Are AI-based Anti-Money Laundering (AML) Systems Compatible with European Fundamental Rights?

Winston Maxwell, Astrid Bertrand, Xavier Vamparys

Research paper by

OPERATIONAL AI ETHICS

telecom-paris.fr/en/ai-ethics

Are AI-based Anti-Money Laundering (AML) Systems Compatible with European Fundamental Rights?

Astrid Bertrand[1], Winston Maxwell[2] and Xavier Vamparys[3]

▶ To cite this version:

Astrid Bertrand, Winston Maxwell, Xavier Vamparys. Are AI-based Anti-Money Laundering (AML) Systems Compatible with European Fundamental Rights? Telecom Paris Research Paper Series November 2020.

Revised version November 10, 2020

[1] PhD candidate, Télécom Paris, Institut Polytechnique de Paris

[2] Director of Law and Digital Technology Studies, Télécom Paris, i3, Institut Polytechnique de Paris

[3] Manager of AI Ethics, CNP Assurances; visiting researcher, Télécom Paris, Institut Polytechnique de Paris

Are AI-based Anti-Money Laundering (AML) Systems Compatible with European Fundamental Rights?

By Winston Maxwell, Astrid Bertrand and Xavier Vamparys

Abstract

Anti-money laundering and countering the financing of terrorism (AML/CFT) systems must comply with the GDPR and the proportionality test under European fundamental rights law, as most recently expressed by the Court of Justice of the European Union (CJEU) in the *Digital Rights Ireland, Tele2 Sverige - Watson* and *Quadrature du Net* cases. The objective of this paper is to present how AML/CFT laws and systems work, how artificial intelligence (AI) can enhance those systems, and examine whether these systems comply with the European proportionality test. We conclude that current AML/CFT systems violate the proportionality test in several ways: AML/CFT laws are not specific enough to satisfy the 'provided by law' test, and the lack of feedback on the utility of suspicious activity reports sent by banks to law enforcement authorities means that it is impossible to determine if AML/CFT systems are 'genuinely effective'. We propose a mechanism that would permit law enforcement authorities to provide feedback on what alerts are 'genuinely effective' in fighting money laundering, feedback that would help banks adjust their monitoring systems to be more efficient and proportionate. We also propose (i) that banks be required to inform their customers when they have been targeted by a suspicious activity report, as soon as doing so would no longer compromise an investigation, and (ii) that an independent regulatory authority verify the proportionality of the system on a regular basis. The institutional oversight structure could be inspired by the regulatory structure used to oversee intelligence gathering in France.

Keywords: *anti-money laundering, artificial intelligence, European fundamental rights, monitoring systems, proportionality.*

TABLE OF CONTENTS

I. AML/CFT SYSTEMS WILL SOON INCORPORATE AI

1. AML/CFT Monitoring Raises Serious Fundamental Rights Issues

Imagine that your telecommunications operator is required by law to monitor your call records to detect abnormal or suspicious calling patterns and report any suspicious activity to authorities without informing you. This kind of systematic monitoring of telecommunications data would be manifestly illegal under the Court of Justice of the European Union's (CJEU) *Digital Rights Ireland*[1] and *Tele2 Sverige – Watson*[2] cases, among others. Laws on anti-money laundering and countering terrorist financing (AML/CFT) are quite similar to this fact pattern. They oblige banks to implement transaction monitoring systems to ferret out suspicious activity by customers, and report any suspicious activity to financial intelligence units (FIUs) within governments. These systems currently operate on static, rule-based, models, but they will soon be enhanced by artificial intelligence (AI), making the fundamental rights concerns even more important. There are differences between the telecommunications example and bank AML/CFT measures: the relationship between banks and their customers is different than the relationship between telecom operators and their customers, and there are economic incentives explaining why banks need a strong legal obligation to actively look for criminal activity.[3] Nevertheless, the telecom analogy is sufficiently close to AML/CFT to raise red flags: the fundamental rights issues related to AML/CFT are serious.

The original purpose of our research was to identify the specific aspects of AI that raise fundamental rights issues in the AML/CFT context. However, we quickly realized that the whole AML/CFT system itself needs to be analyzed, particularly in light of the CJEU's *Digital Rights Ireland* and *Tele2 Sverige – Watson* cases. Can AML/CFT be reconciled with those cases? AI adds new challenges to the fundamental rights questions, such as explainability, but many of the fundamental rights problems are not linked to AI *per se*, but to the AML/CFT system itself.

2. Current AML/CFT Systems Are Relatively Ineffective

Interference with privacy rights can be justified when necessary to fight crime. The proportionality test is there to help strike the right balance between fighting crime and privacy. But current AML/CFT approaches are surprisingly ineffective in apprehending criminal funds. Europol estimates that approximately €200 billion in criminal funds circulate every year in Europe, yet only 1% of criminal funds are confiscated.[4] Banks have invested in costly transaction monitoring systems that generate alerts and help bank employees report unusual or suspicious activity. But current detection systems generate 90% of false positives, which must then be reviewed by bank employees. After this review, banks send suspicious activity reports (SARs) to FIUs, which then decide whether

[1] Digital Rights Ireland v. Minister for Communications, Joined Cases C-293/12 and C-594/12, 8 April 2014 (ECLI:EU:C:2014:238).

[2] Tele2 Sverige v Post- och telestyrelsen, and Secretary of State v Watson, Joined Cases C-203/15 and C-698/15, 21 December 2016, (ECLI:EU:C:2016:970).

[3] Lucia Dalla Pellegrina and Donato Masciandaro 'The risk-based approach in the new European anti-moneylaundering legislation: A law and economics view' (2009) 5 Review of Law & Economics 2, 931.

[4] Europol, 'Does Crime Still Pay?' (2016) Survey of statistical information.

or not to alert the prosecutor's office, intelligence authorities or the tax, social or customs administration. FIUs typically review only 10% to 20% of the SARs they receive,[5] and only a portion of those are forwarded to other law enforcement agencies for action.[6] The bottom line is that most SARs contribute little or nothing to the objective of detecting and apprehending criminal funds. According to the former director of Europol: 'We have created a whole ton of regulations ... the banks are spending $20 billion a year to run the compliance regime ... and we are seizing 1 percent of criminal assets every year in Europe'.[7]

AML/CFT laws have created a large industry of compliance technology, processes, regulators and professionals, who focus almost exclusively on better compliance processes and technology. Like many regulatory structures, AML/CFT has taken on a life of its own, with a tendency to look for ever bigger and better technological solutions to detect illegal activity. AI is the next major revolution in AML/CFT compliance. AI can automate data collection, enhance client risk scoring and alert prioritization, leverage link analysis, improve segmentation and sharpen anomaly detection. But several obstacles stand in the way of AI adoption, among them fundamental rights concerns and in particular the proportionality test examined in this article. Our focus on proportionality for AML/CFT will shed light on many other use cases involving AI to fight crime: fraud detection, predictive policing, airline passenger screening, cyber-security, and counterterrorism to name a few. The proportionality test will apply to any government-imposed measure designed to fight crime which at the same time adversely impacts fundamental rights, particularly the right to privacy. The contribution of this article extends beyond AML/CFT.

The remainder of this article is structured as follows: Part II provides an overview of the European proportionality test; Part III gives an overview of AML/CFT regulation and processes; Part IV applies the proportionality test step by step to AML/CFT systems, suggesting cures to certain proportionality problems; Part V concludes.

II. THE PROPORTIONALITY TEST

1. The Purpose and Origin of the Proportionality Test

The proportionality test ensures that powers given to government in a particular law, for example a law on electronic surveillance, do not unduly restrict privacy or other fundamental rights such as non-discrimination, freedom of expression, or a right to a fair trial.[8] The proportionality test exists at two levels. The first level consists of the proportionality requirement under EU Charter of Fundamental Rights (Charter)[9], the European Convention of Human Rights (ECHR)[10], as well as many national constitutions.[11] We will call this the "constitutional"

[5] Europol, *From Suspicion to Action: Converting Financial Intelligence into Greater Operational Impact.* (Publications Office of the European Union 2017).

[6] TRACFIN (2018) Activity Report.

[7] Giulia Paravicini, 'Europe Is Losing the Fight against Dirty Money' *Politico* (4 February 2018).

[8] Tom Hickman, 'The Substance and Structure of Proportionality' [2008] Public Law 694.

[9] Charter of Fundamental Rights of the European Union, OJ 2010 C 83/389.

[10] European Convention for the Protection of Human Rights and Fundamental Freedoms, Sept. 3, 1953, ETS 5, 213 UNTS 221 (ECHR).

[11] Jean-Marc Sauvé, 'Le principe de proportionnalité, protecteur des libertés' (Intervention at the Portalis Institute, Aix-en-Provence 2017).

proportionality level. The second level consists of the proportionality requirements contained in specific legislation, including the General Data Protection Regulation 2016/679 (GDPR)[12], the Police-Criminal Justice Directive 2016/680[13], and national legislation. We will call this the "statutory" proportionality level. Most European cases examine whether laws adopted by Member States to fight crime conform to the proportionality test at the constitutional level, although the CJEU may also be required to interpret proportionality wording at the statutory level.

