COMPUTER LAW & SECURITY REVIEW 35 (2019) 306–329

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/CLSR

Computer Law & Security Review

Profiling tax and financial behaviour with big data under the GDPR

*Eugenia Politou, Efthimios Alepis, Constantinos Patsakis**

Department of Informatics, University of Piraeus, Greece

A B S T R A C T

Big data and machine learning algorithms have paved the way towards the bulk accumulation of tax and financial data which are exploited to either provide novel financial services to consumers or to augment authorities with automated conformance checks. In this regard, the international and EU policies toward collecting and exchanging a large amount of personal tax and financial data to facilitate innovation and to promote transparency in the financial and tax domain have been increased substantially over the last years. However, this vast collection and utilization of "big" tax and financial data raise also considerations around privacy and data protection, especially when these data are fed to clever algorithms to build detailed personal profiles or to take automated decisions which may exceptionally affect people's lives. Ultimately, these practices of profiling tax and financial behaviour provide fertile ground for discriminating processing of individuals and groups.

In light of the above, this paper aims to shed light on the following four interdependent and highly disputed areas: firstly, to review the most well-known profiling and automated decision risks emerged from big data technology and machine learning algorithmic processing as well as to analyse their impact on the tax and financial privacy rights through their immense profiling practices; secondly, to document the current EU initiatives toward financial and tax transparency, namely the AEOI, PSD2, MiFID2, and data retention policies, along with their implications for personal data protection when used for profiling and automated decision purposes; thirdly, to highlight the way forward for mitigating the risks of profiling and automated decision in the big data era and to investigate the protection of individuals against these practices in the light of the new technical and legal frameworks; in this respect, we finally delve into the regulatory EU efforts towards fairer and accountable profiling and automated decision processes, and in particular we examine the extent to which the GDPR provisions establishes a protection regime for individuals against advanced profiling techniques, enabling thus accountability and transparency.

1. Introduction

Big data analytics, used to infer and predict behaviours, trends, choices, and preferences, lie at the heart of modern algorithmic data processing and facilitate advanced profiling and automated decisions processes employed extensively by both the private and the public sectors. Although private corporations are far ahead as far as their technological means are concerned, public administrations are closely following their best practices. For instance, while private companies are now able to monitor people's consumption patterns to predict future trends and to provide personalized advertisement, in public administration the cases of profiling citizens are increasingly emerging, and automated decision algorithms are more and more employed to substitute previously human undertaken interventions and decisions.

Against this background, international and European Union (EU) policies for promoting innovation and transparency in the financial and tax domain have been increased

* Corresponding author to: Department of Informatics, University of Piraeus, Greece.
E-mail addresses: epolitou@unipi.gr (E. Politou), talepis@unipi.gr (E. Alepis), kpatsak@unipi.gr (C. Patsakis).

https://doi.org/10.1016/j.clsr.2019.01.003
0267-3649/

substantially over the last years. The Organisation for Economic Co-operation and Development (OECD) and the EU have been engaged in an unprecedented effort towards the Automatic Exchange of Information (AEOI) of tax-related data among their jurisdictions whereas the recent EU legislations on open banking and financial services are challenging the status quo of the traditional bank sector. Although ultimately the purpose of those EU initiatives is to promote financial and tax transparency as well as innovation and new individualized services by exploiting "big" tax and financial data, the vast collection of personal financial and tax-related data and their processing by clever algorithms in order to either build detailed personal profiles or to take automated decisions that may exceptionally affect people's lives, raise concerns around privacy and data protection. In a recent report[1] published by the Ethics Advisory Group, a group set up by the European Data Protection Supervisor[2] (EDPS), the interactions based on algorithmic profiling described as exacerbating information imbalances between decision-making governments and companies on the one hand and individuals on the other hand. Indeed, big data analytics along with machine learning (ML) algorithms enable, now more than ever, the extensive profiling, namely the construction of detailed personal profile of one's life ready to be used and exploited either by private firms for profit, most commonly through advertisements or domain-specific scoring systems, or by public authorities for accomplishing their duties regarding conformance audits and controls.

In this work, we delve into the technical, social and legal aspects surrounding the above described phenomena. More precisely, the contributions of this work are summarized as follows. Firstly, we review the most well-known profiling, and automated decision risks emerged from big data technology and machine learning algorithmic processing. Moreover, we analyse their impact on the tax and financial privacy rights through their immense profiling practices. Secondly, we document the current EU initiatives toward financial and tax transparency, namely the AEOI, PSD2, MiFID2, and data retention policies, along with their implications for personal data protection when used for profiling and automated decision purposes. Based on the above, we highlight possible ways to mitigate the risks of profiling and automated decision in the big data era and investigate methods for the protection of individuals against these practices in the light of the new technical and legal frameworks. Finally, we analyse the regulatory EU efforts towards fairer and accountable profiling and automated decision processes. To this end, we examine the extent to which the GDPR provisions establish a protection regime for individuals against advanced profiling techniques, enabling accountability and transparency. To the best of our knowledge, this is the first work which tries to provide a holistic picture of the problem, highlighting not only the issues that are raised but discussing the possible countermeasures and the efforts that are being made to address them in terms of legal frameworks.

In this regard, the rest of this work is structured as follows. In Section 2, after introducing the concept of profiling in the big data era, we discuss the threats of big data and algorithmic processing techniques to tax and financial privacy and data protection rights, especially when these techniques are used to profile and to make automated decisions that affect people's lives. In Section 3, following the analysis of the current EU initiatives toward financial and tax transparency, we discuss their impact on people's profiling and privacy, while in Section 4 we identify and discuss the necessary mitigation strategies and the way ahead. Then, in Section 5 we present the General Data Protection Regulation (GDPR) provisions regarding profiling and automated decision making, and we discuss the extent to which these provisions protect individuals against advanced profiling practices and discriminatory automated decisions. Moreover, we discuss implementation issues for addressing, even partially, the raised issues. Finally, Section 6 concludes the paper by highlighting the current state of affairs and the future directions on the corresponding fields of law and practice.

2. Profiling

Profiling, according to its many definitions in various business and grammar dictionaries, is "*the recording and analysis of a person's psychological and behavioural characteristics, so as to assess or predict their capabilities in a certain sphere or to assist in identifying a particular subgroup of people*"[3] or "*the act or process of extrapolating information about a person based on known traits or tendencies*".[4] In other words, profiling is all about *personalization*, which according to Cohen (2012) is the new religion of the information society whose high priests are the "*quant jocks*" of big data. Below we will shortly examine the academic literature of profiling and the impact of big data in building predictive personalized profiles.

2.1. *Profiling in the literature*

The academic literature on the definition of profiling is prolific with diverse interpretations under various technical, social and legal contexts, mainly because profiling is a highly evocative term with multiple meanings, used in both specialist and non-specialist contexts (Bosco et al., 2015). Hildebrandt defines profiling as a process of "discovering" correlations between data in databases that can be used to identify and represent a human or non-human 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

[1] https://edps.europa.eu/data-protection/our-work/publications/ethical-framework/ethics-advisory-group-report-2018_en (last access 27/12/2018).

[2] The European Data Protection Supervisor (EDPS) is an independent supervisory authority responsible for advising EU institutions on privacy related policies and legislation.

[3] https://www.igi-global.com/dictionary/profiling/23752 (last access 27/12/2018).

[4] https://www.merriam-webster.com/dictionary/profiling (last access 27/12/2018).

a member of a group or category (Hildebrandt, 2008). In other words, profiling is a technique to automatically process personal and non-personal data, aimed at developing predictive knowledge from data in the form of constructing profiles, that is discovering unexpected patterns and probabilities between data in large datasets that can subsequently be applied as a basis for decision-making (Savin, 2014). In technical terms, profiling can be understood as a specific data mining method. In this perspective, profiling is regarded as an automated process to examine large data sets to build classes or categories of characteristics. These can be used to generate profiles of individuals, groups, places, events or whatever is of interest aiming at generating prognostic information to anticipate future trends and to forecast behaviour, processes or developments (Bosco et al., 2015) as well as to assess the risks and/or opportunities of individual subjects (Hildebrandt, 2008).

Several distinctions of profiling have emerged such as the distinction between group and individual, or personalized, profiling. The first categorizes and classifies groups of people based on specific characteristics, such as the affective classifications described in our survey in which state-of-the-art mobile computation methods of classifying people to specific emotional states or personality traits are analysed (Politou et al., 2017), whereas the second mines the data of one individuated subject, such as the case of behavioural biometrics (Hildebrandt, 2008). Furthermore, Hildebrandt defines distributive profiling (as opposed to non – distributive) as the case where a group of which all members share all the attributes of the group's profile and hence the group profile can be applied without any problem to a member of the group, building thus a kind of personal profile (Hildebrandt, 2008). Likewise, she describes automated profiling, where profiles are generated and applied in the process of data mining after which human experts filter the results before making decisions, versus autonomic machine profiling, where the decisions routinely follow the machine's "advice" without requiring a human intervention (Hildebrandt, 2008).

2.2. *Profiling and big data*

Undeniably, the emergence of big data contributed greatly in the data mining techniques, especially those aiming towards profiling. As reported by the UK ICO (UK ICO, 2017), big data analytics, characterized mainly of the use of ML algorithms and huge collections of either new types of data or often repurposed, can bring benefits to business, to society and individuals as consumers and citizens. Big data analytics can also help the public sector to deliver more effective and efficient services and to produce positive outcomes that improve the quality of people's lives (UK ICO, 2017). At the same time however, big data analytics create a new digital landscape since the predictive nature of the extracted inferences as well as the complexity and the obscurity of data processing distinguish them from previous profiling solutions (Mantelero, 2016). Indeed, big data analytics, by combining algorithms and information from large and diverse datasets, create a new kind of knowledge as they locate unexpected and previously unknown structures, correlations and patterns (Hildebrandt, 2009; Pasquale, 2015). Thereby, person's online and offline activities are turned into profiling scores whereas predictive algorithms mine personal information to make guesses about individuals' likely actions and behaviours.

The risks of the profiling opportunities arisen from the big data technology are being discussed increasingly throughout the academic literature. Although Hildebrandt identified almost a decade ago the threats of profiling as commonly related to the key aspects of fundamental citizen rights, such as the rights to privacy, data protection and non-discrimination, democracy, autonomy, and self-determination, big data and algorithmic processing technologies expanded further these threats to caveats pertaining to dependence, fairness, due process, auditability, transparency, and knowledge asymmetries (Hildebrandt, 2008; Gutwirth and Hildebrandt, 2010). In this regard, many computer and human scientists assert that the discriminatory nature of ML algorithms used by big data technologies "*prioritize information in a way that emphasizes or brings attention to certain things at the expense of others*" (Diakopoulos, 2016; Crawford and Schultz, 2014). However, as private and public entities worldwide rely more and more on predictive algorithmic assessments and profiling methods to make important decisions about individuals and to steer social and technological processes, the need of dealing with these threats grows dramatically (Bosco et al., 2015; Citron and Pasquale, 2014). Therefore, we analyse below the most well-known risks arising from the big data and algorithmic processing techniques when used to profile and to make automated decisions that affect people's lives.

2.2.1. *Biased information*

While advocates of automated processing applauded the removal of human beings and their flaws from the assessment process claiming that automated systems rate all individuals in the same way, thus averting discrimination, Citron and Pasquale (Citron and Pasquale, 2014; Pasquale, 2015), among other scholars, argued that this account is misleading because when humans program predictive algorithms, their biases and values are embedded into the software's instructions. Indeed, often profiling and automated decision-making systems mine large and diverse datasets containing inaccurate and biased information to create derived or inferred data about people. But when these data are inaccurate may lead to incorrect predictions and scores about their behaviour, health, creditworthiness, or insurance risk, challenging thus the general fairness of the system. Fairness in decisions making systems has been first explored and formalized as a generalization of the notion of differential privacy (Dwork, 2006) by Dwork et al. (2012) who defined fairness as the extent to which similar individuals are treated similarly by the system. An example of such unfair processing is when algorithms place a low score on occupations like migratory work or low-paying service jobs, resulting thereby, even with no discriminatory intent, to unfairly impact consumers' loan application outcomes if a majority of those workers are racial minorities (Citron and Pasquale, 2014). Hence, it is argued that in relation to the fair processing when profiling methods are used, it is important to distinguish between the concept of unintentional discrimination as classification or prioritization of information and unfair discrimination as a conscious choice

that leads to prejudicial treatment (Barocas and Selbst, 2016; Kamarinou et al., 2016).

2.2.2. *Algorithmic opacity*

The autonomous and opaque nature of ML algorithms signifies that decisions based on their outputs may only be identified as having been discriminatory afterwards – when the impacts have already been felt by the people discriminated against (UK ICO, 2017). As Burrell describes in Burrell (2016), the opacity of ML algorithms may stem either from intentional corporate or state secrecy or from technical reasons such as technical illiteracy or the characteristics of ML algorithms and the scale required to apply them usefully. The latter is the case when in certain algorithms the number of possible features to include in a classifier rapidly grows way beyond what can be easily grasped by a reasoning human, or when machine optimizations are employed based on training data which do not naturally accord with human semantic explanations. As she explains, this is the reason why ML is applied to the kind of problems for which encoding an explicit logic of decision-making performs very poorly (Burrell, 2016).

2.2.3. *Misrepresented data*

Beyond the biases embedded into the systems, hidden bias may be produced when misrepresented data are fed into these systems, questioning thereby the general fairness of the processing (UK ICO, 2017). Yet human beings, due to the complexity of the applied algorithms, cannot always properly intervene in repairing the possible original bias occurred in the data collection phase of the decision-making process (Gutwirth and Hildebrandt, 2010). This is the case reported in ProPublica's study[5] [6] where 7000 risk scores, produced by a ML tool used in some US states to predict the future criminal behaviour of defendants, were analysed and the findings revealed discrimination based on race, with black defendants falsely classified as future criminals on nearly twice as many occasions as white defendants. Admittedly, as authors in (Houser and Sanders, 2016) note, it is difficult to challenge inaccuracies caused by misrepresented data that are used to predict an algorithmic profile because these inaccuracies are not about the individual's actual behaviour, but rather the reported behaviour that may have been, either intentionally or unintentionally, misrepresented. Such misrepresentations may further affect people's recruitment prospects in cases where big data profiling based on candidates' choice of installed browsers is used for recruitment purposes.[7]

2.2.4. *Correlation instead causation*

Big data analytics accuracy also suffers from a logical fallacy when their results offer insights into irrelevant factors. It has been thoroughly outlined that big data processes, due to the use of the employed analysis methods which are picking up things that are not there, deal with correlation rather than causality (Taylor et al., 2014; Viktor and Kenneth, 2013; Calude and Longo, 2017). Hence, the distinction between correlation and causation is very important to overcome this fallacy because correlation indicates only a probability, not a certainty. Therefore, it has been argued that organisations using ML algorithms to discover associations need to appropriately consider this distinction and the potential accuracy (or inaccuracy) of any resulting decisions (Houser and Sanders, 2016). Along the same lines, Diakopoulos (Diakopoulos, 2016) points out that "*one issue with the church of big data is its overriding faith in correlation as king. Correlations certainly do create statistical associations between data dimensions. But despite the popular adage, "correlation does not equal causation", people often misinterpret correlational associations as causal*". A good example demonstrating this fallacy is the spurious correlations project[8] where various deceptive correlations are depicted, e.g. the correlation of per capita consumption of mozzarella cheese with the civil engineering doctorates awarded.

2.2.5. *The tyranny of the minority*

Beyond the aforementioned issues, Barocas and Nissenbaum have indicated (Barocas and Nissenbaum, 2014) that profiling and automated decision-making processes based on big data suffer from the "*tyranny of the minority*" where the willingness of few individuals to disclose information about themselves may implicate others who happen to share the same group profile with them, and particularly the same observable traits that correlate with the traits disclosed. This is commonly referred to as the case of "*creditworthiness by association*" which has been recently reported in an FTC report (Federal Trade Commission, 2016). In that report, several commenters explained that some credit card companies had lowered a customer's credit limit, not based on the customer's payment history, but rather based on analysis of other customers with a poor repayment history that had shopped at the same establishments where the customer had shopped.

2.2.6. *Exposure of sensitive data*

Problems of service discrimination or exclusion are also arising when profiling reveals sensitive personal information, such as the medical condition of a user or her propensity to develop a certain disease, out of seemingly harmless information. Therefore, although the issues of discrimination are most often associated with the impacts of the processing of personal data, Hildebrandt argues that the difference between data and personal data becomes unimportant when profiling infers highly sensitive information out of seemingly trivial and/or anonymous data (Hildebrandt, 2009). As Zarsky concisely summarizes, discrimination carried out nowadays is data-driven, often does not involve intent, and is not split along the simple clear lines of the noted special categories of data (Zarsky, 2016).

