M. Talamo, H. Roßnagel, D. Hühnlein, C. Schunck (Ed.): Open Identity Summit 2016,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2016 – 1

What's in a name: the conflicting views of pseudonymisation under eIDAS and the General Data Protection Regulation

Niko Tsakalakis,[1] Sophie Stalla-Bourdillon[2] and Kieron O'Hara[3]

Abstract:

Pseudonymisation is gaining traction among modern electronic identification systems as a privacy enhancing technique that can significantly reduce risks of personal data misuse. The recently agreed General Data Protection Regulation (the GDPR) encourages the use of pseudonymisation to comply with its requirement of privacy-by-design. Art. 5 of the European Regulation on electronic identification and trust services (eIDAS) on data processing and protection simply allows the use of pseudonyms in electronic transactions although the facilitation of the implementation of the principle of privacy by design is clearly among the aims listed by Art. 12 of eIDAS. This paper examines the concept of pseudonymisation under eIDAS and the GDPR and suggests that the two Regulations employ two very different, if not incompatible, notions of pseudonymisation. It concludes that a common terminology and approach would be preferable in order to ensure consistency and legal certainty.

Keywords: eIDAS, GDPR, pseudonymisation, e-ID, electronic identity, eIDM

1 Introduction

This paper focuses on pseudonymisation, a concept that was only recently formally introduced in the EU policy landscape. In particular it attempts to derive the effects of the introduction of pseudonyms (or pseudonymous credentials) as part of the Regulation on electronic identification and trust services (eIDAS). It examines how eIDAS conceives pseudonymisation and explains how this interpretation would translate in practical uses in the context of a pan-European interoperability framework. It then assesses this approach in the light of the newly adopted data protection rules, i.e. the final version of the General Data Protection Regulation (the GDPR). The GDPR, as adopted on the 27th of April 2016, introduces a new category of data, between personal and data that has undergone anonymisation, that allows flexible processing of the data beyond the initial collection purposes. Finally it concludes with some useful insights about the intersection of the two Regulations and possible inconsistencies. The paper is organised as follows: Sect. 2 offers an overview of eIDAS Regulation and its requirements in sect. 2.1. Sect. 2.2 examines how eIDAS conceives pseudonymisation and a parallel is drawn with the definition and approach of the GDPR in sect. 3. Sect. 3 combines both eIDAS and the GDPR to assess their compatibility.

[1] http://orcid.org/0000-0003-2654-0825, University of Southampton, Web Science Doctoral Training Centre, Building 32, Highfield Campus, Southampton SO17 1BJ, UK, N.Tsakalakis@southampton.ac.uk
[2] University of Southampton, Institute for Law and the Web, Faculty of Business and Law, Highfield Campus, Southampton SO17 1BJ, UK, S.Stalla-Bourdillon@soton.ac.uk
[3] University of Southampton, Web and Internet Science Group, Electronics and Computer Science, Highfield Campus, Southampton SO17 1BJ, UK, kmo@ecs.soton.ac.uk

Finally sect. 4 offers arguments for the need to redress how pseudonymisation is currently defined essentially in the GDPR and how privacy by design is promoted in eIDAS.

2 eIDAS Regulation

In 2012 the Commission decided that a revision to the eSignatures Directive[4] was necessary. The proposed draft was adopted by the Parliament and the Council in 2014.[5] eIDAS' aim is to create a uniform legal framework under which the various national electronic identification schemes and trust services will be able to interoperate in cross-border transactions. In line with most recent policy initiatives, eIDAS follows a technology-agnostic approach in regulating the interoperability framework. The Regulation sets out the goals of the policy but does not define any system specifications. According to eIDAS member states that wish to operate cross-border authentication should notify their national eID Management systems (eIDM) to a designated body. Voluntary notification applications are accepted since September 2015 and by 2018 all member states are required to start accepting identifications from notified schemes.[6]

2.1 eIDAS minimum requirements

As already mentioned, the Regulation and implementing acts try to keep requirements to a minimum.[7] eIDAS defines three Levels of Assurance that define a risk based identification, requiring higher levels of identification certainty depending on the importance of the transaction about to be performed. The levels are named 'Low', 'Substantial' and 'High' and their requirements are defined through an implementing act.[8] A specific mention is made to data protection, as all components of the framework are required to fully comply with the Data Protection Directive (DPD),[9] which was in effect when the Regulation was published.[10] All services also need to be built around 'Privacy-by-design' principles.[11] Though not legally binding, the recitals advise that services follow the data minimisation principle and request and process only data strictly necessary for each transaction.[12] The goal of the framework is identification 'uniquely representing' a natural or legal person.[13]

