EJPLT

# Non-personal data processing – why should we take it personally?

NICOLETA CHERCIU
Research Fellow on the Regulation of Robotics and AI at Sant'Anna School of Advanced Studies of Pisa
Attorney at Law
TEODOR CHIRVASE
LL.M. in Law & Technology at Tilburg University
Legal Counsel

European Journal of Privacy Law & Technologies • 2020/2 • ISSN 2704-8012

**Abstract**

*Data processing technologies have become more pervasive than ever in recent years, making EU citizens subject to profiling and automated decisions. This article aims to address the issues related to the widely ranging interpretations of the notion of personal data, and to evaluate the extent to which the existing regulatory framework in the EU provides adequate safeguards to the protection of individuals' rights and freedoms in the context of both personal and non-personal data processing operations, identifying the key gaps in the current legal framework that the European Commission and other relevant bodies will need to address to adequately protect EU consumers.*

**Keywords:** personal data - non-personal data - GDPR - consumer profiling.

**Summary:** Introduction. – 1. The notion of personal data. Uncertainty and fragmentation. – 2. Profiling based on non-personal data. Transparency and accountability gap. – Conclusions.

## Introduction.

'Data is the lifeblood of the economy'[1] and it is a very important asset for the future economic and strategic development of the EU market and economy. However, the use of data, when not properly scrutinized and regulated can also have a negative impact on EU citizens' rights and freedoms, as data is being intensively used for profiling, microtargeting, analyzing and influencing the behavior of individuals. An accurate and concise legal framework governing the collection and use of data is thus essential for a democratic and prosperous society. In the EU, data is divided into three categories, namely personal data, anonymized data, and non-personal data. The latter is governed and subject to Regulation (EU) 2018/1807 on the free flow of non-personal data.[2] The first two are regulated by the General Data Protection Regulation ('GDPR').[3] Two years after its enactment, the GDPR was recently subject to the European Commission's evaluation report ('EC GDPR Report'), which found, among others, that although it may be too early to draw definite conclusions

[1] European Commission, *Communication from the Commission. A European strategy for data. COM(2020) 66 final*, 2020) 2.
[2] Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union [2018] OJ L 303/59.
[3] 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 L 119/1.

on its success, the GDPR is an important and adequate regulatory tool for the creation of a level playing field for companies operating in the EU and for the strengthening of individuals' right to personal data protection. However, the EC GDPR Report has also identified inconsistencies in the implementation and application of the GDPR, generated mainly by the Member States' use of facultative specification clauses,[4] which may lead to fragmentation and therefore to discrimination with regards to the degree in which citizens of different EU Member States are able to enjoy their rights and freedoms in the digital single market. This paper aims to complement the EC GDPR Report and contribute to the debate on the efficiency of the extant data processing legal framework on safeguarding EU citizens' rights.

To this end, we will first show that the level playing field and protection of EU citizens fundamental rights are hindered by the uncertainty and fragmentation created by the lack of uniformity with respect to the interpretation of the notion of personal data and by the difficulty to distinguish in practice between the three categories of data (personal, anonymized and non-personal). Second, we will show that, although the strengthening of individuals' right to personal data protection has, to a certain extent, been achieved by the GDPR, the emphasis on the protection of personal data unjustly omits to tackle the risks posed by the processing of non-personal data, which mainly arise from profiling based on said data, which can in turn have an impact on individuals rights to self-determination, equal treatment and privacy. Given the risks posed by the processing of non-personal data, the current legal framework focusing exclusively on the protection of personal data and the 'all or nothing' solution enshrined in the GDPR seem to be no longer able to adequately address the risks arising from non-personal data processing. Finally, we propose some remedies to the current accountability and transparency gap related to non-personal data processing and recommend the implementation of procedural and substantive safeguards at EU level for to the processing of such data.