[5] https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing (last access 27/12/2018).

[6] http://www.dailymail.co.uk/sciencetech/article-3606478/Is-software-used-police-identify-suspects-racist-Algorithm-used-predict-likelihood-reoffending-biased-against-black-people-investigation-claims.html (last access 27/12/2018).

[7] https://www.economist.com/blogs/economist-explains/2013/04/economist-explains-how-browser-affects-job-prospects (last access 27/12/2018).

[8] http://www.tylervigen.com/spurious-correlations (last access 27/12/2018).

2.3. *Profiling tax and financial behaviour*

In recent years, the exploitation of "big" tax and financial data, which entails the vast collection of personal financial and tax-related data and their processing by clever algorithms to either build detailed personal profiles or to take automated decisions that may significantly affect people's lives, raise also concerns around privacy, data protection and other fundamental rights. To elaborate better on these issues, we define hereafter the notions of privacy and data protection rights in the tax and financial context, and we document the harms arising from aggressive financial and tax profiling practices.

2.3.1. *The notions of tax and financial privacy*

Up to recently, tax privacy in the extensive literature[9] has been discussed as synonymous with "tax confidentiality" with references to privacy harms when tax information is disclosed to the public or sent for secondary uses to other agencies (Hatfield, 2016; Schwartz, 2008). On the other hand, financial privacy is usually taken as the right of individuals to determine what financial information about them should be known to others (Sharman, 2009) and most commonly refers to the maintenance of confidentiality of customer information about financial transactions. The right to data protection[10] on the contrary, which is enshrined in the data protection laws,[11] is not simply about the confidentiality of the data being gathered and exchanged but it gives the data subject far more extensive rights[12] (Calo, 2015; Baker, 2016). Taking into account the principles of data protection, tax and financial privacy nowadays have a broader meaning which includes the adverse privacy implications of the extended collection of information by both competent authorities and corporations to build detailed profiles of people's tax and financial behaviour either for profit or for auditing.

2.3.2. *Are tax and financial data "sensitive"?*

In terms of their sensitivity, tax and financial information are not specified in either the Data Protection Directive 95/46/EC (DPD) or the GDPR to belong to the special categories of data labelled as "sensitive"[13] requiring stricter protections than for those anticipated for other types of personal data.[14] However, as it has been already pointed out[15] (Hildebrandt, 2009), the regime for special categories of data is no longer adequate in the era of big data analytics because the practice has shown that the same data may be sensitive in one context but not in another (particularly where data are combined). Therefore, it is becoming more and more unclear whether specific categories of data are sensitive as the use of these data may or may not be sensitive depending on each context. For instance, companies and researchers use potentially "innocent" data to make sensitive distinctions between individuals (Hildebrandt, 2009; Kosinski et al., 2013; Politou et al., 2017). A notorious example is the story of a retailer shop, Target, which managed to identify pregnant customers based on their shopping habits and became a top story when predicted the pregnancy of a teenager before even her father knew about it.[16]

In the light of the above, privacy scholars have repeatedly highlighted that tax and financial data are considered to be among the most sensitive forms of personal information, as they may reveal, among others, information about income, spending and savings, employment status, person's health, marital status, lifestyle, hobbies, personal belongings, and disability status (Cockfield, 2015). This detailed and fine-grained personal information may be used to build a concrete profile of individuals' identity, including religious and political beliefs, political alliances, and personal behaviour, thereby offering an important picture of who they are (Sharman, 2009; Cockfield, 2015, 2008). To this end, and given the possible criminal nature of tax evasion in some states, Article 29 Data Protection Working Party[17] (WP29) asserted back in 2012 (Article 29 Data

[9] Historically, tax privacy related to the non-disclosure of tax information has occupied substantially authorities, society and citizens. Since the late 19th century, where tax returns were considered to be public documents deliberately disclosed in order to increase social compliance (Schwartz, 2008), until the 21th century, where tax data are combined with other financial and personal information in order to assess the potential tax evaders, tax privacy has stimulated long debates among tax and legal scholars. As Schwartz describes (Schwartz, 2008, Schwartz and Solove, 2014), in the US, due to the fact that privacy law is regulated through narrow sectoral laws, it was not until 1975 that the Congress established the principle of tax privacy in a statute for the first time. In Europe however, the rights of tax and financial privacy derive from the fundamental rights of privacy enshrined in European Convention of Human Rights (ECHR) and entered into force in 1953.

[10] The right to data protection, which is a central concept in privacy regulation around the world, is foreseen in the 1981 Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (Convention 108) and it was recognised as a fundamental, autonomous right in the Charter of Fundamental Rights of the European Union enacted by Lisbon Treaty in 2009.

[11] The current legal framework in the EU for protecting personal information is based on the GDPR which is enforced on the 25th of May 2018 and replaced the 1995 European Data Protection Directive (DPD).

[12] such as the right for personal data to be collected and exchanged only for lawful clearly identified purposes and not to be retained longer than necessary for the identified purposes.

[13] The GDPR defines as sensitive the personal data that reveal racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, the genetic data and biometric data, data concerning health or data concerning a natural person's sex life or sexual orientation.

[14] The US law, although extends heightened protection to certain data through specific laws and regulations, does not globally recognize types of data that receive heightened protection across various laws akin to EU-style "sensitive" data (Schwartz and Solove, 2014).

[15] https://iapp.org/news/a/gdpr-conundrums-processing-special-categories-of-data/ (last access 27/12/2018).

[16] https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#72a898e66686 (last access 27/12/2018).

[17] The Article 29 Working Party, set up under Article 29 of Directive 95/46/EC (DPD), is an independent European advisory body on data protection and privacy bringing together the European Union's national data protection authorities. As from 2018, under the newly adopted GDPR, the Article 29 Working Party (WP29) has been transitioned into a new legal framework, the European Data Protection Board (EDPB).

Protection Working Party) that personal data linked to tax "*may be deemed as sensitive data and therefore care should be taken to afford it higher standards of data protection*".

2.3.3. *The rights to tax and financial privacy*

While the amount of the private non-financial information, and especially sensitive, that compose the tax and financial information is extraordinary, Hatfield (2016) underscores that the identification of some details as private does not mean necessarily that it is unjust to collect them because privacy concerns do not always outweigh other factors. Instead, they should be weighed against these factors. Indeed, as other scholars explain, the right to tax and financial privacy is not absolute since an absolute right to privacy would make any modern tax system unworkable (Sharman, 2009; Avi-Yonah and Mazzoni, 2016; Schwartz, 2008). Hence, it is uncontroversial that privacy can be compromised to defend other rights or social interests. However, the presumption is that the onus is on authorities to provide a compelling reason why privacy should be compromised, rather than the onus being on citizens to show why privacy should be upheld. In the context of taxation, in particular, individuals only have procedural safeguards (e.g. notification, consultation or intervention) and not a substantive right to privacy. Yet the absence of those procedural rights might constitute an infringement of the substantive right to privacy (Avi-Yonah and Mazzoni, 2016; Baker, 2000).

2.3.4. *Privacy harms from tax and financial profiling practices*

Major technology companies and intelligent public authorities already have access to a lot of online data such as search terms, blogs and social connections as well as payment transactions and tax-related information. By accessing such fine-grained information which eventually describes a detailed people's behaviour, companies and authorities can develop highly precise individual profiles which can be later used either for influencing the consumers' choices or for predicting future legal and tax liabilities. Still, as it is highlighted by many studies (Einav and Levin, 2014; Veale et al., 2018), while many government agencies are increasingly smart about using data analytics to improve their operations and services, most agencies lag behind the best private sector firms and face challenges related to resource and infrastructure constraints as well as poor initial scoping.[18] These shortcomings, along with the fact that corporations do not have the same mandate for public accountability, led some scholars to argue that much of what private companies are best at doing would not be easily transferred to tax or public context in general (Hatfield, 2015; Diakopoulos, 2016).

Private companies, by exploiting big data and ML capabilities, can recommend customer-specific products either for increasing their returns or for replacing existing products for new ones. However, even when profiled for marketing and advertising purposes most consumers do not like the fact of being monitored or identified. According to a case reported in Reijers et al. (2016), the Dutch bank ING was planning "*to explore if customers would be interested in receiving tailored discounts from third parties in line with their spending behaviour*", an intention that raised many negative reactions from customers and media and subsequently compelled the bank not to move forward with its plans.[19] Behavioural profiling commenced by Facebook which in 2015 changed its terms of services to allow the use of its customer data for commercial purposes, such as targeted advertisement, has also provoked legal actions of German, French and Dutch authorities against the firm[20] [21] [22] (Van Alsenoy et al., 2015). Yet, profiling practices can also become even more intrusive, like in the cases where people's credit limits are being lowered based on an analysis of the poor repayment histories of other people who shopped at the same stores as them (Federal Trade Commission, 2016) or the previously described case of the retailer who managed to predict customers' pregnancy based on the consumption of just 25 products (Pasquale, 2015). Nevertheless, while commonly profiling based on consumer payment data can be used for harmless causes like marketing, personalized advertising and price discrimination, there are cases where can be used for more malicious ones like identity theft and social engineering (Reijers et al., 2016). The research literature is full of cases where online tools are using sophisticated algorithmic scoring techniques to target on consumers at moments when they are likely to be especially vulnerable to low-value, short-term credit products with usurious interest rates and highly unfavourable terms (Hurley and Adebayo, 2016).

In the tax domain, profiling refers to the categorization of taxpayers into risk profiles based on the utilization of big data technologies where a vast amount of data about them are collected through various sources (Taylor et al., 2014). In fact, tax information can be cross-indexed by the public revenue authorities against other digital personal information maintained by domestic and foreign governments (e.g., customs, criminal or immigration data) or by the private sector (e.g., records of consumer purchases) to allow for a detailed profile of an individual to be put together from formerly discrete bodies of data. This detailed profile can be used for purposes outside of traditional tax concerns such as a part of an investigation for terrorist financing schemes (Cockfield, 2008). In the US,

[18] A recent study (Veale et al., 2018) of 27 public sector ML practitioners across 5 OECD countries about the faced challenges of understanding and instilling public values into their work revealed that there is a disconnection between institutional realities and research outcomes toward transparent and non-discriminative ML systems. Researchers concluded that for transferring the values of fair and accountable ML into public sector, the respective processes should be studied in vivo, in the messy, socio-technical contexts in which they inevitably exist since issues like fairness have been shown to come with technically difficult to reconcile, or even irreconcilable, trade-offs—or concerns raised that explanation facilities might work better for some outputs than for others (Veale et al., 2018).

[19] https://www.ing.com/About-us/ING-and-the-use-of-customer-data.htm (last access 27/12/2018).

[20] https://autoriteitpersoonsgegevens.nl/en/news/dutch-data-protection-authority-facebook-violates-privacy-law (last access 27/12/2018).

[21] https://www.theguardian.com/technology/2017/may/16/facebook-facing-privacy-actions-across-europe-as-france-fines-firm-150k (last access 27/12/2018).

[22] https://www.theguardian.com/technology/2018/feb/12/facebook-personal-data-privacy-settings-ruled-illegal-german-court (last access 27/12/2018).

the Inland Revenue Service (IRS) is entitled to collect an enormous amount of private information, such as sleeping habits, individuals' hobbies, reading preferences (where and for how long a taxpayer's gaze falls on certain screens), religious affiliations, travel plans, medical conditions, weight and doctor's recommendations about it, to name a few (Hatfield, 2015; Thimmesch, 2017). Therefore, the IRS has been described by its Commissioner[23] as *"an information intensive enterprise"* which works on *"the organization of data and ultimately the knowledge and intelligence we extract from the information"*. Yet it has been argued that there are some IRS methods, mostly unknown to the general public, which violate fair information practices. According to Houser and Sanders (Houser and Sanders, 2016), the IRS is reported to have used automated computer programs (known as spiders) and big data analytics to sort through and mine social media sites[24] not only about a taxpayer who is being audited but even for potential tax violators not selected for audit.

Likewise, in the UK the integration of predictive analytics tool with big data warehouses has built the *"all-seeing eye"* of the HMRC (Her Majesty Revenue and Customs) that targets taxpayers' online information and enables drilling down into over one billion pieces of data, analysing the digital patterns of behaviour, payments and money flows of individuals and businesses.[25] HMRC tools interrogate 30 databases in total, containing not only information spontaneously available in government departments but data also found online as well, like on the Airbnb and the e-bay,[26] to apply sophisticated profiling and modelling techniques and to search for patterns and behaviours that signify tax anomalies.[27] [28] It also scrutinizes the digital footprint that people leave when they use the internet, searching social media for holidays and luxury items information which then is used to build a lifestyle profile of individuals who are under investigation for tax or benefits fraud.[29] HMRC, utilizing the well-known industry model of "understanding customers' behaviour",[30] is further requesting bulk data from third parties, like insurance companies and hospitals or payments to general practitioners and dentists, when there is evidence of widespread tax evasion or under-reporting. Within this scope of exchanging data with other agencies, the recently enacted Digital Economy Act 2017,[31] which regulates matters of information sharing between public bodies in respect of Public Service Delivery, Debt and Fraud, allows the HMRC and other public sector authorities to exchange and disclose personal data for the prevention of fraud and the recovery of debts. Nevertheless, while the Digital Economy Act 2017 maintains safeguards on privacy and anticipates various provisions and codes of practice relating to the confidentiality of personal information,[32] has also been widely criticized for excessive disclosure risks as it provides a number of exemptions that permit the disclosure of confidential information.[33] [34] [35]

3. Current EU initiatives towards tax and financial transparency

As it has been analysed in the previous section, modern methods of profiling have so far contributed considerably to the risks associated with people's discrimination and breaches of privacy. Nowadays, however, there is also an increased dominance of the principle of transparency for personal financial and tax information in the expense of privacy. In fact, following the 2008 global financial crisis, numerous international policymakers argued that the need for free and unfettered access to personal financial and tax data to combat criminals and terrorists supersedes the principle of a right to privacy[36] (Sharman, 2009). In this respect, transparency in the public domain has been presented as the solution and privacy as an obstacle to policy success (Sharman, 2009; Hatfield, 2016).

Similarly, in the private sector, privacy policies and regulations have been commonly linked to inhibiting innovation and directly affecting the economic growth and the efficacy of emerging technologies (Goldfarb and Tucker, 2012; Einav and Levin, 2014). Under this perspective, tax authorities are currently engaged in automatic ways of exchanging information for combating tax evasion and fraud, while at the same time private corporations are exploiting innovative ways to hold and capitalize on personal financial information; all within the EU regulatory framework. These EU regulatory efforts and their implications to people's privacy through big data profiling techniques will be discussed hereafter.

3.1. *EU initiatives toward big tax data exchange*

Over the last years, an increase of international policies towards the exchange of tax-relevant information between competent authorities to fight tax fraud and evasion has

[23] https://www.irs.gov/newsroom/commissioner-doug-shulman-speaks-at-aicpa-meeting (last access 27/12/2018).
[24] http://washington.cbslocal.com/2014/04/16/report-irs-data-mining-facebook-twitter-instagram-and-other-social-media-sites/ (last access 27/12/2018).
[25] https://www.computing.co.uk/ctg/feature/2244719/connecting-the-dots-at-hmrc (last access 27/12/2018).
[26] https://www.telegraph.co.uk/tax/return/taxman-unleashes-snooper-computer-information-does-have/ (last access 27/12/2018).
[27] https://perma.cc/F33W-M9FL (last access 27/12/2018).
[28] https://www.accountancylive.com/hmrcs-connect-targets-taxpayers-online-information (last access 27/12/2018).
[29] https://www.ft.com/content/0640f6ac-5ce9-11e7-9bc8-8055f264aa8b (last access 27/12/2018).
[30] https://www.capgemini.com/wp-content/uploads/2017/07/ss_hmrc_adept.pdf (last access 27/12/2018).
[31] http://www.legislation.gov.uk/ukpga/2017/30/contents/enacted (last access 27/12/2018).
[32] https://www.gov.uk/government/publications/digital-economy-act-2017-part-5-codes-of-practice (last access 27/12/2018).
[33] https://www.lexology.com/library/detail.aspx?g=f137a29a-4145-4bf9-bd49-ae9e4c77a1fc (last access 27/12/2018).
[34] https://www.twobirds.com/en/news/articles/2017/uk/very-latest-data-protection-changes (last access 27/12/2018).
[35] https://www.theguardian.com/commentisfree/2017/feb/05/the-guardian-view-on-the-digital-economy-bill-a-last-chance-to-get-it-right.
[36] This tendency is very often justified on the grounds of a "nothing to hide, nothing to fear" logic arguing that only the guilty have secrets to hide, an argument which Solove efficiently confronts in Solove (2007).

emerged. Hereafter we will analyse in brief the basic features of these initiatives as well as their impact on taxpayers' privacy and data protection rights and eventually their contribution to big data profiling through their mandate for fine-grained tax-related data retention.

