling of disillusion and frustration.

### 6.6.2 Disclosure

**Description:** *"The revelation of truthful information about a person that impacts the way others judge his/her character"* (Robin Mansell, 2015).

*"Disclosure" occurs when certain true information about a person is revealed to others. Disclosure differs from breach of confidentiality because the harm in disclosure involves the damage to reputation caused by the dissemination; the harm with breach of confidentiality is the violation of trust in the relationship. [...] Although protecting against disclosure does limit freedom of speech, disclosure can inhibit the very interests free speech protects. Protection from disclosure, like free speech, promotes individual autonomy. The risk of disclosure can prevent people from engaging in activities that further their own self-development. Second, as with free speech, disclosure protections further democratic self-governance. [...] Disclosure can inhibit people from associating with others, impinging upon freedom of association, and can also destroy anonymity, which is sometimes critical for the promotion of free expression. Disclosure can also threaten people's security. [...] People want to protect information that makes them vulnerable or that can be used by others to harm them physically, emotionally, financially, and reputationally."* (Solove, 2006)

Disclosing a private matter can be perceived as highly offensive when it is not legitimate concern to the public.

### 6.6.3 Exposure

**Description:** *"revealing another's sensitive or personal activities such as nudity, grief, or bodily functions"* (Robin Mansell, 2015).

*"These are all illustrations of a disruption I call "exposure." Exposure involves the exposing to others of certain physical and emotional attributes about a person. These are attributes that people view as deeply primordial, and their exposure often creates embarrassment and humiliation. Grief, suffering, trauma, injury, nudity, sex, urination, and defecation all involve primal aspects of our lives—ones that are physical, instinctual, and necessary. We have been socialized into concealing these activities.*

*Although exposure is similar to disclosure—both involve the dissemination of true information—they diverge in an important respect. Exposure is related to disclosure in that concealed information is revealed to others, but the information is not revealing of anything we typically use to judge people's character. Unlike disclosure, exposure rarely reveals any*

significant new information that can be used in the assessment of a person's character or personality.

*Exposure creates injury because we have developed social practices to conceal aspects of life that we find animal-like or disgusting. Further, in certain activities, we are vulnerable and weak, such as when we are nude or going to the bathroom. [...] The need for privacy, and therefore the prevention of exposure, is created by the fact that we have social relationships and concomitant norms of dignity and decorum. [...] When these practices are disrupted by exposure, people can experience a severe and sometimes debilitating humiliation and loss of self-esteem. Exposure thus impedes a person's ability to participate in society.*" (Solove, 2006)

**Examples:** Should we consider the case of a data breach involving the passwords used for one's e-mail server and should we assume that same password(s) are used for both e-mail box and user's mobile device (as is the case for iPhone and Android). In case of data breach, the user should react as fast as possible to change that/those passwords /s to avoid any risk related to the devices' content as well as for the future e-mail traffic. Indeed, the user should be aware of the meaning beyond the concept of *Data Breach*.

While a data breach is a big problem for the spread of personal information, there are several other ways by which this could happen: for example secrets told to a friends and not maintained so secret, or pictures posted or phrases written online thinking to be anonymous and that in the end are associated back to the posting individual.

#### **Possible means for prevention and-or resolution:**

Again, for what concerns the prevention, most of the work is left to be done to the data controller.

For what concerns the user, he/she should carefully evaluate the party with whom he is going to share information.

Should the user realise that data and information have been already made available to the public, thus circulated, he/she could refer and apply the Right of erasure foreseen by the General Data Protection Regulation (Art. 17 of the GDPR). The user could also undertake legal measures at national level to limit damages resulting from the *Data Breach*.

#### **Digital Competences involved:**

We consider these three problems together, since the digital competencies involved, both in the management and in the prevention, are the same.

The digital competences involved in the management of this problem include 1.3, 2.2, 2.3, 2.4, 2.6, 4.2, 4.3, 5.4 (see paragraph 6.1).

Furthermore:

- "1.2 *Evaluating data, information and digital content*" the user should be able to critically evaluate if the third party processing his/her personal data can be trusted,
- "4.1 *Protecting personal data and privacy*",
- "5.1 *Solving technical problems*" if the user detects a problem related to his/her devices, he/she should be able to solve it or ask somebody (trusted) with a sufficient proficiency level to solve it,
- "5.2 *Identifying needs and technological responses*" the user should be able to evaluate if he/she needs further protection while browsing online and apply countermeasures if needed.