## 1. The notion of personal data. Uncertainty and fragmentation.

### 1.1 The notion of personal data.

*Personal data* is defined under art. 4(1) GDPR as 'any information relating to an identified or identifiable natural person'. Per the same, 'an identifiable natural person is one who can be identified, directly or indirectly, in particular 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'.

As per Recital 26 of the GDPR, *anonymized data* is 'information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable' by taking into account all the means reasonably likely to be used for identification, based on the following exemplificative and objective factors: the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments.[5]

*Non personal data*, and more broadly data, is defined under art. 3 (1) of Regulation (EU)

[4] European Commission, *Communication from the Commission. Data protection as a pillar of citizens' empowerment and the EU's approach to the digital transition - two years of application of the General Data Protection Regulation. COM(2020) 264 final*, 2020) 7.

[5] Recital 26 of the 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 L 119/1.

2018/1807 on the free-flow of non-personal data,[6] as any data other than personal data as defined by the GDPR.

The GDPR and the safeguards provided therein apply only when *personal data* is processed.[7] Conversely, the GDPR does not apply to the processing of anonymized or non-personal data.[8] However, the line between these three categories of data is blurred, and it is often very difficult in practice to draw a clear-cut distinction between them, because (i) the criteria for classifying a set of data as personal or non-personal (ie. content, purpose or element of processing),[9] may be applied differently or ignored altogether by data protection authorities ('DPAs'); and (ii) it is becoming increasingly easier for non-personal and anonymized data to be rendered into personal data by way of combining and/or de-anonymizing data sets.[10]

With respect to *sub (i)*, above, the GDPR definition of personal data, which in essence is similar to the one provided by Directive 95/46/EC.[11] was interpreted by the Article 29 Working Party ('WP') in its 2007 opinion and guidelines ('Guidelines').[12] The Guidelines are the most referenced document when the notion of personal data is analyzed by scholars, data protection authorities ('DPA'), national courts and the CJEU, even after the enactment of the GDPR.

As per the WP Guidelines, *personal data* is any information and statement about a person, whether it is objectively verifiable or subjective in nature, false or true (e.g. subjective opinions and assessments), deducted or inferred;[13] that either: (a) refers to a particular person or it is about that person (content),[14] (b) is likely to have an impact on a certain person's rights and interests, including a minor one (element),[15] or (c) is used or likely to be used to treat in a certain way or influence the status or behavior of an individual (purpose);[16] and the

---

[6] Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union [2018] OJ L 303/59.

[7] See in this respect art. 1(1) Recital 26 of the 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 L 119/1 whereby it is stated that the 'regulation lays down rules relating to the protection of natural persons with regard to the processing of personal data and rules relating to the free movement of personal data'.

[8] As per Recital 26 of the Recital 26 of the 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 L 119/1, "the principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes".

[9] Article 29 Data Protection Working Party, *Opinion 4/2007 on the concept of personal data* (WP 136, 2007) 10.

[10] Article 29 Data Protection Working Party, *Opinion 03/2013 on purpose limitation* (WP 203, 2013) 31.

[11] See 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 281/31. According to art. 2 (a) of Directive 95/46/EC, "'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".

[12] See Article 29 Data Protection Working Party, *Opinion 4/2007 on the concept of personal data* (2007).

[13] Ibid 6.

[14] Ibid 10.

[15] Ibid 11.

[16] The broad interpretation of the notion of personal data was upheld by CJUE in the Nowak case, where the Court held that data relates to an individual when either an element relating to content, purpose, or a result are

(d) processing of said data results in the identification of a person or in the possibility to identify a person. The WP introduced the three-pronged test (i.e. content, element, and purpose) to be applied by controllers and DPAs for determining whether data is personal or not.