3.1.1. *US FATCA*

As it is commonly acknowledged among the relevant literature, the catalyst for the worldwide expansion of the AEOI was a piece of legislation adopted in the US in 2010 called the Foreign Account Tax Compliance Act (FATCA) (Gadžo and Klemenčić, 2017; Tello, 2014; De Simone et al., 2017). As a consequence of the financial crisis of 2008 and the various banking scandals over the world, FATCA was ratified by the US Congress as part of the HIRE Act on the grounds that many Americans were holding offshore accounts that they were not disclosing to the IRS, resulting in millions of dollars in unreported income each year (Christensen and Tirard, 2016; Grinberg, 2012; Gadžo and Klemenčić, 2017). In this regard, FATCA aimed to decrease tax evasion by identifying undisclosed bank accounts of US citizens held outside the US, given the fact that Foreign Financial Institutions (FFIs), such as banks, did not have an obligation to report income earned on accounts held by US taxpayers, in contrast with the US Financial Institutions (FI) that had such obligation (Gadžo and Klemenčić, 2017; Baker, 2016; Grinberg, 2012; Christensen and Tirard, 2016; Tello, 2014).

Originally under FATCA, the FFIs must be entered into an agreement with the IRS according to which they had to report, directly to the IRS, information[37] on financial accounts of US persons and foreign entities with significant US ownership (HJI Panayi, 2016). This raised great criticism due to the fact that local FFIs would be in violation of their local data protection law if they were to disclose this information to the IRS since generally the disclosure of these reports to foreign governments is not permitted (Tello, 2014; Cockfield, 2014; HJI Panayi, 2016; Brodzka, 2013; Grinberg, 2012; Schaper, 2016). Taking into account that the choice of not complying with FATCA leads to a 30% withholding tax (HJI Panayi, 2016; Grinberg, 2012), critics pointed out that, among others,[38] FATCA is blatantly extraterritorial in application as it essentially represents an exertion of US law into the jurisdictional realm of foreign countries without their consent (Gadžo and Klemenčić, 2017). Furthermore, in terms of data privacy breaches it was pointed out by the WP29 that against this bulk transfer and screening of data, an examination of alternative, less privacy-intrusive means must be carried out to demonstrate FATCA's necessity (Article 29 Data Protection Working Party). These concerns, along with the compliance costs involved, caused intense lobbying by many countries and FIs worldwide and eventually led to the adoption in 2012 of the multilateral approach for implementing FATCA according to which inter-governmental agreements (IGAs) between the US and various foreign governments need to be signed. According to the IGAs (specifically Model 1 IGA), reporting FFIs have to gather the relevant data and provide reporting information to its own tax authority who would then oversee the automatic transmission of the data to the US IRS on an annual basis (Gadžo and Klemenčić, 2017; Baker, 2016; Christensen and Tirard, 2016; Tello, 2014; Grinberg, 2012; Brodzka, 2013). The IGAs can also be reciprocal to allow for a mutual exchange of information and to permit each country's FFIs to collect the necessary information (Christensen and Tirard, 2016; Tello, 2014). Still, the effectiveness of the FATCA legislation (Dharmapala, 2016; De Simone et al., 2017) and the legal status of the IGAs are being broadly questioned (Tello, 2014; Christians, 2013; Morse, 2013; Christians and Cockfield, 2014; Cockfield, 2014). On top, IGA implementation raised a number of serious concerns about taxpayers' privacy rights and international law violations while the "sensitive" type of information shared as well as the scale of sharing has given rise to disputes over its "fishing expedition" surveillance techniques (Cockfield, 2014; Christians and Cockfield, 2014) and resulted in US campaigns to repeal FATCA.[39]

3.1.2. *OECD CRS*

The enactment of FATCA and the consequent spurt that provoked accelerated and advanced the progress towards an automatic global tax information exchange (Tello, 2014) and resulted in the US and the EU G5 group to commit to work together towards common reporting and due diligence standards to support a global system for combating offshore tax evasion (HJI Panayi, 2016). Although, substantial amounts of some types of data have been already subject to automatic exchange for decades based on two OECD instruments[40] (Baker, 2016) (the Model Tax Convention[41] [42] and the Multi-

[37] In order FFIs to know which bank accounts are held by US customers, and hence are reportable under the FATCA regime, they are examining account information for indicia of US status including ownership of the account by a US person, or US telephone numbers or addresses associated with the account (Christensen and Tirard, 2016; Tello, 2014; Gadžo and Klemenčić, 2017; Brodzka, 2013).

[38] There has been extensive criticism also about the potential negative consequences on the US economy brought by FATCA which is discouraging investment in US assets (Gadžo and Klemenčić, 2017, Brodzka, 2013).

[39] http://repealfatca.com/ (last access 27/12/2018).

[40] Actually the OECD had long ago initiated processes for implementing full tax transparency through automatic information exchanges. Indeed, in 2005 OECD expanded the scope of the OECD Model Tax Convention amending its Article 26 to introduce the concept of "foreseeably relevant" exchanged information to replace the previous wording of "necessary" in order for the treaties to provide for exchange of information to the widest possible extent (Article 29 Data Protection Working Party). Later, in 2012, Article 26 amended again to introduce the opportunity for the competent authorities to use the received information for other purposes than tax matters, like in criminal cases, given that these purposes are allowed under the laws of both countries and the competent authority of the supplying country authorizes such use (Noseda, 2017, Van Alsenoy et al., 2015).

[41] http://www.oecd.org/ctp/treaties/model-tax-convention-on-income-and-on-capital-condensed-version-20745419.htm (last access 27/12/2018).

[42] Under the article 26 of the Model Tax Convention, besides the information exchange upon request, also the spontaneous and automatic information exchange of financial information were allowed (Article 29 Data Protection Working Party, Christians, 2013, Noseda, 2017).

lateral Convention[43] Baker, 2016; OBERSON, 2015), in 2014 the OECD published the so-called Global Standard for AEOI, the key component of which is the Common Reporting Standard (CRS) for automatic exchange of financial account information. The CRS, which is based on FATCA Model 1 IGA, constitutes the new international standard for cooperation between revenue authorities as it has been endorsed by the G20 Finance Ministers (Baker, 2016; OECD 2014; Somare and Wöhrer, 2015; OBERSON, 2015) and aims at putting an end to evasive tax practices[44] by giving governments an instrument for retrieving information on the assets their tax residents hold with FFIs, including all types of investment income[45] and account balances (Gadžo and Klemenčić, 2017; Diepvens and Debelva, 2015). However, for the automatic exchange of information between tax authorities to be activated, an agreement between those authorities was required. Hence, in 2014 a Multilateral Competent Authority Agreement (MCAA) based on Article 6 of the Multilateral Convention was signed by 51 jurisdictions who committed to having their first information exchange by September 2017 (Baker, 2016). While the OECD's work created a progressively effective technical platform and a viable legal framework for multilateral information exchange and produced substantial positive spillovers in terms of public finance and welfare (Grinberg, 2012; Marchiori and Pierrard, 2017), it was also criticized heavily due to its current legal and technical loopholes and deficiencies that may prevent its effectiveness[46] [47] [48] (Noseda, 2017; Arbex and Caetano, 2016; Diepvens and Debelva, 2015).

In terms of taxpayer protection, although the OECD was always promoting consistent principles of good tax administration as well as effective taxpayer rights[49] and obligations (Cockfield, 2008), the multilateral nature of the CRS which obliges the systematic and periodic transmission of "bulk" taxpayer information about various categories of income and wealth to be exchanged on a nearly global scale and on an automatic basis between all member jurisdictions has raised increased concerns regarding the effective protection of taxpayers' rights and especially the right to personal security regarding the exposure of charitable trusts[50] and the right to financial privacy[51] in relation to the proportionality principle (Lotmore, 2017; Avi-Yonah and Mazzoni, 2016; Noseda, 2017; Noseda, 2017). Although the Multilateral Convention makes specific reference to the protection of personal data (Articles 21, 22)[52] and specifically mentions that the use of information for purposes other than stated in the Convention could lead to a breach of privacy and a class with the protection of personal data, many transparency advocates explain[53] that tax authorities already regularly share information with colleagues in other law enforcement agencies and beyond. Indeed, according to a survey sent by the Tax Justice Network to the administrations of more than 130 jurisdictions, 83% of 30 tax authorities throughout the globe are in favour of sharing information with other local authorities to also tackle non-tax issues (Knobel, 2017).

3.1.3. *EU DAC1/2/3/4/5*

Within the EU the first legislation for facilitating the exchange of tax-related information to combat international tax evasion and avoidance was the 1977 Mutual Assistance Directive (77/799/EEC). However, the first time that the EU Member States (MS) implemented an AEOI was in 2003 through the Savings Directive (Council Directive 2003/48/EC) which anticipated the automatic information exchange for savings and interest (Christensen and Tirard, 2016; Baker, 2016; Schaper, 2016; Meinzer, 2017). Nevertheless, in 2009 under the pressure of the successive financial crises and aiming to enhance the correct assessment of taxes in cross-border situations and to fight fraud, the EU Commission put forward a proposal for administrative cooperation in the field of taxation, the Directive on Administrative Cooperation (DAC1) (2011/16/EU) which introduced broad exchange of information without prior request (Schaper, 2016). As of 1 January 2015, DAC1 provides for the exchange[54] of information on five non-financial categories of income and capital: employment income, director's fees, life insurance products, pensions, ownership and income from immovable property (Christensen and Tirard, 2016; Somare and Wöhrer, 2015; Schaper, 2016).

In April 2013, and following the developments taken under the OECD and the US, six major EU MS announced their intention to exchange FATCA type information amongst themselves[55] (Brodzka, 2013). Additionally, the EU endorsed the OECD CRS and aligned AEOI within the EU in a way that is uniform and coherent with it (Somare and Wöhrer, 2015). Accordingly, the DAC1 was amended by the new Directive on Admin-

[43] The Article 6 of the Convention on Mutual Administrative Assistance in Tax matters (Multilateral Convention) provides for the exchange of information which is "foreseeable" relevant for the tax administration or enforcement of domestic laws in any of the three forms: on request, spontaneously, and automatically (Noseda, 2017).

[44] https://www.oecd.org/g20/topics/taxation/oecd-secretary-general-tax-report-g20-finance-ministers-april-2016.pdf (last access 27/12/2018).

[45] interest, dividends, income from certain insurance contracts and other similar types of income.

[46] https://www.taxjustice.net/2016/10/25/oecd-information-exchange-dating-game/ (last access 27/12/2018).

[47] https://www.taxjustice.net/wp-content/uploads/2013/04/TJN-141124-CRS-AIE-End-of-Banking-Secrecy.pdf (last access 27/12/2018).

[48] http://www.the-best-of-both-worlds.com/support-files/26-loopholes-report.pdf (last access 27/12/2018).

[49] http://www.oecd.org/tax/forum-on-tax-administration/publications-and-products/Taxpayers'_Rights_and_Obligations-Practice_Note.pdf (last access 27/12/2018).

[50] https://www.linkedin.com/pulse/did-you-know-charities-caught-crs-before-lives-danger-filippo-noseda/?lipi=urn%3Ali%3Apage%3Ad_flagship3_profile_view_base_post_details%3BwiLdZCyyT1m9MVk%2FxWccPA%3D%3D (last access 27/12/2018).

[51] http://www.theworldin.com/article/12770/too-much-light?fsrc=scn/fb/wi/bl/ed/ (last access 27/12/2018).

[52] http://www.oecd.org/ctp/exchange-of-tax-information/keeping-it-safe.htm (last access 27/12/2018).

[53] https://financialtransparency.org/information-exchange-needs-go-beyond-tax/ (last access 27/12/2018).

[54] automatic, on request, or spontaneous.

[55] http://europa.eu/rapid/press-release_MEMO-13-533_el.htm (last access 27/12/2018).

istrative Cooperation (DAC2) (Council Directive 2014/107/EU) which extended the cooperation between tax authorities to the automatic exchange of information on additional categories of income (dividends, capital gains, account balances, etc.) held by non-residents (Christensen and Tirard, 2016; Somare and Wöhrer, 2015; OBERSON, 2015). The enactment of DAC2 was followed by a series of new amendments known as DAC3, DAC4 and DAC5 for automatically exchanging information on advance cross-border rulings and pricing agreements, country-by-country reporting and anti-money-laundering information amongst tax authorities. Similar to FATCA Model 1 IGA, under DACs the financial information is exchanged between the MSs by a two-step reporting mechanism where in the first one, the FIs must perform the due diligence rules and report to the competent authority of the state of establishment, while in the second step the information has to be transferred on an annual basis to other MS in which the data subjects to whom the data relate are residents (Schaper, 2016; Somare and Wöhrer, 2015). While DAC2 and CRS are similar in their general approaches, the US has indicated that it does not intend to adopt CRS and will continue to follow FATCA arrangements which differ from DAC/CRS in many respects (Baker, 2016). In the EU, the first AEOI in accordance with the DAC2 took place in September 2017[56] (Somare and Wöhrer, 2015).

While the systematic exchange of financial information may enhance transparency and tackle tax avoidance and evasion, the effectiveness of DAC initiatives in terms of its implementation and its consequences to taxpayers' rights has been under wide critical examination (Baker, 2016; Somare and Wöhrer, 2015). Although DAC1 Article 25 acknowledges the data protection obligations imposed by the relevant legislation, it restricts some basic subjects' rights[57] to safeguard important economic interests of the MS. These restrictions are legitimate inasmuch as are provided under appropriate legislative measures and were considered necessary and proportionate given potential revenue losses (Baker, 2016; Schaper, 2016; Somare and Wöhrer, 2015). Still, the DAC1 data protection concerns were clearly reflected in DAC2 which specifies explicitly that the reporting FIs and the competent authorities of each MS are data controllers, ensures that the data subject has a right to be informed by the reporting FI that financial account information will be collected and transferred in sufficient time to exercise his data protection rights, and finally, enforces that data shall be retained for no longer than necessary and in accordance with each data controller's domestic rules. These amendments were deemed a critical change toward enhancing data protection rights, and they are found neither in FATCA nor in CRS (Baker, 2016; Somare and Wöhrer, 2015; Schaper, 2016; Baker, 2016). However, while DAC2 brought additional data protection safeguards to the AEOI framework in taxation matters, these were not deemed good enough for the AEOI to be in line with the right to data protection[58] (Schaper, 2016; Somare and Wöhrer, 2015). Next, we will examine some of the AEOI's pitfalls in terms of its privacy and data protection implications.

3.1.4. *AEOI and data protection rights*

Despite AEOI's well-establishment as a prerequisite for effective taxation of foreign-sourced income and assets (Gadžo and Klemenčić, 2017; Johannesen and Zucman, 2014), it has been simultaneously highly criticized. On the one hand, due to its obligations of exchanging not only information relating to a single taxpayer (or a specific group of taxpayers) but of a bulk information without any indications of non-compliant behaviour of the taxpayers[59] (Baker, 2016; Somare and Wöhrer, 2015; Debelva and Mosquera, 2017; Bessard, 2017; Rocha, 2016); on the other hand, due to the removing of several existing safeguards to improve the efficiency of the exchange process (HJI Panayi, 2016; Diepvens and Debelva, 2015; Rocha, 2016; Baker and Pistone, 2016). Noseda (2017) points out that the fundamental problem with the CRS is that it was developed between 2009 and 2013 when governmental policies in collecting massive quantities of data about their citizens felt that they were justified. However, this ended with the revelations of Edward Snowden in June 2013[60] and the subsequent adoption by the EU in 2016 of the GDPR. Legal scholars raised concerns that the transmission of such detailed financial information to other countries may entail inherent legal risks associated with different law, policies and practices with respect to taxpayer rights and may lead to the information be treated in ways deemed unacceptable by certain countries[61] (Cockfield, 2008; Lotmore, 2017). Furthermore, in order AEOI to be in line with data protection regulation within the EU, data collected and transferred needs to be adequate, justified, relevant and not disproportionate whereas revenue authorities may not retain the information indefinitely, but data must be destroyed once the purpose for which they have been gathered is completed (Baker, 2016a; Baker, 2016b). These safeguards had already been highlighted in 2013 by the EDPS when DAC2 was first proposed.[62] Besides, the decision of the Court of Justice of the European Union (CJEU) in the 2014 Digital Rights

[56] DAC2 being a directive had to be transposed in national legislation by all EU MS by the end of December 2015 and its effective date set to be the 1st January 2016. The DAC2 requires for the information obtained by the authorities of the MS to be exchanged on an automatic basis annually within nine months following the end of the calendar year or another appropriate reporting period to which the information relates..