## 6.6.4 Increased Accessibility

**Description:** "The amplification of the accessibility of information" (Robin Mansell, 2015).

*"Increased accessibility does not involve a direct disclosure. Secret information is not disclosed. Rather, information that is already available to the public is made easier to access. Unlike disclosure, the harm is not a direct revealing of information to another. Confidentiality is not breached; the cat is already out of the bag. With increased accessibility, a difference in quantity becomes a difference in quality—it enhances the risk of the harms of disclosure. Increased accessibility to personal information has many benefits. It enhances openness, allowing people to locate information that they are seeking more easily. [...] Increased accessibility, however, creates problems such as the increased possibility of disclosure. Information can readily be exploited for purposes other than those for which it was originally made publicly accessible."* (Solove, 2006)

**Example:** The problem seems to be similar to that of *aggregation*, even though in this case we are considering specific data stored in specific records. It can happen that those records are made too easily accessible.

For example, let's assume that a bankruptcy is recorded for an individual, who meanwhile has repaid the outstanding debt and has nowadays a stable financial situation. The recorded bankruptcy and the increased accessibility by financial institution, which might have access to that judgment, could invalidate the accessibility to a loan by the said individual.

**Possible means for prevention and-or resolution:**

...

**Digital Competences involved:**

In this specific case, users' digital competences do not contribute significantly in the solution of the problem, however we think that the digital competencies 2.2, 2.3, 2.4 could be helpful to raise awareness about the issue.

## 6.6.5 Blackmail

**Description:** "The threat to disclose personal information" (Robin Mansell, 2015).

*"Blackmail allows a person to be dominated and controlled by another. With blackmail, the harm is not in the actual disclosure of the information, but in the control exercised by the one who makes the threat over the data subject. In some cases, blackmail can also involve information more akin to exposure than disclosure. Breach of confidentiality is also related to blackmail, as a confidant can threaten to disclose a secret in return for money. Blackmail differs from disclosure, exposure, and breach of confidentiality in that it involves a threat of disclosure rather than an actual disclosure."* (Solove, 2006)

**Example:** Let's suppose that following a data breach, somebody got some sensitive information related to us that we don't want to be spread. The person holding this information, could try to obtain some profit from us through blackmail.

**Possible means for prevention and-or resolution:**

Blackmailing is a crime, regardless if it happens online or outside the digital world.

In case of *Blackmail*, the best way to react is to seek help from a legal point of view; nevertheless, we believe that users can protect themselves from threatening to reveal personal data by way of prevention.

### **Digital Competences involved:**

Since Blackmail could be performed through the threat to disclose stolen information, most of the digital competencies involved are related to the protection of the data:

The digital competences involved in the management of this problem include 1.3, 2.2, 2.3, 2.4, 2.6, 4.2, 4.3, 5.4 (see paragraph 6.1).

Furthermore, since Blackmail could be performed through the threat to disclose stolen information, most of the digital competencies involved are related to the protection of the data:

- "1.2 *Evaluating data, information and digital content*" the user should be able to critically evaluate if the third party processing his/her personal data can be trusted,
- "4.1 *Protecting personal data and privacy*",
- "5.1 *Solving technical problems*" if the user detects a problem related to his/her devices, he/she should be able to solve it or ask somebody (trusted) with a sufficient proficiency level to solve it,
- "5.2 *Identifying needs and technological responses*" the user should be able to evaluate if he/she needs further protection while browsing online and apply countermeasures if needed.

### **6.6.6 Appropriation**

**Description:** "*The use of the data subject's identity to serve the aims and interests of another*" (Robin Mansell, 2015).

*"Appropriation" is the use of one's identity or personality for the purposes and goals of another. Appropriation, like the privacy disruptions of disclosure and distortion, involves the way an individual desires to present herself to society. [...] The interest safeguarded by protections against appropriation is control of the way one presents oneself to society. The products and causes people publicly endorse shape their public image. When people are associated with products, they become known in terms of these products. [...] Thus, appropriation can be harmful even if it is not humiliating, degrading, or disrespectful. Being unwillingly used to endorse a product resembles, in certain respects, being compelled to speak and to represent certain viewpoints.*

*Protection against appropriation establishes what society considers appropriate for others to do in shaping a person's identity. The harm, then, is an impingement on the victim's freedom in the authorship of her self-narrative, not merely her loss of profits."* (Solove, 2006)

**Example:** A malicious individual that obtains personal information (through a data breach or social engineering) related to another person, could use them for his/her benefits. Personal information gathered could be of any kind:

- Login credentials (user name and password) on a social network/e-mail/instant messenger/shop online,
- Credit card number/Online payment service account,
- Personal information related to the real life that could be used to pretend to be that person while being online.