Although useful, the assessment criteria proposed by the WP in the test (ie. content, element or purpose) are in themselves a source of uncertainty as to whether a set of data constitutes personal data or not, since the way in which they are formulated leaves room for subjectivity and because they are not applied consistently throughout Member States.[17]

To exemplify, in a recent analysis of digital billboards used to provide tailored advertisement by using inferences (such as gender, age, mood) drawn from anonymized data gathered by cameras equipped with analysis software, the Dutch DPA considered that, since the purpose of processing is ultimately to influence the shopping behavior of passers-by, the anonymized data should be treated as personal data and that the processing operation falls under art. 22 of the GDPR on automated individual decision-making.[18]

On the contrary, when confronted with the assessment of a similar computer vision solution, used for gathering customers' in-store statistical demographics (age, sex, mood, time spent in front of the shelve, etc.), the Bavarian DPA found the data processed to be non-personal, not by applying the three-pronged test but only by looking at the at the risk of de-anonymization. The DPA's argument relied upon the fact that the processing lasted only for 150 milliseconds, after which the images were discarded, meaning that there was a very low risk for the controller or a third party to be able to cross-reference such images with additional data so as to render a specific person identifiable.[19]

Furthermore, the diverging conclusions of the two DPAs in applying the GDPR go even further, since while the Bavarian DPA considered that a 150 milliseconds processing is secure enough to avoid the risk of de-anonymization, the Dutch DPA stated that regardless of the duration of the processing, 'one of the technological possibilities to make data identifiable is to change software settings, (sometimes also remotely) and thus the risk for identification through additional coupling of data still exists'.[20]

Since the Bavarian DPA found that the risk of de-anonymization is low, it has reached the conclusion that the data in question is non-personal, without ever applying the WP test, which, if applied (as shown in the case of the Dutch DPA) may have led to an entirely

present. The purpose element is present when data is used with the intent of evaluating or influencing the status or behavior of an individual or treating the subject in a certain way. See Judgment of 20 December 2017, Nowak, C-434/16, EU:C:2017:994, paragraphs 34-40.

[17] Article 29 Data Protection Working Party, *Opinion 4/2007 on the concept of personal data*, 6.

[18] The Dutch DPA stated that the data can be used to make automated decisions to show certain ads to certain groups of people or individual persons and the latter may be a decision which affects parties concerned to a significant extent. See in this respect Dutch Data Protection Authority AP, *Digital billboards standards framework*, (2018) available at https://autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/brief_branche_normkader_digitale_billboards.pdf (last accessed 16 th November, 2020).

[19] The computer vision solution was used for gathering customers' in-store statistical demographics collecting data such as age, sex, mood, time spent in front of the shelve, etc. Said solution is a technology that enables real-time analysis of customers' faces with regards to their gender, estimated age (range), and emotional state (eg. happy, sad, surprised, angry). The original pictures are neither stored permanently nor processed in the cloud, rather only a transient copy is kept in the camera system memory. The camera software detects the characteristics (metadata) of the person and generates a hash value of the metadata. This process takes fractions of a second. See in this respect Damian George, Kento Reutimann and Aurelia Tamò-Larrieux, 'GDPR Bypass by Design? Transient Processing of Data Under the GDPR' (2019) 9 International Data Privacy Law 285

[20] See in this respect Dutch Data Protection Authority AP, *Digital billboards standards framework,* available at https://autoriteitpersoonsgegevens.nl/sites/default/files/atoms/files/brief_branche_normkader_digitale_billboards.pdf (last accessed 16th November, 2020).

opposite conclusion.

With respect to *sub (ii)* above, the risk of de-anonymization has been emphasized many times.[21] This risk emerges in the context of big data analytics and powerful AI algorithms, which are able to process, mix and merge different unstructured sets of raw and anonymized data and create second or third generation of data that reveal intimate insights and predictions about individuals (such as political affiliations, sex orientation and purchasing behaviors). The personal patterns and correlations inferred by the algorithms may even be unknown and unpredictable to both the controllers and the data subjects at the time of collection, since these algorithms are characterized most of the time by opacity and self-learning capabilities ('black-box algorithms'). Thus, anonymized data that is linked with additional anonymized data can become personal data.[22]

The two use-cases above show that when deciding on the applicability of the GDPR based on the notions of personal and/or anonymized data, national DPAs need to determine whether a certain data set presents the risk of de-anonymization based on the state of the art. Although the state of the art is an objective criterion, as seen above, it may still be applied differently. Furthermore, even if two DPAs reach the same result in the sense that a data set must be considered anonymized, the final result of the analysis may still be different depending on whether the DPAs apply the WP test or not.