[57] the right of information about the purpose of processing, the identity of the data controller, the possible recipients of the data, and the existence of the right of access to and the right to rectify the data to the extent required.

[58] For instance, there is still no EU law obligation on a tax authority of a MS to inform a taxpayer ex ante or ex post that it exchanges tax data relating to that taxpayer with another MS. Also, Recital 9 states that MS should be prevented from engaging in "fishing expeditions", but Article 1 of DAC1/2 anticipates for the exchange of information among MS which are "foreseeably relevant" to the administration and enforcement of the domestic laws of the MS concerning the taxes.

[59] http://freedomandprosperity.org/2017/blog/new-tax-oppression-index-shows-grim-toll-of-oecds-statist-agenda/ (last access 27/12/2018).

[60] https://www.theguardian.com/world/2013/jun/06/nsa-phone-records-verizon-court-order (last access 27/12/2018).

[61] https://www.gov.uk/hmrc-internal-manuals/international-exchange-of-information/ieim406010 (last access 27/12/2018).

[62] https://edps.europa.eu/sites/edp/files/publication/13-11-05_taxation_cooperation_en.pdf (last access 27/12/2018).

Ireland Case[63] [64] to declare the blanket data collection under the EU Data Retention Directive (2006/24/EC) illegal as it violates the EU Charter of Fundamental Rights and in particular the right of privacy, imposed high uncertainties and questions respecting the principle of proportionality and specifically on whether the significant amount of personal data required to be exchanged under the DAC2 is the minimum necessary to reach the goal of fighting cross-border tax fraud and tax evasion (Noseda, 2017; Somare and Wöhrer, 2015; Baker, 2016; Avi-Yonah and Mazzoni, 2016). In 2015 under its famous *Schrems* vs *Facebook* judgment[65] the CJEU again declared a warning to EU governments against "*legislation that is not limited to what is strictly necessary where it authorises, on a generalised basis, storage of all the personal data of all the persons …. without any differentiation, limitation or exception being made*".[66] In parallel, the results of a survey conducted by Baker and Pistone in 2015 to identify minimum standards and best practices in the protection of taxpayer rights found that not much attention has been paid by countries promoting AEOI to the respective data protection issues (Baker and Pistone, 2015). Against the above background, the WP29 also issued strong concerns in its 2016 letter to the OECD[67] regarding "*repercussions on fundamental rights of mechanism entailing major data processing and exchange operations such as those envisaged by the CRS*".

Taking into account the above concerns and in the light of the then forthcoming GDPR, in 2015 the WP29 published a statement (Article 29 Data Protection Working Party, 2015) on the automatic inter-state exchanges of personal data for tax purposes in order to underline that the bilateral/multilateral agreements and the European and national laws implementing such instruments need to ensure appropriate and consistent safeguards at data protection level. Later that year the WP29 published its analytical guidelines for MS to ensure compliance with data protection requirements in the context of AEOI for tax purposes (Article 29 Data Protection Working Party, 2015). The guidelines provided a number of safeguards that should always be included in the context of the automatic exchange of personal data for tax purposes between competent authorities of different countries. Additionally, they aimed at providing indications as to the data protection safeguards to apply when personal data exchange is performed between EU MSs as well as when personal data exchange takes place between an EU MS and a third country which may or may not has been the subject of an adequacy decision by the EU Commission.

During the same time, the Commission established an expert group on Automatic Exchange of Financial Information (the AEFI group)[68] with the purpose of providing advice and ensuring that EU legislation on AEOI is effectively aligned and fully compatible with the OECD CRS on automatic exchange of financial account information. One of the first recommendations of the AEFI's group (European Commission 2015) was to highlight the need for a careful and thorough legal analysis of the compatibility of DAC2 and the rights to privacy and personal data protection.[69] Yet, it has been noted that none of the AEFI group, EDPS or WP29 recommendations have been satisfactorily implemented within the DAC2 text (Schaper, 2016; Noseda, 2017).

3.2. *EU initiatives towards big financial data exchange*

Apart from the AEOI framework, there are initiatives in the EU financial domain which may challenge individuals' data protection rights in terms of profiling and automated decisions processes. Two of the most prominent ones are the 2nd Payment Services Directive (EU) 2015/2366 (PSD2) and the 2nd Markets in Financial Instruments Directive 2014/65/EU (MiFID2) which along with the EU data retention policies, contribute to the profiling risks associated with big financial data collection and exchange.

3.2.1. *PSD2*

PSD2 is designed based on the original Payment Services Directive which introduced in 2007 to regulate payment services and payment service providers throughout the EU and European Economic Area (EEA) and thereby to create a single market for payments and to protect consumers' rights. In this regard, the PSD2 widens the scope of the first directive by covering new services and players and by identifying additional business models to encourage the development of a highly competitive market for e-payments (Giambelluca and Masi, 2016). In particular, PSD2 aims to increase the pan-European competition and participation in the payments industry also from non-financial institutions and to promote the development and use of innovative online and mobile payments, such as through open banking, while harmonizing rights and obligations for payment providers and users and making payments safer and more secure for consumers. PSD2 requires banking institutions, also known as Account Servicing Payment Service Providers (ASPSP), to open access to personal

[63] http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:62012CJ0293&from=EN (last access 27/12/2018).

[64] https://curia.europa.eu/jcms/upload/docs/application/pdf/2014-04/cp140054en.pdf (last access 27/12/2018).

[65] https://www.cnbc.com/2015/10/06/top-eu-court-backs-student-in-facebook-privacy-case.html (last access 27/12/2018).

[66] http://curia.europa.eu/juris/document/document.jsf?docid=169195&doclang=en and https://curia.europa.eu/jcms/upload/docs/application/pdf/2015-10/cp150117en.pdf (last access 27/12/2018).

[67] http://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=610127 (last access 27/12/2018).

[68] https://ec.europa.eu/taxation_customs/business/tax-cooperation-control/administrative-cooperation/commission-expert-group-automatic-exchange-financial-account-information_en (last access 27/12/2018).

[69] The group specifically identified that a legal challenge might arise from the current version of DAC2 mainly because of the magnitude of the data to be collected and reported and the fact that it does not guarantee taxpayers a permanent access to their data and a mandatory notification in case of breach. The AEFI Group expresses concerns that the information exchanged may happen to be irrelevant for taxation purposes in the receiving jurisdiction under domestic law and that reporting in such cases might be considered as being in breach of data protection law. It also identifies a risk that the validity of DAC2 might be challenged before the CJEU, due to potential violation of the proportionality principle similar to Digital Rights Ireland case.

information related to customer accounts to Third Party Payment service providers (TPPs) with which the institution has no contractual agreement. TPPs fall into one of two groups: (a) Payment Initiation Services Providers (PISPs) who initiate payments on behalf of customers and they give assurance to retailers that the money is on its way; and (b) Account Information Service Providers (AISPs) who give an overview of available accounts and balances to their customers. As these new market players need specific requirements to comply with the new obligations in PSD2, new Regulatory Technical Standards (RTS) are to be adopted by the EU on the basis of a draft submitted by the European Banking Authority (EBA). Although PSD2 came into force on 13 January 2016, anticipated a two years period for MS to incorporate the directive into their national laws and regulations. Accordingly, PSD2 entered into application on the 13th of January 2018 across all MSs. Nevertheless, the RTS[70] that define requirements for strong customer authentication and secure communication is to become applicable 18 months after its entry into force date, namely on September 2019.[71] [72]

Inevitably, PSD2 will revolutionize the payments industry affecting everything around electronic payments. For instance, it will disrupt banks' monopoly on their customers' data since it will allow businesses, such as Facebook, to retrieve individuals' account data from their bank with their permission and to make payments for their behalf. Beyond any doubt, the PSD2 legislation is an important step toward the open banking regime as facilitates data sharing across all payment's stakeholders and allows for the data-rich consumer information to pass to third parties who can use it to create new products. In this regard, banks will be required to build Application Programming Interfaces (APIs) to give TPPs secure access to their back-end data to build their own products and services around them. Profiling consumers based on their financial transactions for businesses to understand things such as customers' spending habits or credit history, is considered a new highly regarded service to which PSD2 will unavoidably contribute.

As far as data protection implications are concerned, PSD2 foresees in its text (Article 94) that any processing of personal data shall be carried out in accordance with the EU and national data protection laws. Furthermore, PSD2 provides (Article 67) that all services should be based on a user's explicit consent and they should be in accordance with data protection rules. Providers are also prohibited from requesting sensitive[73] payment data and from using, accessing or storing any data for purposes other than the provision of the account information service explicitly requested by the user (Donnelly, 2016). Following the PSD2 original proposal in 2013, the EDPS provided its opinion[74] for a number of required changes so as the directive to meet the requirements of both the DPD and the (at that time) proposed GDPR. Nevertheless, the final text of the PSD2 has given rise to speculations regarding its regulatory coexistence with the GDPR, and in particular its provisions on consent, its withdrawal and the "Right to be Forgotten", and caused concerns on its applicability (Fuster, 2016). For instance, concerns were raised regarding the cases where consumers are withdrawing consent they gave earlier, thereby requesting the removal of their personal data held by a bank or a third party, or when they are requesting removal of their personal data from a data processing or storage facility regarding specific payments. Furthermore, PSD2 Article 94(1) states that "*Member States shall permit processing of personal data by payment systems and payment service providers when necessary to safeguard the prevention, investigation and detection of payment fraud*" while the following paragraph of the Article 94(2) requires that the "*Payment service providers shall only access, process and retain personal data necessary for the provision of their payment services, with the explicit consent of the payment service user*". As Fuster in her thorough analysis explains (Fuster, 2016), "*Article 94(2) triggers a set of questions related to its consistency with the wider EU personal data protection legal framework, as well as with the very Article 94(1) that precedes it*". She further highlights that PSD2 opens the door, in the name of payment fraud prevention, investigation or detection, "*to the processing of personal data of persons completely unrelated to payment fraud, for instance through data mining or profiling techniques that would generally aim at automatically distinguishing fraudulent from non-fraudulent payments*" (Fuster, 2016). Overall, while PSD2, due to its extensive scope, is regarded by many industrial and banking stakeholders[75] as a game-changing initiative establishing a baseline for the future of banking in general, it is also being looked critically by banking industry due to, on the one hand, the low security standards the nonfinancial companies are being aligned thus far, and on the other, the data protection issues it raises by opening up banking transactions data and contributing to the building of a complete financial profile (Mansfield-Devine, 2016).

3.2.2. *MiFID2*

MiFID2 is the successor of the original Markets in Financial Instruments Directive (MiFID) 2004/39/EC which almost ten years ago led to a major shift in the cash equity markets. While MiFID intended to remove barriers to cross-border financial services within Europe for a safer, more transparent and evenly balanced marketplace, MiFID2 has an even more pronounced impact as it affects everyone engaged in the dealing and processing of financial instruments, from business and operating models to data, people and processes. MiFID2, along with its accompanying regulation MiFIR, have been applied since 3 January 2018[76] to strengthen investor protection and to improve the functioning of financial markets in a

[70] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018R0389&from=EN (last access 27/12/2018).

[71] http://europa.eu/rapid/press-release_MEMO-17-4961_en.htm (last access 27/12/2018).

[72] http://www.mondaq.com/uk/x/686420/Financial+Services/EU+Regulatory+Technical+Standards+for+Strong+Customer+Authentication+Enter+Into+Force (last access 27/12/2018).

[73] Interestingly enough, PSD2 classifies the user credentials data as "sensitive" payment data, a type of information which has not been characterized as "sensitive" under the GDPR.

[74] https://edps.europa.eu/sites/edp/files/publication/13-12-05_opinion_payments_en.pdf (last access 27/12/2018).

[75] https://blogs.sas.com/content/sascom/2017/08/18/psd2-demystifying-beast/ (last access 27/12/2018).

[76] https://ec.europa.eu/info/law/markets-financial-instruments-mifid-ii-directive-2014-65-eu_en (last access 27/12/2018).

more efficient, resilient, fairest and transparent way possible. Therefore, it has been characterized as a way to "democratize financial markets".[77] Its impact on banks, asset managers and other financial institutions is huge as the legislative framework covers basically all aspects of trading across the EU. According to its regulators, apart from protecting investors and boosting transparency, it will rebuild the trust that was tarnished by the 2008 global crash.[78]

Since MiFID2 requires early risk detection and immediate reconstruction of events when something suspicious happens, it forces the investment community to keep tabs on almost everything, requiring all communications, including personal data, which could lead to transactions to be stored for up to five years,[79] a requirement that greatly impacts personal data protection across Europe. According to the GDPR, personal data should only be kept for as long as it is necessary, but there is not a prescribed time frame in its text. Hence, if clients and employees wish to exercise their GDPR rights such as the "Right to be Forgotten" or the right to object to the processing of their personal data, the firms must carefully analyze and balance their obligations regarding their conformance to both legislations. Due to these tendencies, MiFID2 has been characterized as one of the EU's most ambitious, yet controversial, packages of financial reforms.[80]

3.2.3. *Other data retention policies*

The effective regulation of data retention policies within the EU is a controversial issue among European lawmakers, privacy advocates and legal scholars. In the UK, the Data Retention and the Investigatory Powers Act (DRIPA) commonly known as the Snoopers' Charter, allowed government bodies to continue to have access to phone and internet records of individuals following the previous repeal of these rights by the CJEU *Digital Rights Ireland* case. According to DRIPA, the telecom providers should have to keep records for at least a year of every website every citizen visits, with this information also including the apps they use on their phone and the metadata of their emails and calls,[81] to be accessed by the authorities. However, in 2016 the CJEU found that the DRIPA's powers on data retention were unlawful in all cases, except serious crimes,[82] while in 2018 the UK Appeal Court ruled this UK's mass digital surveillance regime unlawful and "inconsistent with EU law".[83] [84] [85]

The long-term storing of personal data has also been challenged by the 2016 CJEU decision in the *Tele2 Sverige* case[86] which outlawed the general and indiscriminate obligations to retain traffic and location data covering all persons, all means of electronic communication and all data without any distinctions, limitations or exceptions to combat crime.[87] This decision sought to clarify the impact of the two previous judgments on the domestic regimes covering the retention of and access to communications metadata, namely the *Digital Rights Ireland* and *Schrems* vs *Facebook* cases, and their relation to the ePrivacy Directive 2002/58 which regulates the processing of personal data and the protection of privacy in electronic communications. On a side note, the court stressed that "metadata" even though not revealing the content of the communications could be highly intrusive into the privacy of users of communications services.

The aforementioned series of the CJEU decisions against indiscriminate retention of personal data provoked the proposal[88] of the Regulation on Privacy and Electronic Communications[89] (ePrivacy regulation) aiming to replace the outdated ePrivacy Directive which up to now ensures the right to privacy with regards to communications. The proposal for the new ePrivacy regulation complements the GDPR which only applies to the processing of personal data of individuals. In addition, ePrivacy covers the business-to-business communication and the communication between individuals which may not be limited only to personal data.[90] It also introduces some radical changes[91] in the privacy of the telecommunications within Europe such as stricter retention rules and protection for metadata information. In particular, the ePrivacy proposal explicitly accepts metadata as a cause of potential privacy harm (Recital 2) and in terms of data retention recognizing the validity of targeted retention obligations, invites MS to create national data retention frameworks provided that they comply with the recent CJEU rulings.[92] Still, the WP29

[77] http://www.independent.co.uk/news/business/analysis-and-features/mifid-ii-2018-what-is-how-effect-financial-investments-markets-in-financial-instruments-directive-a8139361.html (last access 27/12/2018).

[78] https://www.bloomberg.com/news/articles/2018-01-02/no-idea-what-mifid-stands-for-here-s-what-you-need-to-know (last access 27/12/2018).

[79] https://www.euractiv.com/section/economy-jobs/opinion/gdpr-a-challenge-for-the-financial-services-industry/ (last access 27/12/2018).

[80] https://www.ft.com/content/ae935520-96ff-11e7-b83c-9588e51488a0 (last access 27/12/2018).

[81] https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/668943/Response_to_the_IPA_codes_consultation.pdf (last access 27/12/2018).

[82] https://www.theguardian.com/law/2016/dec/21/eus-highest-court-delivers-blow-to-uk-snoopers-charter (last access 27/12/2018).

