D3.1

Fundamental Rights Protection by Design for OSN

v 1.4 / 2015-03-31

Katja de Vries, Sari Depreeuw, Mireille Hildebrandt (ICIS-RU)

This document presents an overview of the possibilities for fundamental rights protection by design (FRPbD) in the context of behavioural tracking and personalized advertising based on the digital trail created by the use of Online Social Networks (OSNs) and browsers. We focus on the OSN end users' rights derived from European privacy, data protection and anti-discrimination law. Combining the legal analysis of these fundamental rights with a critical reflection on the architectural design of the USEMP system, this report provides a set of practical design implications for the USEMP tools (notably the OSN Presence tool and the OSN Economic Value Awareness tool) based on legal design requirements which drive, frame and complement the technical and social requirements. The main contribution outside legal research is the development of the so-called Data Licensing Agreement that enables OSN users to license the processing of their personal data in compliance with current EU Data Protection Law. This is the first step towards a modular version that should allow for more granular licensing of personal data processing.

Project acronym	USEMP
Full title	User Empowerment for Enhanced Online Presence Management
Grant agreement number	611596
Funding scheme	Specific Targeted Research Project (STREP)
Work program topic	Objective ICT-2013.1.7 Future Internet Research Experimentation
Project start date	2013-10-01
Project Duration	36 months

Workpackage	WP3
Deliverable lead org.	iCIS (RU)
Deliverable type	Report
Authors	Katja de Vries, Sari Depreeuw, Mireille Hildebrandt (iCIS)
Reviewers	Adrian Popescu (CEA) Marita Holst (LTU)
Version	1.4
Status	Final
Dissemination level	**PU: Public**
Due date	2014-09-30
Delivery date	2014-10-24
Revision date	2015-03-31

Version	**Changes**
1.0	Initial Release by Katja de Vries
1.1	Adjustments by Sari Depreeuw
1.2	Adjustments by Mireille Hildebrandt
1.3	Adjustments after internal review, by Hildebrandt & De Vries
1.4	Revised release after EC review

Table of Contents

1.Structure of the legal deliverables in WP3[1]

1.1. Empowerment and compliance

The overall goal of the legal input in Work Package 3 (*Legal Requirements and the Value of Personal Data*) is to elicit/engineer legal requirements that should inform the development of the various USEMP tools. Thus, the legal deliverables in WP3 are not just theoretical legal treatises on data protection, anti-discrimination, and intellectual property rights in relation to the profiles built in and through OSNs, but they aim to provide hands-on input. The first step ("finding the applicable law") in providing legal input is *descriptive*: it is an inventory of the applicable law and how it applies in the case of USEMP. This first step can be further subdivided in three sub-steps:

(i) A concise description of the applicable law;
(ii) An inventory of how the various rights at stake (that is, privacy, data protection, anti-discrimination, copy- and portrait rights of the user and the copy- and database rights of the OSN's and profile building companies) interact with each other;
(iii) An inventory of how the (interactive) functioning of these various rights could affect tools that aim to empower users who are tracked and profiled when browsing the internet and acting in OSNs.

The second step ("putting the law to work to create tools that make the user more empowered while also being compatible with the various rights at stake") of the legal input in WP3 is *constructive*, in that it aims to translate the legal conditions into legal requirements which specify:

(i) how the USEMP tools can contribute best in the effectuation of privacy, data protection, non-discrimination, profile transparency and (possibly) portrait rights. This is about empowerment.

(ii) how to make sure that the USEMP tools are compatible with the legal fields of privacy, data protection, anti-discrimination and intellectual property law. This is about compliance.

The two steps (descriptive and constructive) are not always explicitly distinguished, but they have an implicit structure in writing the legal deliverables of WP3.

1.2. Original legal research and legal coordination support

Despite the fact that all of the legal input in WP3 is quite hands-on, there are some deliverables which provide cutting-edge legal research (D3.1-3.3 and D3.6-3.8; the latter set of deliverables builds on the former) on the operationalization of "legal empowerment" from a multiple rights perspective (see Table 1). The integration of the legal requirements is taken

[1] Because this chapter discusses the overall structure of all the legal deliverables in WP3, it is repeated in the beginning of each of the legal deliverables (currently: D3.1, D3.2 and D3.3).

up in deliverables D3.4 and D3.9 that report on how the legal requirements are interfaced with the tasks at hand in the other WPs (see Figure 1).

	Version 1	Version 2
Fundamental Rights Protection by Design for OSNs	D3.1 (delivery date: M12)	D3.6 (delivery date: M21)
Profile transparency, trade secrets and Intellectual Property rights in OSNs	D3.2 (delivery date: M12)	D3.7 (delivery date: M24)
Copyrights and portrait rights in content posted on OSNs	D3.3 (delivery date: M12)	D3.8 (delivery date: M24)

Table 1: Overview of the deliverables in WP3 containing original legal research

As shown in Figure 1, the legal research (D3.1-3.3 and D.3.6-3.8) and the integration of the legal requirements into the design of the USEMP tools (D3.4 and D.3.9) are intertwined with each other. D3.1-3.3 and D.3.6-3.8 reflect the work done in T3.1-3.5 [M1-M24]. D3.4 and D.3.9 reflect the work done in T3.6, which implements legal coordination.

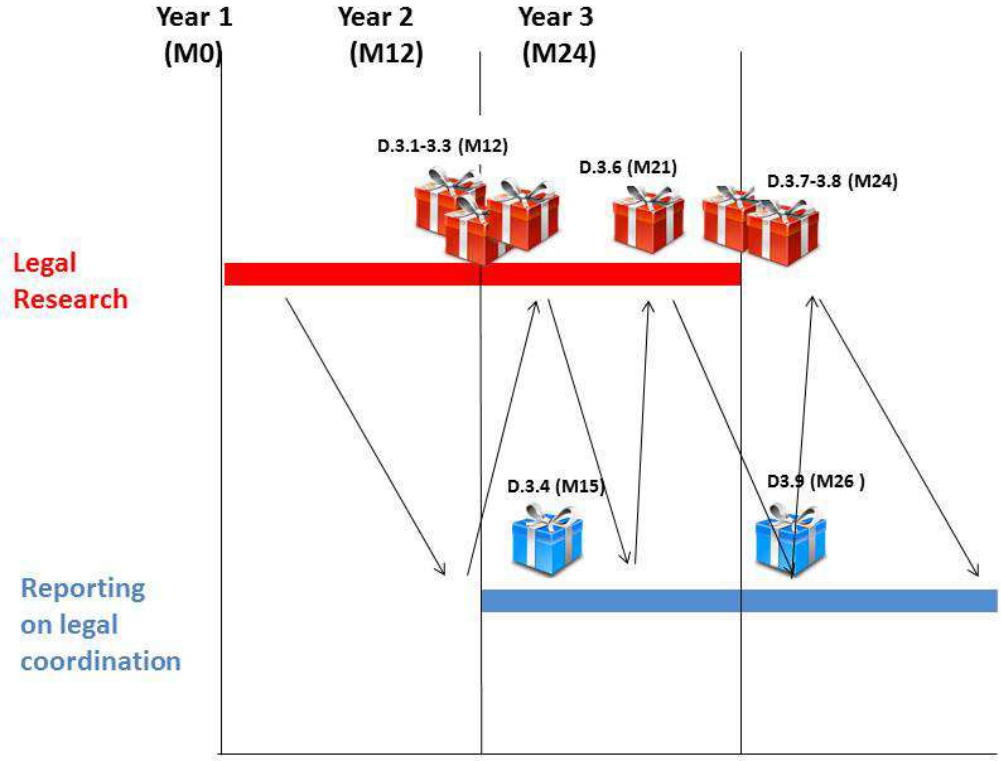


Figure 1. Timeline legal input in USEMP WP3.

1.3. Interaction between the three strands of legal research: the logic of rights trumping each other

With regard to the three strands of legal research ((a) “Fundamental Rights Protection by Design for OSNs”; (b) “Profile transparency, trade secrets and Intellectual Property rights in

OSNs"; and (c) "Copyrights and portrait rights in content posted on OSNs") it is good to mention that these, despite the fact that they are dealt with in separate deliverables, are intertwined as well. They relate to each other as a sequence of cards, where each consecutive card could trump the previous one. Thus, one could say that the *basic* legal compatibility assessment of OSNs is based on a check against data protection, privacy and anti-discriminatory requirements. When creating an application on the internet which tracks and profiles its users, the *first* question to ask is: does it infringe on data protection, privacy and anti-discriminatory requirements by doing so? And if yes: how could one adjust the *design* of the system or practice to prevent this (i.e. fundamental rights protection by design)? These are questions explored in the first step of the legal analysis (D3.1 and D3.6). The *second* question is how the outcome of the first legal step is affected when the rights of others are also taken into account. In the context of USEMP this second step is in particular interesting when profile transparency (a requirement from data protection, i.e. the "first step") is confronted with trade secrets and intellectual property rights (copy- and database rights) of the creators of the system or practice which tracks and profiles its users. With regard to this possible clash of rights, *Data Protection Directive 95/46* states in Recital 41 that:

> Whereas any person must be able to exercise the right of access to data relating to him which are being processed, in order to verify in particular the accuracy of the data and the lawfulness of the processing; whereas, for the same reasons, every data subject must also have the right to know the logic involved in the automatic processing of data concerning him, at least in the case of the automated decisions referred to in Article 15 (1); whereas this right must not adversely affect trade secrets or intellectual property and in particular the copyright protecting the software; whereas these considerations must not, however, result in the data subject being refused all information;

And in Recital 51 of the proposed *General Data Protection Regulation* one can find a similar call for a balanced approach:

> Any person should have the right of access to data which has been collected concerning them, and to exercise this right easily, in order to be aware and verify the lawfulness of the processing. Every data subject should therefore have the right to know and obtain communication in particular for what purposes the data are processed, for what estimated period, which recipients receive the data, what is the general logic of the data that are undergoing the processing and what might be the consequences of such processing. This right should not adversely affect the rights and freedoms of others, including trade secrets or intellectual property, such as in relation to the copyright protecting the software. However, the result of these considerations should not be that all information is refused to the data subject.

Can you have your cake and eat it too? Is it possible for the right to profile transparency to have some bite, if it "should not adversely affect the rights and freedoms of others, including trade secrets or intellectual property"? And what does it mean that the protection of trade secrets or intellectual property rights should not result in the data subject being refused *all* information? Is there indeed a nuanced approach possible where trade secrets or intellectual

property rights only *partly* trump the right to profile transparency? These are questions explored in the second step of a legal compatibility check D3.2 and D3.7). Finally, there is the third step of a legal compatibility check (D3.3 and D3.8), which looks at the copyrights and portrait rights in content of the end-users of OSNs and browsers. In the same way as fundamental rights can be curtailed by trade secrets or intellectual property rights of an OSN, browser or third party tracker-profiler, the protection of the latter could be curtailed ("trumped") by copy-, personality and portrait rights of the end-users of these systems.

The three-fold structure of how the various legal deliverables in WP3 build on each other implies that the *interactive* functioning of the various rights (see above, the first paragraph of this chapter) will *not* be discussed in the first step of the legal analysis (D3.1 and D3.6), but only in the second (D3.2 and D3.7) and third (D3.3 and D3.8).

1.4. A legal compatibility analysis of *what*? The double bind of the USEMP tools as both the subject and the mouthpiece of the law

In constructing the various USEMP tools, end-users are able to gain knowledge about which data are part of their digital trail, what knowledge could be inferred from such data, who is tracking them, to which actors this knowledge could be of interest and what economic value this knowledge could approximately represent. As such the information provided to the end-user of USEMP is one possible example of how legal protection by design could be implemented with regard to systems and practices which track and profile their end-users. The USEMP tools can thus be understood as supportive tools which try to embody *legal protection by design*: not only the requirement of profile transparency as formulated in EU data protection law, but also other legal requirements.

However, the USEMP project and its tools are also a research project which processes many (sensitive) data and which faces the same legal issues as any other data processor. As such, the USEMP consortium is bound by all data protection requirements: it needs to have a proper ground and purpose for the processing of data, process the data in an appropriately secure way, notify the supervisory authority of the processing (at least, if this is required by national data protection law), provide the data subject with all the necessary information about the processing of the data, etc.

Thus, from a legal perspective the USEMP project operates on two levels. On the one hand it tries to embody "legal protection by design" and as such aims to act as the *mouthpiece of the law* (or at least as a technological translation of the law) where OSNs, browsers and third-party profilers are the legal subjects addressed by the law. On the other hand USEMP is also itself a legal subject addressed of the law (at least each and every individual USEMP partner is addressed as such). As a result of this double bind (USEMP is both a translation of and a legal subject addressed by the law), the legal analyses in WP3 operate on two conceptual levels:

(a) the legal compatibility of the tracking and profiling practices performed by OSNs, browsers and third-parties, and the possibility of legal protection by design by tools such as the ones developed by USEMP, and
(b) the legal compatibility of the tracking and profiling practiced by the USEMP tools themselves.

Operating constantly on these two levels of analysis resolves the paradoxical problem that by informing the end-user about the possible “risks” of certain data (showing how sensitive metadata can be inferred: e.g., health or sexual preference from a seemingly “innocent” holiday picture), the USEMP tool itself enters in a field where one has to trod carefully, not to end up infringing fundamental rights while trying to point out (in speculative manner) how such metadata could be extracted by *other* players.

The two levels of the legal analyses in WP3 are nicely exemplified by what was mentioned above (section 1.1) as the two *constructive* forms of legal input, namely that that we need to specify both:

(i) How the USEMP tools can contribute best in the effectuation of privacy, data protection, non-discrimination, profile transparency and (possibly) portrait rights (empowerment).

(ii) How to make sure that the USEMP tools are compatible with the legal fields of privacy, data protection, anti-discrimination and intellectual property law (compliance).

Finally it should be noted that when looking at the legal compatibility between (a) the tracking and profiling practices the USEMP tool and (b) the requirements following from privacy, data protection, anti-discrimination and intellectual property law, the legal analyses also give insight about how legal compatibility would be affected if tools similar to those created by the USEMP project would be commercialized. Within the USEMP project much of data processing and profiling is allowed precisely because the purpose of the processing is purely scientific – but what would happen if (after the end of the project) these tools would still be used and they would be no longer fall under exemptions of scientific research? On top of distinguishing the two aforementioned conceptual levels of legal analysis, we should add that there are two sub-levels which can be distinguished within the second level:

(a) the legal compatibility of the tracking and profiling practices performed by OSNs, browsers and third-parties, and the possibility of legal protection by design by tools such as the ones developed by USEMP, and
(b) the legal compatibility of the tracking and profiling practices of the USEMP tools themselves.
    (i) the legal compatibility of the tracking and profiling practices of the USEMP tools as they are now, that is: processing data with the sole purpose of scientific research;
    (ii) the legal compatibility of the tracking and profiling practices of the USEMP tools as they could hypothetically be in the future, that is: commercialized and no longer part of a research project.

1.5. First and second versions of the legal research deliverables

As shown in table 1 the three strands of legal research result in six deliverables. After the first year each strand of legal research results in intermediate reports (D3.1-3.3), that will be further developed into three final reports in the second version at the end of the second year, taking into account the progression on the technical side (D3.6-3.9). D3.4 and 3.9 form the interface between the legal requirements and the technical specifications of the DataBait

tools. In the current deliverable 3.1 we have added an annex with a first version of the integration tables[2] that will be presented in 3.4.

[2] The integration tables in D3.4 contain legal qualifications of the data/content processed within the USEMP project, the requirements which are derived from these qualifications and their embodiment in the technical specifications of the DataBait tools. The qualifications and requirements follow from the various legal fields studied with regard to the USEMP project (notably data protection, antidiscrimination, copyright, sui generis database right and portrait rights derived from Art. 8 ECHR). The preliminary integration tables in annex B of this deliverable only regard requirements following EU data protection requirements and a little bit of EU antidiscrimination law.

2. The Data Licensing Agreement: a new way to empower end-users of OSNs

2.1. Introduction and objectives

In this deliverable we focus on three fundamental rights that are at stake in the context of OSNs, notably the right to data protection, non-discrimination and privacy. In section 3, we flesh out the right to profile transparency as it has been articulated in the DPD, discussing how USEMP DataBait tools can contribute to provide such transparency. This is followed, in section 4, by a more general discussion of the extent to which USEMP tools can be seen as a form of Legal Protection by Design.