The Charter's proportionality test is expressed in two sentences: "*Any limitation on the exercise of the rights and freedoms recognised by this Charter must be provided for by law and respect the essence of those rights and freedoms. Subject to the principle of proportionality, limitations may be made only if they are necessary and genuinely meet objectives of general interest recognised by the Union or the need to protect the rights and freedoms of others."*[14]

Whether or not specific laws, European or national, contain express references to proportionality, the proportionality principle is necessarily present at the statutory level. A law, whether European or national, that does not comply with the proportionality principle at the constitutional level will be invalid. In fact, most European legislation includes language that expressly imports the proportionality concept into the legislation. For example, the GDPR requires that measures interfering with privacy rights be "necessary and proportionate in a democratic society".

2. Proportionality and the GDPR

Many provisions of the GDPR make express reference to the proportionality principle. One of the GDPR provisions that is the most relevant for AML/CFT is Article 23, which foresees the possibility for the European Union or Member States to enact laws (like AML/CFT laws) that restrict data protection rights in order to protect national security or detect criminal offenses. Article 23 provides that any legal restriction must "respect[ ] the essence of fundamental rights and freedoms and [be] a necessary and proportionate measure in a democratic society".[15] Article 23 says that the relevant law must be specific, describing among other things the purpose of the processing, the categories of data processed, and who may access them.[16] Article 23 imports into the GDPR the constitutional proportionality test that we will examine in this article. A number of other articles in the GDPR may also come into play for AML/CFT measures. For example, when banks analyze their customers' transaction data to detect suspicious activity, they are processing customer data for a new, secondary, purpose linked to the detection of criminal offenses. Article 6.4 of the GDPR together with Article 23.1 require that such processing be

[12] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), OJ 2016 L 119/1 (GDPR).

[13] Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA, OJ L 119/89 (Police–Criminal Justice Directive)

[14] Charter, art 52(1).

[15] GDPR, art 23(1).

[16] GDPR, art 23(2).

provided by a law that is 'necessary and proportionate' in a democratic society. Article 9.1 g) of the GDPR on special categories of personal data, which are sometimes involved in AML/CFT, permits processing on the basis of law that is 'proportionate' to the aim pursued. Article 10 of the GDPR permits processing related to criminal offenses when authorized by law providing for "appropriate safeguards for the rights and freedoms of data subjects" (appropriate safeguards is an element of the proportionality test, as we will see below). Article 35.7 b) of the GDPR requires that the data controller assess the 'necessity and proportionality' of high-risk processing operations (AML/CFT processing is generally considered high-risk). Recital 4 of the GDPR says that privacy rights are not absolute and may be balanced against other fundamental rights in accordance with the principle of proportionality. Recital 19 of the GDPR, which refers specifically to AML/CFT, says that Member State law may restrict certain privacy rights when such restrictions are 'necessary and proportionate' in a democratic society. Article 22 and recital 71 of the GDPR, which deal with the creation of automatic profiles (AML/CFT systems create profiles), require 'suitable safeguards' and a law including 'suitable measures to safeguard the data subject's rights and freedoms'.[17] Suffice it to say that proportionality goes to the heart of the GDPR's approach to balancing data protection rights with other competing rights.

3. Simplifying Assumptions

The two-level nature of proportionality can get confusing, which is why we propose to simplify matters by looking at the proportionality test only at the level of the ECHR and the Charter. We will disregard for the sake of simplicity the specific articles of the GDPR dealing with proportionality, and simply assume that those provisions import into the GDPR the proportionality principle as expressed in the case-law of the CJEU and the European Court of Human Rights (ECtHR). We will also assume for the sake of simplicity that the proportionality principle is the same under the Charter and the ECHR, which is not technically correct.[18] But the differences are not important enough to affect our conclusions. We will present a single proportionality test and see if AML/CFT systems satisfy the test.

4. The Three Steps of the Proportionality Test

The proportionality test can be broken down into three steps that must be cumulatively satisfied. The first step consists in asking whether the measure has been provided for in a law or regulation that is precise, understandable and has been adopted pursuant to democratic processes[19]. The second step consists in asking whether the measure pursues a legitimate objective such as a fundamental right or another pressing social need. The third step consists in asking whether the measure is 'necessary in a democratic society.' This third step is the most difficult one, and involves several sub-tests, including whether the measure is 'genuinely effective', whether it is the 'least intrusive measure' available, whether there is a 'fair balance' between the rights at stake, and whether there 'adequate safeguards'.

[17] There is meaningful human review for the generation of SARs. Therefore, it is uncertain whether Article 22 of the GDPR applies.

[18] David Hart, 'Supreme Court on EU and ECHR Proportionality - Back to Basics' (*UK Human Rights Blog*, 27 June 2015, accessed 9 September 2020).

[19] Ahmet Yildirim v. Turkey, App no 31111/10 (ECtHR, 18 December 2012).

The three steps (and related sub-tests) are summarized in Figure 1.

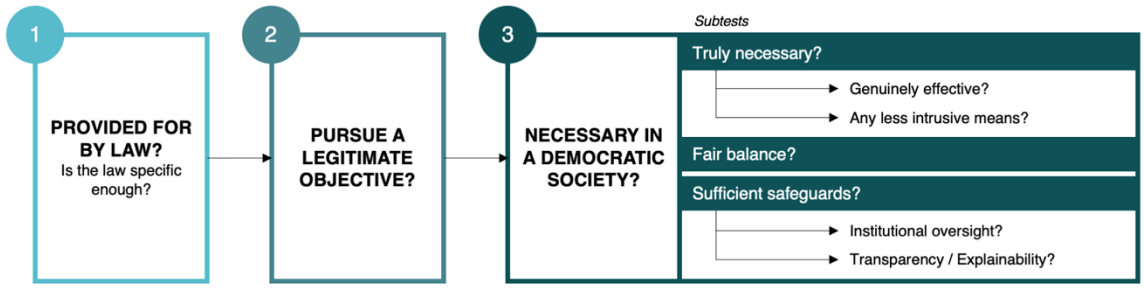

Figure 1 - The proportionality test can be broken down into three tests. To satisfy the proportionality test, the answers to all the questions in the graph must be yes.

5. The CJEU Digital Rights Ireland and Tele2 Sverige - Watson cases

Several CJEU cases shed light on how the proportionality test would apply to AML/CFT systems. In 2014, the CJEU annulled the 2006 EU Data Retention Directive[20] requiring Member States to ensure that telecommunications operators retain traffic and location data for up to two years in order to assist law enforcement authorities and intelligence agencies in crime investigation or terrorism prevention.[21] The CJEU found that the level of interference with privacy rights was particularly high, because the directive required operators to store all traffic and location data of users, regardless of whether the users posed a particular risk, and regardless of whether the data had a particular link to an investigation. The 2006 Data Retention Directive also failed to provide limits on the persons who could access the data. Because of its generality and lack of adequate safeguards, the Data Retention Directive went beyond what was strictly necessary and failed the proportionality test. Two years later, the CJEU declared illegal two national laws containing similar provisions requiring telecom operators to store all connection data, in case the data are needed by law enforcement or national security agencies.[22] The CJEU found that the general and indiscriminate retention of all traffic and location data exceeded what was strictly necessary in a democratic society. In order to be acceptable, *'the national legislation must be based on objective evidence which makes it possible to identify a public whose data is likely to reveal a link, at least an indirect one, with serious criminal offences, and to contribute in one way or another to fighting serious crime or to preventing a serious risk to public security. Such limits may be set by using a geographical criterion where the competent national authorities consider, on the basis of objective evidence, that there exists, in one or more geographical areas, a high risk of preparation for or commission of such offences'*.[23]

[20] Directive 2006/24/EC of the European Parliament and of the Council of 15 March 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC, OJ L 105/54 (Data Retention Directive).

[21] Digital Rights Ireland (n 1).

[22] Tele2 Sverige and Watson (n 2).

[23] Ibid. at para. 111.

This wording suggests that wholesale analysis of all customer transaction data by banks for AML/CFT purposes would be excessive, and that there would have to be some objective link between the population whose data is being analyzed and relevant money-laundering risks. The CJEU also said that there must be safeguards, such as informing the relevant individuals as soon as it is possible to do so without jeopardizing an ongoing criminal investigation. The system would also require institutional oversight. These lessons are directly applicable to AML/CFT.

6. The CJEU Ministerio Fiscal case

Another recent CJEU case[24] involved a Spanish law that required operators to collect and store identification data relating to persons who purchase SIM cards, and make the identification data available to law enforcement authorities if so ordered. The case involved a relatively minor offence, theft of a mobile phone. The CJEU had to determine whether the Spanish law preserved a 'fair balance' under the proportionality test insofar as the law did not only apply to serious crimes, but to more minor offences such as theft of a mobile phone. The CJEU found that the 'fair balance' sub-test was respected because the data involved in this case, only the name and address of the customer, was less intrusive than the detailed traffic and location data involved in the *Digital Rights Ireland* and *Tele2 Sverige-Watson* cases. According to the CJEU, the more detailed and intrusive the data, the more serious must be the crime: '*In accordance with the principle of proportionality, serious interference can be justified, in areas of prevention, investigation, detection and prosecution of criminal offences, only by the objective of fighting crime which must also be defined as 'serious'. By contrast, when the interference that such access entails is not serious, that access is capable of being justified by the objective of preventing, investigating, detecting and prosecuting 'criminal offences' generally*.'[25] When applied to AML/CFT, this means that the analysis of detailed transaction data which entails a high level of interference with individual rights would only be permitted for detection of *serious* crimes.