### **Possible means for prevention and-or resolution:**

A user victim of a crime of appropriation, as first step should seek help from a legal point of view, while, at the same time, he/she should understand how the criminal was able to obtain those information. Once the user understood what is the source of those data, he/she should take countermeasures either if it is related to insecurity of his/her own devices or if it is a problem of insecurity related to a data controller that was hosting his personal data.

### **Digital Competences involved:**

The digital competences involved in the management of this problem include 1.3, 2.2, 2.3, 2.4, 2.6, 4.2, 4.3, 5.4 (see paragraph 6.1).

Furthermore:

- "1.2 *Evaluating data, information and digital content*" the user should be able to critically evaluate if the third party processing his/her personal data can be trusted,
- "4.1 *Protecting personal data and privacy*",
- "5.1 *Solving technical problems*" if the user detects a problem related to his/her devices, he/she should be able to solve it or ask somebody (trusted) with a sufficient proficiency level to solve it,
- "5.2 *Identifying needs and technological responses*" the user should be able to evaluate if he/she needs further protection while browsing online and apply countermeasures if needed.

### **6.6.7 Distortion**

**Description:** "*The dissemination of false or misleading information about individuals*" (Robin Mansell, 2015).

*"I refer to these harms as "distortion." Distortion is the manipulation of the way a person is perceived and judged by others, and involves the victim being inaccurately exposed to the public. I include distortion in the taxonomy of privacy because of its significant similarity to other privacy disruptions. Distortion, like disclosure, involves the spreading of information that affects the way society views a person. Both distortion and disclosure can result in embarrassment, humiliation, stigma, and reputational harm. They both involve the ability to control information about oneself and to have some limited dominion over the way one is viewed by society. Distortion differs from disclosure, however, because with distortion, the information revealed is false and misleading. [...] Reputation is not merely an individual creation. Although it is true that people work very hard to build their reputations, one's reputation is the product of the judgment of other people in society. Reputation is a currency through which we interact with each other. Protection against distortion structures our interactions because it protects this currency. Distortion not only affects the aggrieved individual; it also affects the society that judges that individual: it interferes with our relationships to that individual, and it inhibits our ability to assess the character of those that we deal with."* (Solove, 2006)

This harm refers to one's reputation as indispensable element to self-identity and the ability to engage in public life. Nowadays, in the digital society we assist to the transposition of this concept to the web reputation and digital identity that both deserve acknowledged social regard, acceptance and respect.

**Example:** This problem is extremely hard to prevent and to resolve.



There is a twofold origin for this problem: an individual spread false information either (in good faith) because he/she hasn't verified sources, or intentionally to cause bad reputation to somebody else.

**Possible means for prevention and-or resolution:**

For what concerns the distortion, there is no real countermeasure that a user could take to prevent to incur in that problem, but the user himself/herself could signal to authorities, or the victim if he recognize any false or misleading information about somebody else. In this way, the user behaves as a good digital citizen, reducing the spread of those false information.

**Digital Competences involved:**

While digital competencies like 2.2, 2.3, 2.4 could be helpful to raise awareness about the issue, dissemination of distorted information might take place regardless user's digital competences. Legal measures can be the problem solution and the application of the "right to rectify" of the GDPR (Article 16) or the right to erasure (Article 17) where the "distortion" are online.

