IDENTIFYING THE IDENTIFIED: UNRAVELING THE THIRD ELEMENT OF PERSONAL DATA IN EU LAW

Yacine Ouarab[1]

DOI: https://doi.org/10.33344/vol17iss1pp64-80

Helsinki Law Review, 1/2023, pp. 64–80
© 2023 Pykälä ry, Mannerheimintie 3 B, 5th floor, 00100 Helsinki, Finland, and the author.

VERTAISARVIOITU
KOLLEGIALT GRANSKAD
PEER-REVIEWED
www.tsv.fi/tunnus

ABSTRACT

The notion of 'personal data' is a well-established concept within European Union legislation, having been defined and interpreted through various legal texts and court cases over the past two decades. However, the third element of this definition, which pertains to the identification or identifiability of an individual, continues to generate considerable ambiguity. The crux of this uncertainty lies in determining the circumstances under which an individual can be deemed 'identified'. This interpretation is of paramount importance, as data that cannot be associated with an identified or identifiable individual is not classified as personal data, thereby falling outside the scope of the General Data Protection Regulation (GDPR).

Historically, the Court of Justice of the European Union (CJEU) has not provided a clear stance on the threshold of identifiability. The Working Party, established under Article 29 of the Data Protection Directive, has offered its own interpretation, providing a detailed perspective on what constitutes an 'identified' individual. Despite its non-binding nature, this opinion is frequently employed by legal scholars as a foundation for defining personal data. However, the absence of references to this opinion in CJEU judgements, coupled with the fact that the European Data Protection Board (the successor of the Working Party 29) has not officially endorsed or adopted the WP 136 opinion, leaves the question of what can be considered as 'identified' open to interpretation and debate.

This article posits that 'identification' has to be construed as the process of distinguishing an individual from a larger group. The article argues that alternative interpretations could undermine the fundamental objectives of the GDPR. Thus, this paper seeks to contribute to the ongoing discourse surrounding the definition of personal data within the context of EU data protection law.

1 The article was written with the help of a large language model.

1. INTRODUCTION

The right to personal data protection is enshrined as a fundamental right within the European Union, standing alongside other pivotal rights such as freedom of expression and freedom of religion.[2] Despite its current prominence, this right is relatively recent, having been ratified only at the end of 2009 with the introduction of the Lisbon Treaty.[3] This new right created a legal basis for the European Commission to introduce comprehensive legislation to ensure its effective safeguarding and regulation[4], culminating in the creation of the General Data Protection Regulation 2016/679 (GDPR).

At the heart of the GDPR lies the concept of personal data, a notion largely inherited from the Data Protection Directive (DPD) of the 1990s.[5] The application of the GDPR's provisions hinges on whether the data in question is classified as personal data. Consequently, the precise definition of personal data is of paramount importance for the protection of this fundamental right. The current definition, however, is not devoid of ambiguity, as evidenced by the terse case-law of the Court of Justice of the European Union (CJEU).

According to the GDPR's definition, for information to be classified as personal data, it must pertain to an identified or identifiable natural person.[6] The regulation, however, does not provide a clear definition of when a person should be considered as identified. The CJEU, the ultimate authority on the interpretation of EU laws, has not thoroughly dissected the identified criterion, thereby leaving its definition shrouded in uncertainty.

This article aims to delve into this complex issue, exploring the different interpretations of 'identified' and the ongoing debate surrounding its definition. We will examine the CJEU's stance regarding identification, the Working Party's opinion on the matter, and the impact of these interpretations on the broader understanding of personal data.

Ultimately, this article argues that the concept of identification should be understood as singling a person out of a wider group as it aligns more closely with the purpose of the GDPR.

2 Charter of Fundamental Rights and Freedoms of the European Union [2012] OJ C 326/02, art 8; However, the fundamental right is not an absolute one, see C-184/20 para 70.
3 Treaty of Lisbon Amending the Treaty on European Union and the Treaty Establishing the European Community [2007] OJ C 306/01, art 16 b.
4 European Commission, 'European Commission sets out strategy to strengthen EU data protection rules' (IP/10/1462, 2010).
5 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 [1995] OJ L 281/31 (Data Protection Directive), art 2(a).
6 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, art 4.