7. The Canadian PNR opinion

The European Parliament asked the CJEU to review the draft agreement to be entered into between the EU and Canada relating to the exchange of airline passenger name records (PNRs) for the purpose of preventing serious transnational crime and terrorism. PNR legislation is similar to AML/CFT legislation in that it requires private entities, in this case airlines, to share information about their passengers in order to help government authorities detect individuals potentially involved in criminal or terrorist activity. Once collected and transferred to Canada, the passenger data is analyzed by algorithms. In applying the proportionality test, the CJEU first found that the list of PNR data covered by the agreement was not sufficiently precise, because it included vague terms such as "all available contact information", or the term "etc."[26] The Court found that some of the PNR data could reveal religious beliefs or health status, and that the transfer of such sensitive data should be prohibited.[27] As regards the algorithms, the Court said that the models and criteria must be "specific and reliable", "non-discriminatory", and

[24] Ministerio Fiscal, Case C-207/16, 2 October 2018 (ECLI:EU:C:2018:788).

[25] Ibid at paras. 56 and 57.

[26] Canadian PNR Agreement, Opinion 1/15 of the Court (Grand Chamber), 26 July 2017 (ECLI:EU:C:2017:592), at paras. 157 and 158.

[27] Ibid at para. 165.

permit the identification of persons "who might be under a reasonable suspicion of participation in terrorist offences or serious transnational crime".[28] Any positive result must be analyzed by human reviewers before action is taken, and the data and algorithms, says the Court, must be reviewed regularly to ensure they are non-discriminatory, strictly necessary, and sufficiently reliable.[29] Finally, the Court said that passengers must be individually notified when the analysis of their data shows a positive risk, and gives rise to further action. The notification must occur as soon as it is possible to do so without jeopardising an ongoing investigation.[30]

8. The Quadrature du Net case

After the 2016 *Tele2 Sverige – Watson* case, several countries, including France, maintained their data retention laws, arguing that national security falls outside the scope of EU law, and that the *Tele2-Sverige – Watson* decision did not take sufficient account of the threat of terrorism. In its 6 October 2020 decision, the CJEU found that national security does not fall outside the scope of EU law when the measure involves processing of personal data by private entities such as telecommunications operators.[31] Second, the Court found that general and indiscriminate storage of traffic and location data by telecommunications operators can be allowed for a temporary period during if there exists a serious threat for national security.[32] Under the 'fair balance' test, national security threats can justify more intrusive measures than do other crimes, including serious crimes. For serious crimes, retention of traffic and location data cannot be general and indiscriminate, but must target only certain categories of the population, such as persons already identified as potentially involved in criminal activity, or persons in the immediate proximity of a train station or airport.[33] However, any segmentation must be non-discriminatory.[34] The Court said that the general and indiscriminate retention of IP addresses could be justified for serious crimes because the IP address is less intrusive than all traffic and location data.[35] For non-serious crimes, only the name and address of the subscriber may be retained.[36]

In the second part of the *Quadrature du Net* case, the Court examined the use by French intelligence agencies of algorithms to analyze traffic data in real time to detect potential terrorist threats. The Court concluded that such monitoring could be justified for limited times during which there exists a serious threat of terrorism. However, the threat must be "genuine and present or foreseeable". [37] The use of the algorithm must be surrounded by safeguards, including ongoing institutional review to ensure that the system is non-discriminatory, that the data and models are reliable, and are limited to what is strictly necessary for preventing terrorist activities presenting a serious risk for national security.[38] Any recommendation made by the algorithm must be reviewed by human analysts before further action.[39] In the last part of the *Quadrature du Net* case, the Court found that a separate

[28] Ibid at para. 172.
[29] Ibid at paras. 173 and 174.
[30] Ibid at para. 224.
[31] Quadrature du Net, Joined Cases C-511/18, C-512/18, and C-520/18, 6 October 2020 (ECLI:EU:C:2020:791) paras. 103 and 104.
[32] Ibid at para. 137.
[33] Ibid at para. 150.
[34] Ibid.
[35] Ibid at para. 155.
[36] Ibid at para. 157.
[37] Ibid at para. 177.
[38] Ibid at para. 182.
[39] Ibid.

French law requiring hosting providers to retain detailed logs and identification information about persons contributing content on the platform violated European proportionality rules.[40] The Court referred to the proportionality rule in Article 23.1 of the GDPR read in conjunction with Article 52(1) of the Charter, showing that the two instruments, GDPR and Charter, work hand-in-hand when it comes to proportionality.[41]

9. The Netherlands Social Security Fraud case

The proportionality test under the ECHR was applied by a lower court in the Hague,[42] Netherlands, with regard to an algorithm used by the social security authorities to detect potential social security fraud. The Netherlands adopted a law authorizing the government to collect data from various government databases in order to create risk profiles showing the likelihood that a given individual is cheating on social security benefits. The District Court of the Hague found that the measure was 'provided for by law' (the first test) because the law was sufficiently detailed, and that the objective of fighting welfare fraud was a compelling social interest, satisfying the second test. When it applied the 'necessary in a democratic society' test, the court found that the system was surrounded by inadequate safeguards and in particular that it lacked transparency. First, individuals were not informed of the existence of the system, nor given a general understanding of how data about them were being used. Second, neither the court nor individuals targeted by algorithmic risk profiles were able to understand how the proprietary model operated and how it reached a particular score. This kind of understanding (often referred to as explainability) is necessary to permit individuals to defend themselves, and to permit courts to verify the presence or absence of discrimination.[43] Because the system lacked this transparency, it failed the proportionality test. The risk score reports did not lead to automatic decisions and therefore did not fall under Article 22 of the GDPR. The automatic reports were reviewed by employees who then decided whether to investigate. The Netherlands court did not examine the case under the GDPR or the Police-Criminal Justice Directive but relied solely on the proportionality text of the ECHR. There are numerous similarities between the Netherlands social security fraud case and AML/CFT measures. One of the functions of bank monitoring systems is to build risk profiles, similar to the profiles created by the Netherlands social security authorities using the SyRI algorithm. In both situations, human reviewers evaluate the output of the algorithm, and decide whether to take further action. Yet in the Netherlands case, the existence of human review did not remove the need for appropriate human rights safeguards surrounding the algorithm itself, nor the need for a full application of the proportionality test under the ECHR.

---

[40] Ibid at para. 212.
[41] Ibid at para. 202.
[42] Netherlands Legal Committee for Human Rights and others v. The Netherlands, C-09-550982-HA ZA 18-388 ECLI:NL:RBDHA:2020:865 5 February 2020 (The Hague District Court, NL).

[43] Ibid; Valérie Beaudouin, Isabelle Bloch, David Bounie, Stéphan Clémençon, Florence d'Alché-Buc, James Eagan, Winston Maxwell, Pavlo Mozharovskyi, Jayneel Parech, 'Flexible and Context-Specific Explainability: a Multidisciplinary Approach', Télécom Paris – Institut Polytechnique de Paris Working Paper, 23 March 2020, SSRN: https://ssrn.com/abstract=3559477.

III. AML/CFT MEASURES AND THE IMPACT OF AI

Before applying the proportionality test, let us present AML/CFT systems in more detail.

1. AML/CFT Regulations

AML/CFT is regulated by a series of European directives[44] and countless national laws and regulations. AML/CFT laws impose a number of duties, including 'know your customer' (KYC) verifications, identification of politically exposed persons (PEP), monitoring to detect transactions to high-risk countries, and reporting suspicious or unusual activity based on the organization's knowledge of the customer.[45] Banks and insurers must also implement processes to ensure that no payments are made to persons or entities subject to international sanctions.[46]

AML/CFT laws are unusual because they require private-sector actors actively to look for suspicious activity by their customers and report the activity to FIUs without informing their customers. Laws requiring private sector entities to ferret out criminal activity by their customers are rare. Compliance laws (e.g. anticorruption) often require companies to put into place measures (e.g. whistleblowing) to detect and report criminal activity *within the company*, but laws requiring companies to put into place measures to detect and report criminal activity *by customers* are more unusual. AML/CFT laws put bank compliance departments in the uncomfortable position of becoming government informants, a role that seems opposed to the bank's traditional duty of secrecy to its customers.[47] There are good reasons why AML/CFT law put this obligation on banks, and more generally on financial entities like payment or credit institutions and insurance companies. The level of sophistication of money laundering is so high and so dependent on the context of the customer relationship that only financial institutions have access to the information that would permit the identification of suspicious patterns. Second, financial institutions have strong economic incentives to turn a blind eye to criminal funds.[48] If financial institutions were not required by law to detect and report criminal activities, economic incentives would push such institutions to accept deposits of criminal funds in all but the most manifestly illegal cases.