Ultimately, the notion of personal data is always dependent upon 'all the circumstances of the case' and data can be considered anonymized or not only by reference to state of the art anonymization techniques[23] (which is a poor standard considering that de-anonymization techniques evolve at the same pace with data processing methods). While this approach is meant to render the GDPR as technologically neutral and future-proof as possible, the constantly morphing notions of personal data and non-personal data and their different interpretation and application create a certain level of uncertainty, which in turn, raises important concerns, as shown below.

---

[21] Paul Ohm, 'Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization' (2010) 57 UCLA Law Review 1701, Article 29 Data Protection Working Party, *Opinion 03/2013 on purpose limitation* 31. Gloria González Fuster and Amandine Scherrer, *Big Data and smart devices and their impact on privacy*, 2015) 18. European Data Protection Supervisor, *Opinion 3/2020 on the European strategy for data*, 2020) 8.

[22] For example, in 2006 Netflix released "100 million supposedly anonymized movie ratings. Each included a unique subscriber ID, the movie title, year of release and the date on which the subscriber rated the movie. Contestants were asked to develop an algorithm that was 10% better than Netflix' existing in predicting how subscribers rated other movies. Just 16 days later, two University of Texas researchers announced that they had identified some of the Netflix users in the data set". See in this respect Julianne Pepitone, *5 data breaches: From embarrassing to deadly* (CNN 2010) (last accessed 16th November 2020).
Also, as exemplified in John Bohannon, 'Credit card study blows holes in anonymity' (2015) 347 Science 468, a team of engineers analyzed 3 months of credit card transactions, chronicling the spending of 1.1 million people in 10,000 shops in a single country. The data was obtained from a major bank. Although the bank stripped away names, credit card numbers, shop addresses, and even the exact times of the transactions and all that remained were the metadata: amounts spent, shop type—restaurant, gym, or grocery store, for example—and a code representing each person, by simulating a correlation attack on the credit card metadata and by feeding the algorithm with a collection of random observations about each individual in the data, the algorithm generated certain clues. The computer used those clues to identify some of the anonymous spenders. The researchers then fed a different piece of outside information into the algorithm and tried again, and so on until every person was de-anonymized.

[23] See Inge Graef, Raphael Gellert and Martin Husovec, *Towards a Holistic Regulatory Approach for the European Data Economy: Why the Illusive Notion of Non-Personal Data is Counterproductive to Data Innovation* (TILEC Discussion Paper No 2018-029, 2018) 8. European Commission, *Communication from the Commission. A European strategy for data. COM(2020) 66 final*, 6.

### 1.2 Uncertainty and fragmentation.

The uncertainty surrounding the limits of the notions of personal data and anonymized data may hinder achieving the very aims of the GDPR, namely those of ensuring economic and social progress and of strengthening the convergence of national economies within the internal market, while preserving the well-being of natural persons.[24]

Firstly, different interpretations of the notion of personal data create market fragmentation.[25] While in certain Member States some technologies will be GDPR compliant and thus ready to market, in others they will not. If data processed by such technologies is to be considered personal, the processing operation must be based on a lawful ground,[26] such as consent, which is most likely impossible to obtain for the processing carried out by certain technologies (eg. solutions gathering in-store customers demographics, like those discussed above). Conversely, an easier implementation of such technologies will be possible in Member States where such data is not considered personal data and consent is not required.