Table 5: Information Dissemination

Information dissemination					
Privacy harm	Description	Possible means for prevention and-or resolution		Digital competencies involved	
		Legal resources	Technical resources	Area (dimension 1)	Competence (dimension 2)
<b>Breach of Confidentiality</b>	<i>"The breaking of a promise to keep a person's information confidential"</i>	-General Data Protection Regulation (e.g. Art. 32, Art. 33, Art. 34).		1. Information and data literacy	1.2 Evaluating data, information and digital content 1.3. Managing data, information and digital content
				2. Communication and collaboration	2.2. Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies 2.6. Managing digital identity
				3. Digital content creation	-
				4. Safety	4.1. Protecting devices 4.2. Protecting personal data and privacy 4.3 Protecting health and well-being
				5. Problem solving	5.1 Solving technical problems 5.2 Identifying needs and technological responses 5.4 Identifying digital competence gaps
<b>Disclosure</b>	<i>"The revelation of truthful information about a person that impacts the way others judge his/her character"</i>	-General Data Protection Regulation (e.g. Art. 32, Art. 33, Art. 34).		1. Information and data literacy	1.2 Evaluating data, information and digital content 1.3. Managing data, information and digital content
				2. Communication and collaboration	2.2. Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies 2.6. Managing digital identity
				3. Digital content creation	-
				4. Safety	4.1. Protecting devices 4.2. Protecting personal data and privacy 4.3 Protecting health and well-being
				5. Problem solving	5.1 Solving technical problems 5.2 Identifying needs and technological responses 5.4 Identifying digital competence gaps

<b>Exposure</b>	<i>"revealing another's sensitive or personal activities such as nudity, grief, or bodily functions"</i>	-General Data Protection Regulation (e.g. Art. 32, Art. 33, Art. 34).		1. Information and data literacy	1.2 Evaluating data, information and digital content 1.3. Managing data, information and digital content
				2. Communication and collaboration	2.2. Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies 2.6. Managing digital identity
				3. Digital content creation	-
				4. Safety	4.1. Protecting devices 4.2. Protecting personal data and privacy 4.3 Protecting health and well-being
				5. Problem solving	5.1 Solving technical problems 5.2 Identifying needs and technological responses 5.4 Identifying digital competence gaps
<b>Increased Accessibility</b>	<i>"The amplification of the accessibility of information"</i>	-General Data Protection Regulation (e.g. Art. 10).		1. Information and data literacy	-
				2. Communication and collaboration	2.2. Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies
				3. Digital content creation	-
				4. Safety	-
				5. Problem solving	-
<b>Blackmail</b>	<i>"The threat to disclose personal information"</i>	See end of page 34		1. Information and data literacy	1.2 Evaluating data, information and digital content 1.3. Managing data, information and digital content
				2. Communication and collaboration	2.2. Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies 2.6. Managing digital identity
				3. Digital content creation	-
				4. Safety	4.1. Protecting devices 4.2. Protecting personal data and privacy 4.3 Protecting health and well-being

				5. Problem solving	5.1 Solving technical problems 5.2 Identifying needs and technological responses 5.4 Identifying digital competence gaps
<b>Appropriation</b>	<i>"The use of the data subject's identity to serve the aims and interests of another"</i>	See end of page 34		1. Information and data literacy	1.2 Evaluating data, information and digital content 1.3. Managing data, information and digital content
				2. Communication and collaboration	2.2. Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies 2.6. Managing digital identity
				3. Digital content creation	-
				4. Safety	4.1. Protecting devices 4.2. Protecting personal data and privacy 4.3 Protecting health and well-being
				5. Problem solving	5.1 Solving technical problems 5.2 Identifying needs and technological responses 5.4 Identifying digital competence gaps
<b>Distortion</b>	<i>"The dissemination of false or misleading information about individuals"</i>	-General Data Protection Regulation (e.g. Art. 16, Art. 17).		1. Information and data literacy	-
				2. Communication and collaboration	2.2. Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies
				3. Digital content creation	-
				4. Safety	-
				5. Problem solving	-

## 6.7 Invasion

### 6.7.1 Intrusion

**Description:** *"Invasive acts that disturb one's tranquility or solitude"* (Robin Mansell, 2015).

*"Intrusion involves invasions or incursions into one's life. It disturbs the victim's daily activities, alters her routines, destroys her solitude, and often makes her feel uncomfortable and uneasy. Protection against intrusion involves protecting the individual from unwanted social invasions, affording people what Warren and Brandeis called "the right to be let alone." [...] While many forms of intrusion are motivated by a desire to gather information or result in the revelation of information, intrusion can cause harm even if no information is involved. In particular, intrusion often interferes with solitude, the state of being alone or able to retreat from the presence of others."* (Solove, 2006)

**Example:** Let's assume that the wireless domestic network of your house is not secured enough and that a malicious user in range of your network connection has access to it. If that user starts to download big amount of data, you could experience a reduction of your bandwidth and your experience online could be unpleasant. Intrusion on private accounts of an individual (following a data breach) or on the personal device itself (because it wasn't properly secured) can be also more annoying and cause additional problems.