[4] Council Directive 1999/93/EC [1999] OJ L013/0012

[5] Regulation (EU) No 910/2014 of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC [2014] OJ L257/73

[6] eIDAS n 5 arts 6, 7 and 9

[7] Specific technologies are acceptable only when absolutely essential for the security of the system or the users. Note that the first drafts of the Regulation were even stricter, prohibiting member states to operate any scheme that would require extra hardware or software to be implemented by other member states. This wording was toned down in the final text after objections: *See* [CS14]. However, the implementing acts point towards specific implementations, creating therefore de facto standards.

[8] Commission Implementing Regulation (EU) 2015/1502 [2015] OJ L235/7 ANNEX 2 pp 7-20

[9] Directive 95/46/EC [1995] OJ L281/0031

[10] eIDAS n 5 art 12(3d)

[11] eIDAS n 5 art 12(3c)

[12] eIDAS n 5 recital 11

[13] eIDAS n 5 art 3(1)

Implementating Regulation 2015/1501 clarifies this further by providing a minimum dataset (MDS) that should be transmitted at every identification.[14] The MDS defines four compulsory and four optional identifiers that need to be included (table 1). Note that the presence of a unique identifier is mandatory. The format of the identifier is up to the Member States, but it should be as persistent in time as possible. This decision of the regulating committee, that points to centralised architectures with unique links to users, which have lately been substituted with more privacy-preserving methods, was questioned for designing a framework that offers less privacy than what is currently technically possible [MG13].

	mandatory	optional
attributes	current family name(s)	first and family name(s) at birth
	current first name(s)	place of birth
	date of birth	current address
	unique identifier	gender

Tab. 1: Minimum dataset for a natural person

2.2 Pseudonyms under eIDAS

Art. 5(2) of eIDAS allows the use of pseudonyms for electronic transactions including electronic identification:

> Without prejudice to the legal effect given to pseudonyms under national law, the use of pseudonyms in electronic transactions shall not be prohibited

Pseudonyms are artificial identifiers constructed to replace some information in a dataset in order to make attribution of a record to a person harder. Pseudonyms are not included in the definitions of eIDAS and the only other mentions in the Regulation are in relation to qualified certificates. At first instance inclusion of pseudonyms in art 5(2) could appear like an attempt to encourage the use of privacy-enhancing technologies in national eID systems (eIDM). Pseudonyms offer better privacy protection [PH10, Br14, Ve15] and have been used by national eIDM already to enhance user privacy. In Austria, eIDs use a Unique Identifier (UID) that derives from the Central Residents Register. The architecture resembles in some respects Estonia's, which also uses central Resident Numbers. In contrast to Estonia though, where Residents Numbers are publicly available information, in Austria it is prohibited by law to share this number with the services.[15] Instead the system employs Sector-specific PINs (ssPin) – pseudonyms constructed partly from citizen number and partly from the services requiring them [Rö08]. This way every department receives different credentials for the same user and the eID is cross-sector unlinkable which does not allow to trace the transactions of a user across the system, i.e. the eIDM does not allow the tracing of

[14] Commission Implementing Regulation (EU) 2015/1501 [2015] OJ L235/1 ANNEX 1 pp 1-6

[15] Art. 6, Bundesgesetz über Regelungen zur Erleichterung des elektronischen Verkehrs mit öffentlichen Stellen (E-Government-Gesetz-E-GovG) [2008] BGBl. I Nr. 7/2008, in combination with Bundesgesetz über das polizeiliche Meldewesen (Meldegesetz 1991 – MeldeG) [1992] BGBl. Nr. 9/1992

a user's transactions across services. Pseudonyms are employed by the German nPA as well. The German system revolves around the eID card, the nPA, since there is no Identity Provider present. Instead the user is in charge of their eID and where it will be disclosed. Pseudonymisation is built directly into the system: every time a nPA is used to a service a specific pseudonym is created that is unique to this particular use [Po12]. Pseudonymisation as a requirement is actually found in various pieces of German legislation.[16]

It seems, thus, that inclusion of pseudonyms by eIDAS would afford similar data protection of the personal data exchanged as in the national systems mentioned above. This expectation might seem reasonable also by the fact that pseudonyms are mentioned immediately after the requirement in art 5(1) for all services to comply with the DPD. At the time of publication of eIDAS the DPD was still the prevalent data protection instrument. Although the DPD did not specifically acknowledge the usefulness of pseudonyms in data protection, the GDPR that will soon replace it does.