In section 5 we discuss the legal requirements for making USEMP tools compatible with EU Data Protection legislation, followed by a more in-depth analysis of the rights to privacy and non-discrimination in section 6. In view of the issue of compatibility with the current legal framework (both regarding empowerment and compliance) one of the main concerns during the first year on the legal side has been to develop so-called a Data Licensing Agreement (DLA) for those participating as end-users of the USEMP platform. The aim in developing the DLA is threefold:

1. To provide a legitimate legal ground for the processing of personal data, notably also sensitive data and for downloading the USEMP DataBait tools;
2. To engage the end-user (data subject) by asking her to enter into an obligatory agreement with the USEMP partners (joint data controllers), clarifying mutual rights and obligations;
3. To present the end-user (data subject) with a clear, concise transparent agreement that is legible for lay people and covers all the relevant issues of compliance on the side of the USEMP service providers (joint controllers)

In this section 2 we therefore present the DLA and the underlying personal data processing agreement that has been concluded between the USEMP Consortium Partners (as joint controllers), thus binding the partners to provide some form of profile transparency in exchange for a specified license to process the user's (data subject's) personal data. We will explain the relationship between the DLA and the consent requirement for processing sensitive data (art. 8 DPD) and between the DLA and the consent requirement for storing tracking mechanisms on the user's (subscriber's) device (art. 5.3 ePrivacy Directive).

The idea of employing a data licensing agreement is new and hopes to provide for a new way of addressing the power imbalances between users and providers of OSNs3. It is based

[3] The fact that the service of an OSN is rendered at no cost does not justify a weak position of the user in terms of consumer and data protection. Moreover, the notion of "service at no cost" must be nuanced. See e.g. Wauters e.a. 2014, p. 10: "*Since most SNS do not require an actual payment of a fee, we wonder if SNS can fall under the scope of the Consumer Rights Directive. […] However, it is often stated that personal data is the new currency of the Internet. A SNS offers*

on the fact that data subjects have a bundle of rights with regard to the processing of their personal data. This allows them to contract about such processing to the extent that processing is not e.g. mandatory for reasons of public security or necessary for the legitimate interest of the data controller. For further explanation of the various conditions for lawful processing of personal data, we refer to section 5.

2.2. USEMP Data Licensing Agreement (DLA)

As indicated, the data licensing agreement (DLA) will be concluded between the USEMP Consortium Partners (as joint data controllers) and the end-users of the USEMP tools. It clearly defines the mutual legal obligations, taking the end-users seriously as participants in the research that is conducted. It is also the legitimizing ground for the data processing in USEMP ("contract" as decribed in Art. 7b of Data Protection Directive 95/46).

The DLA is implemented in the USEMP graphic user interface (GUI) and will be part of the sign-up procedure. Each article of the DLA will be presented as a separate screen. The underlying Personal Data Processing Agreement (PDPA, see below) can be seen as an offer made by all each of the USEMP Consortium Partners to conclude the DLA; when the end-user clicks accepts this offer by clicking the button at the end, each USEMP Consortium Partner is bound by the DLA. We note that:

- Arts. A and F clearly define the obligations of DataBait users, while also specifying the consequences in terms of which data will be tracked.
- Art. B clarifies that this agreement entails that the DataBait users license the usage of their personal data for a specified purpose.
- Art. C provides for the consent required in art. 5(3) of the ePrivacy Directive.
- Arts. D and E further specify the purpose for which the USEMP Consortium Partners will use and process the data, notably in terms of the OSN presence management tool and the monetization tool.
- Art. G provides for the consent required for the processing of sensitive data (art. 8 DPD)
- Art. H, I, J further specify the duty of care for the USEMP Consortium Partners when they process the personal data of the DataBait users, stipulating the ilfe cycle management of the involved personal data (collection, usage, deletion or full anonymisation). Art. H also emphasizes the reason for processing sensitive data.

---

its service to users and in exchange, they gather (explicitly through registration forms or 'secretly' via cookies) personal data of their users. Because of this personal data, they are able to offer personal advertisements in order to make a profit. Another indication may be found in the definition of information society services under the e-Commerce Directive (above), which includes service which are financed by advertising." Following Wauters it might be argued that based on the Consumer Rights Directive the license granted by the users to Facebook is too broad and not legally valid. The PDPA which is signed between the USEMP consortium and the users of the DataBait tool is a first step to a more balanced approach, and which can form the basis for a more granular licensing approach. This entails that a later, modular version of the DLA should include licensing of copyrighted material posted on the OSN.

In the context of D3.6 (the next version of this deliverable) we will present a modular version of the DLA, enabling data usage licensing via DataBait tools for profile transparency, with other service providers that may have a commercial interest in providing the tools. This will entail that the purpose is extended or adapted.

Screen 1:

USEMP Data License Agreement

The parties:

(1) […….], user of the USEMP platform and services, from hereon called 'You' and

(2) [CEA-France / iMinds-Belgium/ CERTH-Greece / HWC-UK/ LTU- Sweden /VELTI-Greece/ SKU Radboud University-the Netherlands]4, provider of the USEMP platform and services, joint data controllers, from hereon called 'USEMP consortium partners' 5.

Hereby agree:

Screen 2:

(A) You will install the USEMP DataBait tools, the DataBait-Facebook app and the DataBait web browser plug-in and the DataBait graphic user interface (GUI). The DataBait-Facebook app and the DataBait web browser plug-in will provide access to Your Facebook profile and Your browsing behaviour on Your device(s). These tools will be used by the USEMP consortium partners to collect data that You share on Facebook as well as data collected by the web browser. This data can be data You posted (volunteered data), or data captured by the USEMP tools (observed data). The latter concerns online behavioural data (storing what You did on the Internet and on FaceBook).

This article defines the obligation to install the DataBait tools, which is pertinent for participation in the USEMP research. It clarifies upfront that both volunteered (declared) data will be processed and observed (behavioural) data. In a later, modular version of the DLA, not necessarily focused on scientific research, the same article can be used.

[4] The name of each partner will contain a hyperlink, such that users can click on it and check the organisational website of which is involved.

[5] This text will contain a hyperlink, stating:"*The USEMP consortium partners have entered a separate agreement between themselves, obliging themselves to act in accordance with this contract, their national data protection law and EU data protection law, in which agreement they clarify which partners processes what personal data. This contract can be accessed here.*" When one clicks one "here" this will lead to the *USEMP Personal Data Processing Agreement* (see Appendix 3 of this Deliverable).

Screen 3:

(B) You license the use of Your volunteered and observed personal data by the USEMP consortium partners, as gathered by the the DataBait-Facebook app and the DataBait web browser plug-in for the sole purpose of scientific research and – within that context – to provide You through the DataBait graphic user interface (GUI) with information about what third parties might infer based on Your sharing of information, and on Your online behaviour. The said data may be combined with publicly available personal data gained from other sources to infer more information about Your habits and preferences (inferred data).

This article, first, makes clear that this is a quid quo pro agreement, creating legal obligations on the side of the user (data subject) in the form of licensing the use of the data that will be processed by the USEMP consortium, and on the side of the service provider (data controller) in the form of providing a form of profile transparency. Second, it determines the specific purpose of processing. In a later, modular version of the DLA, not necessarily focused on scientific research, this article will have to be adapted. In principle the article will be replaced with the relevant specific purposes and the relevant consideration (performance on both sides).

Screen 4:

(C) This license agreement confirms Your explicit consent to store the DataBait tools on Your devices.

This article provides the consent required on the basis of art. 5.3 ePrivacy Directive for all and any tracking mechanisms to be stored on the user's (data subject's) device. That such tools contain tracking mechanisms is clarified in the previous articles A and B – the consent thus includes any cookies that are stored on the device, which are – in this case – necessary to fulfil the functionality of the service that is provided. This means that consent may not be required, since – according to the art. 29 WP consent is not required for functional cookies. To be on the safe side we have included this consent. We advise that this article is part of later, modular versions of the DLA.

Screen 5:

(D) The USEMP consortium partners will do scientific research to predict what kind of information Facebook or other third parties with access to Your postings and online behavioural data could or might infer from the said data. These inferences will be shared with You in an intuitive manner, thus providing an online presence awareness tool, embedded in the "DataBait-GUI".

This article further explains the obligation on the side of the service provider, and the purpose of processing, highlighting that the profile transparency which will be provided is

based on statistical inferences by others than OSN providers, meaning that the user is made aware of the fact that the USEMP Consortium partners are not reverse engineering algorithms of the OSN provider and cannot in any way provide certainty about how one may be targeted. This article also ensures that the transparency is provided in a user-friendly manner. This article is crucial in any DLA for DataBait tools.

Screen 6:

(E) The USEMP consortium will also do scientific research to estimate the monetary value of Your data, based on the said data and their inferences. The "DataBait-GUI" will alert You that some of Your online behaviours may be monetisable, for example in the context of personalized advertising or in the context of selling Your data or profile to data brokers, credit rating companies or others willing to pay for access to the data or inferred profiles. This way the DataBait-GUI also acts as an economic value awareness tool.

As with the previous article, this article highlights that the monetary value is an estimation and in no way a claim as to the actual monetary value that may be generated with the DataBait tools. In fact the consortium has decided to refrain from providing an estimate of the monetary value, instead developing an estimate of the added value for the OSN provider or third parties with whom data may be shared. This article is not necessary of a later, modular version of the DLA.

Screen 7:

(F) You agree to participate in surveys and/or focus groups, to enable the consortium to gain insights in how users engage with social networking sites and how they evaluate (1) various scenarios regarding the use of their personal data and targeted profiles and (2) the effectiveness, usability and utility of the USEMP tools.

This article clarifies that the user will participate in the research that enables to correlate their declared preferences or personality traits with the inferences drawn from behavioural data or the mining of multi-media content. This article may be part of a later modular version of the DLA.

Screen 8:

(G) You hereby grant Your consent to process Your sensitive personal data, notably those revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, and those concerning health or sex life.

Since consent is required for processing art. 8 DPD types of data, this article stipulates such consent. It highlights the intrusive nature of the processing of such data. It should be part of a later, modular version of the DLA.

Screen 9:

(H) The USEMP consortium partners will treat all Your personal data, especially Your sensitive data, with care and delete or anonymize them as soon as possible. Because one of

the main goals of the USEMP project is to create awareness about the possibility to infer sensitive data from trivial data trails, it is important to alert You to such inferences and thus to process them.

Screen 10:

(I) The USEMP consortium partners will process Your personal data in a secure way and not keep them any longer than necessary for the purpose of the USEMP study. In order to provide You with access to Your personal data and the inferences drawn from them, the data may be kept until the end of the project. Within 3 months of the ending of the research project all personal data will be either deleted, anonymised or processed for related scientific research. In the latter case the relevant USEMP consortium partner will ask You for Your consent.

Articles H and I confirm the legal obligation for the USEMP partners (joint controllers) that the relevant data will be processed in accordance with the data minimisation principle, stipulating deletion or anonymisation as soon as possible (including a clear deadline) and security by design, while also explaining that to provide profile transparency the processing of both personal and sensitive personal data is necessary. These article should be part of a later, modular version of the DLA, considering that this is a confirmation and reminder of the legal obligations of the service provider (data controller).

Screen 11:

(J) The USEMP consortium partners will not provide Your personal data to any third party other than the Future Internet Research and Experimentation Initiative (FIRE) infrastructure, which is a multidisciplinary scientific infrastructure funded by the EU in which novel internet related tools can be tested and validated. The transfer of the data will happen in a secure way and only in as far as strictly necessary for the scientific goals of the USEMP project.

This article is pivotal to ensure that in the context of USEMP data are not processed beyond the explicitly specified purpose, by the parties to the contract, simply prohibiting any transfer to third parties other than the FIRE infrastructure. The article can be modulated depending on the specifics of a later version of the DLA, for instance allowing to share data with specified third parties and/or specified types of third parties.

Screen 12:

(H) The national law of Your country of residence (at the moment of registration) is applicable to this contract, assuming you are a resident of the EU.

By clicking the box below You become a party to this agreement:

To prevent any confusion about the applicable national law, and to accommodate the natural person whose personal data are being processed we confirm that the national law of the end-user (data subject) of the USEMP platform is applicable. Under current EU Data

Protection Law this seems the most apt, also for a later, modular version of the DLA. This may change under the proposed General Data Protection Regulation (pGDPR).

2.3. USEMP Personal Data Processing Agreement

This internal agreement between the USEMP Consortium Partners specifies which partner will do what kind of processing of personal data, and determines that and how the Consortium Partners are legally bound to treat the personal data they are processing. We note:

- The USEMP partners act as joint data controllers because they have jointly determined the purpose of the processing of personal data within the USEMP project, namely scientific research as explicated in the DOW, the DLA and the PDPA.
- The DLA is part and parcel of this contract; the PDPA is an irrevocable offer to DataBait end-users to conclude the DLA contract. A link will be placed in the DLA to the DPDA contract.
- The PDPD contains strict obligations in terms of the appropriate security measures regarding the capture, storage and transmission of personal data, based on a risk assessment performed by each partner.
- The PDPA thus clarifies to the end-users of the USEMP tools which partner does what kind of processing of data and, finally exonerates partners from liability for data processing performed by other partners over which they have no actual control.
- The PDPA also addresses the user-friendly, layered and precise information to which end-users of the USEMP platform (data subjects) are entitled by stipulating that two bottons will be visible and operational on the platform's website: (1) to obtain more detailed information about the way USEMP Consortium Partners are bound to deliver on the contract, by showing the PDPA contract and by adding a table which shows in even more detail what data are processed how and for what reasons in the design of the USEMP architecture; and (2) to obtain from the USEMP Consortium Partners the erasure of their sensitive data or the removal of the DataBait tools.

USEMP Personal Data Processing Agreement (PDPA)

The parties:

(1) CEA-France,
(2) iMinds-Belgium
(3) CERTH-Greece
(4) HWC-UK
(5) LTU- Sweden
(6) VELTI-Greece
(7) SKU Radboud University-the Netherlands

having concluded the USEMP Consortium Agreement, being providers of the USEMP platform and the DataBait tools and services, and being joint data controllers,

Hereby agree:

(A) Each party will comply with and perform in accordance with the USEMP Data Licensing Agreement (DLA, as attached to this contract) when processing the personal data of DataBait Users, who are defined as the USEMP end-users who have signed the Data Licensing Agreement with the USEMP Consortium Partners.

(B) Each party will comply with their national and EU data protection law, including notification of their national Data Protection Authority if necessary under their national law, when processing the personal data of DataBait Users or any other personal data processed in the context of USEMP.

(C) Each party will provide precise information on what type of personal data they process concerning DataBait users, how it is processed and which data-flows they enable. This information will be available for DataBait users after clicking the button on the USEMP platform, and include an email address for each partner that processes personal data, to make further inquiries. The information will be updated whenever the relevant processing of personal data change. Each party will also provide an email address to be contacted in case a user wants to withdraw her consent for processing her sensitive data; this is preferably the same email address as the one used to gain further information, but will be available behind a separate button on the USEMP platform.

(D) All parties shall carry out a personal information assurance risk assessment from their own context concerning their own collection, storage and/or processing of personal data, prior to deployment of the live service when personal data will be collected, and at any point through the operation of the system where there is a relevant change to either hardware installation, software versions, and/or software interfaces. Such a risk assessment shall follow information assurance principles covering, at least, hardware installation, software development processes, software validation and approval, software execution and backup processes. Each partner is liable for inappropriate security at its own premises.

(E) Parties agree that the following processing of personal data will be performed by the following parties:

> CEA-France will conduct the following processing of personal data: via image recognition and text mining techniques CEA will infer potential preferences for specific objects, places and brands. No personal data of DataBait Users will be stored at the premises of CEA, that will be authorized to run its algorithms on the data stored at HWC.
>
> iMinds Belgium will conduct the following processing of personal data: together with CERTH and LTU, iMinds will prepare a survey asking registered users of the USEMP platform and the DataBait tools to answer a set of questions about their lifestyle preferences, selected health issues and personality traits, religious and political beliefs, sexual orientation, gender, age, place of residence and ethnic background. iMinds will conduct the survey to enable testing of how the inferences drawn from DataBait Users' postings, social graphs and behavioural data match their real preferences and background. The outcome of the survey feeds into the database that is stored at HWC. iMinds can access the result of the survey based on secured authorization. The transmission of these sensitive data will be done in a secure way

by means of appropriate security protocols. iMinds will also conduct user interviews which contain personal user's information. Interviews will be anonymized, transcribed and stored in an appropriately secured server, only accessible to authorized iMinds personnel.

CERTH-Greece will conduct the following processing of personal data: via image, text mining and behavioural profiling techniques (involving the 'likes' and sharing of Facebook pages and visits to URLs) CERTH will make inferences about undisclosed demographic characteristics (gender, age, origin), place of residence, sexual orientation, personality and health traits, as well as potential lifestyle preferences, including those that may interest specific types of brands and enterprises. When developing the DataBait tools, a small portion of DataBait User data will be stored at CERTH. In that case appropriate security protocols will be in force, considering the nature of the data. Data will be deleted or fully anonymized once they are no longer necessary for developing the DataBait tools. CERTH will be authorized to run its algorithms on the data stored at HWC.

HWC-UK will conduct the following processing of personal data: all data collected through the DataBait tools are directed to and stored at HWC, who will secure the data and provide secure access to the USEMP partners for the sole purpose of scientific research as specified in the DLA contract and the description of work that is part of the Grant Agreement with the EU. During storage at HWC appropriate security protocols will be in force concerning storage and access. Data will be deleted or fully anonymized as soon as the scientific purpose as stated in the DLA agreement is fulfilled.