2. BACKGROUND

2.1. Brief History of the Concept of Personal Data in EU legislation

The concept of personal data has served as a cornerstone in the EU legislation for over two decades. However, the genesis of this concept can be traced back to the Council of Europe's Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data, commonly referred to as Convention 108. The Convention, which was opened for signature in 1981, holds the distinction of being the first legally binding international instrument in the data protection domain.[7] It defined personal data as "any information relating to an identified or identifiable individual".[8]

Subsequently, this definition was adopted and integrated into the DPD in 1995. The directive elucidated the concept of personal data as

> "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."[9]

GDPR, which came into effect in 2018, essentially retained the definition established by the DPD. Nevertheless, the GDPR expanded on the concept by elaborating the parameters of identifiability. It asserts that an identifiable natural person is one who can be identified, directly or indirectly, by reference to an identifier such as a name, an identification number, location data, an online identifier, or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural, or social identity of that natural person.[10]

Although the definition has evolved, it has done so by expanding its scope. Consequently, the jurisprudence of CJEU developed during the DPD era remains pertinent and applicable under the current regulatory framework.[11]

7 Greenleaf, Graham: The influence of European data privacy standards outside Europe: implications for globalization of Convention 108. International Data Privacy Law 2012, vol. 2(2), p. 68.
8 Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (adopted 28 January 1981, entered into force 1 October 1985) ETS No 108, art 2(a).
9 Data Protection Directive, art 2(a).
10 General Data Protection Regulation, art 4.
11 Case C-40/17 Fashion ID [19 December 2018] ECLI:EU:C:2018:1039, Opinion of AG Bobek, para 87.

2.2. GDPR's Article 4 and its Interpretative Recitals

The preeminent legal text pertaining to the definition of personal data is of course the GDPR, which delineates the parameters for the majority of personal data processing activities.[12] As explained previously, the definition of personal data is in effect the same as it was for over a decade before the enactment of the GDPR. Beyond article 4, which explicitly defines personal data, the GDPR also incorporates a series of recitals or preambulatory clauses. While these recitals are not endowed with legal enforceability, they are instrumental in elucidating the legislative intent and in interpreting the substantive provisions of the regulation.

The recitals of the regulation articulate that the processing of personal data should be designed to serve mankind and that the right to the protection of personal data should be balanced against other fundamental rights in accordance with the principle of proportionality.[13] The legislator further acknowledges the exponential escalation in data collection and sharing, attributable to rapid technological advancements and advocates for the facilitation of free data flow both within the Union and beyond its borders, while concurrently upholding a stringent standard of personal data protection.[14]

In terms of determining when a natural person could be considered identifiable, the legislator declares that account should be taken of all the means reasonably likely to be used, such as singling out, to identify the natural person directly or indirectly. However, to understand whether identification would be reasonably likely, one should also take into account all objective factors, such as the costs of and amount of time required for the identification.[15] This nuanced perspective suggests that data may be categorized as 'personal' for one entity while remaining 'anonymous' for another, contingent upon these objective factors. It is noteworthy that 'singling out,' mentioned in the recitals and to be further dissected in subsequent chapters, is identified as one among various means of identification.

12 While the GDPR can be seen as primary instrument, it is not the sole regulatory framework governing personal data processing activities, see for example Regulation (EU) 2018/1725 of the European Parliament and of the Council of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC (Regulation 2018/1725) [2018] OJ L295/39.
13 General Data Protection Regulation, recital 4.
14 General Data Protection Regulation, recital 6.
15 General Data Protection Regulation, recital 26.

2.3. The Four Quintessential Elements of Personal Data

In 2007, under the now-superseded DPD, the Working Party established by Article 29 of the Directive (WP29) issued an opinion elucidating the concept of personal data.[16] The opinion dissected the definition into four principal elements: 'any information, 'relating to', 'identified or identifiable', and 'natural person'.[17] Although this opinion is not legally binding - as the CJEU holds exclusive authority to interpret EU legislation and thus determine the ultimate understanding of personal data - it has, in practice, provided the most comprehensive description of personal data outside of case law.[18] The division of the definition of personal data into the aforementioned elements has also become an established practice in legal literature for addressing the concept of personal data.[19]

2.4. The Uncertainty of the Third Element

The three words of the third element, 'identified or identifiable', ostensibly appear unambiguous. However, considerable uncertainty prevails concerning the criteria under which an individual may be deemed identified. The GDPR refrains from providing an explicit definition of when an individual is identified, although it enumerates a non-exhaustive list of data points – or identifiers – which may culminate in an individual being classified as identified or identifiable. The text of the regulation therefore leaves room for interpretation, something legal scholars excel at.