3 Combining eIDAS and the GDPR

Recently the Council, Commission and Parliament agreed on the final text of the GDPR.[17] The GDPR is set to replace the existing DPD on 25th of May 2018. As this is a Regulation, it will have direct effect within all Member States.

The GDPR includes for the first time a definition for pseudonymisation: Art 4(5) defines 'pseudonymisation' as *"the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information"*. These additional sources need to be kept separately by the data controller, who should have taken all appropriate technical and organisational measures to ensure non-attribution.[18] Pseudonymised data are not exempt from the GDPR. Although recital 28 accepts that pseudonymisation can reduce risks of personal data breaches, according to recital 26 pseudonymised data should still be considered to include information relating to identifiable natural persons.[19]

Unfortunately the GDPR's definition of pseudonymisation seems to be too strict to be practically attainable. And, even though 'pseudonymisation' is expressly defined, not once a reference to what a pseudonym is given. An illustrative example of the difficulty to meet the GDPR's definition of pseudonymisation is to be found in eIDAS Minimum Data Set.

[16] For example, *see* the 'telemedia act' that mandates users use and pay for services pseudonymously: Telemediengesetz vom 26. Februar 2007 (BGBl. I S. 179). *See* also the Constitutional Court decision that bans the use of any UID in Germany: Volkszählung Urteil des Ersten Senats vom 15. Dezember 1983 auf die mündliche Verhandlung vom 18. und 19. Oktober 1983, BVerfGE 65, 1, 1 BvR 209, 269, 362, 420, 440, 484/83 in den Verfahren über die Verfassungsbeschwerden [in German].

[17] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119/1

[18] Another innovation of the GDPR is the introduction of privacy-by-design and data-protection-by-design as the appropriate minimum technical requirements to ensure data protection: n 17 art 25

[19] The GDPR proposes a risk-based analysis to mitigate risk of re-identification, so that data controllers judge how 'reasonably likely' it is for attempts to use indirect identifiers to re-identify the dataset to come into fruition, but it has been criticised as too subjective [Es16].

As aforementioned, in order for any dataset to be considered pseudonymised by the GDPR any attribution of information to identifiable persons should be excluded. This is perhaps possible in systems that allow for selective disclosure. Selective disclosure allows the system to transmit only those attributes of the eID that are absolutely necessary for the needs of each service [PH10], i.e. if a service needs only to know if a user is a citizen or not, the system can reply with only a pseudonym for the particular user and a Yes/No answer to their citizen status. This is a function that can be observed in modern privacy-preserving systems, such as in the German nPA [Fe11]. The nPA could potentially transmit completely pseudonymous datasets (by the GDPR's definition) when, apart from the pseudonym, no other information could potentially identify a person. For example, if a service only requires a user to be over a certain age, the system will send only a pseudonym for this use (see sect. 2.2) and a calculation of the user's age based on their date of birth. [20] In this case the chance of re-identifying the person based on indirect identifiers is minimal – and therefore the test of 'reasonably likely' of the GDPR can be satisfied while the data are in transit.[21] On the other hand, in a dataset where the pseudonym is accompanied by a date of birth, address and gender the possibility of re-identification is significantly higher.[22]

Evidently the MDS would fail this test: As required by Implementing Regulation 2015/1501, it is mandatory for the MDS to *always* include all of the identifiers needed to uniquely single out a person. And even though the Regulation introduces the possibility of pseudonyms, no other mention of them exists in either the rest of the main text or the implementing acts. It seems, therefore, that pseudonyms were never conceived as a privacy-enhancing tool,[23] meaning that the interoperability framework will likely offer less privacy protection than the individual national systems are capable of.

Omission of pseudonyms as privacy-preserving features begs the question why the definition was included in eIDAS in the first place. A practical application of pseudonyms under the current regime can be seen when considering eIDAS' requirement for a UID. The framework needed to be able to accommodate all available national architectures. Detailing the differences in eIDM architecture is beyond the scope of this paper, but implementation varies from systems with central governmental databases that use the Central Residents Register number as a UID, like Estonia's [Ma10], to systems employing a varying UID like Austria [Rö08] and systems designed specifically to disallow the existence of such UID, like Germany's nPA [Po12] and UK's Verify [BD11].