Moreover, said differences may impact the implementation of the principles of privacy by design and privacy by default, forcing manufacturers to comply with different standards across the EU.[27]

Further to the above, this uncertainty leads to companies having the space to circumvent the applicability of the GDPR through fictitious constructs and arbitrary standards as to what turns personal data into anonymized/non-personal data (e.g. by measuring the duration of the processing in milliseconds). The lack of foreseeability may also have a potential chilling effect on the willingness of businesses to use data and invest in new technologies.

Additionally, the very purpose of the Regulation (EU) 2018/1807 to ensure the free flow of data may be adversely impacted, as the latter is directly dependent upon the notion of non-personal data, which, if different across Member States, will hinder the possibility to create a uniform data trade environment in the EU. Uncertainty may also create a barrier for the free movement of services, as different interpretations will determine different prevailing technologies across Member States and will distort competition as manufacturers may be incentivized to choose to offer goods and services in some Member States to the detriment of others.

Lastly and most importantly, the same will create different levels of protection across geographies, which will in turn decrease users' trust to use certain technologies and reduce

---

[24] See Mario Viola de Azevedo Cunha, 'Review of the Data Protection Directive: Is There Need (and Room) For a New Concept of Personal Data?' in Serge Gutwirth and others (eds), *European Data Protection: In Good Health?* (Springer 2012).

[25] As stated by the EC, 'fragmentation between Member States is a major risk for the vision of a common European data space and for the further development of a genuine single market for data'. See European Commission, *Communication from the Commission. A European strategy for data. COM(2020) 66 final*, 6.

[26] See Art. 5 of the 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 L 119/1.

[27] As per art. 25 of the 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 L 119/1, the controllers must implement privacy by design and by default measures. However, art. 4 (7) of the 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 L 119/1, defines controller as 'the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data'. Thus, only when personal data is processed, these safeguards must be complied with.

their willingness to adopt and purchase new technologies.[28] For instance, although the in-store demographics technology implemented in Germany was found to be privacy compliant by the Bavarian DPA, the company had to stop using the system because in-store customers did not accept it.[29]

## 2. Profiling based on non-personal data. Transparency and accountability gap.

### 2.1 Profiling based on non-personal data. Risks.

Besides the negative social and economic impact of this current state of uncertainty surrounding the very applicability of the GDPR, we can also identify a potential regulatory gap in protecting EU citizens against the risks of profiling based on anonymized and non-personal data.

Profiling is can be defined as 'the process of 'discovering' correlations between data in databases that can be used to identify and represent a human or nonhuman subject (individual or group) and/or the application of profiles (sets of correlated data) to individuate and represent a subject or to identify a subject as a member of a group or category'.[30] Thus, profiles can be either created for an individual or for a group and can be based on any type of data, be it personal or non-personal.

The risks of profiling to the EU citizens' rights and freedoms are manifold, as acknowledged by the WP in its 2018 Guidelines on automated individual decision-making and profiling, which underlines that profiling can perpetuate existing stereotypes and social segregation, lock a person into a specific category and restrict them to their suggested preferences, and lead to inaccurate predictions, denial of services and goods and unjustified discrimination in some cases.[31] Nevertheless, these risks seem to be acknowledged only for individual profiling based on personal data, although – as we will further show below – the risks of profiling based on non-personal data or anonymized data are as significant.[32]

Thus, in the age of interconnectivity and big data, large unstructured non-personal data sets may be linked together and processed by algorithms in ways that can reveal intimate human traits, and even seemingly trivial data (when processed in the same way) may result in the creation of patterns which can be used to predict behavior and to reveal sensitive attributes such as political affiliation, sex orientation, race and purchase behaviors of

---

[28] Similarly, the EC stated that 'citizens will trust and embrace data-driven innovations only if they are confident that any personal data sharing in the EU will be subject to full compliance with the EU's strict data protection rules'. See European Commission, *Communication from the Commission. A European strategy for data. COM(2020) 66 final*, 1.

[29] See in this respect von WIRED Staff, *Real beendet die Gesichtserkennung in Supermärkten* (2017) (last accessed 16th November 2020).

