4

AI Opacity in the Financial Industry and How to Break It

Zofia Bednarz and Linda Przhedetsky[*]

4.1 INTRODUCTION

Automated Banks – the financial entities using ADM and AI – feed off the culture of secrecy that is pervasive and entrenched in automated processes across sectors from 'Big Tech' to finance to government agencies, allowing them to avoid scrutiny, accountability, and liability.[1] As Pasquale points out, 'finance industries profit by keeping us in the dark'.[2]

An integral part of the financial industry's business model is the use of risk scoring to profile consumers of financial services, for example in the form of credit scoring, which is a notoriously opaque process.[3] The use of non-transparent, almost 'invisible' surveillance processes and the harvesting of people's data is not new: financial firms have always been concerned with collecting, aggregating, and combining data for the purposes of predicting the value of their customers through risk scoring.[4] Automation[5] introduces a new level of opacity in the financial industry, for example through the creation of AI models for which explanations are not provided – either deliberately, or due to technical explainability challenges.[6]

[*] The authors would like to thank Arundhati Suma Ajith for excellent research assistance.

[1] Frank Pasquale, *The Black Box Society* (Cambridge: Harvard University Press, 2015) 187.

[2] Ibid.

[3] Janine S Hiller and Lindsay Sain Jones, 'Who's Keeping Score?: Oversight of Changing Consumer Credit Infrastructure' (2022) 59(1) *American Business Law Journal* 61, 104.

[4] Pernille Hohnen, Michael Ulfstjerne, and Mathias Sosnowski Krabbe, 'Assessing Creditworthiness in the Age of Big Data: A Comparative Study of Credit Score Systems in Denmark and the US' (2021) 5(1) *Journal of Extreme Anthropology* 29, 34–35.

[5] Solon Barocas and Andrew D Selbst, 'Big Data's Disparate Impact' (2016) 104(3) *California Law Review* 671, 673–77.

[6] Alejandro Barredo Arrieta et al, 'Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI' (2020) 58 *Information Fusion* 82, 99–101.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

In this chapter we argue that the rise of AI and ADM tools contributes to opacity within the financial services sector, including through the intentional use of the legal system as a 'shield' to prevent scrutiny and blur accountability for harms suffered by consumers of financial services. A wealth of literature critiques the status quo, showing that consumers are disadvantaged by information asymmetries,[7] complicated consent agreements,[8] information overload,[9] and other tactics that leave consumers clueless if, when, and how they have been subject to automated systems. If consumers seek to access a product or service, it is often a requirement that they be analysed and assessed using an automated tool, for example, one that determines a credit score.[10] The potential harms are interlinked and range from financial exclusion to digital manipulation to targeting of vulnerable consumers and privacy invasions.[11] In our analysis we are mostly concerned with discrimination as an example of such harm,[12] as it provides a useful illustration of problems enabled by opacity, such as significant difficulty in determining if unfair discrimination has occurred at all, understanding the reasons for the decision affecting the person or group, and accessing redress.

The rules we examine will differ among jurisdictions, and our aim is not to provide a comprehensive comparative analysis of all laws that provide potential protections against scrutiny and increase the opacity of ADM-related processes of Automated Banks. We are interested in exploring certain overarching tendencies, using examples from various legal systems, and showing how financial firms may take advantage of the complex legal and regulatory frameworks applicable to their operations in relation to the use of AI and ADM tools.

As the use of AI and ADM continues to grow in financial services markets, consumers are faced with the additional challenge of knowing about, and considering how their ever-expanding digital footprint may be used by financial institutions. The more data exists about a person, the better their credit score (of course within certain limits, such as paying off debts on time).[13] The exact same mechanism may

[7] Peter Cartwright, 'Understanding and Protecting Vulnerable Financial Consumers' (2014) 38 (2) *Journal of Consumer Policy* 119, 121–23.

[8] Frederik Borgesius, 'Consent to Behavioural Targeting in European Law: What Are the Policy Implications of Insights from Behavioural Economics?' (Conference Paper for Privacy Law Scholars Conference, Berkeley, CA, 6–7 June 2013).

[9] Petra Persson, 'Attention Manipulation and Information Overload' (2018) 2(1) *Behavioural Public Policy* 78.

[10] Andrew Grant and Luke Deer, 'Consumer Marketplace Lending in Australia: Credit Scores and Loan Funding Success' (2020) 45(4) *Australian Journal of Management* 607.