As a result, requiring a mandatory UID at the national level would have prohibited some member states from notifying their schemes. It seems that the use of a pseudonym can rectify this issue. This has already been the accepted solution in the design of the nPA

[20] The German system is capable of performing calculations based on the date of birth. Every identifier of an eID resides inside the eID card of the user and the system reads and transmits only the necessary information for each use [Po12].

[21] The GDPR employs a risk-based approach to data protection (n 19 above). Organisations are encouraged to implement organisational and technical measures suitable for the activities they engage in, following a risk assessment. See n 17 arts. 30-34 and rec. 26.

[22] According to [Sw00] these three indirect identifiers are enough to identify 87% of the American population.

[23] See, for example, [ZS13] where the authors explain how pseudonyms and selective disclosure can enhance system privacy.

interoperability functionality [BS15] and it has been proposed as a necessary architectural amendment of UK's Verify [Ts16]. Such a view seems compatible with art 87 of the GDPR that implicitly states that member states are free to determine any identifier in place of a national identification number.[24]

4 Conclusion

According to the above it seems that eIDAS and the GDPR do not have the same approach to pseudonymisation. The GDPR promotes pseudonymisation as a useful compliance tool to mitigate personal data risks and effect data minimisation and data-protection-by-default [Es15]. Yet, its proposed definition appears over-restrictive [SK16] and will create confusion between pseudonymous datasets (that still fall within the scope of the GDPR) and anonymous datasets (that are exempted). On the other hand, eIDAS regards pseudonyms only as a means to construct (perhaps safer) unique identifiers, failing to incorporate pseudonyms in a way that would allow the system to enjoy selective disclosure functionality – a fact that was highlighted during the drafting phase of eIDAS by the Article 29 Working Party [Ar15].

In an effort to give an incentive to data controllers to pseudonymise datasets the GDPR seems to facilitate the further processing of pseudonymised datasets. Art 5 of the GDPR provides a waiver of the requirement for a legal basis to process data where, in conjunction with art 6(4), datasets that have been pseudonymised can be further processed if the controller deems the processing 'compatible' with the initial purpose(s). This could potentially be troublesome (assuming the definition of the GDPR is met), as many services connected to a national eID scheme collect a wealth of personal data, e.g. tax offices or healthcare providers, and eIDAS does not allow the use of a plurality of pseudonyms according to the service used each time.

Finally, in a recent communication about the future of online platforms, the Commission expresses the opinion that online platforms should be encouraged to start accepting eIDs notified under eIDAS as a means to authenticate their users [Eu16].[25] The possibility of applying a system that was designed to offer officially approved national identities to private service providers could be problematic. By not allowing for selective disclosure, eIDAS limits pseudonymisation in identifications to serve as a replacement for UID. If, by the GDPR, pseudonymisation is regarded as a key means of ensuring privacy- and data-protection- by default, eIDAS has missed a chance to further data privacy of citizens, which should be worrying if eIDAS is used in commercial online platforms. At the same time, the Commission and its policy making bodies should work on how to make its definition of pseudonymisation less obscure before it can fully live up to its envisaged potential as a means of compliance with data protection obligations.

[24] "Member States may further determine the specific conditions for the processing of a national identification number or any other identifier of general application": n 17 art 87

[25] "In order to empower consumers and to safeguard principles of competition, consumer protection and data protection, the Commission will promote interoperability actions to encourage on-line platforms to recognise other elD means, in particular those notified under eIDAS Regulation, that offer the same reassurance as their own."Please note that this is a leaked draft; the official document has not been released at the time of writing.

Acknowledgements
This research was funded by the Research Councils UK Digital Economy Programme, Web Science Doctoral Training Centre, University of Southampton, EP/L016117/1, partly supported by the SOCIAM Project, funded by the UK Engineering and Physical Sciences Research Council under grant number EP/J017728/2, and was partly funded the European Union's Horizon 2020 research and innovation programme under grant agreement No 700542. This paper reflects only the authors' views; the Commission is not responsible for any use that may be made of the information it contains.

Literature

[Ar15] Article 29 Data Protection Working Party: Letter of the WP29 to eIDAS. Available from: http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2015/20150622_letter_of_the_art_29_wp_to_eidas.pdf [Accessed 02 May 2016], 22 June 2015.

[BD11] Beynon-Davies, P.: The UK national identity card. J Inf technol, 1(1):12–21, 2011.

[Br14] Bringer, J.; Chabanne, H.; Lescuyer, R.; Patey, A.: Efficient and Strongly Secure Dynamic Domain–Specific Pseudonymous Signatures for ID Documents. In (Christin, Nicolas; Safavi-Naini, Reihaneh, Eds.): Financial Cryptography and Data Security: 18th International Conference, FC 2014, Christ Church, Barbados, March 3–7, 2014, Revised Selected Papers. Springer Berlin Heidelberg, Berlin, Heidelberg, S. 255–272, 2014.