#### **Possible means for prevention and-or resolution:**

The victim of an intrusion should try to understand what has been done by the intruder and evaluate the seriousness of the actions performed. In the better case it should be enough that the user fix (if he's able to do that, or otherwise ask for support) the "entry point" of the intruder, while in the worst case it could be needed the intervention of law enforcement agencies.

#### **Digital Competences involved:**

The digital competences involved in the management of this problem include 1.3, 2.2, 2.3, 2.4, 2.6, 4.2, 4.3, 5.4 (see paragraph 6.1), furthermore the following digital competencies are useful to prevent intrusion caused by insecurity:

- "4.1 Protecting personal data and privacy",
- "5.1 Solving technical problems" if the user detects a problem related to his/her devices, he/she should be able to solve it or ask somebody (trusted) with a sufficient proficiency level to solve it,
- "5.2 Identifying needs and technological responses" the user should be able to evaluate if he/she needs further protection while browsing online and apply countermeasures if needed.

### 6.7.2 Decisional Interference

**Description:** *"The government's incursion into the data subject's decisions regarding his/her private affairs"* (Robin Mansell, 2015).

*"[...] what I call "decisional interference"—that is, governmental interference with people's decisions regarding certain matters of their lives. [...] Many commentators have argued that the language of privacy is inappropriate for decisional interference cases, since they primarily concern a harm to autonomy and liberty, not to privacy. [...] What*

*relationship does decisional interference have with the other forms of privacy in the taxonomy?*

*The decisional interference cases are deeply connected to information privacy. In particular, [...] the constitutionally protected "zone of privacy" extends not only to the "interest in independence in making certain kinds of important decisions" but also to the "individual interest in avoiding disclosure of personal matters." " (Solove, 2006)*

**Example:** It is considered decisional interference the restriction or incursion, by the government, into the data subject's decisions regarding fundamental rights. As an example, in the online world, it could be taken the one done for the surveillance where "a government" (or better "a regime") regulate the freedom of expression and opinion of the individuals. In that case, the citizens probably will change their behaviour online to be compliant with the regime.

**Possible means for prevention and-or resolution:**

**Digital Competences involved:**

For the decisional interference (as for increased accessibility), users' digital competences do not contribute significantly in the solution of the problem, however we think that the digital competencies 2.2, 2.3, 2.4 could be helpful to raise awareness about the issue.

Table 6: Invasion

Invasion					
Privacy harm	Description	Possible means for prevention and-or resolution		Digital competencies involved	
		Legal resources	Technical resources	Area (dimension 1)	Competence (dimension 2)
<b>Intrusion</b>	<i>"Invasive acts that disturb one's tranquility or solitude"</i>	See end of page 34		1. Information and data literacy	1.3. Managing data, information and digital content
				2. Communication and collaboration	2.2. Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies 2.6. Managing digital identity
				3. Digital content creation	-
				4. Safety	4.1. Protecting devices 4.2. Protecting personal data and privacy 4.3 Protecting health and well-being
				5. Problem solving	5.1 Solving technical problems 5.2 Identifying needs and technological responses 5.4 Identifying digital competence gaps
<b>Decisional Interference</b>	<i>"The government's incursion into the data subject's decisions regarding his/her private affairs"</i>			1. Information and data literacy	-
				2. Communication and collaboration	2.2. Sharing through digital technologies 2.3 Engaging in citizenship through digital technologies 2.4 Collaborating through digital technologies
				3. Digital content creation	-
				4. Safety	-
				5. Problem solving	-

THIS PAGE IS INTENTIONALLY LEFT BLANK



## 7 Conclusions

Nowadays, digital citizens, actively or passively accept to provide personal information in their everyday online activities. Some of the information shared are personal data as defined by the General Data Protection Regulation (Regulation (EU) 2016/679 of the European Parliament and of the Council, 2016) in its Article 4(1). Some of those data can be used to identify uniquely an individual, whereas others, even if they are not strictly "personal data", are still related to the individual and can still lead to the identification. Some issues on privacy of the data subject can appear in specific areas of the data flow, as identified by Solove (Solove, 2006) in its paper "A Taxonomy of Privacy". We have used it as the base for our study.

To perform the analysis of those privacy issues, we chose as methodology a multidisciplinary approach, looking at technical tools, legal tools and educational aspects needed to prevent and/or resolve them.