[83] https://www.theguardian.com/uk-news/2018/jan/30/uk-mass-digital-surveillance-regime-ruled-unlawful-appeal-ruling-snoopers-charter?CMP=twt_gu (last access 27/12/2018).

[84] https://www.theverge.com/2018/1/30/16949520/uk-mass-surveillance-illegal-dripa-court-of-appeal (last access 27/12/2018).

[85] https://www.computerworlduk.com/security/draft-investigatory-powers-bill-what-you-need-know-3629116/ (last access 27/12/2018).

[86] http://curia.europa.eu/juris/celex.jsf?celex=62015CJ0203&lang1=en&type=TXT&ancre= (last access 27/12/2018).

[87] https://ccdcoe.org/cjeu-declares-general-data-retention-unlawful-tele2-sverige.html (last access 27/12/2018).

[88] http://europa.eu/rapid/press-release_IP-17-16_en.htm (last access 27/12/2018).

[89] http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52017PC0010&from=EN (last access 27/12/2018).

[90] http://europa.eu/rapid/press-release_MEMO-17-17_en.htm (last access 27/12/2018).

[91] https://peepbeep.wordpress.com/2017/01/12/the-proposed-eprivacy-regulation-when-the-ec-dialogues-with-the-cjeu/ (last access 27/12/2018).

[92] Paragraph 1.2: "Member States are free to keep or create national data retention frameworks that provide, inter alia, for tar-

in its opinion regarding the ePrivacy proposal (Article 29 Data Protection Working Party, 2017) sounded the alarm to the fact that the regulation suggests that non-targeted data retention measures are still acceptable. According to a Council's working paper,[93] most MSs are looking into ways to include mandatory data retention rules in the ePrivacy regulation.

Ultimately, both the relevant CJEU decisions and the ePrivacy regulation whose final text is still under negotiation.[94] [95] [96] raise questions over the definitions and distinction between targeted and untargeted data retention, a subject that touches upon the sensitive question of profiling. Because while in the context of combating criminal activities the data retention of defined user groups or populations is considered the only feasible, efficient and less infringing alternative to untargeted retention, and therefore authorities should determine which groups or areas are particularly prone to criminal connections and hence subject to data retention, this fact of targeted data retention, per se, acts as a gateway to ethnic, religious and social profiling8986[97] [98] (Pap, 2008).

3.3. *Impact on profiling and privacy*

As already mentioned, the impact on the EU policies on individuals' privacy through the facilitation of building profiles of consumers and taxpayers based on payment and tax-related data is overwhelming. Unavoidably, opening and sharing banking transactions data will provide a huge amount of payment data to companies which, by knowing the spending behaviour of individuals, they will be able not only to analyse the data and guide them to better decisions regarding their money spending[99] but also to construct a full spending and consuming profile. Since electronic payments, unlike cash, link a particular person with a particular purchase, the monitoring of consumption patterns, as well as the tracking of a person's movements becomes possible. As demonstrated earlier, people's spending patterns comprise valuable information precisely because it is possible to extrapolate inferences about the individuals in question as payment data are the exact image of their behaviour, choices and preferences, all so far considered as private (Sharman, 2009).

Under the new rules imposed by the PSD2, the ownership of these data will be essentially transferred to the consumer, meaning that account holders will be able to give companies, other than their own bank, permission to access their details. Obviously, this granting should be accomplished easily and securely. Therefore, strong customer authentications, like two-factor authentication, are specified under the RTS as a way of ensuring that data can be shared securely. Two-factor authentication specifies that, apart from using the first knowledge factor (e.g. PIN) for accessing a service, a second factor based on either possession (e.g. a token) or inherence (e.g. biometrics) is needed as well. Yet the use of inherence as the second factor to cater for both the security requirements and user experience priorities of PSPs enables the incremental collection of big biometric data that can be later used for profiling inferences. Apart from biometrics, which is already in widespread use, another important subset of inherence is behavioural profiling. By assessing the customer's location and behaviour against their usual patterns, corporations can gain a clearer view of the risks and the level of authentication required. Even though behavioural profiling is a comparatively new mechanism that is currently being used by the industry as an augmentation to strengthen fraud controls,[100] its future contribution to the construction of an integrated individual profile, when combined with other personal data, is indisputable.

The AEOI framework under which governments automatically exchange cross-border big data consisting of bulk taxpayer information to combat international tax evasion and better target audits of aggressive international tax planning (Cockfield, 2015), while being potentially revolutionary (Taylor et al., 2014) it also facilitates the construction of detailed taxpayers profiles that may be used for purposes beyond tax context. For instance, aiming at fighting offshore tax fraud, tax authorities are inclined to use phone records which may reveal whether an individual is contacting an offshore service provider based in a tax haven[101] (Cockfield, 2015). In that respect, AEOI along with national laws for data retention such as DRIPA, which is currently under revision,[102] and for data sharing such as the Digital Economy Act 2017, may impact hugely on citizens' privacy.

4. Toward mitigating risks of profiling and automated decision making

In the era of artificial intelligence and machine learning, the accountability of algorithmically automated decision systems occupies increasingly the legal and technical research community who call for automated decisions to be accountable

geted retention measures, in so far as such frameworks comply with Union law, taking into account the case-law of the Court of Justice on the interpretation of the ePrivacy Directive and the Charter of Fundamental Rights.".

[93] http://www.statewatch.org/news/2017/dec/eu-council-ms-papers-data-retention-eprivacy-reg-wk-9374-17-rev1.pdf (last access 27/12/2018).

[94] https://iapp.org/media/pdf/resource_center/ePriv-reg_03-2018.pdf (last access 27/12/2018).

[95] http://www.statewatch.org/news/2017/dec/eu-data-ret-ms-positions.htm (last access 27/12/2018).

[96] https://www.euractiv.com/section/digital/news/member-states-ask-for-new-eu-data-retention-rules/ (last access 27/12/2018).

[97] Understanding and Preventing Discriminatory Ethnic Profiling, FRA, 2010, http://www.statewatch.org/news/2010/oct/eu-fra-profiling.pdf (last access 27/12/2018).

[98] Data collection in the field of ethnicity, DG JUST, 2017, http://ec.europa.eu/newsroom/just/document.cfm?action=display&doc_id=45791 (last access 27/12/2018).

[99] https://www.theguardian.com/money/2018/jan/08/open-banking-bank (last access 27/12/2018).

[100] https://www.accenture.com/_acnmedia/PDF-40/Accenture-PSD2-Open-Banking-Security-Fraud-Impacts.pdf (last access 27/12/2018).

[101] Because residents, in order to avoid paper trails when set up offshore trusts, proceed to oral instructions regarding disbursements. (last access 27/12/2018).

[102] https://www.computerworlduk.com/security/draft-investigatory-powers-bill-what-you-need-know-3629116/ (last access 27/12/2018).

to the public and individuals to have the right to inspect, correct, and dispute inaccurate data, to know their sources, or, at the very least, to have a meaningful form of notice and a chance to challenge predictive decisions that harm their ability to obtain credit, jobs, or other important opportunities (Burrell, 2016; Citron and Pasquale, 2014; Diakopoulos, 2016; Zarsky, 2013; Richards and King, 2013; Schermer, 2011). Yet according to some scholars, detecting discriminatory decisions in hindsight is not sufficient, and hence they urge big data analysts to find ways to build discrimination detection into their systems to prevent such decisions being made in the first place. This can be achieved by introducing and implementing appropriate algorithmic tools and interventions to both identify and rectify cases of unwanted bias so as, apart from making more accurate predictions, to offer increased transparency and fairness as well (UK ICO, 2017; Goodman and Flaxman, 2016). Diakopoulos (2016) on the other hand explains that transparency, as a medium that facilitates accountability, should be demanded from the government and should be exhorted from the industry, whereas Zarsky (2013) identifies three stages in which the transparency requirements should be met: in the data collection stage, in the data analysis stage, and the final ex-post usage stage of the produced decision. While many scholars have pointed out that obviously in the collection stage legitimate arguments for some level of big data secrecy commonly related to corporate intellectual property and national security secrets may be raised (Citron and Pasquale, 2014; Zarsky, 2013; Richards and King, 2013), Richards and King (2013) identify this secrecy as a big data "*transparency paradox*" since even though big data promises to make the world more transparent, its collection is invisible, and its tools and techniques are opaque.

Pursuing transparency, Citron (2007) called more than a decade ago for a "technological due process" in the automated decisions context to underline that these decisions cannot be made within black boxes, but due processes are needed to entail limits on fine-grained personalization in a range of public administrative processes (Cohen, 2012; Hatfield, 2015). In the big data context, these "due processes" should apply to both government and corporate decisions derived from big data analytics, and when these decisions affect individuals, those people should have a right to know on what basis those decisions were made (Citron, 2007). In public administration, in particular, there are additional ethical, social and legal constraints that probably will rule out a range of "private sector-like" uses of predictive modelling, profiling and algorithmically automated decisions practices on similar targeted public services, e.g. for tax or health relief (Einav and Levin, 2014). For instance, while private companies aim to monitor, predict, and change consumer behaviour, and hence their analysis does not require legal-standard accuracy, when profiling and automated decision-making is used in the tax or other public-related context, the results must interpret the law and would need to analyse consequences within legal standards of accuracy since any errors may violate citizens legal rights. Furthermore, as Hatfield notes (Hatfield, 2015), any system to "automate" tax or other legal decision-making would be tremendously complex as it would have to reveal how the decision was made and how the legal values were interpreted and applied in a way that the taxpayer could understand and respond. In that respect, Cohen (2012) proposes the concept of "semantic discontinuity", as opposed to "seamless continuity", "*as a function of interstitial complexity within the institutional and technical frameworks that define information rights and obligations and establish protocols for information collection, storage, processing, and exchange*". She also notes that "semantic discontinuity" can be conceptualized more generally as a right to prevent precisely targeted individualization and continuous modulation, and serves similar ends as what the legislators of the GDPR are indented to deal with when introducing the "Right to be Forgotten" (Politou et al., 2018).

A well-known application of introducing transparency and accountability in public administration domain is found at the city council of New York which in 2017 introduced a bill[103] that would require the city to make public the up to then invisibly used algorithms in all kinds of government decision-making systems used for detecting and addressing financial fraud, crimes, as well as public safety and quality of life issues.[104] [105] Supported by many transparency and privacy advocates along with social and computer scientists, the bill passed, albeit amended.[106] [107] The amendment foresees for an experts task force to be created in order to review city agencies' use of algorithms and respective policies and to develop a set of recommendations on a range of issues, including which types of algorithms should be regulated, how citizens can meaningfully assess the algorithms' functions and gain an explanation of decisions that affect them personally, and how the government can address cases in which a person is harmed by algorithmic bias.[108] Although far behind what the original bill anticipated, the amendment is considered to have a significant impact on the automated decision-making by public authorities. As of May 2018, the Automated Decision Systems Task Force, the first of its kind in the US, was announced with the task to develop a process for reviewing New York City's algorithms and automated decision systems through the lens of equity, fairness and accountability.[109] While the Task Force is to produce its first report in December 2019 recommending procedures for reviewing and assessing City algorithmic tools to ensure equity and opportunity, in August 2018 experts in the field of civil rights and artificial intelligence co-signed a letter to the task force providing recommendations such as creating a publicly accessible list of all the automated

[103] https://www.nytimes.com/2017/08/24/nyregion/showing-the-algorithms-behind-new-york-city-services.html (last access 27/12/2018).
[104] http://www1.nyc.gov/site/analytics/index.page (last access 27/12/2018).
[105] https://www.oreilly.com/ideas/predictive-data-analytics-big-data-nyc (last access 27/12/2018).
[106] https://laws.council.nyc.gov/legislation/int-1696-2017/ (last access 27/12/2018).
[107] https://www.aclu.org/blog/privacy-technology/surveillance-technologies/new-york-city-takes-algorithmic-discrimination (last access 27/12/2018).
[108] https://www.newyorker.com/tech/elements/new-york-citys-bold-flawed-attempt-to-make-algorithms-accountable (last access 27/12/2018).
[109] https://www1.nyc.gov/office-of-the-mayor/news/251-18/mayor-de-blasio-first-in-nation-task-force-examine-automated-decision-systems-used-by.

decision systems in use, consulting with experts before adopting an automated decision system, and creating a permanent government body to oversee the procurement and regulation of automated decision systems.[110] [111]

Despite the above, transparency has been critically seen by many scholars as an inadequate measure for accountability of modern algorithmic systems (Ananny and Crawford, 2018; Kroll et al., 2016). Kroll et al. (2016) have thoroughly demonstrated that transparency is not enough and not even possible in automated decision-making systems based on ML algorithms. In this regard, they introduce computational methods that can provide accountability for procedural regularity even when some information is kept secret. These methods can be used alongside transparency and auditing and can be applied to all computer systems (Kroll et al., 2016). In addition, interpretability, that is providing ex-ante and ex-post explanations on the inferred decisions as a mean of accountability holds the attention of a big part of the scientific and legal community in terms of its effectiveness in the algorithmically decision supported systems and its benefits compared to its cost (Doshi-Velez et al., 2017). As Lipton analyses in Lipton (2016), the term interpretability does not refer to a monolithic concept, but it can be addressed within the context of various model properties and techniques. On the other hand, however, Hildebrandt remarks (Hildebrandt, 2017) that explanation, as a notion of interpretability, in itself does not imply justification since a decision of an automated system should be justifiable independently of how the system came to its conclusion.

Depending on the stage of which the lack of transparency and accountability may be identified in a particular algorithmic application, a different course of actions, ranging from legislative, to organizational and technical, are likely to mitigate its problems (Burrell, 2016). While the research work in the technical layer is booming as various techniques and methods for interpretable ML algorithms which provide explanations on the derived profiling classifications and decisions, have been proposed (Ribeiro et al., 2016; Lakkaraju et al., 2016; Letham et al., 2015; Wachter et al., 2018), in the legislative layer the advancements are quite reserved. It has been suggested that automated decision-making systems should be subject to licensing and audit requirements when they enter critical settings like employment, insurance, and healthcare (Citron and Pasquale, 2014; Diakopoulos, 2016), whereas other scholars proposed for an oversight board or a federal agency to ensure that algorithms produce accurate, fair and effective decisions (Houser and Sanders, 2016; Tutt, 2017). The idea of regulators to be able to test automated decision-making systems to ensure their fairness and accuracy had also presented by Citron and Pasquale who argued that individuals should be granted meaningful opportunities to challenge adverse decisions based on scores or decisions miscategorizing them (Citron and Pasquale, 2014). In this respect, proposals have been made toward regulations that compel information with at least, and always depending on the context of each algorithm, five broad categories of information: human involvement, data, model, inference, and algorithmic presence (Diakopoulos, 2016). The recent EU regulatory effort to minimize the unwanted implications of big data profiling and automated decision-making are analysed in the following chapter.

5. Regulating profiling and automated decision making under the GDPR

5.1. *GDPR provisions*

As Hildebrandt noted back in 2008, for a long time, profiles, as opposed to personal data, didn't have a clear legal status and therefore the protection against profiling was very limited (Hildebrandt, 2008). In 2010, the Council of Europe published its recommendation on "*the protection of individuals with regard to automatic processing of personal data in the context of profiling*" (Council of Europe 2010). Therein, the notion of profile was defined as "*a set of data characterizing a category of individuals that is intended to be applied to an individual*" and the profiling was referring to "*an automatic data processing technique that consists of applying a "profile" to an individual, particularly in order to take decisions concerning her or him or for analysing or predicting her or his personal preferences, behaviours and attitudes*". The GDPR, the newly enforced EU regulation on data protection, largely inspired by this definition provides a similar term for profiling: *"profiling" means any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements"* (Article 4). Yet, according to WP29 guidelines on the automated individual decision-making and profiling under the GDPR (Article 29 Data Protection Working Party, 2017), the two definitions are not identical to the fact the recommendation excludes processing that does not include inference.

The GDPR in Article 22 specifically provides for people's right not to be subject to a decision based solely on automated processing, including profiling, if this profiling "*significantly affects*" them. While a corresponding definition in Article 15[112] of the DPD had been criticized by scholars as providing limited protection against application issues of profiling (Bygrave, 2001; Schermer, 2011), the scope of the GDPR Article 22 is much broader in terms of the rights of the data subjects when their personal data are being processed for profiling purposes (Mendoza and Bygrave, 2017; Kaminski, 2019). Although the choice of the term "*right*" in the provision suggests that the Article applies when it is actively invoked by the data subject, the WP29 guidelines clarify that the article "*establishes a general prohibition for decision-making based solely on automated processing. This prohibition applies whether or not the data subject takes an action regarding the processing of their personal data*" (Article 29 Data Protection Working Party 2017). In other words, this prohibition, which is also suggested by

[110] https://ny.curbed.com/2018/8/24/17775290/new-york-city-automated-decision-systems (last access 27/12/2018).