The third and fourth European directives on AML/CFT changed the approach from a rules-based "check the box" approach to a risk-based approach, leaving banks free to develop detection rules to fit their clientele and business, using 'evidence-based decision-making in order to target the risks of money laundering and terrorist financing'.[49] The most relevant risk factors for a financial institutions are its customers, countries of operation, types of products

[44] Directive (EU) 2018/843 of the European Parliament and of the Council of 30 May 2018 amending Directive (EU) 2015/849 on the prevention of the use of the financial system for the purposes of money laundering or terrorist financing, and amending Directives 2009/138/EC and 2013/36/EU, OJ L 156/43; Directive (EU) 2018/1673 of the European Parliament and of the Council of 23 October 2018 on combating money laundering by criminal law OJ L 284/22.

[45] TRACFIN (n 6); Laurent Dupont, Olivier Fliche and Su Yang, 'Governance of Artificial Intelligence in Finance' ACPR discussion document (11 June 2020).

[46] Ibid.

[47] Matthew Hall, 'An Emerging duty to Report Criminal Conduct: Banks, Money Laundering and the Suspicious Activity Report' 84 Kentucky Law Journal 643 (1996).

[48] Pellegrina and Masciandaro (n 3)

[49] Directive 2015/849 (n 44) recital 22.

or services and distribution channels. Accordingly, each reporting entity must determine the appropriate risk profile for each of its customers and transactions, and apply the appropriate level of scrutiny (simplified, standard, reinforced or complementary).[50] Based on the financial institution's knowledge of its customer, and the relevant risk score, the institution must detect and report abnormal or suspicious transactions.

2. Transaction Monitoring Step-by-Step

Figure 2 explains the usual frameworks deployed by banks to comply with their AML/CFT obligations.

Every incoming transaction is first screened by an automated review system that uses the information gathered on the client profile through KYC and client due diligence (CDD), the transaction details, information concerning watch-listed countries and other triggers, such as key words, for alerts. Other data sources such as market activities, trade-based data or even social media and news feed may be used, subject to limits imposed by data protection laws. This automated review system will decide whether to generate an alert for a given transaction, either based on deterministic 'if-then' rules or based on more sophisticated AI models. All the generated alerts then follow a two- or three-step review by compliance experts who decide either to escalate the alert to the next review level or to close it. Alerts that go through all review steps are then consolidated into SARs (suspicious activity reports) that are forwarded to the FIUs (financial investigation units) for investigation. Banks often terminate the accounts of customers subject to an SAR, but are not allowed to inform the customer that an SAR has been generated.[51] The FIUs that receive the SARs generally provide no feedback about individual SARs, and investigate only a small portion of the SARs they receive.

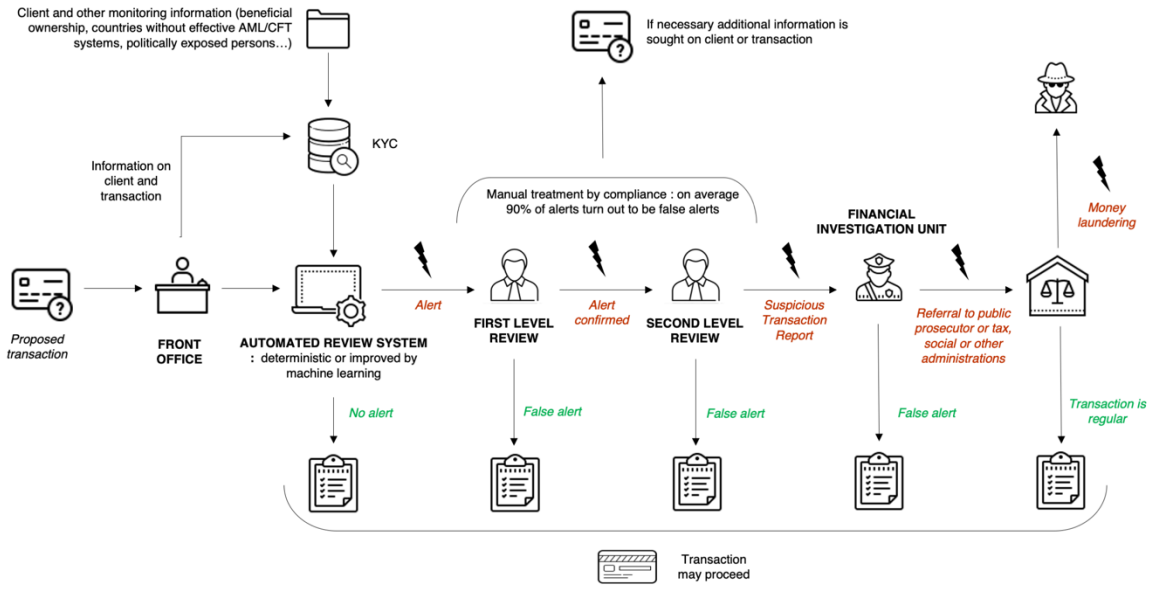

Figure 2 - Diagram representing the Anti-Money Laundering Process

[50] Monetary and Financial Code *(Code Monétaire et Financier*) (FR) art 561; Department of Financial Services Regulations (NY) Part 504: Transaction Monitoring and Filtering Program Requirements and Certifications.

[51] ACPR, 'Principes d'application sectoriels relatifs aux obligations de lutte contre le blanchiment des capitaux et le financement du terrorisme dans le cadre du droit au compte', ACPR Explanation Document, 25 April 2018, para. 42 ; N v. Royal Bank of Scotland, [2019] EWHC 1770 (Comm) (England and Wales High Court, 8 July 2019).

Bank monitoring systems perform several upstream tasks to decide whether or not to generate an alert for a particular transaction. The first task consists of dividing customers into specific risk segments based on common behavioral attributes. This process is usually driven by specific industrial knowledge and static (and sometimes outdated) rules and thresholds applied to the industry of the clients, the client's size, countries of operation, etc. After that, the review system will execute an anomaly detection task to see if the incoming transaction falls outside the 'normality' scenario for that customer based on its segment profile. This is usually set up by creating thresholds that represent a certain level of deviation from the typical client behavior pattern in the segment. In some banks, alerts are then prioritized for human review. Alerts generated by the system are ranked according the likelihood that they actually correspond to a real risk. This is usually done by calculating a risk-score based on rules of amount and type of transaction. For example, transactions involving a round amount, for example €1,000, are often considered more suspicious than transactions involving a non-round amount such as €1,293.88.[52] Lastly, visualization tools gathering all useful data for the human review can help analysts see all the data relevant for the alert. However, these visualization tools are very recent and in many cases gathering all relevant data for human review constitutes a challenge for large banks, where information concerning a specific alert is fragmented, incomplete, and/or inconsistent. Figure 3 presents the different tasks involved in the transaction monitoring system to generate alerts.

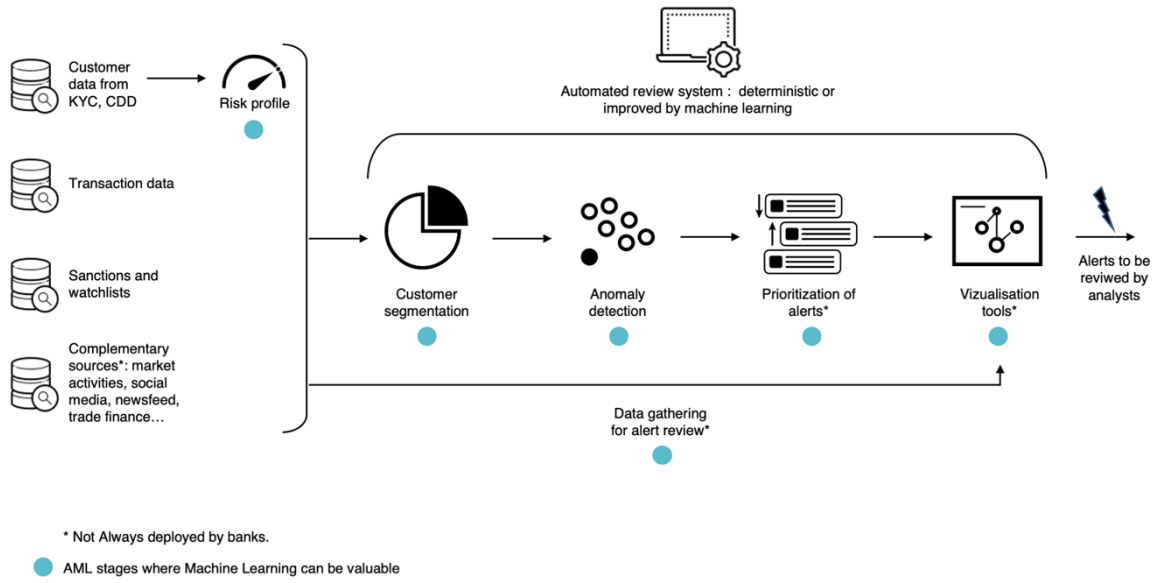

Figure 3 - Diagram representing the potential applications of AI in the Automated Review System of AML processes.