In practice, when talking about whether data is personal or not, the 'identifiable' facet of the third element frequently presents a lower threshold for classification, and thus garners more analytical

16 Article 29 Working Party, 'Opinion 4/2007 on the concept of personal data' (WP 136, 2007).
17 Ibid, p. 6; The opinion can still be considered valid as the definition of personal data has not substantially changed between the Data Protection Directive and General Data Protection Regulation.
18 Purtova, Nadezhda: The law of everything. Broad concept of personal data and future of EU data protection law. Law, Innovation and Technology 2018, vol. 10(1), p. 43. (Purtova 2018); The opinion has also been cited in numerous advocate general opionons i.a. C-245/20, C-40/17, C-131/12.
19 See for example Borgesius, Frederik: Singling out people without knowing their names – Behavioural targeting, pseudonymous data, and the new Data Protection Regulation. Computer Law & Security Review 2016, vol. 32(2), p. 256–271. (Borgesius 2016); Canneyt, Tim and others: Data Protection: CJEU case law review – 1995–2020. Computerrecht 2021, vol. 56, p. 78–144; Davis, Peter: Facial Detection and Smart Billboards: Analysing the 'Identified' Criterion of Personal Data in the GDPR. University of Oslo Faculty of Law Research Paper No. 2020-01, <https://ssrn.com/abstract=3523109> accessed 01 August 2023; Purtova 2018.

attention from legal scholars compared to the 'identified' aspect.[20] The term 'identifiable' suggests that while identification may not have been actualized, it is reasonably plausible in the foreseeable future. Prima facie, concentrating on the lower threshold of potential identifiability might be pragmatically sensible. However, as elucidated by Purtova and Davis in their scholarly work, comprehending the potentiality of identifying an individual necessitates an understanding of the actual implications of being 'identified'.[21] Is a person considered identified when we know what he's wearing and can single him out of the surrounding context or do we need to actually know who that person is in order for him to be identified? The nuances are significant as varying interpretations of 'identified' can either expand or contract the regulatory purview of the GDPR, thereby affecting the fundamental rights of EU citizens.

3. INTERPRETATION OF 'IDENTIFIED'

3.1. Examining CJEU Stance

The CJEU has, over the course of the last two decades, grappled with the definition of the third element, 'identified', in several instances. However, the court's articulations have often been marked by brevity and an absence of comprehensive reasoning. Subsequently, we'll explore the cases in chronological sequence to discern the recurring themes and patterns in the CJEU's interpretation pertaining to the third element of personal data.

At the dawn of the new millennium, the CJEU adjudicated its first case concerning the threshold for deeming a person identified.[22] Based on the facts of the case, the court stated that a person's name in conjunction with other information such as their phone number, employment details, or hobbies was enough to identify a person.[23] However, the court abstained from elaborating on the criteria for identification, and further affirmed that a person could be identified without knowledge of their name.[24]

Nearly a decade later, the CJEU adjudicated the Scarlet Extended case, where it grappled with the question of whether an Internet Service Provider (ISP), Scarlet, could be compelled to institute a monitoring system to filter its customers' network traffic for the purpose of thwarting the illicit sharing of copyrighted content through peer-to-peer networks.[25] The court held that

20 Purtova, Nadezhda: From knowing by name to targeting: the meaning of identification under the GDPR. International Data Privacy Law 2022, vol. 12(3), p. 164. (Purtova 2022)
21 Purtova 2022 p. 164; Davis 2020 p. 14.
22 Case C-101/01 Lindqvist [6 November 2003] ECLI:EU:C:2003:596.
23 Ibid para 27.
24 Ibid.
25 Case C-70/10 Scarlet Extended [24 November 2011] ECLI:EU:C:2011:771.

the imposition of such a system would encroach upon the fundamental right to the protection of personal data, owing to the requisite extensive collection of users' IP addresses. Moreover, the court acknowledged that IP addresses qualify as protected personal data due to their potential to facilitate precise identification.[26] The CJEU did not expand what it meant by being able to identify users accurately, which is most certainly not possible if all you have is an IP-address.[27] If interpreted in a literal sense, the judgement rendered by the CJEU appears to suggest that an IP address, be it static or dynamic, in isolation, suffices to identify an individual, which might imply that the court construes identification as the ability to single out an individual from a broader group. Nevertheless, it is imperative to consider the contextual nuances, as will be elucidated in our examination of the Breyer ruling. As an Internet Service Provider (ISP), Scarlet would have had access to a repository of data regarding the allocation of IP addresses to its clientele, including timestamps of usage, thereby facilitating the correlation of an IP address with the comprehensive personal particulars of a customer.

Couple of years after the Scarlet judgement, the CJEU tackled the Ryneš case, which revolved around the recording of video footage in a partially public area.[28] The CJEU briefly stated that an image of a person captured in video footage qualifies as personal data to the extent that it can be used to identify the person.[29] The judgement by the CJEU was most certainly not drowned in thorough reasoning but luckily the opinion given by the Advocate General did shed some light on when exactly a person could be identified via video footage. Drawing upon antecedent case law, the Advocate General opined that an assemblage of recordings such as video footage could enable drawing detailed inferences regarding individuals' lifestyles, daily routines, social relationships, and more.[30] The Ryneš case therefore revolved around the indirect possibility of identifying a person based on the physical characteristics conveyed through a video footage, combined with other potentially identifiable information found in the footage.