[111] http://assets.ctfassets.net/8wprhhvnpfc0/1T0KpNv3U0EKAcQKseIsqA/52fee9a932837948e3698a658d6a8d50/NYC_ADS_Task_Force_Recs_Letter.pdf (last access 27/12/2018).

[112] DPD Article 15 includes "*automated processing of data intended to evaluate certain personal aspects relating to him, such as his performance at work, creditworthiness, reliability, conduct, etc*".

the text "*should be allowed where expressly authorized*" in recital 71, implies that processing under Article 22(1) is not allowed generally and hence individuals are automatically protected from the potential effects this type of processing may have. Therefore, this right cannot be considered as a special form of opt-out as it has been claimed thus far (Mendoza and Bygrave, 2017; Malekian, 2016). Certainly, this general prohibition is legitimate unless one of the exceptions of Article 22(2) applies, that is when the automated decision making is necessary for the performance of or entering into a contract; or is authorized by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or is based on the data subject's explicit consent. Moreover, as WP29 emphasizes, the Article 22(1) prohibition *only* applies in specific circumstances when a decision based solely on automated processing, including profiling, has a legal effect on or similarly significantly affects someone.[113] While it has been suggested that this right can be circumvented relatively easily by inserting even nominal involvement of a human in the loop (Zarsky, 2016; Wachter et al., 2017) (as the provisions is restricted to "solely" automated processing), the GDPR text as well the WP29 guidance identify that there are still many situations where the right is very likely to apply, such as credit applications, recruitment and insurance (UK ICO, 2017). Nevertheless, WP29 notes that "*targeted advertising based on profiling will not have a similarly significant effect on individuals*", raising thus concerns around cases where targeted advertising relies on highly intrusive profiling based on behavioural observed, inferred or predicted data (Kaltheuner and Bietti, 2018; Pasquale, 2015).

Furthermore, Article 22(3) specifies that "*the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision*". Apart from the fact that the Article does not elaborate on what these safeguards are, beyond "*the right to obtain human intervention*", it has been pointed out that the wording indicates that in the absence of decision-making, profiling alone does not give rise to safeguards under Article 22 (Kaltheuner and Bietti, 2018). Yet, the GDPR still gives rise to safeguards under Articles 13 to 15 to provide information on the processing. Actually, for many scholars the novelty of the GDPR profiling provisions is contained in Articles 13, 14 and 15 which oblige data controllers to provide "*information as to the existence of automated decision-making, including profiling*", and "*meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject*" (Bosco et al., 2015; Zarsky, 2016). Given this wording and the ML algorithmic opacity, scholars have unavoidably prompted the question what is required for data controllers to provide meaningful information to explain not only an algorithm's decision but also the envisaged consequences of its processing (Kamarinou et al., 2016; Burrell, 2016; Goodman and Flaxman, 2016).

As a matter of fact, the discussions on whether the GDPR implements an *ex ante* or an *ex post* right to explanation as a way to achieve accountability and transparency in automated decision-making provoked a heated debate among the legal, privacy and ML community, with some scholars arguing that the GDPR does not, in its current form, implements an *ex-post* right to explanation (Wachter et al., 2017), while others arguing otherwise (Mendoza and Bygrave, 2017; Goodman and Flaxman, 2016; Selbst and Powles, 2017; Kaminski, 2019). Under a third perspective, it has been asserted that a right to an explanation in the GDPR, even if exists, it is unlikely to present a complete remedy to algorithmic harms. Instead, a right to appeal to a machine against a decision made by a human may be proved to be the more effective remedy (Edwards and Veale, 2017; Kamarinou et al., 2016). Considering these arguments, the WP29 in its guidelines (Article 29 Data Protection Working Party, 2017) underline that the GDPR does not require the controller to provide a complex explanation of the algorithms used or disclosure of the full algorithm. Instead, the controller should find simple ways to inform the data subject about the rationale behind, or the criteria relied on in reaching the decision. On top, given that the controller should provide the data subject with information about the *envisaged consequences* of the processing, rather than an explanation of a *particular* decision, the WP29 affirms that information must be provided about the *intended or future* processing and should include general information (notably, on factors taken into account for the decision-making process) useful for challenging the decision. While this reading clarifies that the GDPR specifies a right to an *ex-ante* explanation, still it has been argued that the requirement for data subjects to be provided with "*knowledge of the reasoning underlying data processing*" in the context of decisions taken on the basis of big data-type processing is both unrealistic and deeply paradoxical, especially when they involve self-learning algorithms (Rouvroy, 2016). In this respect, counterfactual explanations have been proposed as a solution that bypasses the current technical limitations of interpretability and strikes a balance between transparency and the rights and freedoms of individuals (Wachter et al., 2018).

The WP29 in its guidelines (Article 29 Data Protection Working Party, 2017) clarifies further that decisions that are not solely automated might also include profiling whereas highlights the distinctions between profiling and automated decisions (Article 29 Data Protection Working Party, 2017): "*Automated decisions can be made with or without profiling; profiling can take place without making automated decisions. However, profiling and automated decision-making are not necessarily separate activities. Something that starts off as a simple automated decision-making process could become one based on profiling, depending upon how the data is used*". It also provides a couple of interesting explanations around the wide disputes provoked by Article's 22 interpretations. Taking into account that profiling practices can create a special category of "sensitive" data by inference from data which are not "sensitive" but become so when combined with other data, as well as the fact that a profile that relates to an individual and makes her identifiable is considered a type of personal data and ought to

[113] However, there are exceptions to these circumstances, such as when the profiling activities are necessary for a contract between the data subject and the data controller, or when the profiling is authorized by Member State law to which the controller is subject, including for fraud and tax-evasion monitoring, or data subjects have given explicit consent.

be protected (Gutwirth and Hildebrandt, 2010; Edwards and Veale, 2017), the WP29 concludes that the rights to rectification and to be forgotten (article 16 and 17 respectively) apply both to the "input personal data" (the personal data used to create the profile) and the "output data" (the profile itself or the "score" assigned to the person) (Article 29 Data Protection Working Party, 2017). However, the precedent WP29 guidance on the right to data portability (article 20) (Article 29 Data Protection Working Party, 2016) specifies that the right does not cover inferences from personal data analysis, like algorithmically or statistically derived categorisations or personalisation profiles,[114] implying thereby that the inferences of a system "belong" to the system that has generated them and not to the users whose personal data feeds this system (Edwards and Veale, 2017). Taking further into consideration the complementary nature of the right of data portability and the right to be forgotten, along with the fact that, as explained in Politou et al., 2018 (), the GDPR explicitly specifies that when the exercise of the right to be forgotten is based on the withdrawal of a previously given consent then the revocation is not retroactive, meaning that it does not apply for the processing that had taken place before withdrawal (Article 7(3)), it is deduced that profiles constructed and decisions previously taken on the basis of this information can therefore not be simply annulled.[115] In our opinion, this conclusion, also supported by the guidelines on the right to data portability as mentioned above (Article 29 Data Protection Working Party, 2016), clearly contradicts the WP29 guidance on applying the right to be forgotten on "output data", and hence it creates a serious loophole (Edwards and Veale, 2017; Urquhart et al., 2018).

As the GDPR foresees and regulates the core feature of big data analytics, namely the ability to profile individuals and to make automated decisions about them when algorithms are applied to large amounts of granular data (UK ICO, 2017), it has been forcefully argued that GDPR's implementation impact on big data practices would be substantial and highly problematic, albeit not prohibitive (Mayer-Schonberger and Padova, 2015; Zarsky, 2016). For big data enthusiasts, the prohibition defined under Article 22 is perhaps the most salient example of the GDPR's rejection of the big data revolution, and it is actually the main reason why its predecessor, Article 15 of the DPD, was either rarely applied or even a dead letter in some MSs (Zarsky, 2016). And given the fact that the GDPR provides persons with stricter protections from such decision making processes than its predecessor did, there have been even greater doubts as to whether it will have a significant practical impact on automated profiling decisional systems that are extremely complex and opaque (Mendoza and Bygrave, 2017; Kamarinou et al., 2016). Therefore, some legal scholars argue that the generic key principles and procedural rights of individuals, already established under the EU data protection law since its inception, are more potent to mitigate the long-term risks of big data and algorithmic decision-making compared to the specialized provisions on automated decision-making and profiling in the GDPR (Oostveen and Irion, 2017). Others scholars, however, such as Hildebrandt, feel confident that the GDPR might allow citizens to "*have their cake and eat it too*" as they will benefit from enhanced data protection while enjoying the innovations advanced data analytics bring about (Hildebrandt, 2015). Yet the fact that the GDPR applies to the profiling of individual data subjects and not of groups (since data that do not pertain to natural persons are beyond the scope of the GDPR) raises many questions on how data subjects are protected against decisions that have significant effects on them and subsequently affect their lives but they are based on group profiling (Oostveen and Irion, 2017; Edwards and Veale, 2017; Kamarinou et al., 2016; Taylor et al., 2016).

Acknowledging the aforementioned limitation of the GDPR as far as the regulation of personal data processing in the era of big data profiling analytics and automated decision is concerned, the Council of Europe published, almost a year after the GDPR's adoption by the EU, its Convention 108 Guidelines *"on the protection of individuals with regard to the processing of personal data in a world of Big Data"* (COUNCIL OF EUROPE 2017). According to Mantelero (2017), the guidelines move away from the EU traditional view in data protection regulation and provide for a more transparent approach towards the use of algorithms in decision-making processes as well as extended data subjects protections against the so-called dictatorship of data. Overall, the case of big data regulation under either the GDPR or the Convention 108 Guidelines, brings forward the discussion for the future of data protection regulation in Europe and contributes to the pursuit of regulating big data technology and specifically invasive and discriminatory profiling and automated decision practices.

5.2. *Implementation challenges and countermeasures*

While the scepticism towards reasoning of automated decision systems and correlations of ML algorithms, and privacy invasion of data mining on Big Data exist, at the end of the day, they are used due to their effectiveness. Therefore, especially for the case of public authorities, which end up performing such tasks for the sake of having a fairer tax system, the major question is whether the goal can be achieved with less privacy invasion and respecting the new GDPR legal framework.

Currently, there is a lot of effort into the integration of privacy enhancing technologies. By leveraging cryptographic primitives such as secure multi-party computation (SMC), order-preserving encryption, functional encryption, and homomorphic encryption one can perform a wide range of queries in a privacy-preserving way along with the training and inference of ML algorithms. However, the major challenge comes from the heterogeneity and sparsity of the data since in this scenario, the goal is not to determine whether an individual belongs in some lists, but, e.g. whether her aggregated deposits or expenses from all banks are beyond a threshold.

[114] In Article 29 Data Protection Working Party (2016) WP29 specifies that *"any personal data which have been created by the data controller as part of the data processing, e.g. by a personalisation or recommendation process, by user categorisation or profiling are data which are derived or inferred from the personal data provided by the data subject, and are not covered by the right to data portability"*.

[115] Still, the GDPR does not clarify what happens when the erasure is not based on the consent revocation but on some other available ground defined in Article 17(1). For these cases, it remains unclear whether a data controller is obliged to stop using the model or to go back and retrain the model either without including the erased data or even not to do anything at all.

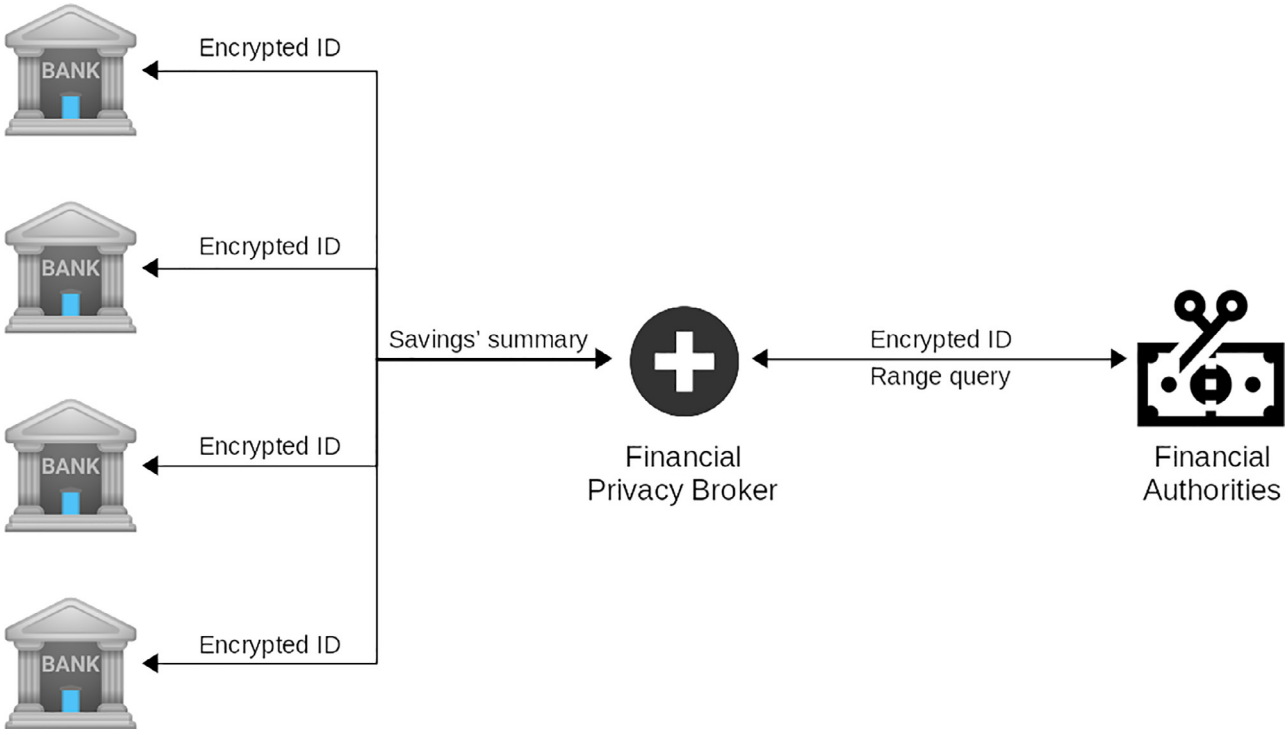

Fig. 1 – Example use case of the financial privacy broker.

For privacy-preserving aggregation of this form, there are several cryptographic solutions (Kursawe et al., 2011; Erkin et al., 2013; Patsakis et al., 2015). Nevertheless, they do not support threshold, but exact sums and they require all parties to be simultaneously online. Similarly, range queries over encrypted data (Boneh and Waters, 2007; Li et al., 2014), have as a prerequisite the use of a common public key.

Based on the above, since SMC only supports computations on data encrypted under the same public key the introduction of an independent semi-trusted third party would facilitate the requested task. In this regard, this entity would act as a broker/intermediator to allow the orchestration of collection of the financial data of individuals in an encrypted form, under a single key, and perform the requested operations of the authorities in an encrypted setup drastically decreasing the privacy invasive methods used so far. Since the data are encrypted, this entity cannot extract sensitive information of individuals or even differentiate them. Therefore, this entity does not need to be fully trusted.

Moreover, there is currently a lot of research effort in the field of privacy-preserving machine learning (Mohassel and Zhang, 2017; Brickell and Shmatikov, 2009; Hesamifard et al., 2018; Lu and Sakuma, 2018; Shokri and Shmatikov, 2015; Samet and Miri, 2012). The recent work of Li et al. (2018) is aligned with the application scenario that we are dealing with in the sense that we have multiple data providers and that the analyst at the end of the protocol performs ML over the joint dataset which contains the data of individuals with minor errors which provides privacy for individuals and does not disclose the operations to the data providers. The seminal work of Graepel et al. (2012), despite its inherent limitations; it allows only two trivial classifiers, showed that one could train an ML algorithm using encrypted data. This work initiated many other works, improving efficiency and including far more ML algorithms (Bost et al., 2015; Bos et al., 2014; Gilad-Bachrach et al., 2016; Ohrimenko et al., 2016). Therefore, while the ML algorithms in the privacy-preserving model might not be very efficient as their non-private counterparts, they can adhere to the privacy regulations of GDPR and provide good yet not so fine-grained results. Again, the introduction of a semi-trusted entity could significantly improve both efficiency and results.