LTU- Sweden will conduct the following processing of personal data: together with CERTH and iMinds, LTU will prepare a survey asking registered users of the USEMP platform and the DataBait tools to answer a set of questions about their lifestyle preferences, selected health issues and personality traits, religious and political beliefs, sexual orientation, gender, age, place of residence and ethnic background. LTU will conduct the survey to enable testing of how the inferences drawn from DataBait Users' postings, social graphs and behavioural data match their real preferences and background. The outcome of the survey feeds into the database that is stored at HWC. LTU can access the result of the survey based on secured authorization. The transmission of these sensitive data will be done in a secure way by means of appropriate security protocols. LTU will also conduct user interviews which contain personal user's information. Interviews will be anonymized, transcribed and stored in an appropriately secured server, only accessible to authorized LTU personnel.

VELTI-Greece will conduct the following processing of personal data: based on the inferences made by CEA and CERTH, VELTI will conduct further processing operations to visualize information on potential inferences to be provided to the DataBait users. Velti will also use historical Facebook and behavioural data of DataBait users, stored at HWC, for the estimation of the (monetary) value of the personal data of the DataBait users. Some of this data may be retrieved from HWC and stored temporarily at VELTI for preliminary testing. In that case appropriate security protocols will be in force, considering the nature of the data. Data will be deleted or fully anonymized as soon as the purpose of such testing is achieved.

SKU Radboud University-the Netherlands will not conduct any processing of personal data.

(F) Each party that processes personal data hereby exempts all other parties from liability for any unlawful processing of personal data, and from processing personal data in violation of the USEMP DLA or this PDPA. Thus parties will not be severely liable for violations committed by other parties.

(G) Belgium law will be applicable to this contract.

3. Prohibited forms of profiling and the right to profile transparency: translating user empowerment in OSNs into the discourse of European fundamental rights (part I[6])

3.1. User empowerment and profile transparency

The USEMP project aims to develop tools that enable users of social networks and browsers to control their digital trail and to understand how their data are used by the providers of social networks and browsers and by third parties piggy-backing on these systems. One of the underlying assumptions of the USEMP project (or, as it will be known to the end-user: the *DataBait tools*) is that *knowledge is power.* The idea is that if light can be shed on the world of tracking and personalized advertising, this will result not only in a better informed user, but also in a more *empowered* user. Currently, the world of tracking and targeting is invisible to the ordinary internet and social network user and its opacity makes it impossible to answer basic questions such as: *What information can be inferred from my data? What is the economic value of my data? What kinds of measures can be taken based on my volunteered, observed and inferred data? Who tracks me? Which commercial actors have access to my data?* Providing relevant knowledge to the user, the USEMP project also assumes that this knowledge is all the more powerful when it is not presented in a generalized way but as knowledge about the particular data trail of a concrete user: instead of knowledge about an abstract average user, the USEMP (aka *DataBait*) tools aim to provide personalized insights with regard to each individual user of these tools.

User empowerment and personalized knowledge about the commercial impact and technological possibilities based one's digital trail are not legal terms as such. In order to know how USEMP relates to the law, we have to translate these terms in legal terminology. Some legal notions, derived from the field of EU data protection law, bear a very obvious relation to the *knowledge is power*-assumption underlying USEMP: the obligation of the data controller to inform a data subject about certain aspects of data processing (Arts. 10 and 11 of the *Data Protection Directive 95/46/EC*[6] [DPD 95/46]) and the data subject's right to access data (Art.12 DPD 95/46) and to object to the processing of them (Art. 14 DPD 95/46). Another legal notion which is relevant to USEMP is *profiling*, that is, a specific kind of data processing which can be described as:

> 'profiling' means any form of automated processing of personal data intended to evaluate certain personal aspects relating to a natural person or to analyse or predict in particular that natural person's performance at work, economic situation, location,

[6] *Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data*, Official Journal L 281, 23/11/1995, p. 31-50. The DPD 95/46/EC is currently the main legal instrument regarding general data protection, but is in the proces of being replaced by the proposed *General Data Protection Regulation*, which will probably enter into force in 2016.

health, personal preferences, reliability or behaviour. (Art. 4-3a of the proposed *General Data Protection Regulation*[7] [GDPR], the successor to DPD 95/46)

The knowledge that the USEMP tools aim to provide is largely about this specific form of data processing: the tools do not only inform the users about which trackers track which of their data ("simple" data processing), but also about which evaluative knowledge could be derived from these data ("profiling"). While the term *profiling* as such is not present in the DPD 95/46, the Directive does contain a specific provision of what can be called *the right to profile transparency*. This right *to obtain knowledge of the logic involved in any automatic processing which significantly affects the data subject* can be derived from Article 15(1) in conjunction with Article 12(a) of the DPD 95/46:

Article 15

Automated individual decisions

1. Member States shall grant the right to every person not to be subject to a decision which produces legal effects concerning him or significantly affects him and which is based solely on automated processing of data intended to evaluate certain personal aspects relating to him, such as his performance at work, creditworthiness, reliability, conduct, etc.

2. Subject to the other Articles of this Directive, Member States shall provide that a person may be subjected to a decision of the kind referred to in paragraph 1 if that decision:

(a) is taken in the course of the entering into or performance of a contract, provided the request for the entering into or the performance of the contract, lodged by the data subject, has been satisfied or that there are suitable measures to safeguard his legitimate interests, such as arrangements allowing him to put his point of view, or

[7] The proposed *General Data Protection Regulation* (GDPR) is currently being created in the so-called *ordinary legislative procedure* (formally known as the *codecision procedure*) of the EU, which is basically a bicameral legislative procedure: it gives the same weight to the European Parliament and the Council of the European Union (consisting of ministers from the 28 EU Member State governments). The GDPR was first proposed on 25 January 2012 by the European Commission (that is, the executive branch of the EU and the only EU institution empowered to initiate legislation) and now has to be jointly adopted by the European Parliament and the Council. The text proposed by the Commission [*Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)*, Brussels, 25 January 2012, COM(2012) 11 final] has been subjected to a first reading by the European Parliament and has been amended the on 12 March 2014 [*European Parliament legislative resolution on the proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)*, Strasbourg, 12 March 2014, COM(2012)0011 – C7-0025/2012 – 2012/0011(COD), online available at : <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2014-0212+0+DOC+XML+V0//EN>]. Currently, the amended text is examined by the Council of the European Union. If Parliament and Council disagree on a proposed legislative text, it can go back and forth between Parliament and Council up to three times. A clear infographic clarifying the ordinary legislative procedure can be found here : <http://www.europarl.europa.eu/aboutparliament/en/0081f4b3c7/Law-making-procedures-in-detail.html> [last accessed 29 September 2014]. Looking at the current status of the proposed General Data Protection Regulation and the steps in the legislative procedures still to be taken, the GDPR will most likely enter into force by 2016.

(b) is authorized by a law which also lays down measures to safeguard the data subject's legitimate interests.

Art. 12(a)

Right of access

Member States shall guarantee every data subject the right to obtain from the controller […] knowledge of the logic involved in any automatic processing of data concerning him at least in the case of the automated decisions referred to in Article 15(1)

With the fast technological development and commercial spread of profiling techniques, *profiling* has developed into a proper legal notion gaining lots of attention. For example, the Council of Europe (Council of Europe 2010) and Working Party 29 (Article 29 Data Protection Working Party 29 2013, 13 May) have devoted quite some attention to the legal definition of *profiling* and to *how the right to profile transparency* could be further developed, and the term is abundantly present in the proposed GDPR (see Annex 1).

Thus, while it is undeniably true that a basic version of *the right to profile transparency* is already present in the current DPD 95/46, it will be present more explicitly and in a stronger and more elaborate way in the future data protection legislation of the GDPR. This becomes particularly clear in Arts. 14ga and 14gb (regarding the information which has to be provided to the data subject when profiling takes place) and Art. 20 (fully devoted to profiling and stipulating when it is allowed and when not) of the pGDPR. Thus when comparing Art. 14ga of the pGDPR to the provisions in the current DPR 95/46, it is clear that the pGDPR is much more specific about the kind of information that has to be provided: the data controller shall provide the data subject with information about the *existence* of profiling, of *measures based on* profiling, and the *envisaged effects* of profiling on the data subject.

It is interesting that the focus is here not just on the profiling as such but also on what the profiling actually *does* in practice. Objecting to profiling should not just be a theoretical possibility but a right that is actually used ("*The data subject shall be informed about the right to object to profiling in a highly visible manner*", Art. 20(1) pGDPR). Moreover, contrary to the current DPD 95/46, which prohibits profiling which has significant or legal effects and is based *solely* on the automated processing of data ("… *the right to every person not to be subject to a decision which produces legal effects concerning him or significantly affects him and which is based solely on automated processing of data intended to evaluate certain personal aspects relating to him*", Art. 15(1) DPR 95/46), the proposed GDPR also prohibits such profiling if it solely *or predominantly* relies on automated processing (Art. 20(5) pGDPR). Furthermore, the proposed GDPR continues, such profiling "shall include human assessment, including an explanation of the decision reached after such an assessment" (Art. 20(5) pGDPR).

Another striking difference is that the proposed GDPR (Art. 20(3) pGDPR) categorically prohibits profiling with discriminatory effects with regard to race or ethnic origin, political opinions, religion or beliefs, trade union membership, sexual orientation or gender identity, and demands of the controller that effective protection against possible discrimination resulting from profiling should be in place. The pGDPR (Art. 20(3)) also prohibits profiling which is *solely* based on data revealing race or ethnic origin, political opinions, religion or

philosophical beliefs, sexual orientation or gender identity, trade-union membership and activities, and the processing of genetic or biometric data or data concerning health or sex life, administrative sanctions, judgments, criminal or suspected offences, convictions or related security measures. These sensitive data are not wholly excluded from being used as input in a profiling process, but they should always be combined with other, non-sensitive, data.

Thus, to summarize: in the pGDPR both the input (the data on which the profiling is *based*) and the output (the *effects*) of profiling are scrutinized to prevent discrimination based on a set of protected grounds. Next to the prohibition of discriminatory profiling, the pGDPR also prohibits profiling in the field of employment (Art. 82(1) pGDPR). Thus, overall the pGDPR will offer a better protection against unwarranted forms of profiling and gives the right to profile transparency more teeth. However, the GDPR also introduces some provisions which could make the protection against unwarranted profiling somewhat weaker. Profiling "which leads to measures producing legal effects concerning the data subject or does similarly significantly affect the interests, rights or freedoms of the concerned data subject" is currently only allowed if there is a law authorizing such processing or if such profiling is necessary for entering or performing a contract lodged by the data subject (Art. 15(2a) DPD 95/46). In the pGDPR such profiling could also be allowed based on the consent of the data subject. Another addition in the GDPR which could weaken the protection against profiling is the presumption (Recital 58(a) pGDPR) that profiling based solely on the processing of pseudonymous data (i.e., personal data that cannot be attributed to a specific data subject without the use of additional information) should be presumed not to significantly affect the interests, rights or freedoms of the data subject. However, this presumption is highly contested and it will be interesting to see what the Council will do with it in the upcoming step of the legislative process.

Notwithstanding the differences between the current and future *right to profile transparency*, the main *rationale* in both of them seems to be very similar to the assumption of user empowerment underlying USEMP: profile transparency aims to prevent that a data subject is confronted with a "Computer says no" in a situation which significantly affects his or her interests.[8]

However, as soon as law is involved the devil is in the details: profile transparency is not something that a data subject can *always* appeal to. The law is more subtle than the straightforward adage that a user, to whom a profile is applied, can always request *full transparency*. Law is a practice of nuance. Even if we follow the rather straightforward formulation in DPD 95/46, the question *if* the right to profile transparency applies and *how to comply* with it, requires that one looks into a set of specifics such as, for example:

[8] It should be noted that the commercial profiling applications studied in the USEMP project seem to be mainly steered by the interest of nudging a consumer into a particular commercial transaction (*Computer says : "Please, do.. "*) and do not primarily aim to take decisions which are contrary to the user's will (*Computer says no*). However, the line between nudging positively ("*Please, do..* "), nudging negatively ("*Please, don't..* ") and denial of service ("*No!*") is often thin an fluid. For example, think of an insurance company nudging a certain type of users to become their customers with specific discounts (positive nudging through price differentiation). One could say that the flipside of this positive nudge is that this company gives a negative nudge to potential customers who do not fit the profile.

- *Can the "profiling" at stake indeed be qualified as the action described in Art. 15* ("automated processing of data intended to evaluate certain personal aspects relating to him, such as his performance at work, creditworthiness, reliability, conduct, etc. ")*?*
- *Is the provision of insight into the profile in accordance with what is required by Art. 12* ("knowledge of the logic involved in any automatic processing of data")?
- *Does the evaluative automated processing of data result in a decision that produces legal effects or significantly affects the data subject?*
- *Is this decision* solely *based on the automated processing of data or is it based on a combination of human and automated decision making?*
- *Is there a legal ground legitimizing the profiling?*
- *Does the right to profile transparency adversely affect trade secrets or intellectual property rights of other actors (*Recital 41 DPD 95/46)*?*

Only by looking at both the legal details and those of the technological architecture, is it possible to answer the question if a particular tool, system or practice is compatible with the right to profile transparency. This is even more so with the elaborate version of the right to profile transparency in the proposed GDPR, where even more aspects have to be considered, such as, for example:

- *Does the profiling process result in discrimination against individuals on the basis of race or ethnic origin, political opinions, religion or beliefs, trade union membership, sexual orientation or gender identity?*
- *Is the profiling based solely on the processing of pseudonymous data?*
- *Can the data controller nevertheless attribute these pseudonymous data to a specific data subject?*
- *Is the profiling used in the context of employment?*

3.2. Issues with regard to profile transparency that need to be further explored

In terms of the compatibility of the feedback about profiling based on one's digital trail [i.e. the information which each USEMP end-user gets through the *DataBait* graphic user interface (GUI)] with the right to profile transparency and the provisions prohibiting certain forms of profiling, the following points are worth exploring in more detail when further developing the USEMP architecture:

a. Do current systems or practices which profile end-users provide any of the following information – in clear and plain language – with regard to the data that they collect:
    - the purpose for which the data are processed
    - what categories of data are processed,
    - for what estimated period,
    - which recipients receive the data,
    - what is the general logic of the data that are undergoing the processing,
    - what might be the consequences of such processing,

- the existence of the right to request rectification or erasure of the data concerning the data subject and of the right to object to the processing,
- the right to lodge a complaint to the supervisory authority and the contact details of the supervisory authority? (See Recital 51 and Art. 15 of the pGDPR)
- If such information is provided – in what format?

b. Do current systems or practices which profile end-users provide any of the following information about the profiling:
   - the existence of profiling,
   - measures based on profiling,
   - the envisaged effects of profiling on the data subject,
   - the right to object to profiling (the latter should be done in a highly visible manner).

   If yes – how? (See Art. 14ga and Art 20(1) of the pGDPR)

c. Does the USEMP GUI provide the user with any of the following information – in clear and plain language – about the data that OSNs, browsers and third parties collect:
   - the purpose for which the data are processed,
   - what categories of data are processed,
   - for what estimated period,
   - which recipients receive the data,
   - what is the general logic of the data that are undergoing the processing,
   - what might be the consequences of such processing,
   - the existence of the right to request rectification or erasure the data concerning the data subject or to object to the processing
   - the right to lodge a complaint to the supervisory authority and the contact details of the supervisory authority? (See Recital 51 and Art. 15 of the pGDPR)

   With regard to which categories is USEMP able to provide the user any information?