The Breyer case served as a pivotal moment in the CJEU's interpretation of the third element of personal data, involving dynamic IP addresses, and serving to augment the insights from the Scarlet Extended case.[31] In the case, an individual by the name of Breyer frequented multiple websites under the auspices of the German state. The websites in question archived Breyer's IP

26 Ibid para 51.
27 IP addresses identify nodes, or devices used within a network but not natural persons directly. Borgesius has provided examples such as the practice of the University of Amsterdam to route all the university's online traffic through a single visible IP address and the practice of a state-owned internet service provider in Qatar to route all network traffic through a few visible IP addresses. In these cases, multiple users could be behind a single IP address, making identification impossible without additional identifiers. For further details, see Borgesius 2016, p. 264 and Zittrain, Jonathan: The Future of the Internet--And How to Stop It. New Haven 2008, p. 157.
28 Case C-212/13 Ryneš [11 December 2014] ECLI:EU:C:2014:2428.
29 Ibid para 22.
30 Case C-212/13 Ryneš [10 July 2014] ECLI:EU:C:2014:2072, Opinion of AG Jääskinen, para 33.
31 Case C-582/14 Breyer [19 October 2016] ECLI:EU:C:2016:779.

address and retained this information post-visit, ostensibly to bolster defenses against potential denial-of-service attacks.[32] Mr. Breyer sought an injunction against the German state's post-visit retention of dynamic IP addresses.[33] The CJEU's judgment contained several notable declarations. It was pronounced that static IP addresses facilitate the sustained identification of a device in connection to a network,[34] whereas a dynamic IP address, in isolation, does not constitute data appertaining to an identified natural person, as it fails to unveil the identity of the individual operating the computer.[35] The court also highlighted a distinction from the earlier Scarlet Extended case, pointing out that in the latter, an ISP, capable of linking IP addresses to individuals, was the entity collecting the data.[36]

Subsequently, the court deliberated whether a dynamic IP address could be construed as data pertaining to an identifiable individual, given that ISPs maintain logs of dynamic IP addresses assigned to particular customers, along with timestamps.[37] The court acknowledged that a person may be indirectly identifiable if a third party possesses additional information that allows for identification.[38] The crux of the matter lay in evaluating the data controller's reasonable capacity to procure ancillary information pertinent to identification from a third party. In the case at hand, the website operator, i.e., the German state, had the legal means to obtain the necessary additional information for identification from the ISP. Therefore, the court ruled that dynamic IP addresses should be considered personal data relating to an identifiable person in the given context.[39] The Breyer judgment has ostensibly been interpreted as constricting the definition of "identified", tethering the threshold for identification to an individual's civil identity.

IP addresses – both static and dynamic – are used to single out computers, or nodes, within a network.[40] The court acknowledged that dynamic IP addresses, devoid of additional context, do not in themselves reveal the identity of the natural person utilizing the node assigned the IP address.[41] The same principle extends to static IP addresses. However, the probability of identification is augmented in the context of static IP addresses due to the continuous nature of data collection, thereby affording the opportunity to accrue additional identifying data regarding

32 Ibid para 14.
33 Ibid para 17.
34 Ibid, para 36.
35 Ibid, para 38.
36 Ibid, paras 33–35; The implication of this is that an ISP would have at hand additional information which they could use to connect an IP address to a specific natural person.
37 Case C-582/14 Breyer [19 October 2016] ECLI:EU:C:2016:779, para 45.
38 Ibid, para 44.
39 Ibid, para 49.
40 Vij, Vikrant: Computer Networks. University Science Press 2018, p. 172; The difference between a static and dynamic IP address can be thought through the analogy of driving either your own car (static) or a rental one (dynamic) which you use temporarily and then return. A static IP address is like a personal car that is always available for use.
41 Case C-582/14 Breyer [19 October 2016] ECLI:EU:C:2016:779, para 38.

the user. This parallels the stance in the Ryneš case, where the Advocate General theorized the potential to draw detailed inferences and eventual identification premised upon ongoing scrutiny of video footage.[42]

Approximately one year after the Breyer decision, the CJEU found itself once again grappling with issues pertaining to personal data and identification in the case of Nowak.[43] The Nowak dispute revolved around whether an examination paper—marked with handwritten annotations from the examiner—could be classified as personal data.[44] While the CJEU primarily scrutinized whether the examiner's annotations constituted personal data vis-à-vis the examinee, it also engaged with the concept of 'identification' in assessing the exam paper's relation to an identifiable individual.[45]