In what follows, we will call this semi-trusted entity Financial Privacy Broker (FPB). As already discussed, the main role of FPB is to collect encrypted data from specific data sources and provided the authorities with a range result of aggregated data. A typical example of how we envision FPB to work is illustrated in Fig. 1. Let us assume that financial authorities (FA) of country A want to determine whether the savings of citizen C are in the range of [m,n]. To this end, FA has to contact all cooperating banks B_1, B_2,…,B_k and request the savings of C and classify C to the corresponding class. Rather than doing this, FA sends the request to FPB who will send the query to B_1, B_2,…,B_k. On receiving this, B_1, B_2,…,B_k start the two round protocol of Kursawe et al. (2011) or for more efficiency Patsakis et al. (2015) (if more summaries have to be performed) and compute the aggregated summary S of C's savings. Now, FPB can easily answer FA the range that S belongs to, without disclosing any data about the savings of C on any of the individual banks. Similarly, no information about C's savings will be disclosed to B_1, B_2,…,B_k. This trivial scheme can be further extended to blind FPB of the ID of C.

Therefore, the actual information that FPB will know would not be linkable to any individual.

6. Conclusion

As we have demonstrated, big data profiling and automated decision practices, albeit powerful and pioneering, they are also highly unregulated and thereby unfair and intrusive. In fact, their regulation is currently in infant stages, allowing thus the vast amount of consumer and taxpayer data collected by public and private bodies to make people more transparent to authorities and corporations without applying the principle of transparency vice versa, to make authorities and corporations more transparent to citizens and consumers (Sharman, 2009). Therefore, it is argued that the principles of accountability, transparency, and interpretability need to be clearly and unambiguously addressed in any big data analytics regulatory framework. Admittedly, the GDPR is a framework that governs algorithmic decision-making and profiling by introducing transparency as a basic element of algorithmic accountability (Kaminski, 2019). In that respect, the GDPR renders private and public sector more accountable to individuals and consequently challenges current industry and state approaches in terms of their privacy intrusive profiling practices. As a matter of fact, at the time of writing this article the European Parliament, taking into account and citing the GDPR principles, published its motion for a resolution[116] on the Cambridge Analytica case in which emphasizes the need for much greater algorithmic accountability and transparency with regard to data processing and analytics by the private and public sectors. It also stresses that profiling based on online behaviour, socio-economic or demographic factors, for political and electoral purposes, should be prohibited.

Yet, the enforcement of the GDPR compelled many thus far established international policies and legislations to be reevaluated regarding their compatibility with its data protection principles. In the tax domain, the GDPR's collision with the AEOI initiatives will certainly occupy the future lawmakers extensively. In fact, the first legal complaint against the HMRC and the OECD CRS for infringing privacy and data protection rights was filed in August 2018.[117] [118] Almost a month before that, legislators in the European Parliament released a resolution,[119] [120] following a respective motion,[121] asking the European Commission to ensure that privacy and data protection rights are respected in the context of FATCA and the automatic exchange of tax data. The resolution asks, among others, the MSs to review their IGAs and to amend them, if necessary, to align them with the rights and principles of the GDPR. It also calls on the Commission to conduct a full assessment of the impact of FATCA and the US extraterritorial practice on EU citizens, EU financial institutions and EU economies, and regrets the inherent lack of reciprocity of IGAs signed by MSs, especially in terms of the scope of information to be exchanged, which is broader for MSs than it is for the US. Remarkably, the resolution calls on all MSs to collectively suspend the application of their IGAs until the US agrees to a multilateral approach to the AEOI, by either repealing FATCA and joining the CRS or renegotiating FATCA on an EU-wide basis and with identical reciprocal sharing obligations on both sides of the Atlantic. This resolution came as no surprise since two months earlier the Parliament's Policy Department for Citizens' Rights and Constitutional Affairs published a study[122] [123] on FATCA's compatibility with the EU legislation, and specifically the GDPR, and raised key points indicating FATCA's violation of the new EU legislation. The study pleaded, among others, for IGAs modification to align with the GDPR and to become truly reciprocal.

The future of PSD2/MIFID2 enforcement in the GDPR era is not reassuring either, as there is currently a lack of guidance on the implementation of both directives to be GDPR-compliant.[124] On top, being both directives, as opposed to the GDPR regulation status and its highly imposed fines, weakens any penalties that are to be determined in case of non-compliance with these initiatives. Hence, unless specific and detailed directions on their implementation are not timely provided as well as their coordination with the GDPR is not carefully regulated, their coexistence with the GDPR is uncertain.

Apart from the effect of the GDPR on tax and financial policies, its extraterritorial impact on the transfer of personal data outside the EU/EEA domain is also substantial. Currently, the Privacy Shield agreement, the framework for regulating transatlantic exchanges of personal data between the EU and US, is widely challenged[125] due to the US failure to protect personal data belonging to EU citizens. On 5 July 2018 the European Parliament adopted a resolution[126] that stresses, among others, its concerns about the lack of specific rules and guarantees in the Privacy Shield for decisions based on automated processing and profiling, and calls on the Commission to consider suspending its validity until the US authorities be fully compliant with the framework, setting a deadline of 1 September 2018 for this to be achieved. However, the deadline has been long missed whereas, at the time of writing, the EU-US

[116] http://www.europarl.europa.eu/sides/getDoc.do?type=MOTION&reference=B8-2018-0480&format=XML&language=EN.
[117] https://www.theguardian.com/money/2018/aug/02/mishcon-de-reya-complains-about-anti-tax-evasion-measures.
[118] https://globaldatareview.com/article/1172676/tax-and-money-laundering-information-schemes-face-gdpr-complaint.
[119] http://www.europarl.europa.eu/oeil/popups/ficheprocedure.do?lang=en&reference=2018/2646(RSP).
[120] http://www.europarl.europa.eu/news/en/press-room/20180628IPR06837/meps-want-to-open-negotiations-on-an-eu-us-fatca-agreement.
[121] http://www.europarl.europa.eu/sides/getDoc.do?type=MOTION&reference=B8-2018-0306&language=EN.
[122] http://www.europarl.europa.eu/RegData/etudes/STUD/2018/604967/IPOL_STU(2018)604967_EN.pdf.
[123] https://iapp.org/news/a/study-examines-fatca-through-the-lens-of-gdpr/.
[124] https://www.insideprivacy.com/financial-institutions/overlap-between-the-gdpr-and-psd2/.
[125] https://www.reuters.com/article/us-eu-dataprotection-usa/eu-u-s-personal-data-pact-faces-second-legal-challenge-from-privacy-groups-idUSKBN12X253?il=0.
[126] http://www.europarl.europa.eu/sides/getDoc.do?type=TA&reference=P8-TA-2018-0315&language=EN&ring=B8-2018-0305.

negotiations on the future of Privacy Shield are intense and ongoing.[127]

Unquestionably, big data algorithmic profiling techniques are a huge step toward knowledge production and innovation. Hildebrandt (2008) had long ago envisioned that "*advanced profiling technologies generate knowledge and since knowledge is power, profiling changes the power relationships between the profilers and the profiled*". According to the Ethics Advisory Group1, digitally generated profiles based on very large quantities of data are powerful and increasingly unaccountable. Furthermore, as Pasquale explained thoroughly in his book, "*profiling is big business*" (Pasquale, 2015). Certainly, a successful one, given the latest revelations in the Cambridge Analytica case regarding the manipulation of 50 million Facebook profiles claimed to have won the 2016 US elections.[128] Therefore, Pasquale concludes that the "*need to anticipate* [and regulate] *how profiling technologies categorize and preempt us is indeed more urgent than the need to prevent identification or to remain anonymous*" (Pasquale, 2015). Solove (2007) prophetically quoted a decade ago that "*protecting individuals from excessive observation, scrutiny, and categorization is not an individualistic agenda, but rather one of promoting societal goods*".

Even though most EU policies aim to promote effective political and legal responses for enabling an innovative, transparent and with equal opportunities economic environment, most of the times they disregard data protection values and their impact to people's privacy, especially when these policies are combined with big data technology and algorithmic processing for profiling citizens and consumers. The widespread belief that public administrations, as they are held to a higher standard than the private sector organisations, can't engage in an identical implementation of big data profiling or automated-decision making systems[129] does not seem realistic anymore following the practices described in this article. Moreover, there is little doubt that as over the next few years big data will change the landscape of economic policy, traditional values such as the right to privacy and data protection will be highly challenged (Einav and Levin, 2014). Indeed, given the amount of personal information collected by the current technology as well as the amount of inferred personal data for which people don't even know their existence, privacy seems to become an obsolete notion, destined to be dropped from our vocabulary.[130] Nevertheless, although up to now the EU Data Protection Authorities have not received any significant number of complaints on profiling, probably due to the novelty of the use of automated profiling and to a general lack of awareness by the citizenry (Bosco et al., 2015), we firmly believe that this will not be the case henceforth. To this end, we believe that the use of concepts like FPB will become relevant in the near future and further research should be made in this field.

[127] https://www.euractiv.com/section/digital/news/eu-us-privacy-shield-review-jourova-to-meet-us-secretary-amid-compliance-concerns/.
[128] https://www.theguardian.com/uk-news/2018/mar/23/leaked-cambridge-analyticas-blueprint-for-trump-victory.
[129] https://bureaudehelling.nl/artikel-tijdschrift/efficiency-vs-accountability.
[130] https://www.nytimes.com/2014/05/21/opinion/friedman-four-words-going-bye-bye.html.

REFERENCES

Ananny M, Crawford K. Seeing without knowing: limitations of the transparency ideal and its application to algorithmic accountability. New Media Soc 2018;20(3):973–89.

Arbex M, Caetano S. Welfare implications of AEoI, No 1608, working papers. University of Windsor, Department of Economics; 2016 https://EconPapers.repec.org/RePEc:wis:wpaper:1608.

Article 29 Data Protection Working Party, Statement of the WP29 on automatic inter-state exchanges of personal data for tax purposes, WP 230, 2015, http://collections.internetmemory.org/haeu/20171003035404/http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2015/wp230_en.pdf.

Article 29 Data Protection Working Party, Guidelines for member States on the criteria to ensure compliance with data protection requirements in the context of the automatic exchange of personal data for tax purposes, WP 234, 2015, http://collections.internetmemory.org/haeu/20171003035404/http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2015/wp234_en.pdf.

Article 29 Data Protection Working Party, Guidelines on the right to data portability, WP242rev.01, Adopted on 13 December 2016. As last Revised and adopted on 5 April 2017, https://ec.europa.eu/newsroom/document.cfm?doc_id=44099.

Article 29 Data Protection Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, WP251rev.01, Adopted on 3 October 2017. As last Revised and Adopted on 6 February 2018, http://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053.

Article 29 Data Protection Working Party, Opinion 01/2017 on the proposed regulation for the ePrivacy regulation (2002/58/EC), WP 247, 2017, http://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=610140.

Article 29 Data Protection Working Party, Article 29 Data Protection Working Party Letter 21/06/2012 to the Director General of Taxation and Customs Union European Commission Ref. Ares (2012) 746461 following a request for assistance by DG TAXUD to evaluate the compatibility of the obligations under US Foreign Account Tax Compliance Act (FATCA) and Directive 95/46/EC, http://collections.internetmemory.org/haeu/20171122154227/http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2012/20120621_letter_to_taxud_fatca_en.pdf.

Article 29 Data Protection Working Party, Article 29 Data Protection Working Party Letter 01/10/2012 to the Director General of Taxation and Customs Union European Commission Ref. Ares (2012) 1148996 regarding FATCA and Model II agreements, http://ec.europa.eu/justice/data-protection/article-29/documentation/other-document/files/2012/20121001_letter_to_taxud_fatca_en.pdf.

Avi-Yonah, Reuven S. and Mazzoni, Gianluca, Taxation and human rights: a delicate balance (2016). U of Michigan Public Law Research Paper No. 520. Available at SSRN: https://ssrn.com/abstract=2834883 or http://dx.doi.org/10.2139/ssrn.2834883.

Baker P, Pistone P. The Practical Protection of Taxpayers. In Fundamental Rights, 'General Report, International Fiscal Association, 2015 Basel Congress (Vol. 100).

Baker P, Pistone P. BEPS Action 16: the taxpayers' right to an effective legal remedy under European law in cross-border situations. EC Tax Rev 2016;25(5):335–45.
Baker P. Taxation and the European convention on Human Rights. EUROPEAN TAXATION-AMSTERDAM-2000;40(8):298–374.
Baker P. CRS/DAC, FATCA and the GDPR. Br Tax Rev 2016(3):249–52.
Baker P. Privacy rights in an age of transparency: a European perspective. Tax Notes Int 2016;82(6):583–6.
Barocas S, Nissenbaum H. Big data's end run around procedural privacy protections. Commun ACM 2014;57(11):31–3.
Barocas S, Selbst AD. Big data's disparate impact. Cal. L. Rev. 2016;104:671.
Bessard P. Inidividual rights and tax oppression in the OECD. Liberales Institut paper 2017;3:1–29. https://www.libinst.ch/publikationen/LI-Studie-Tax-Oppression.pdf.
Big data, artificial intelligence, machine learning and data protection, UK INFO.COMMISSIONER'S OFF 37–39 (2017), https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf, [hereinafter UK ICO Report].
Boneh D, Waters B. Conjunctive, subset, and range queries on encrypted data. Proceedings of the theory of cryptography conference. Springer; 2007. p. 535–54.
Bos JW, Lauter K, Naehrig M. Private predictive analysis on encrypted medical data. J Biomed Inf 2014;50:234–43.
Bosco F, Creemers N, Ferraris V, Guagnin D, Koops BJ. Profiling technologies and fundamental rights and values: regulatory challenges and perspectives from European Data Protection Authorities. Reforming European data protection law. Dordrecht: Springer; 2015. p. 3–33.
Bost R, Popa RA, Tu S, Goldwasser S. Machine learning classification over encrypted data. NDSS; 2015.
Brickell J, Shmatikov V. "Privacy-preserving classifier learning. Proceedings of the international conference on financial cryptography and data security. Springer, 2009.
Brodzka A. FATCA from the European Union perspective. J Gov Regul 2013;2(3).
Burrell J. How the machine 'thinks': understanding opacity in machine learning algorithms. Big Data Soc 2016;3(1).
Bygrave LA. Minding the machine: article 15 of the EC data protection directive and automated profiling'. Comput Law Secur Report 2001;17:17–24.
Calo R. Privacy and markets: a love story. Notre Dame L Rev 2015;91:649.
Calude CS, Longo G. The deluge of spurious correlations in big data. Found Sci 2017;22(3):595–612.
Christensen III H, Tirard JM. The amazing development of exchange of information in tax matters: from double tax treaties to FATCA and the CRS. Trusts Trustees 2016;22(8):898–922.
Christians A, Cockfield AJ. Submission to finance department on implementation of FATCA in Canada (2014). Available at SSRN: https://ssrn.com/abstract=2407264 or http://dx.doi.org/10.2139/ssrn.2407264.
Christians A. The Dubious Legal Pedigree of IGAs (and Why it Matters) (February 11, 2013). Tax Notes Int 2013;69(6). Available at SSRN: https://ssrn.com/abstract=2280508.
Citron DK, Pasquale F. The scored society: due process for automated predictions. Wash L Rev 2014;89:1.
Citron DK. Technological due process. Wash UL Rev 2007;85:1249.
Cockfield AJ. Protecting taxpayer privacy rights under enhanced cross-border tax information exchange: toward a multilateral taxpayer bill of rights. UBC Law Review 2010;42(2):421.
Cockfield AJ. FATCA and the erosion of Canadian taxpayer privacy (April 1, 2014). Report to the Office of the Privacy Commissioner of Canada, 2014. Available at SSRN: https://ssrn.com/abstract=2433198
Cockfield AJ. Big data and tax haven Secrecy. Fla Tax Rev 2015;18:483.
Cohen JE. What privacy is for. Harv L Rev 2012;126:1904.
Council of Europe, The protection of individuals with regard to automatic processing of personal data in the context of profiling Recommendation CM/Rec(2010)13 and explanatory memorandum Council of Europe 2010. https://rm.coe.int/16807096c3.
COUNCIL OF EUROPE, "Guidelines on the protection of individuals with regard to the processing of personal data in a world of Big Data", 2017. Available at https://rm.coe.int/16806ebe7a.
Crawford K, Schultz J. Big data and due process: toward a framework to redress predictive privacy harms. BCL Rev 2014;55:93.
De Simone L, Lester R, Markle K. Transparency and tax evasion: evidence from the foreign account tax compliance act (FATCA) (2017). Stanford University Graduate School of Business Research Paper No. 17-62. Available at SSRN: https://ssrn.com/abstract=3037426.
Debelva F, Mosquera I. Privacy and confidentiality in exchange of information procedures: some uncertainties, many issues, but few solutions. Intertax 2017;45(5):362–81.
Dharmapala D. Cross-border tax evasion under a unilateral FATCA regime. J Public Econ 2016;141:29–37.
Diakopoulos N. Accountability in algorithmic decision making. Commun ACM 2016;59(2):56–62.
Diepvens N, Debelva F. The evolution of the exchange of information in direct tax matters: the taxpayer's rights under pressure. EC Tax Rev 2015;24(4):210–19.
Donnelly M. Payments in the digital market: evaluating the contribution of payment services directive II. Comput Law Secur Rev 2016;32(6):827–39.
Doshi-Velez F, Kortz M, Budish R, Bavitz C, Gershman SJ, O'Brien D. et al., Accountability of AI Under the Law: The Role of Explanation (2017). Berkman Center Research Publication Forthcoming; Harvard Public Law Working Paper No. 18-07. Available at SSRN: https://ssrn.com/abstract=3064761 or 10.2139/ssrn.3064761.
Dwork C, Hardt M, Pitassi T, Reingold O, Zemel R. Fairness through awareness. Proceedings of the 3rd innovations in theoretical computer science conference. ACM; 2012. p. 214–26.
Dwork C. Differential privacy. Proceedings of the 33rd international colloquium on automata, languages and programming, part II (ICALP 2006). Springer; 2006. p. 1–12.
Edwards L. Veale M. (2017). Slave to the Algorithm? Why a 'Right to Explanation'is Probably Not the Remedy You are Looking for.
Einav L, Levin J. The data revolution and economic analysis. Innov Policy Econ 2014;14(1):1–24.
Erkin Z, Troncoso-Pastoriza JR, Lagendijk RL, Pérez-González F. Privacy-preserving data aggregation in smart metering systems: an overview. IEEE Signal Process Mag 2013;30(2):75–86.
European Commission, First Report of the Commission AEFI expert group on the implementation of Directive 2014/107/EU for automatic exchange of financial account information, 2015, https://ec.europa.eu/taxation_customs/sites/taxation/files/resources/documents/taxation/tax_cooperation/mutual_assistance/financial_account/first_report_expert_group_automatic_exchange_financial_information.pdf.
Federal Trade Commission 2016. Big data A tool for inclusion or exclusion Washington, DC: Federal Trade Commission.
Fuster GG. EU data protection and future payment services. Bitcoin and Mobile Payments. London: Palgrave Macmillan; 2016. p. 181–201.