3. The Shortcomings of Current AML/CFT Systems

Although machine learning is gradually being adopted by financial institutions, simple rule-based models and manual filtering of alerts still prevail today. It is estimated that about 50% of banks still use manual methods to

[52] Federal Deposit Insurance Corporation, DSC Risk Management Manual of Examination Policies, Bank Secrecy Act, Anti-Money Laundering, Office of Foreign Assets Control, Section 8.1-40.

comply with anti-money laundering requirements.[53] Yet rule-based AML/CFT mechanisms exhibit significant limitations. Rules are mainly deterministic, static, and difficult to maintain manually. As a result, AML/CFT systems are congested with redundant and obsolete rules which lead to inappropriate alerts, and a huge and unnecessary workload for compliance teams. To cope with the increasing number of alerts within the constraints of human analysis, financial institutions are hiring more and more compliance officers, which significantly increases costs.

4. The Effect of AI on AML/CFT

In recent years, banks have been testing AI to assist analysts in highly repetitive AML/CFT compliance reviews on the one hand, and to improve the performance AML/CFT frameworks to fight crime on the other hand. As shown in Figure 3, machine learning can be introduced in several parts of the monitoring system to improve efficiency. More specifically, we can distinguish five ways in which AI is transforming current AML/CFT processes: automating data collection, redistributing resources through client risk scoring and alert prioritization, leveraging linkage analysis, improving segmentation, and improving anomaly detection either through identifying known suspicious patterns or by discovering new patterns. These mechanisms are briefly described below.

- **Automating data gathering**

Collecting data to fuel AML/CFT systems is a tedious and error-prone task. Using machine learning, financial institutions can detect errors in the data, tap into external data sources, read unstructured data otherwise too burdensome to cipher manually, and ultimately enrich the context of the alerts. For instance, OCR (Optical Character Recognition) automatically extracts data from paper-based documents, such as trade finance documents, which can represent essential sources to uncover criminal activities. NLP (Natural Language Processing) methods are helpful to provide the context and sentiment of a newspaper article or blog post, for example, which may help a human compliance analyst characterize a risk.

- **Allocating surveillance resources**

A popular application of AI is to optimize alert processing or improve client risk scoring. Alerts, transactions, or clients are prioritized for review or surveillance. By redirecting alerts directly to the right level of human review, or by sorting out alerts by order of relevance, significant amounts of time can be saved and compliance officers' workload is reduced.[54] By issuing client risk scores with machine learning algorithms, surveillance resources can be smartly allocated to focus on the most suspicious profiles.[55] This range of AI use-cases is gaining popularity within banks because their impact on final AML/CFT decision-making is minimal. Human compliance officers remain in control of final decisions. The explainability of these AI-based based routing tools is less of an issue.

[53] ACAMS - LEXIS NEXIS, 'Current Industry Perspectives into Anti-Money Laundering Risk Management and Due Diligence' (2015).

[54] Dupont, Fliche and Yang (n 45).

[55] Su-Nan Wang and Jian-Gang Yang, 'A Money Laundering Risk Evaluation Method Based on Decision Tree', *2007 International Conference on Machine Learning and Cybernetics* (IEEE 2007).

- **Leveraging link analysis**

Social network analysis is an emerging field in AML/CFT useful both for improving the visualization of information during investigations and for automatically analyzing inferences between parties. Graph structures are very powerful in the fight against money laundering because they capture the essence of money laundering schemes i.e. cash flow relationships. This relational information can then be leveraged to improve anomaly detection models.[56]

- **Improving segmentation**

Unsupervised learning can be particularly useful to drive more intelligent segmentation by helping compliance experts detect behavioral patterns otherwise invisible during human review. Classical clustering techniques like K-means algorithm and PCA can enhance customer segmentation and increase the ability to identify abnormal behaviors that diverge from the 'normal' scenario for each cluster. These AI-based techniques are also promising for partially automating the segmentation model and making it more resistant to variations in the data.

- **Improving anomaly detection**

Self-learning models for anomaly detection are increasingly gaining attention and may one day replace or complement banks' programmed-based anomaly detection models, i.e. based on simple rules-based and statistical-based. Due to the great range of applications in real-world situations, anomaly detection is a well-studied problem, which spawned profuse research solutions. According to the availability of labelled data, anomaly detection models can either serve to (1) detect suspicious patterns based on historical ones, or (2) detect anomalous behavior regardless of past suspicious cases, i.e. infer new abnormal behaviors.

- **Detection of known suspicious patterns**

Through supervised learning, models can be trained to distinguish suspicious from normal transactions using previous examples of suspicious activity reports.[57] These algorithms can enhance the anomaly detection phase shown in Figure 3. However, the dependency of supervised algorithm on training data creates challenges. The first is the absence of ground truth on what was in fact a suspicious transaction. The bank will only know what it previously labelled as suspicious, not what actually turned out to be suspicious after investigation by the FIU, or what suspicious activities were entirely missed (false negatives). Learning from prior labels, the algorithm will only be as good as the previous bank compliance team in identifying suspicious transactions. The second challenge is the imbalance in training examples, the proportion of suspicious transactions representing only a tiny fraction of the volume of total transactions in the training data set.

- **Detection of new suspicious patterns**

[56] Mark Weber, Giacomo Domeniconi, Jie Chen, Daniel Karl Weidele, Claudio Bellei, Tom Robinson and Charles Leiserson, 'Anti-Money Laundering in Bitcoin: Experimenting with Graph Convolutional Networks for Financial Forensics' [2019] arXiv:1908.02591; David Savage, Qingmai Wang, Pauline Chou, Xiuzhen Zhang and Xinghuo Yu, 'Detection of Money Laundering Groups Using Supervised Learning in Networks' [2016] arXiv:1608.00708.

[57] Jun Tang and Jian Yin, 'Developing an Intelligent Data Discriminating System of Anti-Money Laundering Based on SVM', *2005 International Conference on Machine Learning and Cybernetics* (IEEE 2005).

Since labels for normal instances are much more available than labels of suspicious cases, semi-supervised learning only uses negative labels, i.e. all the non-suspicious transactions, to infer positive ones, i.e. those that fall outside the learned "normal" type. This range of methods are popular in the anomaly detection domain where historical data of known anomalies is not easily available. [58]

Unsupervised algorithms do not rely on training data and are therefore able to detect anomalous behavior regardless of past suspicious cases. Unsupervised learning is therefore useful to discover new money laundering patterns otherwise too complicated for humans to notice. They may notice a correlation between an account opened in Hong Kong, a transaction in Nigeria and a music festival in France, a correlation human analysts would never detect. Academics and banks are exploring unsupervised learning approaches for this purpose.[59] Many of the machine learning approaches explored for AML/CFT are also being examined by intelligence agencies to identify potential terrorist threats.[60]

5. AI Used by Banks Today

Banks currently use AI to take some of the burden off human reviewers, for example by helping to identify false positives generated by the monitoring system, and prioritizing alerts, sending some alerts directly to the level 2 or 3 review team. Machine learning methods have proven themselves to be valuable by reducing the false positive rate by 20 to 30%.[61] Same banks are also using AI to help gather context data, and to present graphs of customer and transaction relationships to help human reviewers. When it comes to using AI to improve segmentation and detect suspicious patterns, banks are generally in experimental phase only, due to regulatory uncertainties and operational complexities.

IV. APPLYING THE PROPORTIONALITY TEST TO AML/CFT TRANSACTION MONITORING

1. Do AML/CFT Systems Interfere with Fundamental Rights?

A threshold question is whether there is an interference with fundamental rights. Even without machine learning, AML/CFT processes are particularly intrusive: they analyse all transaction data without exception – similar to

[58] Drausin Wulsin and others, 'Semi-Supervised Anomaly Detection for EEG Waveforms Using Deep Belief Nets', *2010 Ninth International Conference on Machine Learning and Applications* (2010).

[59] Nhien An Le Khac and M Tahar Kechadi, 'Application of Data Mining for Anti-Money Laundering Detection: A Case Study', *2010 IEEE International Conference on Data Mining Workshops* (IEEE 2010); Ebberth L Paula, Marcelo Ladeira, Rommel N. Carvalho, Thiago Marzagao, 'Deep Learning Anomaly Detection as Support Fraud Investigation in Brazilian Exports and Anti-Money Laundering', *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)* (IEEE 2016); Rui Liu, Xiao-long Qian, Shu Mao, Shuai-zheng Zhu, 'Research on Anti-Money Laundering Based on Core Decision Tree Algorithm', *2011 Chinese Control and Decision Conference (CCDC)* (2011); Tang and Yin (n 59) .

[60] Damien Van Puyvelde, Stephen Coulthart, M. Shahriar Hossain, 'Beyond the buzzword: big data and national security decision-making' 93 International Affairs 1397 (2017).

[61] Mark Weber, Jie Chen, Toyotaro Suzumura, Aldo Pareja, Tengfei Ma, Hiroki Kanezashi, Tim Kaler, Charles E. Leiserson, Tao B. Schardl 'Scalable Graph Learning for Anti-Money Laundering: A First Look' (2018).