Firstly, the court unequivocally asserted that an individual participating in the examination could be directly identified through a name inscribed either on the exam paper or its accompanying cover sheet or indirectly via an identification number marked on the same documents.[46] Subsequently, the CJEU posited that the examiner's ability or inability to associate the exam paper with a specific individual was immaterial. This is because the entity administering the examination inherently possessed the capability to correlate the identification number marked on the exam paper with the examinee's identity.[47] What renders the Nowak adjudication particularly intriguing is the CJEU's implicit suggestion that a numerical identifier, allocated to a natural person for the purpose of differentiation amongst a cohort of examination participants, may not necessarily serve as a direct means of identification for the individual in question. This perspective ostensibly contravenes the traditional 'singling out' paradigm of identification and instead advocates for an approach rooted in the concept of civil identity.

More recently, the year 2023 has been particularly noteworthy due to the adjudication of two cases that grapple with the concept of identification. The first case, adjudicated by the General Court, involved the European Union's Single Resolution Board (SRB) and its collection of comments from stakeholders—specifically, registered and verified shareholders and creditors—during a "right to be heard" process.[48] Each comment was tagged with a unique 33-digit identifier, thereby enabling the SRB to link each comment to a registered data subject.[49] Subsequently, a subset of these comments, stripped of the registrant information but retaining the unique identifiers, was shared with Deloitte, a third-party consulting firm.[50]

42 Case C-212/13 Ryneš [10 July 2014] ECLI:EU:C:2014:2072, Opinion of AG Jääskinen, para 33.
43 Case C-434/16 Nowak [20 December 2017] ECLI:EU:C:2017:994.
44 Ibid, para 26.
45 Ibid, paras 29–31.
46 Ibid, para 29.
47 Ibid, para 31.
48 Case T-557/20 SRB [26 April 2023] ECLI:EU:T:2023:219.
49 Ibid, paras 14–15.
50 Ibid, paras 22–24.

The crux of the question revolved around whether the data transmitted to Deloitte could be classified as anonymous, given that Deloitte lacked the capability to associate individual comments with the data subjects' registration information—a privilege solely held by the SRB.[51] The European Data Protection Supervisor (EDPS) posited that the data should be considered pseudonymous, and thus personal, as the unique 33-digit identifiers could, in theory, be linked back to the registration information maintained by the SRB.[52]

The General Court's judgment was twofold: First, it opined that the data shared with Deloitte did not pertain to 'identified' individuals, owing to the security protocols and data segregation measures implemented by the SRB.[53] The rationale is readily comprehensible, as the unique comment identifiers were designed to distinguish the comments themselves, rather than the individuals who authored them. Second, the court scrutinized the EDPS's assertion that the data related to an identifiable natural person. The court concluded that the EDPS had failed to adequately assess whether Deloitte possessed any legal avenues to access supplementary information that would enable the re-identification of the comment authors.[54] This line of reasoning was congruent with the precedent set by the Breyer case, wherein IP addresses were deemed personal data only if the website operator had a legal means to correlate the IP addresses with auxiliary information.

It merits attention that the case will be subject to appellate review by the CJEU and thus, further judicial scrutiny on the matter should be anticipated.[55]

In the recently adjudicated Pankki S case, the primary focus was a bank customer's invocation of Article 15(1) of the GDPR to gain access to his processed personal data.[56] Alongside this central issue, the court also explored a subsidiary question: whether log data, which recorded the identities of bank employees who had accessed the customer's information, should itself be categorized as the customer's personal data.[57] Upholding the expansive legal interpretation of personal data, the court concluded that the log data generated during the processing activities indeed qualifies as information relating to an identified or identifiable individual.[58] While not groundbreaking, the court's decision serves to reinforce the prevailing stance favoring a broad conceptualization of personal data.

Upon reviewing the aforementioned CJEU case law, it is evident that the court has refrained from articulating a definitive criterion for when it deems a person to have been identified. One could confidently argue that the court has internally applied either a 'singling out' approach or a 'civil identity' approach when it has reached its verdicts. Since the court has not taken an explicit stance on the matter, it is necessary to turn to other sources of legal doctrines in order to discern the conditions under which an individual may be deemed identified.

51 Ibid, para 76.
52 Ibid, para 79.
53 Ibid, para 84.
54 Ibid, para 105.
55 Case C-413/23 P – EDPS v SRB was submitted to the CJEU on 4th of August 2023.
56 Case C-579/21 Pankki S [22 June 2023] ECLI:EU:C:2023:501.
57 Ibid, para 28.
58 Ibid, para 45.