[30] Mireille Hildebrandt, 'Defining Profiling: A New Type of Knowledge?' in Mireille Hildebrandt and Serge Gutwirth (eds), *Profiling the European Citizen* (Springer 2008), 34.

[31] Article 29 Data Protection Working Party, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679* (wp251rev01, 2017), 6.

[32] Profiling is defined by art. 4 (4) of the 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 L 119/1 as 'any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements'. As clarified by the WP, 'profiling is composed of three elements: it has to be an automated form of processing; it has to be carried out on personal data; and the objective of the profiling must be to evaluate personal aspects about a natural person'. See ibid, 6-7.

individuals or groups.[33]

These inferences and the knowledge obtained from such data may be clustered, aggregated and associated based on different attributes identified by the algorithm, to create certain groups. These groups, which may refer to different segments of the population, are usually not pre-established by the members of the group (such as in the case of a class of students or members of a political party) but are the creation of the algorithm.[34] In this way, the members of a certain group do not even know that they are being categorized as belonging to that group, or that they have been profiled and that their behavior has been analyzed.[35]

Moreover, these algorithms can create spurious and erred correlations and make inaccurate predictions.[36] For example, groups of persons may be profiled as having a predisposition of engaging in a specific type of dangerous activity based on aggregated and anonymous data from a search engine, even though said searches were the result of mere curiosity or popularity of a given topic at a certain moment in time.

This type of information, inferred from non-personal data, can be used in abusive ways absent any safeguards, and may affect individuals even when they are not identified or identifiable. For example, when harmful profiling targets a group, and not an individual, the risks of behavioral manipulation and social sorting arising from profiling are present for the group as a whole and for all its individual members, without ever the data controller having or wanting to know who are the members of the group.[37]

We will provide a couple of examples on the negative effects of profiling based on non-personal data applied at group-level. First, processing data about the purchasing behaviors of a country's citizens to sell lower quality goods in that territory may lead to unfair and misleading consumer practices against all the members of the group.[38] Secondly, allowing advertisers to exclude, based on ethnicity or gender, certain groups from their offerings on housing, employment, and credit history may lead to the discrimination of the natural persons clustered based on these attributes.[39] Thirdly, profiling groups based on algorithmic inferred preferences (ie. affinity profiling) may lead to their members being presented only with certain information (eg. such as in the case of social media networks news feed personalization), thus impairing their access to information. Finally, profiling can exacerbate the phenomenon of digital market manipulation whereby firms exploit consumers' biases and vulnerabilities and use these profiles to dynamically set prices (sometimes to the

---

[33] See Mireille Hildebrandt, 'Profiling: From data to knowledge' (2006) 30 Datenschutz und Datensicherheit - DuD 548 549: 'the data are recorded as a type of brute facts, decontextualized and – as far as the machine is concerned – without meaning. Most data may be trivial in themselves, acquiring new meaning after having been connected with other trivial data and used for decision making'.

[34] Ibid.

[35] Ibid, 550.

[36] Article 29 Data Protection Working Party, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, 6 and 12.

[37] "A firm does not need to know specifically who the consumer is to extract rent by exploiting her biases or creating a persuasion profile, only that the observed behavior is that of the same person buying the good or visiting the website". See Ryan Calo, 'Digital Market Manipulation' (2014) 82 George Washington Law Review , 1030.

[38] A recent investigation showed that popular multinational food and drink companies have 'cheated and misled' shoppers in eastern Europe for years by selling them inferior versions of well-known brands compared to products sold in the West. See Daniel Boffey, *Food brands 'cheat' eastern European shoppers with inferior products* (The Guardian 2017) (last accessed on 16th November, 2020).

