

IP PARIS <u>telecom-paris.fr/en/ai-ethics</u>

AML-CFT, AI and human rights

Academy of European Law (ERA) workshop Artificial intelligence and financial transparency as a national security priority January 26, 2022

Winston Maxwell, Director, Law and Digital Technology Studies Télécom Paris - Institut Polytechnique de Paris winston.maxwell@telecom-paris.fr

Models of public-private cooperation to detect terrorism

Ask private sector to send customer data to intelligence authorities for analysis (PNR model)

Allow intelligence authorities to connect directly to customer data in the network (telecoms model)

Ask private sector to analyze data on behalf of government and send reports to intelligence authorities (AML-CFT model)



AML-CFT model

Fight criminality and terrorism by blocking the money. But how do you do it?

(A) make a list of high-risk persons and countries and block any transaction involving them

(B) report "suspicious" operations

Challenges of delegating this job to financial institutions:

- Misalignment of incentives.
- Is it massive surveillance?
- Can banks do what the state could not do directly?
- Inadequate information.
- Institutional oversight.
- Imprecise legal standards.
- The problem of false positives and effectiveness.



The base rate fallacy (adapted from Wikipedia)

An algorithm has **99% accuracy rate** in detecting a potential terrorist, which means that:

when presented with true profiles of terrorists, the algorithm will correctly classify as terrorists 99% of the time (true positive rate).

when presented with profiles of non-terrorists, the algorithm will correctly classify as non-terrorists 99% of the time (true negative rate).

The bank has 1,000,000 customers, 100 of them are "true" potential terrorists

What's the false positive rate when the algorithm is applied to all the bank's customers?

99% x 100 = 99 terrorists detected and 1 false negative

99% x 999,900 = 989,901 true negatives, and 9,999 false positives.

9,999 false positives for 99 true positives, and 1 false negative

Out of 10,098 positive alerts, 9,999 (99%) will be false positives, and 99 (1%) will be true positives

False positive rate = 99%



References

W. Maxwell, The GDPR and private-sector measures to detect criminal activity, Revue des Affaires Européennes, 2021 n°1, p. 103

A. Bertrand, W. Maxwell, X. Vamparys, Do Al-based anti-money laundering (AML) systems violate European fundamental rights ? 11 Int'l Data Privacy Law, No. 3, 2021, p. 276

<u>Blog entry</u> on July 13, 2021 hearing in CJEU PNR case (C-817/19)