CJEU's *Tele2 Sverige - Watson* case – and they create risk profiles, similar to the Netherlands social security fraud case. The objective of processing is to detect and report criminal activity, which normally requires special safeguards under Article 10 of the GDPR.[62] Political, religious or health information may be inferred from bank transaction data, and in some cases AML/CFT actually requires processing of ethnic or religious data protected under Article 9 of the GDPR. AML/CFT systems create risk profiles, a high-risk processing under Article 35 and Recital 91 of the GDPR. As we saw from the Netherlands social security fraud case, the use of machine learning techniques can exacerbate the problem because of the opacity of the algorithms. AML/CFT systems interfere with privacy rights, but they may also interfere with the right to non-discrimination, because the system may generate more alerts for certain nationalities, or other groups of the population. They may also interfere with the right to free expression: if customers feel that their transaction data are being monitored, they may refrain from making payments to certain charities or subscribing to certain publications.[63] Finally, AML/CFT systems may affect the right to an effective remedy because individuals will not be informed of the processing, and even if they are informed, they may be unable to challenge the system due to its opacity. The very fact that these data are processed creates a high level of interference with fundamental rights even if bank customers never suffer any actual direct consequences.[64]

2. 'Provided by Law'

The first proportionality test, 'provided by law', requires a law or regulation describing in reasonable detail the type of tool required, the type of data processed, and the safeguards surrounding the use of the tool, so that citizens and other political stakeholders can understand and react, holding the government accountable.[65] Current monitoring requirements are described in general terms by regulations, such as those of the French Monetary and Financial Code or the guidelines from ACPR and TRACFIN,[66] but the specificity is left to the banks, who develop complex systems to satisfy the expectations regulatory authorities. There is considerable leeway in the interpretation of AML/CFT rules on transaction monitoring, the main requirements being *first* that the systems be 100% effective in blocking payments toward entities covered by international sanctions, and *second* that the systems designed to detect suspicious transactions take into account precisely defined risk scenarios. In practice, regulators will expect systems deployed by banks to conform to best industry practice, which will constantly evolve, as AML/CFT vendors propose ever more sophisticated and intrusive systems, without any clear upper limits set by law.[67] This upward spiral is fueled by banks' impression that investing in more technology and more

[62] Article 10 of the GDPR provides that processing of personal data relating to offences shall be carried out only under the control of official authority or when the processing is authorized by Union or Member State law providing for appropriate safeguards for the rights and freedoms of data subjects.

[63] La Quadrature du Net (n 31) at para. 113.
[64] Ibid at para. 115.
[65] Scarlet v. SABAM, Case n° C-70/10, Opinion of Advocate General (14 April 2011), at para 95; Canadian PNR Agreement (n 26) at para. 155.

[66] ACPR and TRACFIN, 'Publication des Lignes directrices conjointes de l'Autorité de contrôle prudentiel et de résolution et de TRACFIN sur les obligations de déclaration et d'information à TRACFIN' (2018).

[67] AML laws state that TMS systems must always comply with the GDPR, but the GDPR does not set hard limits. Banks justify their TMS under the 'required by law' legal basis of the GDPR. If banks believe that the regulator requires more intrusive processing to be compliant with law, the processing will appear compliant under the GDPR. The reasoning becomes circular.

compliance employees will reduce the risk of regulatory sanctions. AML/CFT sanctions, which can include revocation of a banking license, are perceived by banks as bigger threats than data protection sanctions. This perception can lead banks to err on the side of deploying ever more sophisticated compliance technology.
The lack of specificity of current AML/CFT regulations is striking when compared with regulations imposing data storage obligations on telecom operators[68], the Canadian PNR agreement, or the Netherlands law authorizing the use of algorithms to predict social security fraud, all of which provide a high degree of specificity on the data that may be processed and the safeguards that must be applied.[69] More specific regulation on AML/CFT systems would be required not only under the 'provided by law' branch of the proportionality test, but also from the standpoint of compliance with the GDPR. Article 23 of the GDPR contains its own version of the 'provided by law' rule, stating that privacy rights may be interfered with to the extent necessary for the prevention, investigation, detection or prosecution of criminal offences, but only in a law balancing the competing rights in accordance with the proportionality test. Article 23.2 of the GDPR requires that any such law contain specific provisions on the categories of data that may be processed, the purposes of processing, the maximum storage periods, and the risks to the rights and freedoms of individuals. These details are missing from current AML/CFT laws.

One of the criticisms made by data protection authorities of current AML/CFT systems is that regulators encourage banks to go beyond what is strictly required by law, a phenomenon known as 'gold plating'.[70] Banks may not go beyond what is strictly required by law yet current AML/CFT law is unclear on what is required. This puts banks in the untenable position of determining what level of monitoring goes far enough to satisfy imprecise AML/CFT laws, yet not so far as to violate the GDPR. Under the current system, application of the proportionality test has been delegated to banks, whereas the balancing of competing societal rights should be made by the legislature and government, as required by Article 23 of the GDPR.

To cure this defect, precise requirements on transaction monitoring for AML/CFT should be defined in a law or in a decree subject to constitutional review by an institution such as the Conseil Constitutionnel or Conseil d'État in France. The law or decree would define the extent to which machine-learning algorithms may be used, and what safeguards should be attached. One objection to publishing requirements in a law or decree is that doing so would help money launderers game the system to avoid detection. However, this problem also exists in cybersecurity regulations, and can be overcome by describing the overall functionalities of the system banks must deploy in a public regulation, the categories of data that may be used, and keeping the specific details of the algorithm in a confidential governmental ruling. In France, the cybersecurity requirements for the banking sector are published in a regulation, but the specific requirements applicable to a particular bank of critical national importance will be kept in a classified ruling.[71] It would be possible to use this approach for AML/CFT, which

[68] For example, the French Decree n° 2006-358 of 24 March 2006 which describes exactly what data must be retained by telecommunication operators in order to assist with potential law enforcement investigations.

[69] Netherlands Legal Committee on Human Rights vs. The Netherlands, (n 42).

[70] Article 29 Working Party, 'Opinion on data protection issues related to the prevention of money laundering and terrorist financing' (WP 186, 13 June 2011).

[71] French Defense Code, art R1332-41-1, which provides that the details of cybersecurity rules for a particular installation of vital importance may be issued in a confidential ruling.

would contribute greatly to curing the 'provided by law' problem.

3. 'Pursuing a Legitimate Objective'

The second proportionality step, pursuing a legitimate objective and pressing social need, is easy to satisfy for AML/CFT measures, because AML/CFT is part of the broader fight against serious crime and terrorism, which the CJEU and the ECtHR case law recognize as a legitimate objective. Recital 19 of the GDPR expressly mentions AML/CFT as a legitimate case in which privacy rights can be interfered with, subject to proportionality.

4. 'Necessary in a Democratic Society'

4.1. 'Genuinely effective'

A key element of the third step of proportionality is whether a given approach is genuinely effective in achieving the desired objective and is the least intrusive means.[72] But how do you measure effectiveness for AML/CFT? As noted above, banks currently measure effectiveness based only on their own past reports. For banks, alerts are effective if they resemble situations that bank compliance teams previously labelled as suspicious. But banks do not know whether their own reports of suspicious activity are useful to law enforcement. Law enforcement authorities will not confirm or deny whether what appeared suspicious to the bank was in fact linked to a criminal offense. This leaves banks in the dark. On one level, this is understandable: the existence of criminal activity is highly sensitive information that should not be shared with banks.[73] But on another level, the absence of feedback means that banks are labelling people as potential criminals without any verification of whether the suspicions are justified. The banks' lists of suspicious transactions also remain hidden from the customer, and may lead the bank to terminate the customer relationship entirely, without actually knowing if the suspicion was justified.[74] Without feedback, banks are unable to assess whether their reports are genuinely effective in helping to stop crime. Banks try to evaluate effectiveness of their systems using a proxy, which is the number of machine-generated alerts that are subsequently 'converted' by bank reviewers into SARs. But this is an imperfect metric because it does not capture whether a given approach really helps law enforcement authorities identify and confiscate criminal funds and prosecute criminals, which is the objective of the law. According to Europol,[75] FIUs investigate only 10 percent of the SARs they receive[76]. This suggests that 90% of SARs sent to FIUs serve little or no law enforcement purpose. Worse, they may lead to perfectly innocent bank customers having their accounts terminated for no reason other than an unconfirmed suspicion based on a bank process that the customer does not understand. Adding machine learning does not change the nature of the problem. In the absence of reliable metrics showing the utility of the SARs for the pursuit of criminals, it is impossible to affirm that AML/CFT methods, whether

[72] European Data Protection Supervisor (EDPS), 'Assessing the Necessity of Measures That Limit the Fundamental Right to the Protection of Personal Data: A Toolkit' (11 April 2017).

[73] European Data Protection Supervisor (EDPS), 'Opinion 5/2020 on the European Commission's action plan for a comprehensive Union policy on preventing money laundering and terrorism financing (23 July 2020), para 43.