In this report we have presented a survey of the techniques to perform the tracking (browser fingerprinting, ETags, location tracking, etc.) and profiling of the data subjects. We complemented it by analysing some of the Privacy-Enhancing Technologies (PETs) (see 5.1.1) available that could be helpful for the digital citizen to better manage and safeguard his privacy online.

We also considered the recent General Data Protection Regulation in our analysis, as it represents a big step ahead in "the protection of natural persons with regard to the processing of personal data". Indeed, the regulation helps the European citizen to manage better his/her personal data, with specific articles addressing some issues related to the privacy such as, for example, the '*Right of Access by the data subject*' (art.15), to avoid '*Exclusion*' and '*Right of Rectification*' (art.16), '*Right of Erasure*' (art.17), to fight '*Secondary Use*' of personal data (art.5) or '*Distortion*'.

We concluded our analysis summarising our findings related to technical and legal tools to address (prevent and resolve) privacy harms following the categorisation of Solove (Solove, 2006). The analysis was done under the point of view of the data subject (any digital citizen) that could face those harms. This perspective allowed us to identify the digital competences needed by the data subject in order to enact the suggested measures of prevention or resolutions of privacy issues.

Our research have led to the main conclusion that, whilst there exist legal and technical tools to protect privacy online, these cannot be effective unless they are complemented by a proper education of the individuals. Therefore, it is our recommendation to increase user awareness in this topic and put forward initiatives to promote existing tools and education campaigns to improve the well-being of digital citizens.

THIS PAGE IS INTENTIONALLY LEFT BLANK

## References

- Bacigalupo Margherita, K. P. (2017, 05 10). *Entrepreneurship Competence*. Retrieved from European Commission: <https://ec.europa.eu/jrc/en/entrecomp>
- Biermann, K. (2011, MArch 10). Data Protection: Betrayed by our own data. *Betrayed by our own data*. (Z. ONLINE, Ed.) Retrieved from <http://www.zeit.de/digital/datenschutz/2011-03/data-protection-malte-spitz>
- Brecko Barbara, F. A. (2017, 04 08). *The Digital Competence Framework for Consumers*. Retrieved from European Commission: <https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/digital-competence-framework-consumers>
- Brečko, B. F. (2017, 03 20). *The Digital Competence Framework for Consumers*. doi:10.2791/838886
- Carretero Gomez Stephanie, V. R. (2017). *DigComp 2.1: The Digital Competence Framework for Citizens with eight proficiency levels and examples of use*. Retrieved from European Commission: <https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/digcomp-21-digital-competence-framework-citizens-eight-proficiency-levels-and-examples-use>
- Carretero Stephanie, P. Y. (2017, 04 20). *The Digital Competence Framework 2.0*. Retrieved from European Commission: <https://ec.europa.eu/jrc/en/digcomp/digital-competence-framework>
- Charalambos Vrasidas, E. C. (n.d.). *MOOCS4inclusion*. Retrieved from MOOCS4inclusion: <http://moocs4inclusion.org/>
- Control, T. I. (A cura di). (2002). *IEC TC 65/290/DC* (Vol. Device Profile Guideline).
- DG Justice and Consumers. (2015, 06 24). *Data protection Eurobarometer out today*. Retrieved from European Commission: [http://ec.europa.eu/justice/newsroom/data-protection/news/240615\\_en.htm](http://ec.europa.eu/justice/newsroom/data-protection/news/240615_en.htm)
- Eckersley, P. (2010). How Unique Is Your Web Browser? In M. J. Atallah, & N. J. Hopper (Ed.), *Privacy Enhancing Technologies: 10th International Symposium, PETS 2010, Berlin, Germany, July 21-23, 2010. Proceedings* (pp. 1-18). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-14527-8\_1
- European Commission. (2016, 09 21). *European Commission*. Retrieved from INFORMATION PROVIDERS GUIDE: The EU Internet Handbook - Cookies: [http://ec.europa.eu/ipg/basics/legal/cookies/index\\_en.htm](http://ec.europa.eu/ipg/basics/legal/cookies/index_en.htm)
- European Commission. (2016, 07 01). *PACT — Result In Brief*. Retrieved from European Commission: [http://cordis.europa.eu/result/rcn/155988\\_en.html](http://cordis.europa.eu/result/rcn/155988_en.html)
- European Commission. (2017). Joint Communication to the European Parliament and the Council - Resilience, Deterrence and Defence: Building strong cybersecurity for the EU. *European Commission - High Representative of the Union for Foreign Affairs and Security Policy, Final*.
- European Parliament, E. C. (2002, July 12). Data protection in the electronic communications sector. Retrieved from <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=URISERV%3A124120>
- European, E. P. (n.d.). DIRECTIVE 2009/136/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL. Retrieved from <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2009:337:0011:0036:en:PDF>
- FERGUSON Rebecca, B. A. (2017, 01 13). *Research Evidence on the Use of Learning Analytics: Implications for Education Policy*. Retrieved from European Commission: <https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical->