3.2. The Working Party 29 and WP 136

The Working Party 29 was an advisory body that played a crucial role in shaping data protection standards in the EU. Established under Article 29 of the DPD, the WP29 was composed of representatives from the data protection authority of each EU Member State, the European Data Protection Supervisor, and the European Commission. The WP29 was tasked with providing expert advice on data protection matters and promoting the consistent application of the DPD across all EU Member States.[59]

The WP29 was known for its influential opinions and guidelines on various aspects of data protection. These documents, while not legally binding, were highly respected data protection authorities in the EU.[60] In 2018, with the implementation of the GDPR, the WP29 was succeeded by the European Data Protection Board.

In 2007 the WP29 issued an opinion – Working Paper 136 – on the concept of personal data. This document was, and still is, the most comprehensive documentation of the different aspects of personal data.[61] According to the opinion, an individual is considered 'identified' when there are means to distinguish them from other members of a group. This does not necessarily mean knowing the individual's name or other specific details, but rather having the ability to single them out of the surrounding context.[62] For instance, in a database of employees, an individual could be identified by a unique employee number, even if their name is not known.

WP29 highlights that both identifiers and the context in which they are used are critical in determining identification. Identifiers vary widely – from unique ones like social security numbers to a blend of personal traits or actions that can single out someone within a particular context. To grasp this, picture a scenario where you need to pick out one individual from a group using a combination of identifiers. If you describe someone as wearing a black suit, it might be enough to identify them in a small classroom but not in a large hall filled with people in similar attire. The uniqueness of the combination is key and highly dependent on context.[63]

59 Data Protection Directive, art 29.
60 Purtova 2018 p. 59; Gutwirth, Serge, and Yves Poullet. "The contribution of the Article 29 working party to the construction of a harmonised European data protection system: an illustration of reflective governance?" in "Human rights in the web of governance: towards a learning-based fundamental rights policy." Bruylant, 2010, p. 283–284.
61 Purtova 2018, p. 43.
62 WP 136, p. 13–14.
63 WP 136, p. 13.

Additionally, WP29 stresses that the data controller doesn't need to possess the means of identification. If a third party has a reasonable way to identify someone, that individual is 'identifiable'.[64] For instance, if an outside entity has a record linking an employee number to a name, that employee is considered identifiable even though the data controller doesn't have that record.

The opinion of WP29 is, as mentioned, an expert opinion. It tackles the issue of defining 'identified' head on and makes the stance of WP29 clear: a person is identified when he is singled out from a larger group of persons. While influential in the realm of data protection law, some have critiqued it for seemingly expanding the definition of personal data too broadly.[65] The opinion's impact is somewhat diminished as the CJEU has not directly referenced it[66] and the successor of WP29, EDPB has not officially endorsed or adopted the opinion like it has done for other opinions and documents of WP29.[67]

4. ONGOING DEBATE

The crux of the ongoing debate on when a person is considered identified lies in two contrasting interpretations. First, there's the stance taken by WP29, where a person is regarded as identified if they can be singled out from a crowd. We'll call this the 'Singling Out' approach.

On the flip side, there's the 'Civil Identity' approach. Here, a person is only considered identified if we can pin down their real-life identity, meaning we know exactly who they are in the societal context. This interpretation is often favored by companies that stand to gain from a more restrictive definition of identification, but it also has support from other institutions.[68] The next chapters will delve into the methods and effects of these alternative identification schemes.

64 WP 136, p. 16–17.
65 Purtova 2018.
66 A search performed on 13 June 2023 with the CJEU document tracker with a text filter of "Article 29 Working Party" yielded zero results in any judgements but did yield 17 documents, all AG opinions where the working party is mentioned, see Court of Justice of the European Union, 'curia.europa.eu' (2023) <https://curia.europa.eu/juris/documents.jsf?text=%2522Article%2B29%2BWorking%2BParty%2522> accessed 13.06.2023.
67 EDPB has endorsed 16 documents of the Working Party, see European Data Protection Board, 'Endorsement 1/2018' (2018) <https://edpb.europa.eu/sites/default/files/files/news/endorsement_of_wp29_documents_en_0.pdf> accessed 13.06.2023; Though EDPB has referenced the WP 136 opinion and more specifically the singling out portion of the opinion in one of the guidelines it has adopted, see European Data Protection Board, 'Guidelines 8/2020 on the targeting of social media users' (Version 2.0, 2021) < https://edpb.europa.eu/system/files/2021-04/edpb_guidelines_082020_on_the_targeting_of_social_media_users_en.pdf> accessed 13.06.2023, para 19.
68 Institutions such as European Union Agency for Fundamental Rights and Council of Europe have advocated for a civil identity approach, see European Union Agency for Fundamental Rights and Council of Europe, 'Handbook on European data protection law' (Publications Office of the European Union, 2018), p. 89; Borgesius 2016 p. 258.