   With regard to which categories is that impossible?

d. Does the USEMP GUI provide the user with any of the following information – in clear and plain language – about the data that the USEMP tools (the DataBait Facebook app and the DataBait browser plugin) collect:
   - the purpose for which the data are processed
   - what categories of data are processed,
   - for what estimated period,
   - which recipients receive the data,
   - what is the general logic of the data that are undergoing the processing,
   - what might be the consequences of such processing,
   - the existence of the right to request rectification or erasure of the data concerning the data subject and of the right to object to the processing,

- the right to lodge a complaint to the supervisory authority and the contact details of the supervisory authority? (See Recital 51 and Art. 15 of the pGDPR)

Most of this information is included in the Data Licensing Agreement (see Annex 2) which is part of the DataBait GUI and that has to be signed by every user of the DataBait tools.

e. Does the user-feedback in the DataBait GUI include information about:
   - the existence of profiling,
   - measures based on profiling,
   - the envisaged effects of profiling on the data subject? (See Art. 14ga of the pGDPR)

f. Does the user-feedback in the DataBait GUI qualify as "meaningful information about the logic of any automated processing"? (See Art. 14gb of the pGDPR) Here it is particularly interesting to interface with the ongoing work in Task 6.3 ("*Visualisation of and Interaction with user empowerment data*") with regard to good user interfaces that display information in such a way that is does not become too complex or overwhelming. From a legal perspective (see e.g. Wauters, Lievens et al. 2014, p. 292) it is important to study if such (simplified) visualizations still offer enough detail to qualify as "meaningful information about the logic of any automated processing" (Art. 14gb GDPR; see also Art. 12a of the DPD 95/46). There is a fine line between (a) ease of user interface and intuitiveness of the representation, (b) too much simplification. As an aside it should be noted that this point is a good example that a legal compatibility assessment is not always a one-way street where a technological or organizational architecture is simply checked against a set of legal requirements. Because legal terms (e.g., "knowledge of the logic involved in any automatic processing of data", Art. 12 DPD 95/46) do not always have an exhaustive definition, the design solutions in the USEMP project might actually be an inspiration to the lawyer. Legal requirements and technological design can be a two-way street.

g. Profile transparency is not just something that the USEMP tools provide *about* other profiling systems and practices, but is a requirement which also applies to the profiling performed by the USEMP consortium. For example, when providing users with information about the possible economic value of certain parts of their data trail, it is important to also show the DataBait user how monetary value is modelled by the USEMP consortium – because, obviously, there are *many* ways to model economic or monetary value. (See for research on the actual price of data trails: Olejnik, Minh-Dung et al. 2014)

h. Is there any evidence that the profiling performed by the OSNs, browsers and third party trackers results in what could be qualified as "*measures producing legal effects concerning the data subject or [that] […] similarly significantly affect the interests, rights or freedoms of the concerned data subject*"? (Art. 20(2) pGDPR) Can we give examples of such measures? If the USEMP consortium concludes that it is possible that profiling results in such measures, what is the best way to integrate information about this in the DataBait GUI?

i. Can the DataBait GUI, based on the data trail of the user, present the user with an example of such "*measure producing legal effects concerning the data subject or [that] […] similarly significantly affect[s] the interests, rights or freedoms of the concerned data subject*"? Could the DataBait GUI present the user with a (fictional) example of the human assessment and with an explanation of the decision that would be reached after such an assessment, thus clarifying why and how human assessment is required when profiling results in legal or otherwise significant effects?

j. Are the data collected by the DataBait Facebook app and the DataBait browser plugin anonymized, pseudonomized or neither? Is it technologically possible to do so without losing the possibility to provide the end-user with personalized feedback?

k. Does the USEMP project infer any knowledge about the Databait user which is based *solely* on data revealing race or ethnic origin, political opinions, religion or philosophical beliefs, sexual orientation or gender identity, trade-union membership and activities, and the processing of genetic or biometric data or data concerning health or sex life, administrative sanctions, judgments, criminal or suspected offences, convictions or related security measures? (This is prohibited in Art. 20(3) of the pGDPR)

l. Could one envision situations where the digital trail of Databait users would result in discrimination against individuals on the basis of race or ethnic origin, political opinions, religion or beliefs, trade union membership, sexual orientation or gender identity (Art. 20(3) of the pGDPR)? Is there any evidence that this is the case? As an example of such discriminatory effects one could, for example, further explore the case introduced by Sweeney (2013), who showed that an ad for a company selling information on whether individuals had been arrested or convicted, was more likely to show up next to the Google search results if the search term was an Afro-American sounding name. With regard to the possibly discriminatory effects of profiling it should also be explored what would be the best way to inform the DataBait user about this. Because the information is likely to be speculative (the USEMP system has no means of establishing which measures are actually taken based on profiling) it is important to explore how to present this information without slandering the actors that are in fact profiling the end-users (such as the OSN that is involved).
m. What does it mean to "implement effective protection against possible discrimination resulting from profiling"? (This is an obligation imposed on the data controller in Art. 20(3) of the pGDPR) Can the USEMP consortium come up with examples of what such effective protection could entail? What about Discrimination Aware Data Mining (Kamiran 2011, Custers, Zarsky et al. 2012)?

4. The USEMP tools as a form of Fundamental Rights Protection by Design?

The terms “Data Protection by Design” (DPbDesign) and “Data Protection by Default” (DPbDefault), which have a prominent place in the pGDPR (Art. 23), are not explicitly mentioned in the current DPD 95/46. However, Article 17 (*Security of processing*) of the DPD can be seen as a first step towards DPbDesign:

> “*Security of processing***.** Member States shall provide that the controller must implement appropriate technical and organizational measures to protect personal data against accidental or unlawful destruction or accidental loss, alteration, unauthorized disclosure or access, in particular where the processing involves the transmission of data over a network, and against all other unlawful forms of processing. Having regard to the state of the art and the cost of their implementation, such measures shall ensure a level of security appropriate to the risks represented by the processing and the nature of the data to be protected.” (Art. 17(1) DPD)

Next to a very similar requirement of taking “appropriate organizational and technical measures” in the context of the security of the processing (art. 30 of the proposed GDPR), the proposed GDPR also contains a general article on *Data Protection by Design and by Default* (Art. 23 GDPR) which does not merely relate to the *security* of the processing but aims to meet *all* requirements of the proposed GDPR:

> “Having regard to the state of the art and the cost of implementation, the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures and procedures in such a way that the processing will meet the requirements of this Regulation and ensure the protection of the rights of the data subject.” (Art 23(1) of the proposed GDPR)

Despite the seemingly extensive definition of *Data Protection by Design* in Art. 23(1) an exact understanding of this notion is still heavily debated. Article 23(2) obliges the data controller to implement mechanisms to ensure *Data Protection by Default*, which is a certain form of *Data Protection by Design* based on the idea “that privacy intrusive features of a certain product or service are initially limited to what is necessary for the simple use of it”. (European Data Protection Supervisor 2012, 7 March, p. 29-30) However, as the European Data Protection Supervisor (EPDS) argued in his Opinion on the GDPR, Article 23(2) does not give “any clear substance” to “data protection by default”:

> “The first sentence does not add much to the general principles of data processing in Article 5, and the data minimisation principle in Article 5(c) in particular, except from the confirmation that such principles should also be embedded in the design of relevant systems.” (European Data Protection Supervisor 2012, 7 March, p. 29).

Obviously, data protection requirements are not the only legal requirements which could be used to shape the technological and organizational design of IT architectures. Theoretically one could strive to use all kinds of fundamental rights (“Fundamental Rights Protection by Design” or “Legal Protection by Design”) to shape the technological and organizational design of IT architectures. The latter notion, Legal Protection by Design (LPbD), is a term first coined by Hildebrandt (2011) conveying the idea that legal norms can be articulated in

architecture and which is especially concerned with the articulation of *fundamental rights* in *ICT architecture*. LPbD is based on the idea that “the legal requirements of fundamental rights such as privacy and data protection must be translated into computer system hardware, code, protocols and organizational standards to sustain the effectiveness of such right in a changing technological landscape.” (Hildebrandt 2013, p. 10) When we try to imagine how the right to profile transparency could be transposed into the technological and organizational design of systems and practices which profile end-users, tools like the ones developed in the USEMP project could be the answer. When the proposed GDPR comes into force, and DPbDesign becomes an enforceable legal requirement; the USEMP tools can be a good example of how profile transparency could be built into otherwise opaque automated profiling systems. In this sense the USEMP tools can act as the technical “extension” or mouthpiece of data protection law.

5. Are the USEMP tools compatible with all relevant EU data protection requirements?

Profile transparency is far from being the only relevant legal provision in EU data protection law. While profile transparency is very important for the realization of the objectives of USEMP (i.e., a more empowered user in OSN environments), it is not the only data protection requirement which could contribute to this objective. A user who has full knowledge about his or her profile, but who is, for example, subjected to profiling which takes place in an insecure way, of which the purpose is not specified and/or which is not based on any of the legitimizing grounds mentioned in Art. 7 DPD 95/46, can hardly be defined as "an empowered user".

Moreover, data protection requirements are not only important in realizing the scientific objectives of the USEMP project, but they are also important in assessing whether the USEMP project and its tools are themselves compatible with all data protection requirements. The practical aspects of the day-to-day coordination and assessment of the compatibility of the USEMP project with data protection law are discussed in more detail in D3.4. Nevertheless, it is important to explore some of the basic data protection terminology and requirements – not in the least because in an interdisciplinary project like USEMP it is important that all partners have some basic knowledge of the relevant terminology and requirements (especially given the fact that data protection is not just a legal boundary for the development of the USEMP tools but that it is pivotal in realizing the USEMP objective of user empowerment in OSNs).

Before giving a list of the relevant data protection terminology and requirements, we discuss three preliminary caveats about EU data protection law:

(1) EU Data Protection is a field which is currently under revision. Thus we have to look at both current and future legislation. The successor to the aforementioned DPD 95/46 is the proposed *General Data Protection Regulation* (GDPR).

(2) While the DPD 95/46 and the proposed GDPR are the general data protection instruments with regard to the USEMP project (*lex generales*), there are also more specific data protection laws (*lex speciales)* that apply. For example, one of the tools developed by USEMP is an OSN app (the “DataBait Facebook app”) which processes, amongst others, location and traffic data, and thus the *ePrivacy* directive[9] applies (to . One provision from this directive which is particularly relevant to USEMP is the consent requirement stipulated in Article 5(3)*, w*hich states that the DPD 95/46/EC regime applies to location data, traffic data, etc. even if they are not strictly speaking personal data. However, most data protection issues in USEMP can be

[9] *Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications)*, Official Journal L 201 , 31/07/2002 p. 0037 – 0047 ; and the amendments to this Directive in *Directive 2009/136/EC of the European Parliament and of the Council of 25 November 2009 amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws*, OJ L 337, 18.12.2009, p. 11-36.

addressed by looking at the general data protection instruments (DPD 95/46 and the pGDPR).

(3) Currently data protection is not completely uniform in every member state of the EU. This can be partly explained by the fact that currently the main instrument with regard to data protection in the EU is a *Directive*, that is, a legal act which (unlike a Regulation) is not self-executing but requires ember states to create national laws to implement the objectives of the Directive. A Directive thus offers much more leeway to member states than a Regulation: not only because the task of "translating" a Directive offers some room for interpretation but also because some provisions from the Directive are formulated in an optional way: "Member states *may* stipulate that…/provide for an exemption from…/…etc." In the day-to-day handling of legal matters (see D3.4) in the USEMP project this lack of uniformity became for example very apparent with regard to the obligation to notify the national supervisory data protection authority of the data processing: different member states have different exemptions for data processing for research purposes. However, getting into the details of differences between data protection in different member states goes beyond the scope of this deliverable. Moreover, these differences will largely disappear when the pGDPR (which is a *Regulation*) comes into force.

(4) It is a common misconception that EU data protection law only centres on informational privacy and data minimization. Data protection is much more than that: it embodies various (fundamental rights) concerns, including those focused on anti-discriminatory character (this is even more pronounced in the proposed GDPR) and those aiming for due process (or to be more precise: due *processing*, (Coudert, De Vries et al. 2008)) rights. The right to profile transparency is a good example of the latter. Even though the right to profile transparency does not concern due process in a narrow sense (which entails that one should not be tried or imprisoned based on vague or non-existing laws, or without the possibility to counter or appeal the indictment) it is inspired by a similar *rationale*: namely that when decisions are taken that may negatively affect one, there should be accountability, transparency, scrutiny, the right to access the grounds for a decision, to adjust incorrect information, to object to it, etc.

The most basic terms in EU data protection law can be found in Art. 2 (a), (b) and (d) DPD 95/46 (***bold*** *added by us, the authors of this deliverable*):

> (a) **'personal data**' shall mean any information relating to **an identified or identifiable natural person ('data subject')**; an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;
>
> (b) **'processing of personal data**' ('processing') shall mean **any operation or set of operations which is performed upon personal data, whether or not by automatic means**, such as collection, recording, organization, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, blocking, erasure or destruction;
>
> (d) '**controller**' shall mean the **natural or legal person**, public authority, agency or any other body which alone or jointly with others **determines the purposes and means of the processing of personal data**;

With regard to the USEMP project the notion of the "data controller" is particularly interesting. The USEMP consortium has jointly determined what the purpose and means of the processing of personal data will be, and should thus be qualified as a joint data processor:

> "Joint controllers. Where several controllers jointly determines the purposes and means of the processing of personal data, the joint controllers shall determine their respective responsibilities for compliance with the obligations under this Regulation, in particular as regards the procedures and mechanisms for exercising the rights of the data subject, by means of an arrangement between them. The arrangement shall duly reflect the joint controllers' respective effective roles and relationships vis-à-vis data subjects, and the essence of the arrangement shall be made available for the data subject. In case of unclarity of the responsibility, the controllers shall be jointly and severally liable" (Art 24 of the pGDPR)

Because the USEMP consortium does not possess legal personality, it was important to create an internal agreement between the partners (see Annex 3) in which partners commit to implementing relevant data protection law when processing the personal data of USEMP end-users, while each partner exonerates the others from liability for data processing which is not under the actual control of these other partners.

Another data protection notion which is important to the USEMP project is that of *the legal ground*: the process of personal data is only allowed when it is based on one of the grounds mentioned in Art. 7 DPD 95/46:

CRITERIA FOR MAKING DATA PROCESSING LEGITIMATE

Article 7

Member States shall provide that personal data may be processed only if:

(a) the data subject has unambiguously given his consent; or

(b) processing is necessary for the performance of a contract to which the data subject is party or in order to take steps at the request of the data subject prior to entering into a contract; or

(c) processing is necessary for compliance with a legal obligation to which the controller is subject; or

(d) processing is necessary in order to protect the vital interests of the data subject; or

(e) processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller or in a third party to whom the data are disclosed; or

(f) processing is necessary for the purposes of the legitimate interests pursued by the controller or by the third party or parties to whom the data are disclosed, except where such interests are overridden by the interests for fundamental rights and freedoms of the data subject which require protection under Article 1 (1).

As discussed in chapter 2, we have chosen to take the ground from Art. 7(b) as the legitimizing ground for all the processing: a data licensing agreement (DLA) between the USEMP consortium (the joint data controller) and the end-user of the USEMP tools is the legitimizing ground for the data processing in USEMP (see Annex 2). Whereas consent (Art. 7a of *Data Protection Directive 95/46*) is still the most frequently used ground for data processing, USEMP has chosen to take a *contract* as a ground for the data processing instead: such a two-sided legal act with mutual obligations for both parties seems to be a better expression of the objective of USEMP, namely to empower users, than using consent as a legal ground. Moreover, the contract puts data licensing at its centre and as such gives teeth to the data protection requirement of purpose specification ("*personal data must be […] collected for specified, explicit and legitimate purposes and not further processed in a way incompatible with those purposes",* Art. 6b of *Data Protection Directive 95/46*).

It should be noted that this DLA is *not* merely a service license agreement (SLA) in which the part about data processing is only an appendix – i.e., a consent form attached to the main service agreement – but that this contract actually focuses on the *purpose* of data processing within the USEMP project and the mutual obligations between the USEMP consortium partners (the joint data controllers) and the end-user of the USEMP tools. These obligations are created in order to fulfil that purpose. Opting for a DLA, rather than the usual combination of a SLA combined with a privacy policy, user consent and lengthy terms and conditions, also aligns with the USEMP proposal to enable the licensing of the use of personal data by data subjects, as described in the USEMP Description of Work (DOW). Another way in which the DLA embodies the objective of user empowerment, is that it keeps matters as straightforward as possible and puts them in plain language: the DLA avoids any unnecessary "legalese". The DLA is implemented in the USEMP graphic user interface (GUI) and will be part of the sign-up procedure. It will be impossible to sign-up or use the USEMP tools without first signing the DLA. Each article of the DLA will be presented as a separate screen. All the text will fit easily on one screen, making it unnecessary for the user to scroll down.

Building the DLA and the internal agreement into the DataBait GUI can also be considered as a way in which USEMP realizes Data Protection by Design (Art. 23 pGDPR) with regard to Art. 7 DPD 95/46 (legal ground) and the requirement of *data minimization* (Art. 6 DPD 95/46), which is an umbrella term for the requirement of *purpose specification* (that data must be collected for specified, explicit and legitimate purposes and that they must be adequate, relevant and not excessive in relation to the purposes for which they are collected), *use limitation* (that data should not be further processed in a way incompatible with those purposes), that data have to be *accurate and complete*, and that they are deleted or anonymised as soon as they are no longer needed for the purpose that led to their collection. The much debated judgment of 13th May 2014 on the "right to be forgotten", [10] which caused a storm of cross-Atlantic and intra-European confusion, can in fact be simply derived from Art. 7 DPD 95/46. The judgement shows once more how important it is that the legitimizing ground for the processing continues to be valid; this is something which has to be checked

[10] Judgment of the Court (Grand Chamber) in C-131/12, *Google Spain v AEPD and Mario Costeja Gonzalez*, 13 May 2014. Online available at : <http://curia.europa.eu/juris/documents.jsf?num=C-131/12>

on a regular basis. The DLA provides a lawful ground to the processing for the duration of the USEMP project.