research-reports/research-evidence-use-learning-analytics-implications-education-policy

- G. Chen, J. H. (2016, January). In-Depth Survey of Digital Advertising Technologies. *IEEE Communications Surveys & Tutorials*, 18(3), 2124-2148. doi:10.1109/COMST.2016.2519912
- Gutwirth S., H. M. (2010). Some Caveats on Profiling. In P. Y. Gutwirth S., *Data Protection in a Profiled World*. Springer, Dordrecht. doi:https://doi.org/10.1007/978-90-481-8865-9\_2
- Hildebrandt, M. (2008). Defining Profiling: A New Type of Knowledge? In M. G. Hildebrandt (Ed.), *Profiling the European Citizen: Cross-Disciplinary Perspectives* (pp. 17-45). Springer Netherlands. doi:10.1007/978-1-4020-6914-7
- Inamorato dos Santos Andreia, P. Y. (2016, 07 27). *Policy Recommendations for Opening Up Education*. Retrieved from European Commission: <https://ec.europa.eu/jrc/en/open-education>
- ISO/IEC JTC 1. (2015). *ISO/IEC 11889-1:2015: Information technology -- Trusted platform module library -- Part 1: Architecture*. International Organization for Standardization.
- ISO/IEC JTC 1/SC 27. (2013). *ISO/IEC 20008: Information technology — Security techniques — Anonymous digital signatures*. International Organization for Standardization.
- ISO/IEC JTC 1/SC 27. (2013). *ISO/IEC 20009-1:2013: Information technology -- Security techniques -- Anonymous entity authentication -- Part 1: General*. International Organization for Standardization.
- Kampylis Panagiotis, P. Y. (2016, 07 27). *European Framework for Digitally Competent Educational Organisations*. Retrieved from European Commission: <https://ec.europa.eu/jrc/en/digcomporg>
- Kampylis Panagiotis, P. Y. (2016, 07 27). *The Computational Thinking Study*. Retrieved from European Commission: <https://ec.europa.eu/jrc/en/computational-thinking>
- PACT project. (2014, 11 14). *PACT Final Conference - Vienna [13-14 November 2014]*. Retrieved from PACT Final Conference: <http://www.projectpact.eu/>
- Party, A. 2. (2013, October 2). Working Document 02/2013 providing guidance on obtaining consent for cookies. Retrieved from [http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp208\\_en.pdf](http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2013/wp208_en.pdf)
- Požrl, M. K. (2017). Diversity in recommender systems – A survey. *Knowledge-Based Systems*, 123(Supplement C), 154 - 162. doi:https://doi.org/10.1016/j.knosys.2017.02.009
- Punie Yves, R. C. (2017, 12 01). *Digital Competence Framework for Educators (DigCompEdu)*. Retrieved from European Commission: <https://ec.europa.eu/jrc/en/digcompedu>
- Regulation (EU) 2016/679 of the European Parliament and of the Council. (2016, may). *Official Journal of the European Union*, 1-88. Retrieved from <http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2016:119:TOC>
- Robin Mansell, P. H. (2015). *The International Encyclopedia of Digital Communication and Society*. Wiley Publishing.
- Roosendaal, A. (2012). We Are All Connected to Facebook ... by Facebook! In A. Roosendaal, *European Data Protection: In Good Health?* (pp. 3-19). Dordrecht: Springer Netherlands. doi:10.1007/978-94-007-2903-2\_1