4.1. Alternative Methods and Effects of Identification

4.1.1. Civil Identity

Civil identity identification is a concept that revolves around the use of an individual's legal or civil identity for the purpose of identification.[69] The foundation of one's civil identity is typically set by governmental authorities and is characterized by documentation such as birth certificates, identity cards, passports, and social security numbers. This documentation encompasses various attributes including an individual's legal name, date of birth, and nationality, which collectively serve to uniquely identify an individual within official contexts.

The Civil Identity approach to identification implies that an individual can only be identified if their civil identity is known. Inherently, this approach restricts the scope of what is considered 'identified', as it does not recognize other potential identifiers that may not be linked to a civil identity.[70] With the civil identity approach, data is only considered personal if it is, or or can feasibly be, connected to an individual's official identity. For instance, in the realm of internet advertising, companies employing the Civil Identity approach could potentially gather and process extensive data without the constraints of the GDPR, since the data need not be connected to an individual's civil identity for the purpose of targeted advertising.[71] Although this approach may simplify the classification of data into personal and anonymous categories, it may significantly diminish the fundamental rights pertaining to the protection of personal data.[72]

It is noteworthy that the European Agency for Fundamental Rights and the Council of Europe, in their 2018 Handbook on European Data Protection Law, appeared to acknowledge the Civil Identity approach. They articulated that identification necessitates elements which characterize an individual distinctly enough to differentiate them from all others.[73]

4.1.2. Singling Out

In contrast to the Civil Identity approach, the Singling Out approach is a more nuanced, context-sensitive method of identification. Under this approach, an individual is regarded as identified if they can be distinguished or isolated from a group through certain identifiers. These identifiers

69 Leenes, Ronald: Do they Know Me? Deconstructing Identifiability. University of Ottawa Law & Technology Journal 2007, vol. 4(1–2), p. 140 (Leenes 2007); Davis 2020, p. 15.
70 For instance, a person can have multiple separate online as well as social identities that aren't necessarily tethered back to the civil identity of the person in question.
71 Leenes 2007, p. 145.
72 Although Leenes does point out that in the case of internet search engine companies, the more data a company collects, the more likely it becomes for the company to infer a users (civil) identity, see Leenes 2007, p. 144.
73 European Union Agency for Fundamental Rights and Council of Europe, 'Handbook on European data protection law' (Publications Office of the European Union, 2018), p. 89.

may be singular or a combination thereof, sufficient to isolate one individual from others in a given context. Consequently, an understanding of the context becomes almost imperative in determining whether an individual has been successfully singled out.[74]

One of the primary advantages of the Singling Out approach is its inherent flexibility. It does not rigidly adhere to any specific identity construct and adopts an agnostic stance towards identification. As such, it encompasses a wider range of applicability, which, in theory, should bolster the protection of the fundamental right to data protection. This theory derives from the notion that a broader, context-sensitive definition necessitates data controllers to critically assess whether data pertains to an identified individual. However, the lack of a concrete, easily comprehensible definition may lead to inconsistencies and errors in categorizing data. Such mis-categorization could subsequently impact the level of data protection afforded to individuals.

The drawbacks of this approach are therefore evident. The ambiguous nature of the approach complicates the categorization of data, as it requires an understanding of not only the identifiers but also the surrounding context, which may not always be evident. This can create challenges for data controllers in making accurate determinations, and inadvertently result in diminished data protection in practice. Additionally, the subjectivity involved in contextual assessment can introduce elements of uncertainty and inconsistency in the application of data protection standards.

5. THE CASE FOR SINGLING OUT AS THE PREFERRED MODE OF IDENTIFICATION

When delineating a domain-specific definition in the realm of EU jurisprudence, we generally commence with secondary domain law sources - in this instance, the GDPR and its predecessor, the DPD. However, neither of these sources elucidate when a person should be regarded as identified. The GDPR does provide indications within its recitals regarding identification methods, and it is acknowledged by the legislator that Singling Out can be a way of identification but ultimately the recital leaves the door open for other identification methods as well.[75] Given this, we turn to the case law of the CJEU, the ultimate arbitrator of EU law. While the court's rulings have been highly context-specific and sparing in their reasoning, they have proffered indications of what general guidance concerning identification could entail. The context-specificity of these

74 Though it is conceivable that identifiers could form such a unique combination that they would single out a person in any context, eliminating the need to know the context. This concept aligns closely with the Civil Identity approach.