[39] Julia Angwin and Terry Parris Jr, *Facebook Lets Advertisers Exclude Users by Race* (2016), Julia Angwin, Ariana Tobin and Madeleine Varner, *Facebook (Still) Letting Housing Advertisers Exclude Users by Race* (2017), Jeremy B. Merrill and Ariana Tobin, *Facebook Moves to Block Ad Transparency Tools — Including Ours* (2019)

consumer's detriment), draft purchase and/or website's terms and conditions, minimize perceptions of danger or risk, or otherwise attempt to extract as much profit as possible from their consumers, all of which may ultimately lead to a decrease in consumers' autonomy and freedom of choice.[40]

### 2.2 Accountability and transparency gap.

Although group profiling based on non-personal data creates substantial risks, this activity is shielded from any accountability and transparency obligations as it falls outside the scope of the GDPR and it happens in the shadows.[41]

Nevertheless, this paper argues that certain safeguards should be adopted even for the processing of non-personal data and profiling based on such data.

For example, individuals should have a right to be informed about their behavior being analyzed based on automated techniques and about the existence of automated decision-making (similarly to the right to information provided by art. 13 and art. 14 GDPR). When profiles are inaccurately created, individuals should have the right to request the rectification of the profile and of the drawn inferences (similarly to the right to rectification provided by art. 16 GDPR).

Since in practice it may be difficult to enforce such rights at an individual level, as suggested by other scholars, the right holders can be the representatives of a group and not all its individual members.[42]

Also, controllers could be made subject to the obligation of maintaining a record of processing activities under their responsibility and to the obligation of carrying out a data processing impact assessment, which shall at least provide an evaluation on the risks to the rights and freedoms of data subjects.

The above transparency and accountability measures can provide individuals with an understanding of how data is being used for their profiling and of what assumptions or inferences are currently drawn about them.

## Conclusions.

Even though the GDPR has for the most part reached its aims to enhance data subjects' control over their personal data, there is still room for improvement in what concerns the protection of EU citizens' rights and freedoms from the negative side effects of certain data processing operations.

The processing of non-personal data should be taken seriously and, sometimes, personally. Non-personal data can still be processed to influence individuals' behavior, in which case, under the current status quo, certain DPAs may consider the data as personal data given the purpose for which it is being used. Others may apply the purpose criterion lightly or omit it altogether, in which case the same processing operation may fall outside the scope of the GDPR. Furthermore, even when we leave aside the discussion of whether data is personal or not, the processing of any data, even when it does not refer to a specific

---

[40] See Calo, 'Digital Market Manipulation', 1001. Similarly, see Natali Helberger, 'Profiling and targeting consumers in the Internet of Things - A new challenge for consumer law' in Reiner Schulze and Dirk Staudenmayer (eds), *Digital Revolution: challenges for contract law in practice* (Hart Publishing 2016).

[41] See Hildebrandt, 'Profiling: From data to knowledge', 550: 'as a consequence of the focus on data instead of knowledge, the debate seems to be directed to anonymisation, or the use of pseudonyms, in order to protect personal data. However, citizens may rather need protection against the application of profiles, or at least access to such profiles and transparency concerning their use'.

[42] Linnet Taylor, Luciano Floridi and Bart van der Sloot, 'Introduction: A New Perspective on Privacy' in Linnet Taylor, Luciano Floridi and Bart van der Sloot (eds), *Group Privacy New Challenges of Data Technologies* (Springer 2017).

individual, it can still affect him or her.

This current framework creates for companies and data subjects an environment of uncertainty and of lack of sufficient safeguards. It may be that, the 'all or nothing' extant legal framework under the GDPR and the Regulation (EU) 2018/1807 is insufficient, as it does not cover data processing operations which fall under a grey area, namely processing operations where the data does not refer to a specific individual, nor it is about industrial metrics, but to humans in general. Under the current framework, compromises will always be made with respect to humans' data either by considering it as personal, and thus making it subject to the stringent regulatory framework of the GDPR although this data may not always lead to a person being identifiable, or by considering it non-personal, and thus strip its processing of any safeguards.

For these situations, this paper argues for an intermediate level of accountability and for the creation of a regulatory framework for humans' data and for anonymized data which acknowledges the current risks deriving from the processing of non-personal data.