[11] Zofia Bednarz and Kayleen Manwaring, 'Risky Business: Legal Implications of Emerging Technologies Affecting Consumers of Financial Services' in Dariusz Szostek and Mariusz Zalucki (eds), *Internet and New Technologies Law: Perspectives and Challenges* (Baden: Nomos, 2021) 59–74.

[12] Aaron Klein, Brookings Institution, *Reducing Bias in AI-Based Financial Services* (Report, 10 July 2020) <www.brookings.edu/research/reducing-bias-in-ai-based-financial-services/>.

[13] Hohnen et al, 'Assessing Creditworthiness', 36.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

underpin ‘open banking’ schemes: consumers who do not have sufficient data – often vulnerable people, such as domestic violence victims, new immigrants, or Indigenous people – cannot share their data with financial entities, may be excluded from accessing some products or offered higher prices, even if their actual risk is low.[14]

In Australia, consumers have claimed that they have been denied loans due to their use of takeaway food services and digital media subscriptions.[15] Credit rating agencies such as Experian explicitly state that they access data sources that reflect consumers’ use of new financial products, including ‘Buy Now Pay Later’ schemes.[16] As more advanced data collection, analysis, and manipulation technologies continue to be developed, there is potential for new categories of data to emerge. Already, companies can draw surprising inferences from big data. For example, studies have shown that seemingly trivial Facebook data can, with reasonable accuracy, predict a range of attributes that have not been disclosed by users: in one study, liking the ‘Hello Kitty’ page correlated strongly with a user having ‘[d]emocratic political views and to be of African-American origin, predominantly Christian, and slightly below average age’.[17]

Unless deliberate efforts are made, both in the selection of data sets and the design and auditing of AMD tools, inferences and proxy data will continue to produce correlations that may result in discriminatory treatment.[18]

This chapter proceeds as follows. We begin Section 4.2 with discussion of rules that allow corporate secrecy around AI models and their data sources to exist, focusing on three examples of such rules. We discuss the opacity of credit scoring processes and the limited explanations that consumers can expect in relation to a financial decision made about them (Section 4.2.1), trade secrecy laws (Section 4.2.2), and data protection rules which do not protect de-identified or anonymised information (Section 4.2.3). In Section 4.3 we analyse frameworks that incentivise the use of ADM tools by the financial industry, thus providing another ‘protective

[14] Zofia Bednarz, Chris Dolman, and Kimberlee Weatherall, ‘Insurance Underwriting in an Open Data Era – Opportunities, Challenges and Uncertainties’ (Actuaries Institute 2022 Summit, 2–4 May 2022) 10–12 <https://actuaries.logicaldoc.cloud/download-ticket?ticketId=09c77750-aa90-4ba9–835e-280ae347487b>.

[15] Su-Lin Tan, ‘Uber Eats, Afterpay and Netflix Accounts Could Hurt Your Home Loan Application’ (5 December 2018) *Australian Financial Review* <www.afr.com/property/uber-eats-afterpay-and-netflix-accounts-could-hurt-your-home-loan-application-20181128-h18ghz>.

[16] ‘Credit Bureau’, *Experian Australia* (Web Page) <www.experian.com.au/business/solutions/credit-services/credit-bureau>. ‘Secured from critical sectors of the Australian credit industry as well as from niche areas such as Specialty Finance data, short-term loans (including Buy Now Pay Later) and consumer leasing, enabling a more complete view of your customers’.

[17] Michal Kosinski, David Stillwell, and Thore Graepel, ‘Private Traits and Attributes Are Predictable from Digital Records of Human Behavior’ (2013) 110 *Proceedings of the National Academy of Sciences of the United States of America* 5805.

[18] Anya ER Prince and Daniel Schwarcz, ‘Proxy Discrimination in the Age of Artificial Intelligence and Big Data’ (2020) 105(3) *Iowa Law Review* 1257, 1273–76.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

layer' for Automated Banks, again discussing two examples: financial product governance regimes (Section 4.3.1) and 'open banking' rules (Section 4.3.2). The focus of Section 4.4 is on potential solutions. We argue it is not possible for corporate secrecy and consumer rights to coexist, and provide an overview of potential regulatory interventions, focusing on preventing Automated Banks from using harmful AI systems (Section 4.4.1), aiding consumers understand when ADM is used (Section 4.4.2), and facilitating regulator monitoring and enforcement (Section 4.4.3). The chapter concludes with Section 4.5.