Gadžo S, Klemenčić I. Effective international information exchange as a key element of modern tax systems: promises and pitfalls of the OECD's common reporting standard. Public Sector Econ 2017;41(2):207–26.
Giambelluca G, Masi P. The regulatory machine: an institutional approach to innovative payments in Europe. Bitcoin and mobile payments. London: Palgrave Macmillan; 2016. p. 3–25.
Gilad-Bachrach R, Dowlin N, Laine K, Lauter K, Naehrig M, Wernsing J. Cryptonets: applying neural networks to encrypted data with high throughput and accuracy. Proceedings of the international conference on machine learning; 2016. p. 201–10.
Goldfarb A, Tucker C. Privacy and innovation. Innov Policy Econ 2012;12(1):65–90.
Goodman B, Flaxman S. European Union regulations on algorithmic decision-making and a "right to explanation". All Magazine 2017;38(3):50–7.
Graepel T, Lauter K, Naehrig M. ML confidential: machine learning on encrypted data. Proceedings of the International Conference on Information Security and Cryptology. Springer; 2012. p. 1–21.
Grinberg I. The battle over taxing offshore accounts. UCLA L Rev 2012;60:304.
Gutwirth S, Hildebrandt M. Some caveats on profiling. Data protection in a profiled World. Dordrecht: Springer; 2010. p. 31–41.
Hatfield Michael. Taxation and surveillance: an agenda. Yale J. L. & Tech. 2015(17):319–67.
Hatfield Michael. Privacy in Taxation. Florida State University Law Review; 2016 Forthcoming University of Washington School of Law Research Paper No. 2016-15Available at SSRN: https://ssrn.com/abstract=2788238 .
Hesamifard E, Takabi H, Ghasemi M, Wright RN. Privacy-preserving machine learning as a service. Proc Priv Enhanc Technol 2018;2018(3):123–42.
Hildebrandt M. Defining profiling: a new type of knowledge?. Profiling the European citizen. Dordrecht: Springer; 2008. p. 17–45.
Hildebrandt M. Profiling and the rule of law. Identity Inf Soc 2008;1(1):55–70.
Hildebrandt M. Who is profiling who? Invisible visibility. Reinventing data protection?. Dordrecht: Springer; 2009. p. 239–52.
Hildebrandt M. Smart technologies and the end (s) of law: novel entanglements of law and technology. Edward Elgar Publishing; 2015.
Hildebrandt M. Privacy as protection of the incomputable self: from agnostic to agonistic machine learning (2017). Available at SSRN: https://ssrn.com/abstract=3081776 or 10.2139/ssrn.3081776.
HJI Panayi C. Current trends on automatic exchange of information. Singapore Management University School of Accountancy Research Paper 2016:43.
Houser KA, Sanders D. The use of big data analytics by the irs: efficient solutions or the end of privacy as we know it. Vand. J. Ent. Tech. L. 2016;19:817.
Hurley M, Adebayo J. Credit scoring in the era of big data. Yale JL Tech 2016;18:148.
Johannesen N, Zucman G. The end of bank secrecy? An evaluation of the G20 tax haven crackdown. Am Econ J: Econ Policy 2014;6(1):65–91.
Kaltheuner F, Bietti E. Data is power: towards additional guidance on profiling and automated decision-making in the GDPR. J Inf Rights Policy Pract 2018;2(2).
Kamarinou, D. Millard, C. Singh, J., Machine Learning with Personal Data (November 7, 2016). Queen Mary School of Law Legal Studies Research Paper No. 247/2016. Available at SSRN: https://ssrn.com/abstract=2865811.
Kaminski ME. The right to explanation, explained (June 15, 2018). U of Colorado Law Legal Studies Research Paper No. 18-24; Berkeley Technology Law Journal, Vol. 34, No. 1, 2019. Available at SSRN: https://ssrn.com/abstract=3196985 or http://dx.doi.org/10.2139/ssrn.3196985.
Knobel, A. (2017). Findings of the 2nd TJN Survey on Automatic Exchange of Information (AEOI), Tax Justice Network, https://financialtransparency.org/wp-content/uploads/2017/01/Knobel2017_AEOI-Survey-Report.pdf.
Kosinski M, Stillwell D, Graepel T. Private traits and attributes are predictable from digital records of human behavior. Proc Natl Acad Sci 2013;110(15):5802–5.
Kroll JA, Barocas S, Felten EW, Reidenberg JR, Robinson DG, Yu H. Accountable algorithms. U Pa L Rev 2016;165:633.
Kursawe K, Danezis G, Kohlweiss M. Privacy-friendly aggregation for the smart-grid. Proceedings of the international symposium on privacy enhancing technologies symposium. Springer, 2011.
Lakkaraju H, Bach SH, Leskovec J. Interpretable decision sets: a joint framework for description and prediction. Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. ACM; 2016. p. 1675–84.
Letham B, Rudin C, McCormick TH, Madigan D. Interpretable classifiers using rules and bayesian analysis: building a better stroke prediction model. Ann Appl Stat 2015;9(3):1350–71.
Li P, Li T, Ye H, Li J, Chen X, Xiang Y. Privacy-preserving machine learning with multiple data providers. Fut Gener Comput Syst 2018;87:341–50.
Lipton ZC. The mythos of model interpretability. Proceedings of the 2016 ICML workshop on human interpretability in machine learning (WHI 2016), 2016.
Li R, Liu AX, Wang AL, Bruhadeshwar B. Fast range query processing with strong privacy protection for cloud computing. Proc VLDB Endow 2014;7(14):1953–64.
Lotmore KW. The decline of financial privacy and its costs to society. Trusts Trustees 2017;23(9):944–54.
Lu WJ, Sakuma J. More practical privacy-preserving machine learning as a service via efficient secure matrix multiplication. Proceedings of the 6th workshop on encrypted computing & applied homomorphic cryptography. ACM; 2018. p. 25–36.
Malekian Hajar. Profiling under general data protection regulation (GDPR): stricter regime? Malekian, Hajar. 2016. (2016) https://www.linkedin.com/pulse/profiling-under-general-data-protection-regulation-gdpr-malekian/.
Mansfield-Devine S. Open banking: opportunity and danger. Computer Fraud & Security 2016;2016(10):8–13.
Mantelero A. Personal data for decisional purposes in the age of analytics: from an individual to a collective dimension of data protection. Comput Law Secur Rev 2016;32(2):238–55.
Mantelero A. Regulating big data. The guidelines of the Council of Europe in the context of the European data protection framework. Comput Law Secur Rev 2017;33(5):584–602.
Marchiori L, Pierrard O. Unlocking the gates of paradise: general equilibrium effects of information exchange. J Econ Dyn Control 2017;87:152–72.
Mayer-Schonberger V, Padova Y. Regime change: enabling big data through Europe's new data protection regulation. Col Sci Technol Law Rev 2015;17:315.
Meinzer M. Automatic Exchange of Information as the new global standard: the end of (offshore tax evasion) history?. Automatic exchange of information and prospects of Turkish-German cooperation; 2017.
Mendoza I, Bygrave LA. The right not to be subject to automated decisions based on profiling. EU internet law. Cham: Springer; 2017. p. 77–98.
Mohassel P, Zhang Y. "SecureML: a system for scalable privacy-preserving machine learning. Proceedings of the 2017 38th IEEE symposium on security and privacy (SP). IEEE, 2017.

Morse SC. Why FATCA intergovermental agreements bind the U.S. government (April 15, 2013). Tax Notes Int 2013;70(3). Available at SSRN: https://ssrn.com/abstract=2252843.
Noseda F. Common reporting standard and EU beneficial ownership registers: inadequate protection of privacy and data protection. Trusts Trustees 2017;23(4):404–9.
Noseda F. CRS and beneficial ownership registers—what serious newspapers and tabloids have in common: the improbable story of a private client lawyer turned human rights activist: the improbable story of a private client lawyer turned human rights activist. Trusts Trustees 2017;23(6):601–9.
Noseda F. Trusts and privacy: a new battle front. Trusts Trustees 2017;23(3):301–10.
OBERSON X. Towards automatic exchange of information. Revue suisse de droit des affaires et du marché financier 2015;87(2):91–107.
OECD (2014). Standard For automatic exchange of financial account information in tax matters, OECD Publishing. 10.1787/9789264216525-en.
Ohrimenko O, Schuster F, Fournet C, Mehta A, Nowozin S, Vaswani K, et al. Oblivious multi-party machine learning on trusted processors. Proceedings of the USENIX security symposium; 2016. p. 619–36.
Oostveen M, Irion K. The golden age of personal data: how to regulate an enabling fundamental right?. Personal data in competition, consumer protection and IP law - Towards a holistic approach?. Berlin: Springer; 2017 Forthcoming; Institute for Information Law Research Paper No. 2016-06; Amsterdam Law School Research Paper No. 2016-68.
Pap AEthnicity and Race-Based Profiling in Counter-Terrorism, Law Enforcement and Border Control. European parliament's committee on civil liberties. Justice Home Affairs 2008:63. 2008Available at SSRN: https://ssrn.com/abstract=2372777.
Pasquale F. The black box society: the secret algorithms that control money and information. Harvard University Press; 2015.
Patsakis C, Laird P, Clear M, Bouroche M, Solanas A. Interoperable privacy-aware e-participation within smart cities. Computer 2015;48(1):52–8.
Politou E, Alepis E, Patsakis C. A survey on mobile affective computing. Comput Sci Rev 2017;25:79–100.
Politou E, Alepis E, Patsakis C. Constantinos Patsakis; Forgetting personal data and revoking consent under the GDPR: Challenges and proposed solutions. J Cybersecurity 2018;4(1). doi:10.1093/cybsec/tyy001.
Reijers J, Jacobs BPF, Poll IE. Payment Service Directive 2. The Netherlands: Thesis for the Degree of Master of Science in Information Sciences at the Radboud University Nijmegen; 2016.
Ribeiro MT, Singh S, Guestrin C. Why should i trust you?: explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. ACM; 2016. p. 1135–44.
Richards NM, King JH. Three paradoxes of big data. Stan L Rev Online 2013;66:41.
Rocha SA. Exchange of tax-related information and the protection of taxpayer rights: general comments and the Brazilian perspective. Bull Int Tax 2016;70(9):502–16.
Rouvroy A. 'Of Data and Men'. Fundamental Rights and Freedoms in a World of Bid Data. Council of Europe, Directorate General of Human Rights and Rule of Law, T-PD-BUR (2015) 09REV, Strasbourg; 2016.
Samet S, Miri A. Privacy-preserving back-propagation and extreme learning machine algorithms. Data Knowl Eng 2012;79:40–61.
Savin, A. (2014). Profiling and Automated Decision Making in the Present and New EU Data Protection Frameworks. Paper presented at 7th International Conference Computers, Privacy & Data Protection, Brussels, Belgium.
Schaper M. Data protection rights and tax information exchange in the European union: an uneasy combination. Maastricht J Eur Comp Law 2016;23(3):514–30.
Schermer BW. The limits of privacy in automated profiling and data mining. Comput Law Secur Rev 2011;27(1):45–52.
Schwartz PM, Solove DJ. Reconciling personal information in the United States and European Union. Cal L Rev 2014;102:877.
Schwartz P. The future of tax privacy. Natl Tax J 2008:883–900.
Selbst AD, Powles J. Meaningful information and the right to explanation. Int Data Privacy Law 2017;7(4):233–42.
Sharman JC. Privacy as roguery: personal financial information in an age of transparency. Public Admin 2009;87(4):717–31.
Shokri R, Shmatikov V. Privacy-preserving deep learning. Proceedings of the 22nd ACM SIGSAC conference on computer and communications security. ACM; 2015. p. 1310–1321.
Solove DJ. I've got nothing to hide and other misunderstandings of privacy. San Diego L Rev 2007;44:745.
Somare M, Wöhrer V. Automatic exchange of financial information under the directive on administrative cooperation in the light of the global movement towards transparency. Intertax 2015;43(12):804–15.
Taylor L, Schroeder R, Meyer E. Emerging practices and perspectives on big data analysis in economics: bigger and better or more of the same? Big Data Soc 2014;1(2).
Taylor L, Floridi L, van der Sloot B. Group privacy: new challenges of data technologies, 126. Springer; 2016.
Tello CP. FATCA: catalyst for global cooperation on exchange of tax information. Bull Int Tax 2014;68(2):88–102.
Thimmesch Adam B. Tax privacy?. Temple Law Review; 2017 Forthcoming. Available at SSRN: https://ssrn.com/abstract=3039753 .
Tutt AAn FDA for Algorithms (March 15, 2016). Admin. L. Rev. 2017;69:83. Available at SSRN https://ssrn.com/abstract=2747994. or 10.2139/ssrn.2747994.
Urquhart L, Sailaja N, McAuley D. Realising the right to data portability for the domestic Internet of things. Pers Ubiquitous Comput 2018;22(2):317–32.
Van Alsenoy B, Verdoodt V, Heyman R, Wauters E, Ausloos J, Acar G. (2015). From social media service to advertising network: a critical analysis of Facebook's Revised Policies and Terms. A report commissioned by the Belgian Privacy Commission. Section 7. https://www.law.kuleuven.be/citip/en/news/item/facebooks-revised-policies-and-terms-v1-2.pdf.
Veale M, Van Kleek M, Binns R. Fairness and Accountability design needs for algorithmic support in high-stakes public sector decision-making. Proceedings of the ACM conference on human factors in computing systems (CHI'18), April 21–26, 2018.
Viktor M-S, Kenneth C. Big data: a revolution that will transform how we live, work, and think. Houghton Mifflin Harcourt 2013.
Wachter S, Mittelstadt B, Floridi L. Why a right to explanation of automated decision-making does not exist in the general data protection regulation. Int Data Priv Law 2017;7(2):76–99.
Wachter S, Mittelstadt B, Russell C. Counterfactual explanations without opening the black box: automated decisions and the GDPR. Harvard J Law Technol 2018;31(2).
Zarsky, T, Transparent Predictions (September 10, 2013). University of Illinois Law Review, Vol. 2013, No. 4, 2013. Available at SSRN: https://ssrn.com/abstract=2324240.
Zarsky TZ. Incompatible: the GDPR in the Age of Big Data. Seton Hall L Rev 2016;47:995.