

Explainable trade secrets: an oxymoron or fertile ground for innovation?

Nika Mahnič

Amendment to the Article 70a of the Polish banking law that came into effect in June of this year, is the first legislative initiative in European Union that has the aim to empower bank clients seeking an explanation of their credit score, or the prediction of their creditworthiness.¹ Strengthening the protection granted to Polish clients by the General Data Protection Regulation (GDPR), this precedent in national law enables borrowers to request “information on the factors, including personal data, which affected the evaluation of their creditworthiness.”²

Limited right to explanation

Already in 2017, before the GDPR came into force, researchers Sandra Wachter, Brent Mittelstadt and Luciano Floridi recognised serious limitations to the right to be informed granted by the new regulation. Both the Article 15(1) of the 1995 Data Protection Directive³ and Article 22(1) of the GDPR⁴ define(d) automated decisions concerning individuals - producing legal or similarly significant effects - as decisions that are based *solely* on automated processing of data. Wachter and others considered this “a loophole whereby even nominal involvement of a human in the decision-making process allows for an otherwise automated mechanism to avoid invoking elements of the right of access.”⁵

This has important implications for credit scoring, where it is difficult to draw the line between automated and human evaluation.⁶ The amendment to Polish banking law extends a right to explanation to decisions involving a human, as opposed to limiting it to the decisions made through fully automated means, as is the case with the protection offered by the GDPR.

Guidelines no. 251 of Article 29 Working Party, now adopted by the European Data Protection Board, advised - not obliged - the data controllers to inform the data subject about the criteria the data controllers relied on in arriving at the decision, about the

¹ <https://en.panoptikon.org/right-to-explanation>

² <http://www.lexlege.pl/prawo-bankowe/art-70a/>

³ <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A31995L0046>

⁴ <https://eur-lex.europa.eu/eli/reg/2016/679/oj>

⁵ <https://academic.oup.com/idpl/article/7/2/76/3860948>

⁶ <https://en.panoptikon.org/right-to-explanation>

rationale and logic behind it, not necessarily disclosing the full algorithm.⁷ After the implementation of the GDPR, customers are granted an explanation of the system functionality, not the circumstances of *specific* decisions.

Analysing drafts of GDPR and negotiations commentary, Wachter and others recognised that a stricter regulation, a legally binding right to explanation of *specific* decisions, was omitted from the final version of the Regulation.⁸ Efforts to reach greater transparency therefore seem to be coming from sectoral laws, as it is the case with the banking law in Poland.

Battles for interpretation

The fine for breaking this law is imposed by the Polish ombudsman and is proportional to the degree of violation of the provisions, the circumstances of the breach, and the financial capacity of the actor responsible for the infringement, reaching up to 100.000 Polish złoty or around 21.000 Pound sterling.⁹ If the amendment gets interpreted in a manner that will necessitate the provision of meaningful explanations, of *specific* decisions that take into consideration personal data, institutions will be financially punished for not meaningfully - more concretely than envisioned by GDPR - explaining their credit decisions.

Tension between consumer interest and business interest can be already inferred from the Recital 41 of the (now replaced with GDPR) Directive, which states that the data subjects' right to know the logic behind the automatic processing “must not adversely affect trade secrets or intellectual property and in particular the copyright protecting the software.”¹⁰ Proponents of the amendments to the banking law, Polish NGO Panoptikon Foundation, recognise this impediment, referring to the limitations put forward by algorithmic logic considered proprietary information. However, they recognise their ability to “check what data (not only personal) was introduced to the system and what data was generated in the form of the final evaluation,” enabling them to refer to particular cases.¹¹

⁷ https://iapp.org/media/pdf/resource_center/W29-auto-decision_profiling_02-2018.pdf

⁸ <https://academic.oup.com/idpl/article/7/2/76/3860948>

⁹ <https://www.arslege.pl/kary-pieniezne-nakladane-przez-rzeczniaka/k1395/a89941/>

¹⁰ <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A31995L0046>

¹¹ <https://en.panoptikon.org/right-to-explanation>

Similar to the German OpenSCHUFA, another project by non-governmental organisations devoted to a greater explainability and transparency of credit ranking¹², Panoptykon Foundation considers the amendment a good starting point. However, it is to be seen to what extent compliance will be demanded and non-compliance punished, and whether the amendment to the Polish banking law will significantly influence the exercise of the right to explanation granted by GDPR, and consequently shape subsequent regulatory efforts, as well as business response.

A recent article analyses the recently approved data protection laws in EU member states that aim to tackle the explainability of automated decision-making. Acknowledging the different ways in which the right to explanation is interpreted in the legal doctrine, it points to the French data protection law, where in the case of private - not judicial or administrative - decisions, “the subject receives an explanation of the rules defining the data processing and the main features of its implementation.” (Malgieri 2019, p. 15)¹³ The author considers the French law one of the few cases in which the right to explanation - or algorithmic legibility - is granted more significantly. However, Malgieri considers the probable outcome that the Data Protection Authorities will have to balance the degree of explanation with trade secrets that impose limitations to explainability (ibid., p. 25).