- Soghoian, C. (2011, January 21). *The History of the Do Not Track Header*. Retrieved from slight paranoia - Analysis and opinion by Christopher Soghoian, security and privacy researcher.: <http://paranoia.dubfire.net/2011/01/history-of-do-not-track-header.html>
- Solove, D. J. (2006, January). A Taxonomy of Privacy. *University of Pennsylvania Law Review*, 154, 477-560.
- Tuzhilin, G. A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, 17(6), 734-749.
- V. Ferraris, F. B. (2014). *Working Paper Defining Profiling*.
- Veijalainen, D. K. (2016). A survey of serendipity in recommender systems. *Knowledge-Based Systems*, 111(Supplement C), 180-192.  
doi:<https://doi.org/10.1016/j.knosys.2016.08.014>
- Verbert, C. H. (2016). Interactive recommender systems: A survey of the state of the art and future research challenges and opportunities. *Expert Systems with Applications*, 56(Supplement C), 9-27.  
doi:<https://doi.org/10.1016/j.eswa.2016.02.013>
- Vourikari Riina, P. Y. (2017, 07 15). *DigComp 2.0: The Digital Competence Framework for Citizens. Update Phase 1: the Conceptual Reference Model*. Retrieved from European Commission: <https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/digcomp-20-digital-competence-framework-citizens-update-phase-1-conceptual-reference-model>
- W3C. (2015, August 20). *Tracking Preference Expression (DNT)*. Retrieved from [www.w3.org](http://www.w3.org): <https://www.w3.org/TR/tracking-dnt/>
- Warren, S., & Brandeis, L. (1890). The Right to Privacy. *Harvard Law Review*, 193-220.

THIS PAGE IS INTENTIONALLY LEFT BLANK

**List of figures**

**Figure 1: An example of target advertising where the recommended products are suggested through the analysis of the previous browsing history of the user. .... 7**

**Figure 2: The tracking of the users make it possible for a website to offer personalized contents for each one. .... 8**

**Figure 3: Lightbeam is a plugin for several browser that make it possible for the user to know the third party components loaded while browsing a specific website and further information on each of them..... 12**

**Figure 4: Browsing different websites it may happen that the same resource is loaded from the same repository, making it possible for third parties to perform cross-site tracking..... 16**

**Figure 5: ZEIT ONLINE newspaper showing the power of crossing information gathered from different sources. .... 17**

**Figure 6: How to enable Do Not Track in Chrome, Edge, Safari and Firefox. ....Error! Bookmark not defined.**

**Figure 7: Example of website before and after the activation of Adblock with the removed ads highlighted. 27**

**Figure 8: Analysis of the browser performed by Panopticlick with highlighted the specific settings of the browser taken into account for browser fingerprinting. .... 28**

**Figure 9: The flow of a message inside the TOR network from a sender to a recipient. .... Error! Bookmark not defined.**

**Figure 10: Privacy harms as grouped by Solove. .... 36**

THIS PAGE IS INTENTIONALLY LEFT BLANK



**List of tables**

**Table 1: DigComp 2.1 ..... 33**  
**Table 2: Solove's taxonomy ..... 35**  
**Table 3: Information Collection ..... 41**  
**Table 4: Information Processing ..... 48**  
**Table 5: Information Dissemination ..... 56**  
**Table 6: Invasion ..... 61**

THIS PAGE IS INTENTIONALLY LEFT BLANK

***Europe Direct is a service to help you find answers  
to your questions about the European Union.***

**Freephone number (\*):**

**00 800 6 7 8 9 10 11**

(\*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

More information on the European Union is available on the internet (<http://europa.eu>).

## **HOW TO OBTAIN EU PUBLICATIONS**

### **Free publications:**

- one copy:  
via EU Bookshop (<http://bookshop.europa.eu>);
- more than one copy or posters/maps:  
from the European Union's representations ([http://ec.europa.eu/represent\\_en.htm](http://ec.europa.eu/represent_en.htm));  
from the delegations in non-EU countries ([http://eeas.europa.eu/delegations/index\\_en.htm](http://eeas.europa.eu/delegations/index_en.htm));  
by contacting the Europe Direct service ([http://europa.eu/eurodirect/index\\_en.htm](http://europa.eu/eurodirect/index_en.htm)) or  
calling 00 800 6 7 8 9 10 11 (freephone number from anywhere in the EU) (\*).

(\*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

### **Priced publications:**

- via EU Bookshop (<http://bookshop.europa.eu>).

## JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.



**EU Science Hub**  
[ec.europa.eu/jrc](https://ec.europa.eu/jrc)



@EU\_ScienceHub



EU Science Hub - Joint Research Centre



Joint Research Centre



EU Science Hub



Publications Office

doi:10.2760/30934

ISBN 978-92-79-77231-3