75 General Data Protection Regulation, recital 26.

judgements tends to favor a more context-sensitive Singling Out approach over the simpler Civil Identity approach.[76] But as these clues are somewhat nebulous, a persuasive argument requires a more substantive body of evidence. Therefore, there is a need to look at soft law -instruments for further guidance. The WP29, an expert body comprised of data protection professionals, has provided explicit clarification of the approach they employ in identifying individuals.

Yet, in arriving at a conclusive determination, we must also consider the arguments of the opposition. As highlighted, neither the EDPB has officially endorsed the WP 136 opinion, nor has the CJEU directly referenced it in its rulings, barring six separate instances in AG opinions.[77] Furthermore, an argument could be made that in each relevant CJEU judgement – Lindqvist, Scarlet Extended, Ryneš, Breyer, Nowak, SRB, Pankki S – the cases have all pertained to such a combination of identifiers that they would have facilitated identification even under the Civil Identity approach. The insinuation by the CJEU in the Nowak case, regarding an assigned identification number not enabling direct identification of Mr. Nowak, bolsters the credibility of the Civil Identity approach. Thus, championing the Civil Identity methodology over the Singling Out approach is not without its merits.

Nonetheless, interactions with EU law necessitate a consideration of the law's underlying purposes and objectives, facilitating an interpretation that most effectively fulfills the Union's goals. Consequently, when two arguments appear equally compelling, the one that better effectuates EU objectives should take precedence.[78]

The preamble of the GDPR states inter alia that due to the challenges brought by the rapid technological developments and globalization along with the scale of data collection increasing, a strong data protection framework is necessary and that natural persons should have control of their own personal data.[79] Given its narrower scope, the Civil Identity approach would inevitably lead to circumstances where entities process sensitive data that pertains to an individual, but not necessarily to an individual's civil identity. For instance, advertising companies can compile a sufficiently accurate profile from an individual's browsing habits, device information, and

76 The contention here is that it would have been considerably more straightforward for the CJEU to assert in its rulings that identifying a person necessitates knowing their civil identity. Instead, the court chose a more complex path, opting for highly contextual interpretations and judgments.

77 The following AG opinions all mention "WP 136": Case C-245/20 [6 October 2021] ECLI:EU:C:2021:822, Opinion of AG Bobek; Case C-40/17 Fashion ID [19 December 2018] ECLI:EU:C:2018:1039, Opinion of AG Bobek; Case C-582/14 Breyer [12 May 2016] ECLI:EU:C:2016:339, Opinion of AG Sánchez-Bordona; Case C-141/12 YS and Others [12 December 2013] ECLI:EU:C:2013:838, Opinion of AG Sharpston; Case C-131/12 Google Spain and Google [25 June 2013] ECLI:EU:C:2013:424, Opinion of AG Jääskeläinen; Case C-70/10 Scarlet Extended [14 April 2011] ECLI:EU:C:2011:255, Opinion of AG Villalòn.

78 This is also called the "effet utile" principle; Gombos, Katalin: EU Law viewed through the eyes of a national judge 2018, p. 4.

79 General Data Protection Regulation, recitals 6-7.

other data that does not explicitly or implicitly reveal the person's civil identity. If identification were tethered to a Civil Identity threshold, this would not fulfill the objectives and purpose of the GDPR. Moreover, it would dilute the fundamental right to personal data protection as the definition would exclude data that pertains to an individual but not to their civil identity. The conclusion, therefore, is that construing identification as knowing a person's civil identity would contravene the very purpose of the regulation. By negation, we are thus left with the Singling Out approach, with its more comprehensive scope of applicability.

6. CONCLUSIONS AND IMPLICATIONS

The discourse surrounding data protection law, particularly in the context of identification, is intricate and multifaceted. Two primary approaches have emerged: the Civil Identity approach and the Singling Out approach. This article posits that identification should be understood as the act of singling out an individual from their surrounding context. While this perspective doesn't revolutionize the field of data protection law, it offers a fresh lens through which to understand the concept of identification.

Previously, both Purtova and Davis have argued for the Singling Out approach, albeit through different interpretations of the CJEU case law. This article, however, doesn't aim to reinterpret existing case law in the same way. Instead, it asserts that due to the inconclusiveness of the court's earlier judgments, we need to examine the reasoning from a new, teleological perspective. This viewpoint doesn't necessarily contradict the interpretations of Purtova or Davis; rather, it adds another, previously unexplored, layer to the debate.

Looking ahead, the CJEU will have an opportunity to provide more nuanced argumentation when it deliberates on the upcoming cases EDPS v SRB (C-413/23 P) and IAB Europe (C-604/22). Both cases concern the definition of personal data and its interpretation, offering a chance for the court to further clarify this critical aspect of data protection law.