Secrets of the trade

If the consumer credit lenders were to disclose, for example, “statistical values, weighting of certain elements to calculate probabilities (eg the likelihood of loan repayment), and reference or comparison groups,”¹⁴ they could sufficiently reveal details of their proprietary scoring models.

In the era when all data could be credit data, we seek novel proposals and models of providing explanations that will not prove detrimental to businesses, while giving customers more transparency and meaningful access to the algorithmic logic affecting them. Assisting customers in understanding the significance of data influencing their credit scores would bring benefits to both customers and businesses, as it would greatly influence their mutual trust.

¹²

<https://algorithmwatch.org/en/openschufa-shedding-light-on-germanys-opaque-credit-scoring-2/>

¹³ <https://www.sciencedirect.com/science/article/pii/S0267364918303753>

¹⁴ <https://academic.oup.com/idpl/article/7/2/76/3860948>

In the United Kingdom, Statutory Credit Reports are intended to assist individual customers in improving their creditworthiness. However, they do not provide personalised, detailed explanations of the underlying algorithmic logic, nor do they include other factors, or alternative data. The latter include both financial data – such as mobile phone bill payments – and non-financial data – such as social data scraped from social media networks, or behavioural data resulting from consumers’ online habits. Alternative data are specifically important for the so called “thin-file” or “no-file” borrowers, individuals lacking credit histories that would then enable them to score.¹⁵

A recent article voices the need for innovation in customer or borrower support, especially due to the inclusion of social media data. Researcher Nikita Aggarwal proposes replacing the Statutory Credit Report with a ledger harbouring all consumer data, including alternative data that could be accessed in real-time by both borrowers and lenders. Sharing personal data about borrowers, however, invokes issues of reliability and data security,¹⁶ and the rising stakes for achieving compliance.

Need for inclusive innovation

In the *Policy and Investment Recommendations For Trustworthy AI*, published in April 2019, the High-level Expert Group on Artificial Intelligence proposes creating regulatory sandboxes that would encourage the involvement of various public and private stakeholders, including NGOs, allowing for more thorough viability assessments and a faster implementation of new business models.¹⁷

With the increasing volume and variety of stored data collected about customers, however, there is an ever greater need for privacy protection, data security and reliable products that would enable customers greater control over their documentation, over their credit “portfolios.” In the *Policy and Investment Recommendations For Trustworthy AI*, published in April 2019, the High-level Expert Group on Artificial Intelligence voices

¹⁵

<https://www.law.ox.ac.uk/business-law-blog/blog/2018/11/law-and-autonomous-systems-series-algorithmic-credit-scoring-and>

¹⁶

<https://www.law.ox.ac.uk/business-law-blog/blog/2019/07/big-data-and-obsolescence-consumer-credit-reports>

¹⁷

<https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence>

their support for research and development efforts for industrial solutions that enable fast, secure and compliant data sharing, considering, for example, encryption.¹⁸

This July, the UK's Financial Conduct Authority (FCA) and The Alan Turing Institute announced a collaboration focusing on transparency and explainability of AI in the financial sector.¹⁹ The FCA will inquire into two directions: first, interpretability by design would necessitate creating simpler models that guarantee explainability even at the cost of less predictive power. That direction would eliminate the possibility of using deep learning models. Second, reverse engineering explanatory factors, an approach the FCA is currently exploring in collaboration with The Bank of England, would entail applying an additional algorithm that could uncover the factors that influenced the model's individual predictions, and subsequently help interpret its workings to the consumer through a simplified visualisation of the decisions.²⁰

Customers are becoming more aware of the important influence obscure algorithms have on their credit ratings and private finances. Following the Polish precedent, pressure from consumer advocacy groups and regulatory initiatives, we can expect further developments in this field, and in other sectors that employ predictive tools in manners that affect individuals' and businesses' rights and obligations.

Responsible innovation in this sensitive field is much needed, as it affects the most vulnerable customers seeking financial help. It is of utmost importance that the business and regulatory responses consider the needs and the rights of various stakeholders, positioning consumers - as citizens with rights - at the centre of business development.

¹⁸

<https://ec.europa.eu/digital-single-market/en/news/policy-and-investment-recommendations-trustworthy-artificial-intelligence>

¹⁹

<https://www.turing.ac.uk/news/new-collaboration-fca-ethical-and-regulatory-issues-concerning-use-ai-financial-sector>

²⁰ <https://www.fca.org.uk/insight/explaining-why-computer-says-no>