Another important data protection requirement is that of *confidentiality* and *security* of the data processing (Art. 16 and 17 DPD 95/46). This is dealt with in Art. C of the internal Personal Data Processing Agreement (PDPA) which is will be signed by all USEMP partners before personal data are being processed (see Annex 3). Article C of the PDPA requires that each partner which processes personal data undertakes a security risk assessment, sharing the results and updating the assessment in case of changes to hardware, protocols etc. All partners which process personal data are explicitly bound to have appropriate security measures in place.

Finally, it should be noted that the requirements of Art. 8 DPD 95/46 with regard to *sensitive data* (regarding racial or ethnic origin, political opinions, religious or philosophical beliefs, trade-union membership, health or sex life) are very important in assessing the legal compatibility of the USEMP tools: Art. 8 prohibits to process such sensitive data, unless there is explicit consent, if necessary to protect the vital interests of the data subject, or if the processing is part of the legitimate non-profit-seeking body with a political, philosophical, religious or trade-union aim, or if the processing relates to data which are manifestly made public by the data subject. Art. G of the DLA therefore singles out sensitive data, requiring explicit consent for their processing; as explained in H the goal of USEMP requires the processing of sensitive data in order to provide users with profile transparency precisely on that point. Moreover, in the pGDPR a contract *does* count as a legitimizing ground (Art. 9-2aa proposed GDPR) for the processing of sensitive data.

6. Respect for Private Life and Prohibitions of Certain Kinds of Discrimination and Negative Stereotyping: a translation of user empowerment in OSNs into the discourse of European fundamental rights (part II[12])

Data protection is far from the only relevant legal field when considering the compatibility of profiling systems and practices with EU fundamental rights. Many other rights and freedoms can be equally important when assessing their compatibility with European fundamental rights. In this section we will focus on the right to respect for private life (Art. 8 of the *European Convention on Human Rights*[11] [ECHR]), the prohibition of discrimination with regard to the exercise other fundamental rights (Art. 14 ECHR) and a broad variety of anti-discriminatory instruments from EU law. Clearly, there can be other relevant European fundamental rights next to these: for example, a company whose core business is to track and profile internet users could invoke the right to conduct a business (Art. 16 of the *Charter of Fundamental Rights of the EU*[12]) or even the freedom of expression (Art. 10 ECHR).

However, in this deliverable we only look at European fundamental rights that, at least in some way, seem to express a *rationale* of user empowerment with regard to profiling practices on the internet and particularly in social networks. The right to conduct a business is a fundamental right which could *curtail* user empowerment, and is thus better discussed in D3.2 (on trade secrets and intellectual property rights which could trump profile transparency). And, although freedom of expression is important with regard to what users can post on social networking sites (McGoldrick 2013), it does not have a big role to play with regard to profiling practices on such social networking sites. Thus, we will limit ourselves to privacy and anti-discrimination rights in this section.

6.1. User empowerment and profile transparency revisited

The most obvious way to translate the user empowerment objectives of USEMP into fundamental rights concerns might be by using data protection law (and particularly the right to profile transparency). Thus, unsurprisingly, when looking at the first version of the description of the USEMP architecture (D7.1) the main focus seems to be on profile transparency (as a way of realizing user empowerment). The role of profile transparency can be clarified by briefly sketching the main functionality of the USEMP tools. There are three USEMP tools: a Facebook app, a browser plug-in and a graphic user interface (GUI). The first two tools collect real time (the Facebook app and the browser plugin) and historical (the Facebook app) data from one's social network (Facebook) profile and browser (Chrome and

[11] *Convention for the Protection of Human Rights and Fundamental Freedoms as amended by Protocols No. 11 and No. 14,* Rome 4 November 1950, European Treaty Series - No. 5.
[12] *Charter of Fundamental Rights of the EU*, 2000/C 364/01. Entered into force on 1 December 2009, as part of the *Treaty of Lisbon*.

Firefox). These data include volunteered data (e.g., a status update or a posted picture), behavioral data (e.g., time and location of a log-in or search) and data about trackers. In return for sharing all these data, USEMP (DataBait) users are informed through the GUI about their data trail and the trackers that follow them. Part of the user empowerment/profile transparency is that the GUI offers a visually clear and easy way to see this information. Moreover, presenting this information in a visually appealing way helps the user to see the effect of certain actions (blocking a tracker, removing a picture, etc.). On top of information about the actual data of the user, USEMP also makes inferences based on these data. What kind of information about one's political preferences can be derived from a seemingly innocent post? Which inferred data about one's health or sexual orientation can be derived from a holiday picture?

Here the information provided to the user is partly *speculative* – there is no way of determining if this information is actually inferred or used by commercial actors tracking the Databait user. Nevertheless, it is well known that image and text classification are widely used methods in profiling practices. Showing the Databait user which information *could* be inferred and how it *could* be used can provide empowering insights: even the fact that the classification mechanisms will not always make the right inferences (e.g., based on a fuzzy picture you're misclassified as "obese" or "smoker") might be an enlightening insight for the Databait user. The creation of machine learning algorithms that infer information is as much an art as it is a science, and giving a user a sense of how well -or how bad- an algorithm is in predicting certain characteristics can be a very educational experience.

6.2. User empowerment through information about your legal rights (protection against unlawful discrimination and negative stereotyping)

However, next to showing users their data trail, the possibilities for inferring (sensitive) information from it and scenarios about what these data could be used for, it is also important to inform the users through the GUI[13] about their *legal rights* with regard to these data. Just informing a user that one is (or can be) profiled as "low income" or "caucasian woman" might be of limited use; this information becomes much more empowering if it is complemented with information about *who*, *under which conditions* and *in which context* this information can be used. Where does legitimate segmentation (e.g., targeting ads for a theatre play at culture loving people who live in the neighborhood of the theatre is fully permittable practice) end and illegitimate discrimination (e.g. denying someone a job based on data regarding her race) begin? The answer to this question can only partly be derived from data protection law and lies to a large extend in the aforementioned fields: the right to respect for private life (Art. 8 ECHR), the prohibition of discrimination with regard to the exercise other fundamental rights (Art. 14 ECHR) and a broad variety of anti-discriminatory instruments from EU law.

Before getting into the details of these legal provisions and instruments it is useful to give some brief information about the difference between the rights derived from the ECHR, which

[13] The content of the DataBait GUI can be easily adjusted. The (legal) information which will be shown to the user will be further developed during the USEMP project.

is a treaty that belongs to the legal framework of the Council of Europe, and the rights derived from instruments from the legal framework of the EU.

The Council of Europe

The Council of Europe (CoE) is an international organization which has 47 members, stretching out deeply into the Euroasian territory with members such as Turkey, Armenia and the Russian Federation. The CoE produces treaties, officially known as Conventions. The most important and influential Convention of the CoE is the European Convention on Human Rights (ECHR). When the ECHR was signed on 4 November 1950 and entered into force on 3 September 1953, the possibility was created for all individuals in CoE member states to bring a legal action against a member state before the European Court of Human Rights in Strasbourg (ECtHR). The condition for bringing a legal action before the Court is that an individual believes that a member state has violated his or her fundamental human rights as protected by the ECHR and that all national remedies have been exhausted. Thus, the main rationale of the ECHR is to offer individual citizens protection against State power (so-called *vertical* effect) by providing a concrete legal route of redress when fundamental rights have been infringed by the State. However, the ECHR can also have a so-called *horizontal* effect when non-state actors infringe upon these right and the State should have prevented or redressed such an infringement. Thus, in specific circumstances a citizen whose fundamental ECHR rights are infringed upon by a private actor (for example, a company like Facebook, Google or a databroker) can turn to the ECHR to oblige the State to make national laws that oblige this private actor to abstain from infringing the rights protected in the Convention.

The European Union

The European Union is an economic and political union of 28 member states. It is clear that during the last decades the protection of fundamental rights, adjusting the power imbalances within the relationship between individual and State, has gained an increasing importance within the EU, especially since the *Charter of Fundamental Rights of the EU* (CFREU) entered into force in 2009[14] and the EU made a commitment[15] to accede to the ECHR. However, the primary concern of the EU is that of an economic, political and legal *structure* continues to give a distinct flavor to the legal framework of the EU when compared to that of the ECHR (CoE). The legal framework of the EU is mainly focused on *regulating* the internal market of the EU, aiming to remove restrictions of the free movement of goods, services, people and capital. This is clearly visible in the detailed secondary EU laws on data protection and anti-discrimination which aim to regulate, and thus *prevent*, infringements. (cf. Gellert, de Vries et al. 2012). The ultimate aim of data protection has therefor been to

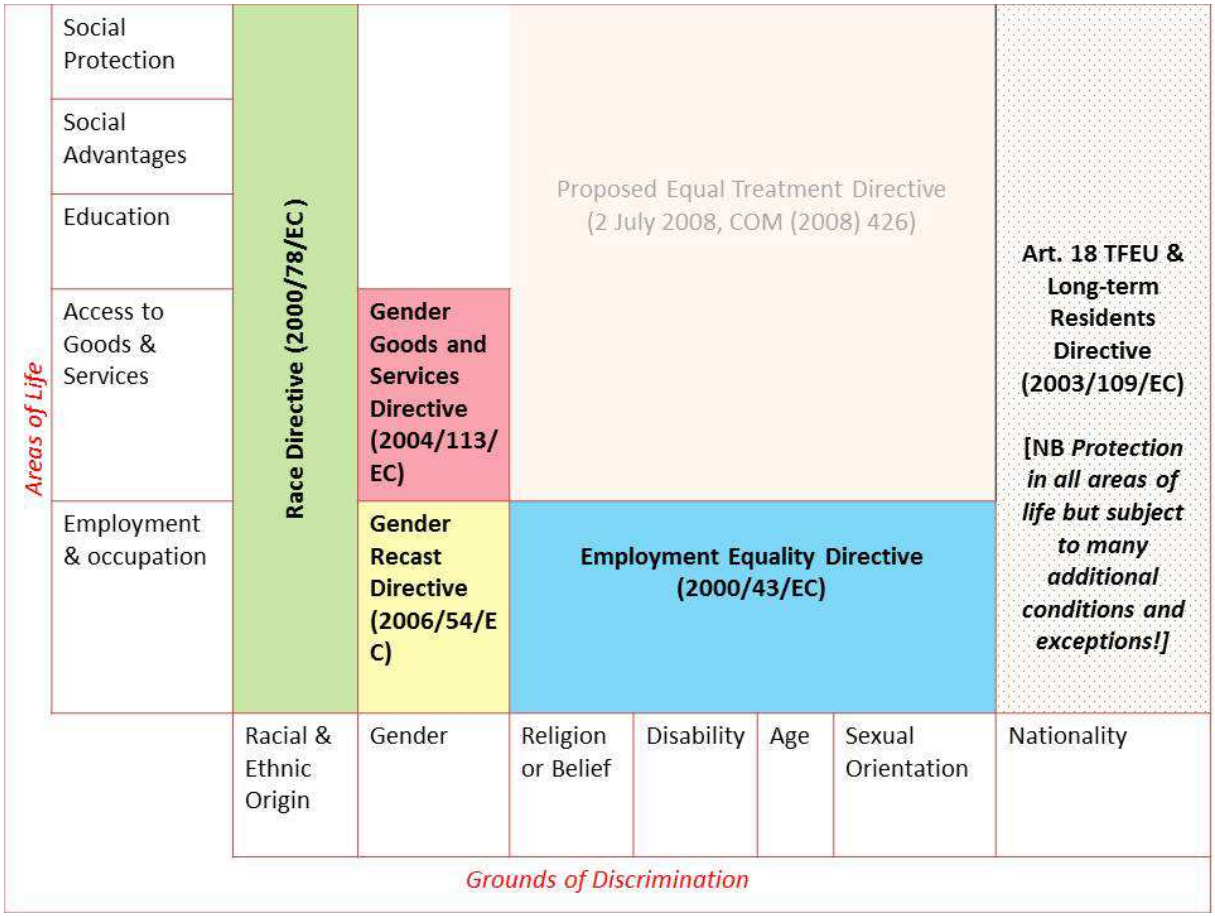
[14] *Charter of Fundamental Rights of the EU*, 2000/C 364/01. Entred into force on 1 December 2009, as part of the *Treaty of Lisbon*.

[15] See for the latest negotiations with regard to the entry of the EU to the ECHR: http://www.coe.int/t/dghl/standardsetting/hrpolicy/Accession/Meeting_reports/47_1(2012)R03_EN_final.pdf (accessed 2 August 2013) See also: Polakiewicz, J. (2013). EU law and the ECHR: Will EU accession to the European Convention on Human Rights square the circle? Fundamental Rights In Europe: A Matter For Two Courts. Oxford Brookes University.

.

harmonize restrictions on – notably – the free flow of information within the internal economic market.

Compared to the legal instruments of the EU, the rights derived from the ECHR often provide a broader but also a fuzzier protection. Not only because the primary goal of the ECHR is to protect the individual citizen against the State (and not against Facebook, Google or a databroker), but also because the route to the Court in Strasbourg (a measure of last resort) is longer than that against EU legislation (or the national implementation thereof). National courts can raise preliminary questions with the Court of Justice of the European Union (CJEU) in Luxemburg about the interpretation of EU law. Nevertheless it is also precisely the broad formulation of ECHR rights which can sometimes provide protection where the more specific provisions of the EU fail to do so. This is particularly clear in the field of anti-discrimination law. As shown in figure 2, the anti-discriminatory law of the EU offers protection with regard to a very specific set of protected grounds (listed in Art. 13 of the *Treaty Establishing the European Community*[16] [TEC, 1997; entry into force in 1999]: sex, racial or ethnic origin, religion or belief, disability, age or sexual orientation) and areas of life. The largest protection is offered with regard to race, and in the field of employment.



[16] Now replaced by Article 19 of the *Treaty on the Functioning of the Union* (TFEU, 2008). The content of Art. 19 TFEU and Art. 13 TEC is identical.

Figure 2: Protected grounds and areas of life in secondary EU anti-discrimination law

When comparing the anti-discriminatory provisions from EU data protection law with those from EU anti-discrimination law, there are some interesting overlaps as well as differences to be pointed out (see table 2).

Data Protection	Art. 9 (1) of the proposed General Data Protection Regulation (GDPR)	The processing of personal data, revealing: **race or ethnic origin, political opinions, religion or beliefs**, trade-union membership, and the processing of ***genetic data*** or data concerning *health or sex life* or criminal convictions or related security measures shall be prohibited.
	Art. 20 (3) of the proposed General Data Protection Regulation (GDPR)	Profiling that has the effect of discriminating against individuals on the basis of **race or ethnic origin, political opinions, religion or beliefs**, trade union membership, *sexual orientation or gender identity*, or that results in measures which have such effect, shall be prohibited. The controller shall implement effective protection against possible discrimination resulting from profiling. Profiling shall not be based solely on the special categories of personal data referred to in Article 9
Anti-Discrimination	Art. 21 Charter of fundamental rights of the European Union (CFREU)	(1) Any discrimination based on any ground such as: sex, **race**, colour, **ethnic** or social **origin, *genetic features***, language, **religion or belief, political or any other opinion**, membership of a national minority, property, birth, *disability*, age or *sexual orientation* shall be prohibited. (2) Within the scope of application of the Treaty […] any discrimination on grounds of nationality shall be prohibited.
	Art. 13 Treaty Establishing the European Community (TEC)	…take appropriate action to combat discrimination based on sex, **racial or ethnic origin, religion or belief**, *disability*, age or *sexual orientation*

Table 2: Discrepancies and overlaps between (1) the data categories classified as sensitive or discriminatory in EU data protection law and (2) the prohibited grounds in EU discrimination law. In this table the bold categories are the ones that overlap, the italic ones partly overlap, and the underlined ones are new additions to the list of sensitive data in Art. 9 (1) of the pGDPR (in comparison to the ones mentioned in Art 8(1) of the current DPD 95/46. This table is an updated and adjusted version of the table in: (Gellert, de Vries et al. 2012)

One way to explain these overlaps and differences is that data protection is more oriented on the *process* of data processing, while the anti-discrimination provisions look at discriminatory

effects. Thus, data such as sex, age, and nationality (which is the kind of basic information which one is required to provide frequently) are not considered to be sensitive data from a data protection perspective, but as soon as one begins to take discriminatory measures based on them, for example in the area of employment, they become "toxic". While the processing of *sensitive* data -for example, data which reveal racial origin or political opinions- requires additional safeguards in comparison to the processing of "ordinary" personal data even when no actual discrimination results from it, data such as sex, age, and nationality are not considered to be sensitive *as such*.

In designing the USEMP tools it is important to keep the specific categories of data in EU anti-discrimination law (protected grounds) and EU data protection law (sensitive data and the protected grounds mentioned in Art. 20(3) GDPR) in mind. Are these categories of data (likely to be) processed by commercial profilers? And if these data are not available as volunteered data – how easy or difficult is it to infer them? How can the user be informed of the relevant legal provisions with regard to these particular kinds of data? Do users indeed feel that the sensitive data and protected grounds deserve a higher level of protection than other data (e.g. income, log-in patterns, educational level, etc.)?