[74] John Binns, 'Customers with blocked accounts are the ones being stung in the fight over money laundering', Euronews.com (28 January 2020, accessed 9 September 2020).

[75] Europol (n 5).

[76] The percentage in France is nearer 20 percent. TRACFIN (n 6).

based on traditional rule-based models or on machine learning, are 'genuinely effective' for purposes of apprehending criminal funds.

Curing this problem requires developing a way to evaluate the utility of the information provided by banks to FIUs. One approach would be for FIUs to attribute scores to SARs to show their actual utility in confiscating criminal funds and prosecuting money launderers. The score should capture the social utility of the SAR using two dimensions: the seriousness of the crime on the one hand, and the relative utility of the SAR in stopping the crime on the other hand. For example, a SAR that made a major contribution to confiscating funds bound for a terrorist organization would score highly on both the seriousness of the crime score and on the level of contribution of the SAR score. That SAR would be highly valuable and receive a high score in light of the public interest objective of AML/CFT. By contrast, a SAR that reveals a possible tax fraud by a local bistro would have a lower score on the seriousness of the crime scale. A SAR that is a false positive or lacks enough information to be useful might have a zero or even negative score. The quality metric could be systematically communicated to banks and/or to a supervisory authority in charge of overseeing the AML/CFT system[77] in order to ensure that the right balance is struck between AML/CFT effectiveness on the one hand, and interference with fundamental rights on the other. The feedback would help train machine learning algorithms to better detect suspicious activity, particularly suspicious activity likely to generate a high-scoring SAR. Reinforcement learning might be used, much like rewarding a police dog that successfully identifies drugs or explosives. Providing detailed law enforcement feedback to banks creates legal issues, because criminal investigations or national security matters are covered by secrecy. Several solutions may exist, one being to designate an employee in the bank who has national defense secret certification. In France this approach is already used for exchanging sensitive cybersecurity information between banks and cybercrime teams in the government. Another approach would be to provide detailed feedback to certain members of the supervisory authority, and less detailed feedback to the banks. Whatever the structure, a solution must be found so that regulators, courts and citizens have confidence that bank detection systems are generating reports that have an important impact on stopping serious crime. The current system, where 90% of notifications are never investigated, would never pass the 'genuinely effective' test.

4.2. 'Less intrusive means'

Related to the 'genuinely effective' question is the question of whether there are less-intrusive means reasonably available that would achieve the same level of effectiveness while creating a lower impact on fundamental rights.[78] This would require analyzing several solutions to addressing the problem and comparing their relative effectiveness and impact on fundamental rights. The optimal scenario would be the one that maximizes both effectiveness and protection of fundamental rights.[79] But as noted above, maximizing effectiveness requires some reliable way of measuring effectiveness, which currently does not exist for AML/CFT since the SARs sent to FIUs fall into a kind of black hole. When it comes to measuring the impact on fundamental rights, qualitative scoring methods exist to evaluate the comparative impact on fundamental rights. In addition to the impact on

[77] On the supervisory authority, see Section 4.4 below.
[78] EDPS (n 77).

[79] The Queen v. Ministry of Agriculture, Fisheries and food, ex parte FEDESA and others, Case C-331/88, ECR 1990 I-04023 ECLI:EU:C:1990:391 (13 November 1990); Charlotte Bagger Tranberg 'Proportionality and data protection in the case law of the European Court of Justice', 1 International Data Privacy Law 239 (2011).

privacy, other fundamental rights would also have to be assessed, including the right to non-discrimination (does the tool create clusters based on the geographic origin of names?), and the right to an effective remedy (the ability to challenge algorithmic outcomes). Each approach to monitoring would be scored in terms of effectiveness and in terms of fundamental rights protection, and in very rough terms, the approach with the highest aggregate score (effectiveness score plus fundamental rights protection score) would emerge as the approach most likely to satisfy the 'least intrusive means' test.[80] We are not aware of a fundamental rights impact assessment of this kind ever being done for AML/CFT, in large part because of the lack of data on the effect of SARs on reducing crime. With the introduction of AI, an impact assessment will be critical, both under Article 35 of the GDPR and under the proportionality test.

4.3. 'Fair balance'

The 'fair balance' enquiry consists of ensuring that weights of the respective rights (fighting crime vs. privacy) are roughly comparable, and that one right does not extinguish the other. The fair balance enquiry ensures that the essence of each right is preserved. In the CJEU's *Ministerio Fiscal*, the court said that intrusive data processing can be justified only for detection of *serious* crimes. In the *Quadrature du Net* case, the CJEU made a further distinction between serious threats to national security, and other crimes including serious crimes. The scope of AML/CFT has been considerably broadened to cover not only drug trafficking, trafficking in human beings, corruption, terrorism (all of which are serious crimes and in some cases threats to national security), to any crime with a potential sentence of more than one year, including tax fraud. The crimes covered by AML/CFT do not reflect the three-level hierarchy fixed by the CJEU between serious threats to national security, serious crimes, and non-serious crimes. Potentially tax fraud by the local bistro could be covered by AML/CFT, which raises the question of whether monitoring systems should look only for serious crimes in line with the CJEU's reasoning in *Ministerio Fiscal* and whether the most intrusive analysis should be reserved solely for threats to national security as indicated in the *Quadrature du Net* case. Another important consideration in the 'fair balance' enquiry is whether AML/CFT automatically fails because of general and indiscriminate processing of all customer data, a fatal defect in the *Tele2 Sverige – Watson* case, but which was nevertheless permitted in *Quadrature du Net* for serious threats to national security. AML/CFT systems analyze all transaction data of all customers in a general and indiscriminate fashion, exactly what the CJEU said was illegal in its case law involving telecommunications operators, even for serious crimes. This problem seems difficult to cure for any AML/CFT system, whether or not it uses AI. It also poses a Catch 22 dilemma: the CJEU says that general and indiscriminate processing is prohibited, and that processing of personal data for law enforcement should be targeted based on risks. However, in AML/CFT, the only way to focus on risks is to define the risky customers groups, which in turn requires processing of all customer data. To overcome this dilemma, it may be possible to divide processing into two phases: systematic processing is in the first phase to identify the risky clusters that merit close monitoring, and then more intrusive monitoring of the risky clusters in a second phase. This approach would be consistent with the risk-based approach in the Third and Fourth AML Directives and in the CJEU's *Tele2 Sverige-Watson* and *Quadrature du Net* cases.

[80] This is an oversimplification. Among other things, each fundamental right would have to receive a minimum score, ensuring that the 'essence' of the relevant right is preserved as well. For a discussion of how such a scoring mechanism might work, see Winston Maxwell, 'Smart(er) Internet Regulation Through Cost-Benefit Analysis', Presses des Mines (2017).

4.4. 'Adequate safeguards'

As part of the 'necessary in a democratic society' test, courts will verify that the measure is surrounded by adequate safeguards. The types of safeguards are generally those specified by the GDPR: data minimization, limitation on who has access to data, limiting the storage time, security measures, transparency, and accountability. For government-imposed measures, such as police or national security surveillance, courts focus on the existence of effective institutional oversight. On the "adequate safeguards" front, existing AML/CFT measures suffer from three weaknesses.

The first is transparency: individuals must be informed that they have been singled out as posing a risk of criminal activity. However, AML/CFT legislation prohibits this, for the understandable reason that tipping off an individual may compromise a criminal investigation. The problem of informing the targets of surveillance plagues many surveillance systems, particularly in national security and anti-terrorism contexts, where secrecy is essential. The middle ground accepted by courts[81] is that individuals must be informed as soon as it is possible without posing a threat for the current investigation. This suggests that AML/CFT regulations should be modified to provide that banks must inform customers that have been the subject of an SAR a certain time after the SAR has been transmitted to the FIU, unless the FIU specifically says that doing so would interfere with an ongoing investigation. Law enforcement authorities would have the responsibility of showing that informing a particular customer would compromise an ongoing investigation. As noted above, up to 90% of SARs are never investigated by FIUs. They sit in a drawer. It seems disproportionate to impose a permanent black-out even for the SARs that are never used.

A second problem is explainability.[82] AI can create new clusters of bank customers based on risks and relationships that are not visible to humans. Customers will have a right to understand why they were put in one cluster and not in another, and why certain scenarios were applied to them. As the Netherlands district court explained in the social security fraud case, individuals have a right to be informed that risk profiles are being created, and they have a right to understand why an algorithm attributed a particular risk score to them and created an alert. This is true even if the algorithm's recommendations are subsequently reviewed and approved by humans[83]. Explainability is a major concern for AI-based algorithms deployed by banks generally[84] and more specifically for AML/CFT systems. As explained recently by France's banking regulator,[85] explainability is necessary for different audiences and different purposes. If a machine learning algorithm creates an alert, explainability will be needed by human reviewers within the bank to decide whether to transform the alert into a SAR. FIUs may need explainability to understand why a particular SAR was generated; banking regulatory authorities to make sure that the bank's monitoring system has no gaps, is robust and auditable. From a proportionality standpoint, explainability is needed for two reasons. First, it helps individuals to have an effective right to challenge what the banks and FIUs are doing. To make an effective challenge, individuals need to be able

[81] Quadrature du Net (n 31) at para. 190.
[82] Beaudouin and others (n 43).