4.2 RULES THAT ALLOW CORPORATE SECRECY TO EXIST

4.2.1 *Opacity of Credit Scoring and the (Lack of) Explanation of Financial Decisions*

Despite their widespread use in the financial industry, credit scores are difficult for consumers to understand or interpret. A person's credit risk has traditionally been calculated based on 'three C's': collateral, capacity, and character.[19] Due to the rise of AI and ADM tools in the financial industry, the 'three C's' are increasingly being supplemented and replaced by diverse categories of data.[20] An interesting example can be found through FICO scores, which are arguably the first large-scale process in which automated computer models replaced human decision-making.[21] FICO, one of the best-known credit scoring companies,[22] explains that their scores are calculated according to five categories: 'payment history (35%), amounts owed (30%), length of credit history (15%), new credit (10%), and credit mix (10%)'.[23] These percentage scores are determined by the company to give consumers an understanding of how different pieces of information are weighted in the calculation of a score, and the ratios identified within FICO scores will not necessarily reflect the weightings used by other scoring companies. Further, while FICO provides a degree of transparency, the ways in which a category such as 'payment history' is calculated remains opaque: consumers are not privy to what is considered a 'good' or a 'bad' behaviour, as represented by data points in their transaction records.[24]

Globally, many credit scoring systems (both public and private) produce three-digit numbers within a specified range to determine a consumer's creditworthiness. For example, privately operated Equifax and Trans Union Empirica score

[19] Eric Rosenblatt, *Credit Data and Scoring: The First Triumph of Big Data and Big Algorithms* (Cambridge: Elsevier Academic Press, 2020) 1.

[20] Hiller and Jones, 'Who's Keeping Score?', 68–77.

[21] Rosenblatt, *Credit Data and Scoring*, 7.

[22] Hohnen et al, 'Assessing Creditworthiness', 36.

[23] 'What's in My FICO® Scores?', *MyFico* (Web Page) <www.myfico.com/credit-education/whats-in-your-credit-score>.

[24] Consumer Financial Protection Bureau, 'The Impact of Differences between Consumer- and Creditor-Purchased Credit Scores' (SSRN Scholarly Paper No 3790609, 19 July 2011) 19.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

consumers in Canada between 300 and 900,[25] whereas credit bureaus in Brazil score consumers between 1 and 1,000.[26] In an Australian context, scores range between 0 and 1,000, or 1,200, depending on the credit reporting agency.[27] By contrast, other jurisdictions use letter-based ratings, such as Singapore's HH to AA scale which corresponds with a score range of 1,000–2,000,[28] or blacklists, such as Sweden's payment default records.[29]

Credit scoring, it turns out, is surprisingly accurate in predicting financial breakdowns or future loan delinquency,[30] but the way different data points are combined by models is not something even the model designer can understand using just intuition.[31] Automated scoring processes become even more complex as credit scoring companies increasingly rely on alternative data sources to assess consumers' creditworthiness, including 'predictions about a consumer's friends, neighbors, and people with similar interests, income levels, and backgrounds'.[32] And a person's credit score is just one of the elements lenders, Automated Banks, feed into their models to determine a consumer's risk score. It has been reported that college grades, and the time of day an individual applies for a loan have been used to determine a person's access to credit.[33] These types of data constitute 'extrinsic data' sources, which consumers are unknowingly sharing.[34]

The use of alternative data sources is purported as a way of expanding consumers' access to credit in instances where there is a lack of quality data (such as previous loan repayment history) to support the underwriting of consumers' loan.[35] Applicants are often faced with a 'Catch-22 dilemma: to qualify for a loan, one

[25] 'What Is a Good Credit Score?', *Equifax Canada* (Web Page) <www.consumer.equifax.ca/personal/education/credit-score/what-is-a-good-credit-score>; 'FICO Score 10, Most Predictive Credit Score in Canadian Market', *FICO Blog* (Web Page) <www.fico.com/blogs/fico-score-10-most-predictive-credit-score-canadian-market>.

[26] Frederic de Mariz, 'Using Data for Financial Inclusion: The Case of Credit Bureaus in Brazil' (SSRN Paper, *Journal of International Affairs*, 28 April 2020).