Anti-discrimination and arts. 14 and 8 ECHR

With regard to anti-discrimination it is also relevant to look at Arts. 14 and 8 ECHR. In those instances where EU law does not offer any remedy, they could both turn out to be useful.

Article 14 ECHR (the prohibition of discrimination with regard to the exercise other fundamental rights) in conjunction with Protocol 12[17] (the prohibition of discrimination with regard to the exercise *any* other right) is -in contrast to anti-discriminatory EU law provisions!- most likely *not* limitative as to the grounds it protects. Art. 14 ECHR enumerates several protected grounds of discrimination: sex, race, colour, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status. However, the formulation "…*on any ground such as*…" does seem to indicate that this list is not limitative. The case law of the European Court of Human Rights is ambivalent and inconsistent as to whether other grounds than the ones named in Art. 14 are also protected (Gerards 2013): sometimes the Court admits cases that concern a non-listed ground, while in other instances cases are declared inadmissible for concerning a non-listed ground. In *Kjeldsen, Busk Madsen and Pedersen*[18] (1976) the ECtHR held that any difference in treatment that was not based on "a personal characteristic", was inadmissible:

---

[17] *Protocol No. 12 to the 1950 European Convention for the Protection of Human Rights and Fundamental Freedoms*, ETS no. 177, adopted on 4 November 2000 (Rome); entry into force on April 1, 2005. Currently (August 2013) 18 member states have ratified the Treaty (from a total of 47 Council of Europe member states).

See:

http://conventions.coe.int/Treaty/Commun/QueVoulezVous.asp?NT=177&CM=8&DF=15/09/2013&CL=ENG

[18] ECtHR, *Kjeldsen, Busk Madsen and Pedersen v Denmark*, no. 5095/71, 5920/72 and 5926/72, judgment of 7 December 1976.

this both extends and limits the list of grounds of Art. 14. Yet, in later cases the Court has dealt with differences in treatment on their merits, without investigating on which grounds they were made and whether this ground would qualify as a personal characteristic or not (Gerards 2013). When interpreted in a non-exhaustive way, Art. 14 can be used to contest *any* discrimination that allegedly lacks reasonableness. Art. 14 would thus be operating according to an "equal treatment rationale", instead of one based on the prohibition of discrimination based on a limited set of protected grounds:

> "Article 14 can be regarded as an expression of the general principle of equality. […] If this perspective is taken, each difference in treatment that affects an applicant's Convention rights should be assessed by the Court for reasonableness and fairness. The ground on which the difference in treatment is based is not relevant to the applicability for a test of justification. The only relevant question is if one group or person is allowed to exercise a certain right or receive a certain benefit, whilst this is not permitted for another person or group. The equal treatment approach is radically different from the nondiscrimination approach, which clearly does have a normative content of its own." (Gerards 2013, p. 118-9)

Interpreting Art. 14 "as an expression of the general principle of equality" gives it the potential to become extremely important in assessing automated profiling practices that are often based on grounds that do neither belong to the limited set of grounds protected by secondary EU legislation nor to the ones explicitly enumerated in Art. 14, but are nonetheless potentially undesirable from a fundamental rights perspective.

Next to Art. 14 (in conjunction with Protocol 12), Art. 8 ECHR (right to respect for private life) could also be useful in combating forms of discrimination, stereotyping and stigmatization that fall outside the protective scope of EU law. That Art. 8 ECHR can be used to combat negative stereotyping can be based on art. 8 ECHR (respect for private life) as was poignantly shown in *Aksu v Turkey* (Applications nos. 4149/04 and 41029/04, EctHR, Judgement of 15 March 2012) and repeated by the Dutch administrative Court in the "Zwarte Piet case"[19] (2014).

> "The Court reiterates that the notion of "private life" within the meaning of Article 8 of the Convention is a broad term not susceptible to exhaustive definition. The notion of personal autonomy is an important principle underlying the interpretation of the guarantees provided for by Article 8. It can therefore embrace multiple aspects of the person's physical and social identity. The Court further reiterates that it has accepted in the past that an individual's ethnic identity must be regarded as another such element (see *S. and Marper v. the United Kingdom* [GC], nos. 30562/04 and 30566/04, § 66, 4 December 2008, and *Ciubotaru v. Moldova*, no. 27138/04, § 49, 27 April 2010). In particular, any negative stereotyping of a group, when it reaches a certain level, is capable of impacting on the group's sense of identity and the feelings of self-worth and self-confidence of members of the group. It is in this sense that it can be seen as affecting the private life of members of the group." (consideration 58, *Aksu v Turkey*)

[19] http://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBAMS:2014:3888

Despite the fact that there is limited case law[20], Art. 8 could also be used in cases of "stigmatisation" (Quinn and Hert 2012); the notion of "private life" is so flexible that it could proof useful in the future in cases relating to profiling.

[20] *Kiyutin v Russia* (Refusal of residence to an HIV+ individual ; applicability of Article 14, taken in conjunction with Article 8 ECHR); *Marper v The United Kingdom* (Retention of DNA samples of arrested individuals); and *A,B and C v Ireland,* application no. 25579/05, Strasbourg, 16 December 2010 (The Irish Republic's ban on abortions on its territory)

7. Conclusion and next steps

Building on the legal analysis in the previous sections the research questions studied in this Deliverable can be answered with the following conclusions and recommendations for further research in D3.6:

(a) What can be said about the legal compatibility with European fundamental rights of the tracking and profiling practices performed by OSNs, browsers and third-parties, and the possibility of legal protection by design by tools such as the ones developed by USEMP?

Especially when considering the stronger data protection regime of the future GDPR and the rather underexplored possibilities of anti-discrimination law with regard to profiling practices, there is still much work to do in making profiling practices compatible with European fundamental rights. The USEMP tools could be a good example of how legal protection by design could be created with regard to profiling practices in social networks and browsers. The USEMP tools could be especially useful with regard to profile transparency, but also with regard to providing users with information about other legal rights (anti-discrimination and respect for private life).

(b) What can be said about the legal compatibility with European fundamental rights of the tracking and profiling practices of the USEMP tools themselves?

(i) The legal compatibility of the tracking and profiling practices of the USEMP tools as they are now, that is: processing data with the sole purpose of scientific research?

The USEMP tools often operate in a tricky area, for example processing significant amounts of sensitive data. In order to show the user what could be inferred from certain data by others, the USEMP project has to infer sensitive data. The consortium is very well aware of the responsibility this brings along and pays great attention to the need to comply with all relevant data protection provisions. The scrutiny that we apply to other actors, applies to ourselves as well.

(ii) The legal compatibility of the tracking and profiling practices of the USEMP tools as they could hypothetically be employed in the future, that is: commercialized and no longer part of a research project?

This is something which needs to be explored in further detail. With regard to European fundamental rights some steps in the data processing within the USEMP project are facilitated by the fact that all the processing has a purely scientific purpose (for example, the notification requirement from Art. 18 DPD 95/46). However, it seems that - given the necessary safeguards and precautions - tools like the ones developed by USEMP could also be developed outside a research context.

Moreover in D3.4, the legal coordination and integration deliverable, we will explain how the data processed in the USEMP project are processed, at who's premises and for what reasons. This deliverable will present a set of tables to clarify in fine grained detail:

- what data types and which data sets are being processed by which partner
- at what premises
- for what reason (as regards the architecture of the Databait tools
- categorizing the legal effect for each data and data set
    - if personal data
    - if also sensitive data
- connecting this with the privacy dimensions, their attributes and so on

A first version of these tables is in the annex of this deliverable D3.1 as far as data protection is concerned. The tables in D3.4 also explore how the data (or: "content") processed in the USEMP project should be qualified in terms of IP rights on user generated content (D3.3) and on OSN databases or software (D3.2).

Bibliography

Article 29 Data Protection Working Party 29 (2013, 13 May). Advice paper on essential elements of a definition and a provision on profiling within the EU General Data Protection Regulation. 00530/12/EN WP 191.

Coudert, F., et al. (2008). Legal Implications of Forensic Profiling: of Good Old Dataprotection Legislation and Novel Legal Safeguards for Due Processing. Forensic Profiling. Deliverable 6.7c of the FIDIS (The Future of Identity in the Information Society) Consortium. Z. Geradts and P. Sommer. http://www.fidis.net, EU Sixth Framework Programme**:** 38-67.

Council of Europe (2010). Recommendation CM/Rec(2010)13 and the Explanatory Memorandum. The Protection of Individuals with regards to Automatic Processing of Personal Data in the Context of Profiling. Strasbourg.

Custers, B., et al., Eds. (2012). Discrimination and Privacy in the Information Society. Effects of Data Mining and Profiling Large Databases. Studies in Applied Philosophy, Epistemology and Rational Ethics. Dordrecht, Springer.

European Data Protection Supervisor (2012, 7 March). Opinion of the European Data Protection Supervisor on the data protection reform package.

Gellert, R., et al. (2012). A Comparative Analysis of Anti-Discrimination and Data Protection Legislations. Discrimination and Privacy in the Information Society. B. Custers, T. Calders, B. Schermer and T. Zarsky. Berlin, Springer**:** 61-89.

Gerards, J. (2013). "The Discrimination Grounds of Article 14 of the European Convention on Human Rights." Human Rights Law Review **13**(1): 99-124.

Hildebrandt, M. (2011). "Legal protection by design: Objections and refutations." Legisprudence **5**(2): 223-248.

Hildebrandt, M. (2013). Legal Protection by Design in the Smart Grid. Privacy, Data Protection & Profile Transparency, Smart Energy Collective.

Kamiran, F. (2011). Discrimination-aware Classification, Technical University of Eindhoven. **Ph.D.**

McGoldrick, D. (2013). "The Limits of Freedom of Expression on Facebook and Social Networking Sites: A UK Perspective." Human Rights Law Review **13**(1): 125-151.

Olejnik, L., et al. (2014). Selling off privacy at auction. Network and Distributed System Security (NDSS) Symposium 2014. San Diego, CA, USA.

Polakiewicz, J. (2013). EU law and the ECHR: Will EU accession to the European Convention on Human Rights square the circle? Fundamental Rights In Europe: A Matter For Two Courts. Oxford Brookes University.

Quinn, P. and P. D. Hert (2012). Stigmatisation and the ECHR. LSTS Monday Research Gatherings. Brussels, Vrije Universiteit Brussel, Center for Law, Science, Technology and Society (LSTS).

Sweeney, L. (2013). "Discrimination in online ad delivery." Queue **11**(3): 10.

Wauters, E., et al. (2014). "Towards a better protection of social media users: a legal perspective on the terms of use of social networking sites." International Journal of Law and Information Technology **22**(3): 254-294.

Annex A – Original and Amended pGDPR

Comparison between the original GDPR (proposed by EU Commission) and the amended GDPR (by the EU Parliament) with regard to profiling

This table compares the original text of the GDPR (proposed by the EU Commission) and the amended GDPR (by the EU Parliament) with regard to profiling. Everything that is **bold** indicates differences between the two versions. The words 'profile' and 'profiling' have been underlined to make it easier to see where they are discussed.

	Text proposed by the Commission, submitted to the European Parliament on 25 January 2012[21]	***Amendments adopted on 12 March 2014 by the European Parliament***[22]
***Monitoring* data subjects: tracking + application of a profile**	(Recital 21) In order to determine whether a processing activity can be considered to 'monitor ***the behaviour' of*** data subjects, it should be ascertained whether individuals are tracked ***on the internet with*** data processing techniques which consist of applying a 'profile' ***to an individual*** , particularly in order to take decisions concerning her or him or for analysing or predicting her or his personal preferences, behaviours and attitudes.	(Recital 21) In order to determine whether a processing activity can be considered to 'monitor' data subjects, it should be ascertained whether individuals are tracked***, regardless of the origins of the data, or if other data about them is collected, including from public registers and announcements in the Union that are accessible from outside of the Union, including with the intention to use, or potential of subsequent use of*** data processing techniques which consist of applying a 'profile', particularly in order to take decisions concerning her or him or for analysing or predicting her or his personal preferences, behaviours and attitudes.

[21] *Proposal for a Regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)*, Brussels, 25 January 2012, COM(2012) 11 final.

[22] *European Parliament legislative resolution on the proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation)*, Strasbourg, 12 March 2014, COM(2012)0011 – C7-0025/2012 – 2012/0011(COD). Ordinary legislative procedure: first reading. Online available at : <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2014-0212+0+DOC+XML+V0//EN>

***Right of access to the logic of the data* in relation to the rights and freedoms of others, such as IP rights.**	(Recital 51) Any person should have the right of access to data which has been collected concerning them, and to exercise this right easily, in order to be aware and verify the lawfulness of the processing. Every data subject should therefore have the right to know and obtain communication in particular for what purposes the data are processed, for what period, which recipients receive the data, what is the logic of the data that are undergoing the processing and what might be***, at least when based on profiling,*** the consequences of such processing. This right should not adversely affect the rights and freedoms of others, including trade secrets or intellectual property ***and in particular*** the copyright protecting the software. However, the result of these considerations should not be that all information is refused to the data subject.	(Recital 51) Any person should have the right of access to data which has been collected concerning them, and to exercise this right easily, in order to be aware and verify the lawfulness of the processing. Every data subject should therefore have the right to know and obtain communication in particular for what purposes the data are processed, for what ***estimated*** period, which recipients receive the data, what is the ***general*** logic of the data that are undergoing the processing and what might be the consequences of such processing. This right should not adversely affect the rights and freedoms of others, including trade secrets or intellectual property, ***such as in relation to*** the copyright protecting the software. However, the result of these considerations should not be that all information is refused to the data subject.
Right to object to profiling ; Prohibition of –especially discriminatory - profiling with legal or similar significant effects.	(Recital 58) Every natural person should have the right ***not to be subject*** to ***a measure which is based on*** profiling ***by means of automated processing***. ***However, such measure*** should be allowed when expressly authorised by law, carried out in the course of entering or performance of a contract, or when the data subject has given his consent. In any case, such processing should be subject to suitable safeguards, including specific information of the data subject and the right to obtain human ***intervention*** and that such measure should not concern a child.	(Recital 58) ***Without prejudice to the lawfulness of the data processing*** , every natural person should have the right ***to object*** to profiling. ***Profiling which leads to measures producing legal effects concerning the data subject or does similarly significantly affect the interests, rights or freedoms of the concerned data subject*** should ***only*** be allowed when expressly authorised by law, carried out in the course of entering or performance of a contract, or when the data subject has given his consent. The In any case, such processing should be subject to suitable safeguards, including specific information of the data subject and the right to obtain human ***assessment*** and that such measure should not concern a child. ***Such measures should not lead to discrimination against individuals on the basis of race or ethnic origin, political opinions, religion or beliefs, trade union membership, sexual orientation or gender identity.***

Presumption that profiling based on *pseudonomous data* does *not* significantly affect the data subject		***(Recital 58a)*** ***Profiling based solely on the processing of pseudonymous data should be presumed not to significantly affect the interests, rights or freedoms of the data subject. Where profiling, whether based on a single source of pseudonymous data or on the aggregation of pseudonymous data from different sources, permits the controller to attribute pseudonymous data to a specific data subject, the processed data should no longer be considered to be pseudonymous.***
Restrictions on data protection rights may be imposed by Union or Member State law, as far as necessary and proportionate in a democratic society to safeguard public security.	(Recital 59) Restrictions on specific principles and on the rights of information, ***access,*** rectification and erasure or on the right to data ***portability*** , the right to object, ***measures based on*** profiling, as well as on the communication of a personal data breach to a data subject and on certain related obligations of the controllers may be imposed by Union or Member State law, as far as necessary and proportionate in a democratic society to safeguard public security, including the protection of human life especially in response to natural or man made disasters, the prevention, investigation and prosecution of criminal offences or of breaches of ethics for regulated professions, other public interests of the Union or of a Member State, in particular an important economic or financial interest of the Union or of a Member State, or the protection of the data subject or the rights and freedoms of others. Those restrictions should be in compliance with requirements set out by the Charter of Fundamental Rights of the European Union and by the European Convention for the Protection of Human Rights and Fundamental Freedoms.	(Recital 59) Restrictions on specific principles and on the rights of information, rectification and erasure or on the right of ***access and to obtain*** data, the right to object, profiling, as well as on the communication of a personal data breach to a data subject and on certain related obligations of the controllers may be imposed by Union or Member State law, as far as necessary and proportionate in a democratic society to safeguard public security, including the protection of human life especially in response to natural or man made disasters, the prevention, investigation and prosecution of criminal offences or of breaches of ethics for regulated professions, other ***specific and well-defined*** public interests of the Union or of a Member State, in particular an important economic or financial interest of the Union or of a Member State, or the protection of the data subject or the rights and freedoms of others. Those restrictions should be in compliance with requirements set out by the Charter of Fundamental Rights of the European Union and by the European Convention for the Protection of Human Rights and Fundamental Freedoms.