[83] The Netherlands scoring algorithm did not lead to an automatic decision – its recommendations were reviewed by humans.

[84] Dupont, Fliche and Yang (n 45).

[85] Ibid.

to understand the tool relied on by banks and the officials. Second, regulators and courts need explainability to ensure that the algorithm is operating as intended, for example it is not creating clusters based on protected features like national origin, religion or gender. Constant verification is needed to make sure algorithms are not learning to discriminate in unpermitted ways. For example, how can we be sure that the algorithms are not segmenting customers based on the geographic origin of their name, or the neighborhood where they live? Explainability was required both in the Netherlands social security welfare fraud case reviewed above, and in a United States federal court decision involving the scoring of teachers[86]. In both cases, explainability was a constitutional requirement. Numerous technical solutions are being proposed to provide both 'global' and 'local' explainability for machine learning models, including explainability by design.[87] These solutions are likely to help, but they need to be designed with fundamental rights objectives in mind.

The third problem relating to adequate safeguards is the lack of institutional oversight. When courts apply the proportionality test to government surveillance measures, an important factor is whether a court or independent commission reviews the process regularly to assess effectiveness and continued respect for fundamental rights. In the case of AML/CFT, the bank's monitoring system is subject to oversight by a bank regulatory authority and by a data protection authority, each authority looking at different aspects of the system. Bank regulatory authorities will ask whether the system is effective compared to industry practice. Data protection authorities will ask whether the system complies with the GDPR and fundamental rights. But the oversight of these two institutions is generally uncoordinated. And the oversight may be limited vis à vis law enforcement authorities who use the SARs to pursue criminals. The independent institutional oversight therefore remains partial and incomplete, broken into silos between data protection and AML/CFT compliance authorities, each authority focusing on different priorities and each having limited visibility and authority over FIUs.

The question of effectiveness and the question of fundamental rights are inextricably linked. An ineffective system cannot be proportionate, and effectiveness must be measured with proportionality cases in mind. While some data protection authorities in theory have power to conduct enquiries on the law enforcement side, this power depends heavily on the existence of specific legislation and the level of staffing of the data protection authorities. In many cases, the data protection authority's powers over police and national security processing are limited.

If AI is introduced to AML/CFT, dedicated institutional oversight will be essential to comply with the CJEU's proportionality requirements, as most recently articulated in the *Quadrature du Net* case. Current AML/CFT laws could be modified to give specific oversight powers to the national data protection authority, or to a newly created supervisory authority to evaluate the whole AML/CFT process, going from the design of the monitoring system by banks to the methods used by law enforcement authorities to examine SARs, to provide feedback to banks and to investigate and prosecute money-launderers. The authority would be required to evaluate the proportionality of the system on an ongoing basis, including the AI algorithms deployed by banks. The authority would have the power to require modifications to monitoring systems in appropriate cases, and to order changes in the quality metrics used by law enforcement authorities. If certain AI-based techniques do not prove effective, the authority would be able to order them to be stopped.

[86] Houston Federation of Teachers, Local 2415 v. Houston Independent School Dist., 251 F. Supp. 3d 1168 (S.D. Tex. 2017).

[87] Dupont, Fliche and Yang (n 45); Beaudouin and others (n 43).

One source of inspiration for such a regulatory authority is the French Commission for the Supervision of Intelligence Gathering Techniques (CNCTR). Created in 2015 as a result of a modernization of French intelligence gathering laws[88], the CNCTR oversees intelligence gathering techniques deployed by various intelligence agencies in France. For the most intrusive techniques, the CNCTR must give a non-binding opinion to the Prime Minister before the Prime Minister may approve the deployment of the technique. The CNCTR's opinion evaluates, among other things, the proportionality of the proposed measure in light of the intelligence-gathering objective. To ensure proportionality, the opinion may impose a time limit on the use of the technique, for example. For less intrusive techniques, the CNCTR does not need to provide an opinion in advance, but may conduct ex post audits. The CNCTR also investigates complaints from citizens about intelligence gathering practices. To promote public accountability, the CNCTR publishes an annual report on its activities containing qualitative and quantitative data on the kinds of intelligence gathering technologies for which the CNCTR issued opinions during the year, and whether the Prime Minister followed the CNCTR's recommendations. According to the CNCTR's latest annual report, the Prime Minister has never authorized an intelligence technique that received a negative opinion from the CNCTR.[89] Some of the CNCTR's opinions are made publicly available. Obviously, none of the materials published by the CNCTR contain intelligence secrets, but they do contain enough information for citizens and parliament to understand the extent of the use of intelligence gathering techniques, their contribution to national security, and the institutional safeguards that surround their use. The CNCTR consists of four judges from France's highest courts, four members from France's parliament, and one expert appointed by the national regulatory authority for electronic communications, ARCEP. All of the members of the commission are subject to "defense secrecy" rules.

The creation of a new authority for AML/CFT would permit many of the proportionality problems highlighted in this article to be addressed. The new authority might have representatives from different stakeholder groups involved in AML/CFT, including the data protection authority, the banking supervisory authority, parliament, and France's supreme courts. Similar to the CNCTR, the new authority would be covered by "defense secret" certification and would have power to conduct audits of AML/CFT systems, going from the bank's transactional monitoring system, to the handling of SARs by FIUs and other law enforcement authorities. The mission of the authority would be to ensure that systems – including new machine learning tools used to detect suspicious activities – are non-discriminatory, explainable and 'strictly necessary' under EU proportionality rules. The authority may have the ability to approve certain particularly massive and intrusive technological measures on a temporary basis, in particular if there is a serious threat to national security.[90] The authority may authorize new technological measures on an experimental basis to test whether they make a significant contribution to AML/CFT objectives, and order the removal of any techniques whose level of interference with privacy is not justified by a marked increase in the level of detection and apprehension of criminal funds. The authority would investigate complaints from citizens, and make an annual report to Parliament. The report would provide qualitative and quantitative information on the technologies used for AML/CFT, the audits conducted and authorizations given

[88] Law n° 2015-912 of 24 July 2015 on intelligence gathering. For a full description of the CNCTR's role and powers, see the CNCTR's website cntr.fr (in French).
[89] Commission national de contrôle des techniques de renseignement (CNCTR), Fourth Activity Report 2019, p. 53.
[90] Readers will recall that in the *Quadrature du Net* case (n 31), the CJEU permits general and indiscriminate processing only temporarily, to deal with serious threats to national security.

by the authority, the complaints received, and the level of contribution of AML/CFT techniques to fighting crime and protecting national security. The public information would obviously not include information covered by defense or criminal procedure secrecy. As required by the *Quadrature du Net* and *Canadian PNR Agreement* cases, the authority would review machine learning models and input data used for AML/CFT to ensure they are non-discriminatory and "limited to that which is strictly necessary in light of the objective of"[91] fighting money laundering and the financing of terrorism.

[91] Quadrature du Net (n 31) at para. 182.

V. CONCLUSION

There is considerable pressure to improve the effectiveness of AML/CFT measures, and AI provides a promising way to apprehend more criminal transactions than current rule-based systems. However, there are numerous barriers to moving to AI-based systems for AML/CFT. One important barrier relates to AI's compatibility with the GDPR and fundamental rights. The biggest fundamental rights question is whether an AI-based system to detect suspicious transactions could satisfy the proportionality test of the CJEU, particularly in light of the *Digital Rights Ireland* and *Tele2 Sverige Watson* decisions. To answer that question, we first unpacked the elements of the proportionality test, and then described exactly how AML/CFT systems, and particularly transaction monitoring, work. We reviewed how AI can help make current monitoring systems more effective, particularly in detecting anomalies. We then applied the proportionality test to AML/CFT systems, and identified five major problems. Most of the problems exist regardless of the presence of AI, but AI makes the problems worse. A major problem is the lack of specificity in AML/CFT laws on the characteristics and limits of the transaction monitoring systems that banks must deploy, a lack of specificity that violates the 'provided by law' step of the proportionality test as well as Article 23 of the GDPR. Another major problem relates to the inability to measure the effectiveness of banks' AML/CFT systems in reducing financial crime. Current reporting systems generate reports of suspicious activity, but up to 90% of these reports are never analyzed by law enforcement authorities, suggesting that they are largely useless. Without detailed feedback from FIUs, banks and regulators remain in the dark on the actual effectiveness of monitoring systems, making it impossible to apply the proportionality test, let alone train a machine learning algorithm to look for the most important criminal transactions. A specific problem relating to AI is the lack of explainability of certain machine learning models. Although providing detailed solutions to these problems is outside the scope of this article, we nevertheless provided some first thoughts on how the proportionality problems might be cured. Important suggestions include the organization of a feedback mechanism so that banks and regulatory authorities can adjust monitoring mechanisms to make them more proportionate, and the creation of dedicated supervisory authority for AML/CFT, potentially modeled after the French Commission for Supervision of Intelligence Gathering Techniques (CNCTR), whose job would be to ensure the proportionality of the AML/CFT system as a whole.