[27] 'Credit Scores and Credit Reports', *Moneysmart* (Web Page) <https://moneysmart.gov.au/managing-debt/credit-scores-and-credit-reports>.

[28] 'Credit Score', *Credit Bureau* (Web Page) <www.creditbureau.com.sg/credit-score.html>.

[29] 'Payment Default Records', *Swedish Authority for Privacy Protection* (Web Page) <www.imy.se/en/individuals/credit-information/payment-default-records/>.

[30] Or even car accidents one will have in the future, Rosenblatt, *Credit Data and Scoring*, 6.

[31] Ibid, 7.

[32] Mikella Hurley and Julius Adebayo, 'Credit Scoring in the Era of Big Data' (2016) 18 *Yale Journal of Law and Technology* 148, 151.

[33] Hiller and Jones, 'Who's Keeping Score?', 68–77.

[34] Zofia Bednarz and Kayleen Manwaring, 'Hidden Depths: The Effects of Extrinsic Data Collection on Consumer Insurance Contracts' (2022) 45(July) *Computer Law and Security Review: The International Journal of Technology Law and Practice* 105667.

[35] 'Examining the use of alternative data in underwriting and credit scoring to expand access to credit' (Hearing before the Task Force on Financial Technology of the Committee on Financial Services, U.S. House of Representatives, One Hundred Sixteenth Congress, First Session July 25, 2019) <www.congress.gov/116/chrg/CHRG-116hhrg40160/CHRG-116hhrg40160.pdf>.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

must have a credit history, but to have a credit history one must have had loans'.[36] This shows how ADM tools offer more than just new means to analyse greater than ever quantities of data: they also offer a convenient excuse for Automated Banks to effectively use more data.

Of course, increasing reliance on automated risk scoring is not the origin of unlawful discrimination in financial contexts. However, it is certainly not eliminating discriminatory practices either: greater availability of more granular data, even when facially neutral, leads to reinforcing of existing inequalities.[37] Automated Banks have been also shown to use alternative data to target more vulnerable consumers, who they were not able to reach or identify when only using traditional data on existing customers.[38] The quality change that AI tools promise to bring is to 'make the data talk': all data is credit data, if we have the right automated tools to analyse them.[39]

Collection, aggregation, and use of such high volumes of data, including 'extrinsic data', also make it more difficult, if not impossible, for consumers to challenge financial decisions affecting them. While laws relating to consumer lending (or consumer financial products in general) in most jurisdictions provide that some form of explanation of a financial decision needs to be made available to consumers,[40] these rules will rarely be useful in the context of ADM and AI tools used in processes such as risk scoring.

This is because AI tools operate on big data. Too many features of a person are potentially taken into account for any feedback to be meaningful. The fact that risk

[36] Hohnen et al, 'Assessing Creditworthiness', 38.

[37] Hiller and Jones, 'Who's Keeping Score?', 87–96; Bartlett et al, 'Consumer-Lending Discrimination in the FinTech Era' (2022) 143(1) *Journal of Financial Economics* 30.

[38] Hiller and Jones, 'Who's Keeping Score?', 92–93.

[39] Quentin Hardy, 'Just the Facts: Yes, All of Them' (25 March 2012) *The New York Times* <https://archive.nytimes.com/query.nytimes.com/gst/fullpage-9A0CE7DD153CF936A15750C0A9649D8B63.html>.

[40] See for example: US: Equal Credit Opportunity Act (ECOA) s 701, which requires a creditor to notify a credit applicant when it has taken adverse action against the applicant; Fair Credit Reporting Act (FCRA) s 615(a), which requires a person to provide a notice when the person takes an adverse action against a consumer based in whole or in part on information in a consumer report; Australia: Privacy Act 1988 (Cth) s 21P, stating that if a credit provider refuses an application for consumer credit made in Australia, the credit provider must give the individual written notice that the refusal is based wholly or partly on credit eligibility information about one or more of the persons who applied; Privacy (Credit Reporting) Code 2014 (Version 2.3) para 16.3 requiring a credit provider who obtains credit reporting information about an individual from a credit reporting bureau and within 90 days of obtaining that information, refuses a consumer credit application, to provide a written notice of refusal, informing the individual of a number of matters, including their right to access credit reporting information held about them, that the refusal may have been based on the credit reporting information, and the process for correcting the information; UK: lenders are not required to provide reasons for loan refusal, even when asked by a consumer, but s 157 Consumer Credit Act 1974 requires them to indicate which credit reporting agency (if any) they used in assessing the application.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

scores and lending decisions are personalised make it even more complicated for consumers to compare their offer with anyone else's. This can be illustrated by the case of Apple credit card,[41] which has shown the complexity of investigation necessary for people to be able to access potential redress: when applying for personalised financial products, consumers cannot immediately know what features are being taken into account by financial firms assessing their risk, and subsequent investigation by regulators or courts may be required.[42] The lack of a right to meaningful explanation of credit scores and lending decisions based on the scores makes consumers facing Automated Banks and the automated credit scoring system quite literally powerless.[43]