The power to adopt more specific acts to fulfill the objectives of the GDPR is delegated to the Commission.	(Recital 129) In order to fulfil the objectives of this Regulation, namely to protect the fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data and to ensure the free movement of personal data within the Union, the power to adopt acts in accordance with Article 290 of the Treaty on the Functioning of the European Union should be delegated to the Commission. In particular, delegated acts should be adopted in respect of ***lawfulness of processing*** ; specifying ***the criteria and conditions in relation to the consent of a child; processing of special categories of data; specifying the criteria and conditions for manifestly excessive requests and fees for exercising the rights of the data subject; criteria and requirements for the information to the data subject and in relation to the right of access;*** the right to ***be forgotten and to*** erasure; ***measures based on profiling*** ; ***criteria and requirements in relation to the responsibility of the controller and to data protection by design and by default; a processor; criteria and requirements for the documentation and the security of processing; criteria and requirements for establishing a personal data breach and for its notification to the supervisory authority*** , ***and on the circumstances where a personal data breach is likely to adversely affect the data subject; the criteria and conditions for processing operations requiring a data protection impact assessment; the criteria and requirements for determining a high degree of specific risks which require prior consultation; designation and tasks of the data protection officer;*** codes of conduct; criteria and requirements for certification mechanisms; criteria and requirements for transfers by way of binding corporate rules; ***transfer derogations;*** administrative sanctions; processing for health purposes; processing in the employment context	(Recital 129) In order to fulfil the objectives of this Regulation, namely to protect the fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data and to ensure the free movement of personal data within the Union, the power to adopt acts in accordance with Article 290 of the Treaty on the Functioning of the European Union should be delegated to the Commission. In particular, delegated acts should be adopted in respect of specifying ***conditions of icon-based mode for provision of information;*** the right to erasure; ***declaring that*** codes of conduct ***are in line with the Regulation*** ; criteria and requirements for certification mechanisms; ***the adequate level of protection afforded by a third country or an international organisation;*** criteria and requirements for transfers by way of binding corporate rules; administrative sanctions; processing for health purposes ***and*** processing in the employment context. It is of particular importance that the Commission carry out appropriate consultations during its preparatory work, including at expert level ***in particular with the European Data Protection Board*** . The Commission, when preparing and drawing-up delegated acts, should ensure a simultaneous, timely and appropriate transmission of relevant documents to the European Parliament and ***to the*** Council.

	and processing for historical, statistical and scientific research purposes . It is of particular importance that the Commission carry out appropriate consultations during its preparatory work, including at expert level. The Commission, when preparing and drawing-up delegated acts, should ensure a simultaneous, timely and appropriate transmission of relevant documents to the European Parliament and Council.	
Definition of profiling		***(Article 4-3a)*** ***'profiling' means any form of automated processing of personal data intended to evaluate certain personal aspects relating to a natural person or to analyse or predict in particular that natural person's performance at work, economic situation, location, health, personal preferences, reliability or behaviour;***
General principles for data subject rights : includes the right to object to profiling		***(Article 10a)*** ***General principles for data subject rights*** ***1. The basis of data protection is clear and unambiguous rights for the data subject which shall be respected by the data controller. The provisions of this Regulation aim to strengthen, clarify, guarantee and where appropriate, codify these rights.*** ***2. Such rights include, inter alia, the provision of clear and easily understandable information regarding the processing of his or her personal data, the right of access, rectification and erasure of their data, the right to obtain data, the right to object to profiling, the right to lodge a complaint with the competent data protection authority and to bring legal proceedings as well as the right to compensation and damages resulting from an unlawful processing operation. Such rights shall in general be exercised free of charge. The data controller shall respond to requests from the data subject within a reasonable period of time.***

Duty to inform about existence of profiling, the measures based on it and the envisioned effects.		**(Article 14-ga)** Information to the data subject. 1. Where personal data relating to a data subject are collected, the controller shall provide the data subject with at least the following information, ***[…]***: […] ***ga) where applicable, information about the existence of profiling, of measures based on profiling, and the envisaged effects of profiling on the data subject;***

Profiling which leads to measures producing legal effects concerning the data subject or does similarly significantly affect the interests, rights or freedoms of the concerned data subject shall not be based solely or predominantly on automated processing and shall include human assessment, including an explanation of the decision reached after such an assessment.	(Article 20) ***Measures based* on** profiling 1. Every natural person shall have the right ***not to be subject*** to ***a measure which produces legal effects concerning this natural person or significantly affects this natural person, and which is based solely on automated processing intended to evaluate certain personal aspects relating to this natural person or to analyse or predict in particular the natural person's performance at work, economic situation, location, health, personal preferences, reliability or behaviour.*** 2. Subject to the other provisions of this Regulation, a person may be subjected to ***a measure of the kind referred to in paragraph 1*** only if the processing: (a) is ***carried out in the course of*** the entering into, or performance of, a contract, where the request for the entering into or the performance of the contract, lodged by the data subject, has been satisfied ***or where*** suitable measures to safeguard the data subject's legitimate interests have been adduced***, such as the right to obtain human intervention*** ; or (b) is expressly authorized by a Union or Member State law which also lays down suitable measures to safeguard the data subject's legitimate interests; or (c) is based on the data subject's consent, subject to the conditions laid down in Article 7 and to suitable safeguards. 3. ***Automated processing of personal data intended to evaluate certain personal aspects relating to a natural person*** shall not be based solely on the special categories of personal data referred to in Article 9. ***4. In the cases referred to in paragraph 2, the information to be provided by the controller under Article 14 shall include information as to the existence of processing for a measure of the kind referred to in paragraph 1 and the envisaged effects of such processing on the data subject.*** 5. ***The Commission shall be***	(Article 20) Profiling ***1. Without prejudice to the provisions in Article 6*** every natural person shall have the right ***to object*** to ***profiling in accordance with Article 19. The data subject shall be informed about the right to object to profiling in a highly visible manner.*** 2. Subject to the other provisions of this Regulation, a person may be subjected to ***profiling which leads to measures producing legal effects concerning the data subject or does similarly significantly affect the interests, rights or freedoms of the concerned data subject*** only if the processing: *(a)* is ***necessary for*** the entering into, or performance of, a contract, where the request for the entering into or the performance of the contract, lodged by the data subject, has been satisfied***, provided that*** suitable measures to safeguard the data subject's legitimate interests have been adduced*;* or (b) is expressly authorized by a Union or Member State law which also lays down suitable measures to safeguard the data subject's legitimate interests; (c) is based on the data subject's consent, subject to the conditions laid down in Article 7 and to suitable safeguards. 3. ***Profiling that has the effect of discriminating against individuals on the basis of race or ethnic origin, political opinions, religion or beliefs, trade union membership, sexual orientation or gender identity, or that results in measures which have such effect, shall be prohibited. The controller shall implement effective protection against possible discrimination resulting from profiling. Profiling*** shall not be based solely on the special categories of personal data referred to in Article 9. 5. ***Profiling which leads to measures producing legal effects concerning the data subject or does similarly significantly affect the interests, rights or freedoms of the concerned data subject shall not be based solely or predominantly on automated***

	empowered to adopt delegated acts in accordance with Article 86 for the purpose of further specifying the criteria and conditions for suitable measures to safeguard the data subject's legitimate interests referred to in paragraph 2.	***processing and shall include human assessment, including an explanation of the decision reached after such an assessment. The*** suitable measures to safeguard the data subject's legitimate interests referred to in paragraph 2 ***shall include the right to obtain human assessment and an explanation of the decision reached after such assessment***. ***5a. The European Data Protection Board shall be entrusted with the task of issuing guidelines, recommendations and best practices in accordance with point (b) of Article 66 (1) for further specifying the criteria and conditions for profiling pursuant to paragraph 2.***
***Risk analysis*: Profiling resulting in measures with legal or similar significant effects is likely to pose a specific risk**		***(Article 32a) Respect to Risk*** ***1. The controller, or where applicable the processor, shall carry out a risk analysis of the potential impact of the intended data processing on the rights and freedoms of the data subjects, assessing whether its processing operations are likely to present specific risks.*** ***2. The following processing operations are likely to present specific risks:*** ***[…]*** ***(c) <u>profiling</u> on which measures are based that produce legal effects concerning the individual or similarly significantly affect the individual;*** ***[…]***

***Binding corporate rules* shall include the right not to be subject to a measure based on profiling.**	(Article 43) Transfers by way of binding corporate rules 1. A supervisory authority shall in accordance with the consistency mechanism set out in Article 58 approve binding corporate rules, […]: […] […] 2. The binding corporate rules shall at least specify: […] (e) the rights of data subjects and the means to exercise these rights, including the right not to be subject to a measure based on profiling in accordance with Article 20, the right to lodge a complaint before the competent supervisory authority and before the competent courts of the Member States in accordance with Article 75, and to obtain redress and, where appropriate, compensation for a breach of the binding corporate rules;	(Article 43) Transfers by way of binding corporate rules ***1. The*** supervisory authority shall in accordance with the consistency mechanism set out in Article 58 approve binding corporate rules, […]: […] 2. The binding corporate rules shall at least specify: […] (e) the rights of data subjects and the means to exercise these rights, including the right not to be subject to a measure based on profiling in accordance with Article 20, the right to lodge a complaint before the competent supervisory authority and before the competent courts of the Member States in accordance with Article 75, and to obtain redress and, where appropriate, compensation for a breach of the binding corporate rules;
No profiling in the employment context	(Article 82) Processing in the employment context […]	(Article 82) ***Minimum standards for*** processing ***data*** in the employment context […] ***1a. The purpose of processing such data must be linked to the reason it was collected for and stay within the context of employment. Profiling or use for secondary purposes shall not be allowed.***

Annex B – Indication of integration of legal requirements into DataBait tools (preliminary versions of tables to be developed in D3.4)

B.1 Data protection requirements based on the legal qualification of data processed in USEMP

If data is *legally qualified* as……,	**….,then the *legal effect* is….**	**…which results in this *legal requirement* :**
PD : Personal data as defined in DPD 95/46	The regime of data protection directive 95/46 applies. I. The DataBait user has the following « Informational rights » (which includes the so-called right to « profile transparency »), which entail he or she should be informed about : • the purpose for which the data are processed • what categories of data are processed, • for what estimated period, • which recipients receive the data, • what is the general logic of the data that are undergoing the processing, • what might be the consequences of such processing, • the existence of the right to request rectification or erasure of the data concerning the data subject and of the right to object to the processing, • the right to lodge a complaint to the supervisory authority and the contact details of the supervisory authority? (See Recital 51 and Art. 15 of the	I. A button which the DataBait user can click with all the information that needs to be given following the informational rights from directive 95/46. The button on the USEMP platform, and include an email address for each partner that processes personal data, to make further inquiries. The information will be updated whenever the relevant processing of personal data change.

	pGDPR) II. A purpose for the processing has to be specified	II. Purpose described in Data Licensing Agreement and also available under informational button
	III. The processing has to be based on a ground legitimizing the processing. The ground used in USEMP is « contract » (Art. 7(b) DPD 95/46)	III. Data Licensing Agreement : The ground used in USEMP is « contract » (Art. 7(b) DPD 95/46), for downloading the DataBait tools consent (art. 6.3 ePrivacy and for processing LPD again consent art. 8 DPD
	IV. The data should not be kept longer than necessary and be deleted or completely anonymized (no re-identification possible) when no longer needed (i.e. at the end of the USEMP project).	IV. The data should not be kept longer than necessary and be deleted or completely anonymized (no re-identification possible) when no longer needed (i.e. at the end of the USEMP project).
	V. Security of the processing needs to be adequate	V. A risk assessment investigating the security of the processing .
	VI. Anticipating the new EU data protection law (Art. 8, pGPDR) : a mechansm which checks the age of DataBait users and does not allow children (below the age of 13 ?) to use it.	VI. Anticipating the new EU data protection law (Art. 8, pGPDR) : a mechanism which inquires after the age of DataBait users and does not allow children

		(below the age of 13) to use it and gives a warning to anyone aged 13-18.
	VII. Anticipating the new EU data protection law (pGPDR) : implement legal protection by default and by design as much as possible	VII. Anticipating the new EU data protection law (pGPDR) : implement legal protection by default and by design as much as possible : all of the above but also [following current law and the principle of data minimization] for example check default settings and try to pseudonymize, anonymize etc. when it is not strictly necessary to have fully identifiable personal data.
	VIII. Anticipating the new EU data protection law (preambule of the pGPDR, stating that data protection is not an absolute right but that it should be balanced with other rights).	VIII. The contract (DLA) provides a more balanced approach – creating mutual duties and rights - than mere consent.
	IX. Notification of national data protection authority of processing of the data	IX. Notification of national data protection authority of processing of the data
LSD : Legal sensitive data as defined in Art. 8 DPD 95/46. Sensitive data are personal data revealing : - racial or ethnic origin, - political opinions, - religious or philosophical beliefs, -trade-union	- Specific consent	- Making sure that the DataBait tool asks the users for explicit consent [Clause G of the DLA takes care of this.] - A button where this consent can be withdrawn : Each party

membership, and -the processing of data concerning health or sex life, and - the processing of data relating to offences, criminal convictions or security measures		will also provide an email address to be contacted in case a user wants to withdraw her consent for processing her sensitive data; this is preferably the same email address as the one used to gain further information, but will be available behind a separate button on the USEMP platform.
	- Exploring whether sensitive data (Art. 8 DPD 95/46) are used as the sole ground for profiling and preferably avoid it [This is not current law and it is up for debate whether a prohibition of such profiling solely based on sensitive data will make it into the pGPDR]	- Check whether any of the inferred data in the USEMP project are solely based on sensitive data
PROFILE-INPUT : Data used as input for profiling	- Exploring whether sensitive data (Art. 8 DPD 95/46) are used as the sole ground for profiling and preferably avoid it [This is not current law and it is up for debate whether a prohibition of such profiling solely based on sensitive data will make it into the pGPDR]	- Check whether any of the inferred data in the USEMP project are solely based on sensitive data
	- Making sure no measures which have a significant or legal impact are taken based on the profiling, unless there is a contract or consent.	- Although the profiling performed through the DataBait tools is not likely to result is measures which have a significant or legal impact in a narrow sense, we interpret "significant" in a broad sense. The DLA (contract) provides a legitimizing ground.
PROFILE-OUTPUT :	- This data subject has the	- The informational

data which result from profiling	right to obtain knowledge of the logic involved in any automatic processing which significantly affects him or her (Art. 15(1) in conjunction with Art. 12(a) of the DPD 95/46). It is not completely clear how "significantly" should be defined, but to be on the safe side we give the term a broad interpretation. - Making sure no measures which have a significant or legal impact are taken based on the profiling, unless there is a contract or consent. It is not completely clear how "significant" should be defined, but to be on the safe side we give the term a broad interpretation.	button and the DataBait GUI should provide insight in the logic involved in the profiling (which knowledge in inferred from which data, how is this done, how reliable is this knowledge, etc.) - Although the profiling performed through the DataBait tools is not likely to result is measures which have a significant or legal impact in a narrow sense, we interpret "significant" in a broad sense. The DLA (contract) provides a legitimizing ground.
LD: location data as defined in e-Privacy Directive 2002/58.	- The legal status of location data is the subject of some controversies, but to be on the safe side we assume that the regime as applicable to personal data (PD) applies. Thus, see above.	- See above, same requirements as with PD.