[BS15] BSI: TR-03110 eIDAS Token Specification. Available from: https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/TechGuidelines/TR03110/BSI_TR-03110_Part-2-V2_2.pdf?__blob=publicationFile&v=1 [Accessed on 05 January 2016], 2015.

[CS14] Cuijpers, C.; Schroers, J.: eIDAS as guideline for the development of a pan European eID framework in FutureID. In (Hühnlein, Detlef, Ed.): Open Identity Summit 2014, Jgg. 237. Bonner Köllen Verlag, S. 23–38, 2014.

[Es15] Esayas, S. Y.: The role of anonymisation and pseudonymisation under the EU data privacy rules: beyond the 'all or nothing' approach. European Journal of Law and Technology, 6(2), 2015.

[Es16] Eskens, S. J.: Profiling the European Citizen in the Internet of Things: How Will the General Data Protection Regulation Apply to this Form of Personal Data Processing, and How Should It? Master thesis, University of Amsterdam. Available from: http://ssrn.com/abstract=2752010 [Accessed 02 May 2016] 2016.

[Eu16] European Commission: Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: Online Platforms and the Digital Single Market, Opportunities and Challenges for Europe, COM (2016) DRAFT. Available from: http://www.politico.eu/wp-content/uploads/2016/04/Platforms-Communication.pdf [Accessed 06 May 2016], 25 May 2016.

[Fe11] Federal Office for Information Security: Technical Guideline TR-03127: Architecture electronic Identity Card and electronic Resident Permit. Available from: https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/TechGuidelines/TR03127/BSI-TR-03127_en.pdf [Accessed 10 January 2016], 2011.

[Ma10] Martens, T.: Electronic identity management in Estonia between market and state governance. Identity in the Information Society, 3(1):213–233, 2010.

[MG13] Massacci, F.; Gadyatskaya, O.: How to get better EID and Trust Services by leveraging eIDAS legislation on EU funded research results. White Paper. Available from: http://www.cspforum.eu/Seccord_eidas_whitepaper_2013.pdf []Accessed 09 January 2016], 27 July 2015 2013.

[PH10] Pfitzmann, A.; Hansen, M.: Anonymity, Unlinkability, Unobservability, Pseudonymity and Identity Management – A Consolidated Proposal for Terminology. Available from: http://www.maroki.de/pub/dphistory/Anon_Terminology_v0.34.pdf [Accessed 23 January 2016], 12 June 2015 2010.

[Po12] Poller, A.; Waldmann, U.; Vowe, S.; Turpe, S.: Electronic Identity Cards for User Authentication – Promise and Practice. 10(1):46–54, 2012. IEEE Security & Privacy.

[Rö08] Rössler, T.: Giving an interoperable e–ID solution: Using foreign e–IDs in Austrian e–Government. Computer Law & Security Review, 24(5):447–453, 2008.

[SK16] Stalla-Bourdillon, S.; Knight, A.: Anonymous data v. Personal data – A false debate: An EU perspective on anonymisation, pseudonymisation and personal data. Wis. Int'l L.J., 2016 [forthcoming].

[Sw00] Sweeney, L.: Simple Demographics Often Identify People Uniquely. Carnegie Mellon University, Data Privacy [Working paper]. Available from: http://dataprivacylab.org/projects/identifiability/ [Accessed 02 May 2016], 2000.

[Ts16] Tsakalakis, N.; O'Hara K.; Stalla-Bourdillon, S.: Identity assurance in the UK: technical implementations and legal implications under the eIDAS regulation. WebSci'16, Hannover, DE, May 22–25 2016. Available from: http://eprints.soton.ac.uk/393204/ [Accessed 05 May 2016], 2016.

[Ve15] Verheul, E. R.: Privacy protection in electronic education based on polymorphic pseudonymization. Report 2015/1228. Available from: http://eprint.iacr.org/2015/1228 [Accessed 04 May 2016], 2015.

[ZS13] Zwingelberg, H.; Schallaböck, J.: H2.4 The Proposal for a Regulation on Electronic Identification and Trust Services under a Privacy and ABC4Trust Perspective. Available from: https://abc4trust.eu/index.php/pub/deliverables/176-h2-4 [Accessed 14 February 2016], 03 August 2015 2013.