Table B.1.1 ***The "answer" to almost all these requirements is the PDPA (which includes the DLA).*** *The legal requirements are based on the legal qualification of data processed in USEMP as* ***personal data*** *– which includes (a) "ordinary" personal data, (b) personal data which are sensitive (Art. 8 DPD 95/46), and (c) personal data which are the input or output to profiling, i.e. data used to infer other data or inferred data; where "profiling" (defined in the pGPDR) is a particular type of "automated processing" (see DPD 95/46) - or* ***location data*** *(as defined in e-Privacy Directive 2002/58)*

B.2 Personal data processed in USEMP, ordered according to source

Personal data processed in the USEMP project, ordered according to source:	**Described in table:**

A. Personal data collected with the DataBait OSN app	B.3.1
B. Personal data collected with the DataBait browser plugin	B.3.2
C. Personal data *inferred* from a subset of the data collected through the OSN app [A] and the browser plugin [B]	B.3.3
D. Personal data in training and testing sets, used to train and test classifiers (i.e., models used to predict and infer data from the. While most data in these training and testing data sets are not personal data (they are anonymized or do not relate to an identified or identifiable person), each data set has to be screened for the presence of personal data. Also, it should be noted, that the fact that most of these data are *not* derived from DataBait users does not mean that the scrutiny in terms of data protection (in as far as these data sets contain personal data) should be any less.	B.3.4

Table B.2.1. Overview of USEMP data ordered according to source

Personal data processed in the USEMP project, ordered according to source:	**Processing premise?**	**What is the technical goal of the processing ?**	**How long are the data stored ?**	**How is the data anonymized/pseusonomized during the USEMP project duration ?**
A. Personal data collected with the DataBait OSN app	HWC	(1) Representing the data in the DataBait GUI to give the DataBait user more insight in her digital trail (2) Inferring other knowledge from the data to give the DataBait user more insight in her digital trail	At most until three months after the end of the USEMP project.	Varying (needs to be further explored)
B. Personal data collected with the DataBait browser	HWC	(1) Representing the data in the DataBait GUI to give the DataBait user	At most until three months after the end of the USEMP project.	Varying (needs to be further explored)

plugin		more insight in her digital trail (2) Inferring other knowledge from the data to give the DataBait user more insight in her digital trail		
C. Personal data collected in the DataBait surveys in the pre-pilot.	HWC	(1) Finding the « true values » (ground truths). These declared data help to assess how well the classifiers developed in USEMP are able to predict/infer these values. (2) Exploring which values users consider to be sensitive.	At most until three months after the end of the USEMP project.	Varying (needs to be further explored)
D. Personal data *inferred* from a subset of the data collected through the OSN app [A] and the browser plugin [B]	HWC	Providing inferred knowledge to the the DataBait user in the GUI to give her more insight in her digital trail and possibilities to control this information.	At most until three months after the end of the USEMP project.	Varying (needs to be further explored)

E. Personal data in training and testing sets, used to train and test classifiers	HWC	Needed to build classifiers which can infer/predict certain attributes and their values based on the data gathered through the DataBait tools	Varying (needs to be further explored)	Varying (needs to be further explored)

Table B.2.2. Detailed usage of USEMP data ordered according to source

B.3 Set of table listing all personal data processed in USEMP

1. Automatically Allowed Permissions	**An app may use this permission without review from Facebook.**	**If available : corresponding numbering in table C of Annex D of Deliverable 7.1**	**Is this data used to infer anything?**	**Legal qualification in terms of EU data protection law and EU anti discrimination law[23]**
public_profile	Access to a subset of items that are part of a person's public profile. A person's public profile refers to the following properties on the user object by default:	This contributes to C7 : User Profile* and Interests ; A user represents a person on Facebook. The /{user-id} node returns a single user.		
	Id (the number of the profile, e.g. ""1424672444497579")		No	PD
	Name (full name of the user)		No	PD ; Does the name reveal race or ethnic origin ? Then it could be LSD; Moreover, if this is the case : differentiation based on race or ethnic origin is

[23] The protected grounds according to EU data protection law are: sex, racial or ethnic origin, religion or belief, disability, age, sexual orientation and nationality. See chapter 6.2.

				prohibited in the fields of employment, access to good and services, social advantages, social protection and education
	first_name (first name of the user)		No	PD ; Does the name reveal race or ethnic origin ? Then it could be LSD; Moreover, if this is the case : differentiation based on race or ethnic origin is prohibited in the fields of employment, access to good and services, social advantages, social protection and education
	last_name (last name of the user)			PD ; Does the name reveal race or ethnic origin ? Then it could be LSD; Moreover, if this is the case : differentiation based on race or ethnic origin is prohibited in the fields of employment, access to good and services, social advantages, social protection and education
	link (link to the Facebook profile, e.g.: https://www.facebook.com/app_scoped_user_id/1424672444497579/)		No	PD

	gender (gender of the user)		No	PD Differentiation based on gender in the field of employment and the access to goods and ervices is prohibited
	locale (locale/language, e.g. "en_GB", which stands for British English)		No	PD
	timezone (timezone of the user)		No	PD
	updated_time (the time of the most recent update) verified (is the Facebook		No	PD
	verified (is the Facebook account linked to a verified phonenumber and/or email address?)		No	PD
user_friends	Access the list of friends that also use your app. (this is commonly used to create a social experience in your app.)	C4; Friends-list or Friends; A person's 'friend lists' - these are groupings of friends such as "Acquaintances " or "Close Friends", or any others that may have been created.	Yes See: C4/D6	PD ; PROFILE-INPUT ; the PROFILE-OUTPUT based on these data *could be* LSD – depending on the content of the inferral made.
Email	Access to a person's primary email address.		No	PD
2. **Requested[24] extended permissions**	These permissions are not optional in the login dialog during the login flow, meaning they are non-optional for people when logging into your app. If you want them to be optional, you should structure your			

[24] Facebook still has to give permission

	app to only request them when absolutely necessary and not during initial login.			
user_about_me	Access to a person's personal description (the 'About Me' section on their Profile) through the bio property on the User object.	This contributes to C7 : User Profile* and Interests ; A user represents a person on Facebook. The /{user-id} node returns a single user.	Maybe	PD ; these data *could be* LSD – depending on the content
User_posts	Access to a person's posts on the User object	Contributes to C6 ; News ; The person's news feed Permission to get 'read_stream' would give full access to a person's newsfeed, but this is unlikely to be granted by Facebook. However 'user_posts' is likely to be granted and returns similar data.	Yes	
user_activities	Access to a person's list of activities as listed on their Profile. This is a subset of the pages they have liked, where those pages represent particular interests.		Maybe	PD ; these data *could be* LSD – depending on the content
user_education_history	Access to a person's education history through the education field on the User object.	This contributes to C7 : User Profile* and Interests ; A user represents a person on Facebook. The /{user-id} node returns a single user.	Maybe	PD
user_hometown	Access to a person's hometown location through the hometown field on the User object. This is set by	This contributes to C7 : User Profile* and Interests ; A	No	PD

	the user on the Profile.	user represents a person on Facebook. The /{user-id} node returns a single user.		
user_interests	Access to the list of interests in a person's Profile. This is a subset of the pages they have liked which represent particular interests[25].	This contributes to C7 : User Profile* and Interests ; A user represents a person on Facebook. The /{user-id} node returns a single user.	Maybe	PD ; these data *could be* LSD – depending on the content
user_likes	Access to the list of things a person likes. Provides access to the list of all Facebook Pages and Open Graph objects that a person has liked.	C2; Likes and Unlikes; The Facebook Pages that this person has 'liked'.	Yes See :C2/D1	PD ; these data *could be* LSD – depending on the content ; PROFILE-INPUT ; the PROFILE-OUTPUT based on these data *could be* LSD – depending on the content of the inferral made.
user_location	Access to a person's current city through the location field on the User object. The current city is set by a person on their Profile.	This contributes to C7 : User Profile* and Interests ; A user represents a person on Facebook. The /{user-id} node returns a single user.	Maybe	PD
user_photos	Access to the photos a person has uploaded or been tagged in. This is available through the photos edge on the User object.	C3; Photos Or Photos Uploaded; Represents an individual photo on Facebook. Contributes to C5 ; Friends' activities upon	Yes See : C3/D5 C5/D7	PD ; these data *could be* LSD – depending on the content ; PROFILE-INPUT ; the PROFILE-OUTPUT based on these data

[25] The user_interests permission is deprecated. On Tuesday, June 23, 2015, this permission request will be silently ignored. Please see Facebook's changelog for more information.

		user's OSN objects ; Represents an action of a friend in one of a user's objects on Facebook.		*could be* LSD – depending on the content of the inferral made.
user_relationships	Access to a person's relationship status, significant other and family members as fields on the User object.	This contributes to C7 : User Profile* and Interests ; A user represents a person on Facebook. The /{user-id} node returns a single user.	No	PD ; LSD
user_relationship_details	Access to a person's relationship interests as the interested_in field on the User object.	This contributes to C7 : User Profile* and Interests ; A user represents a person on Facebook. The /{user-id} node returns a single user.	No	PD ; LSD
user_religion_politics	Access to a person's religious and political affiliations.	This contributes to C7 : User Profile* and Interests ; A user represents a person on Facebook. The /{user-id} node returns a single user.	No	PD ; LSD
user_status	Access to a person's statuses. These are posts on Facebook which don't include links, videos or photos.	C1; Posts Feed; An individual entry in a profile's feed. The profile could be a user, page, app, or group. Contributes to C5 ; Friends' activities upon user's OSN objects ; Represents an action of a friend in one of a user's objects	C1/D2 C5/D7	PD ; *could be* LSD – depending on the content of the status update; PROFILE-INPUT; the PROFILE-OUTPUT based on these data *could be* LSD – depending on the content of the inferral made.

		on Facebook.		
user_tagged_places	Access to the Places a person has been tagged at in photos, videos, statuses and links.			PD ; *could be* LSD – depending on the content of photos, videos, statuses and links in which the user is tagged.
user_videos	Access to the videos a person has uploaded or been tagged in.	Contributes to C5 ; Friends' activities upon user's OSN objects ; Represents an action of a friend in one of a user's objects on Facebook.	C5/D7	PD ; *could be* LSD – depending on the content of the videos
Metadata (which come along with e.g. user_status', 'user_posts' and 'user_tagged_places')				
Location related data e.g. : `"place": place of the user who posted the status update "name":name of the location of the user, e.g. a concert hall or the public library "street":street name "city": city name`				PD ; *could be* LSD – depending on the content of the statuses ; meta data relating to location are LD.

"state":name of state "country":country name "zip":zip code "latitude":latitude "longitude": longitude				
"id": id of the user who posted the status update				PD

Table B.3.1 Personal data collected with the DataBait OSN app

#	**Name**	**Description**	**Is this data used to infer anything?**	**Legal qualification in terms of EU data protection law and EU anti discrimination law[26]**
B1	# of Trackers for Site URL	The number of tracking services when a LIO user visits URL	no	
B2	Tracker	The ID of the tracking services when a LIO user visits a URL	yes	
B2*	TRAINING OR TESTING DATA Tracker	The ID of the tracking services when a LIO user visits a URL		
B3	Tracker email	A Tracker of		

[26] The protected grounds according to EU data protection law are: sex, racial or ethnic origin, religion or belief, disability, age, sexual orientation and nationality. See chapter 6.2.

		users email (e.g., google-mail)		

#	Name	Description	Is this data used to infer anything?	Legal qualification in terms of EU data protection law and EU anti discrimination law[27]
C1	Posts Feed [28]	An individual entry in a profile's feed. The profile could be a user, page, app, or group.	yes	
C1*	TRAINING OR TESTING DATA Posts Feed [29]	An individual entry in a profile's feed. The profile could be a user, page, app, or group.		
C2	Likes and Unlikes [30]	The Facebook Pages that this person has 'liked'.	yes	
C2*	TRAINING OR TESTING DATA Likes and Unlikes [31]	The Facebook Pages that this person has 'liked'.		
C3	Photos Or Photos Uploaded [32]	Represents an individual photo on Facebook.	Yes	
C3*	TRAINING OR TESTING DATA Photos Or	Represents an individual photo on Facebook.		

[27] The protected grounds according to EU data protection law are: sex, racial or ethnic origin, religion or belief, disability, age, sexual orientation and nationality. See chapter 6.2.
[28] https://developers.facebook.com/docs/graph-api/reference/v2.0/user/feed/
[29] https://developers.facebook.com/docs/graph-api/reference/v2.0/user/feed/
[30] https://developers.facebook.com/docs/graph-api/reference/v2.0/user/likes
[31] https://developers.facebook.com/docs/graph-api/reference/v2.0/user/likes
[32] https://developers.facebook.com/docs/graph-api/reference/v2.0/photo/

	Photos Uploaded [33]			
C4	Friends-list or Friends	A person's 'friend lists' - these are groupings of friends such as "Acquaintances" or "Close Friends", or any others that may have been created.	Yes	
C4*	TRAINING OR TESTING DATA Friends-list or Friends	A person's 'friend lists' - these are groupings of friends such as "Acquaintances" or "Close Friends", or any others that may have been created.		
C5	Friends' activities upon user's OSN objects	Represents an action of a friend in one of a user's objects on Facebook.	Yes	
C5*	TRAINING OR TESTING DATA Friends' activities upon user's OSN objects	Represents an action of a friend in one of a user's objects on Facebook.		
C6	News [34] (/home)	The person's news feed.	No (?)	
C7	User Profile and Interests	A user represents a person on Facebook. The /{user-id} node returns a single user.	No (?)	

Table B.3.2 Personal data collected with the DataBait browser plugin (based on annex D of D7.1)

[33] https://developers.facebook.com/docs/graph-api/reference/v2.0/photo/
[34] https://developers.facebook.com/docs/graph-api/reference/v2.0/user/home/

#	Name	Description	Which data from Annex D, D7.1 are used to establish or infer this? (more than one answer is of course possible)	Which method is used if the data are inferred?	Which data are used to train (and/or test) the classifier (model) if data are inferred?	Legal qualification in terms of EU data protection law and EU anti-discrimination law
A	Demographics	1. Age				
		2. Gender				
		3. Nationality				
		4. Racial origin				
		5. Ethnicity				
		6. Literacy level				
		7. Employment status				
		8. Income level				
		9. Family status .				
B	Psychological Traits	1. Emotional stability				
		2. Agreeableness				
		3. Extraversion				
		4. Conscientiousness				
		5. Openness				
C	Sexual Profile	1. Sexual preference				

D	Political Attitudes	1. Parties (Part of list for Belgium: CD&V; Groen!; N-VA; Open VLD /Part of list for Sweden: Centerpartiet; Vansterpartiet; Folkpartiet liberalerna)				
		2. Political ideology (Communist; Socialist; Green; Liberal; Christian democratic; Conservative; Right-wing extremist)				
E	Religious Beliefs	Supported Religion (Atheist, Agnostic, Christian, Muslim, Hinduist, Buddhist, Other, etc.)				
F	Health Factors & Condition	1. Smoking				
		2. Drinking (alcohol)				
		3. Drug use				
		4. Chronic diseases				
		5. Disabilities				

		6. Other health factors (e.g.: Exercise (yes / no); Late night shifts (yes / no); Staying up late)				
G	Location	1. Home				
		2. Work				
		3. Favourite places				
		4. Visited places				
H	Consumer Profile	1. Brand attitude				
		2. Hobbies				
		3. Devices				
I	Digital traces score (How sensitive, uncontrollable and visble are your data?)					
J	Value score (how valuable are your data?)					

Table B.3.3 Personal data inferred from a subset of the data collected through [A] the OSN app and [B] the browser plugin. This table is based on deliverable D6.1. It is not certain that all these data will be inferred. This table will be populated in D3.4

Dataset	Source	Purpose	Inferred attributes	Does the dataset

				contain personal data ?
MyPersonality	http://mypersonality.org/wiki/doku.php	Integration as training set in the behavioral detection module and quite probably also in the topic based attribute detection module.	A1, A2, A9, B, C, D.2, E	Not likely. Anonymized. However, details need to be further explored
Zerr's image privacy dataset	http://l3s.de/picalert/#ustudydata	Integration as training set in a module that assists the user to define his privacy settings. It is used to assist classification of images as private or public. The user is warned when he / she is about to post an image that is classified as private.	None. As mentioned it is not used to infer any profile attributes	Needs to be further explored
Location estimation dataset	http://www.multimediaeval.org/mediaeval2014/placing2014/ Dataset accessible only by competition participants	Integration as training set in the location recognition module.	G (actually I am not sure if we are making a distinction between G1-G4)	Needs to be further explored
Kaggle community detection dataset	https://www.kaggle.com/c/learning-social-circles	Integration as training set in the privacy settings assistance module. It is used in order to help group the friends of a user in circles.	None.	Needs to be further explored
Relevance- and Diversity-based Reranking dataset	http://www.multimediaeval.org/mediaeval2014/diverseimages2014/	Benchmarking of method used for the relevance and reranking module that is used as part of the VIS-REC and PRIV-SCOR modules.	None.	Needs to be further explored
Wikipedia	https://dumps.wikimedia.org/	Creation of a training set that represents different privacy-related dimensions	D.1, D.2, E.1, G.1, G.2, G.3, G.4, H.1, H.3, H.4	Needs to be further explored
SentiWordNet	http://sentiwordnet.isti.cnr.it/	Integration as training set in the opinion mining module	D.1, D.2, E.1, H.1	Needs to be further explored
ImageNet	http://image-net.org/	Training set for the visual concept recognition module	F.1, F.2, H.1, H.3	Needs to be further explored
FlickrLogos-32	http://www.multimedia-computing.de/f	Training set for the logo recognition module	H.1	Needs to be further explored

	lickrlogos			
Yahoo Flickr Creative Commons 100 Million	http://webscope.sandbox.yahoo.com/catalog.php?datatype=i&did=67	Training set for the location recognition and face recognition modules	A.2, A.3, F.1, F2, G.1, G2, G.3, G.4	Needs to be further explored
Pre-pilot / system operation dataset	-	Questionnaire data, OSN data, browsing behavior data. To be obtained and investigated at a later stage.	Most likely all	Needs to be further explored
Twitter data set, derived through public API	http://ceur-ws.org/Vol-1150/overview.pdf;	Used in section 6 of D6.1.	Needs to be further explored	Needs to be further explored

Table B.3.4 Possible personal data in training and testing sets, used to train and test classifiers