4.2.2 *Trade Secrets and ADM Tools in Credit Scoring*

The opacity of credit scoring, or risk scoring more generally, and other automated assessment of clients that Automated Banks engage in, is enabled by ADM tools which 'are highly valuable, closely guarded intellectual property'.[44] Complementing the limited duty to provide explanation of financial decisions to consumers, trade secrets laws allow for even more effective shielding of the ADM tools from scrutiny, including regulators' and researchers' scrutiny.

While trade secrets rules differ between jurisdictions, the origin and general principles that underpin these rules are common across all the legal systems: trade secrets evolved as a mechanism to protect diverse pieces of commercial information, such as formulas, devices, or patterns from competitors.[45] These rules fill the gap where classic intellectual property law, such as copyright and patent law, fails – and it notably fails in relation to AI systems, since algorithms are specifically excluded from its protection.[46] Recent legal developments, for example the European Union

[41] Neil Vidgor, 'Apple Card Investigated after Gender Discrimination Complaints' (10 November 2019) *The New York Times* <www.nytimes.com/2019/11/10/business/Apple-creditcard-investigation.html>.

[42] See e.g. Corrado Rizzi, 'Class Action Alleges Wells Fargo Mortgage Lending Practices Discriminate against Black Borrowers' (21 February 2022) *ClassAction.org* <www.classaction.org/news/class-action-alleges-wells-fargo-mortgage-lending-practices-discriminate-against-black-borrowers> or Kelly Mehorter, 'State Farm Discriminates against Black Homeowners When Processing Insurance Claims, Class Action Alleges' (20 December 2022) *ClassAction.org* <www.classaction.org/news/state-farm-discriminates-against-black-homeowners-when-processing-insurance-claims-class-action-alleges>; Hiller and Jones, 'Who's Keeping Score?', 83–84.

[43] Hiller and Jones, 'Who's Keeping Score?', 65.

[44] Consumer Financial Protection Bureau, 'The Impact of Differences between Consumer- and Creditor-Purchased Credit Scores' (SSRN Scholarly Paper No 3790609, 19 July 2011) 5.

[45] Brenda Reddix-Smalls, 'Credit Scoring and Trade Secrecy' (2012) 12 *UC Davis Business Law Journal* 87, 115.

[46] Katarina Foss-Solbrekk, 'Three Routes to Protecting AI Systems and Their Algorithms under IP Law: The Good, the Bad and the Ugly' (2021) 16(3) *Journal of Intellectual Property Law & Practice* 247, 248.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

Trade Secrets Directive,[47] or the US Supreme Court case of *Alice Corp. v CLS Bank*,[48] mean that to protect their proprietary technologies, companies are now turning to trade secrets.[49] In practice, this greatly reduces the transparency of the ADM tools used: if these cannot be protected through patent rights, they need to be kept secret.[50]

The application of trade secrets rules leads to a situation in which financial entities, for example lenders or insurers, who apply third party automated tools to assess creditworthiness of their prospective clients might not be able to access the models and data they use. Using third party tools is a common practice, and the proprietary nature of the tools and data used to develop and train the models will mean financial entities using these tools may be forced to rely on the supplier's specifications in relation to their fairness as they may not be able to access the code themselves.[51]

Secrecy of ADM tools of course has implications for end users, who will be prevented from challenging credit models, and is also a barrier for enforcement and research.[52] Trade secret protections apply not only to risk scoring models, but often extend also to data sets and inferences generated from information provided by individuals.[53] Commercial entities openly admit they 'invest significant amounts of time, money and resources' to draw inferences about individuals 'using [. . .] proprietary data analysis tools', a process 'only made possible because of the [companies'] technical capabilities and value add'.[54] This, they argue, makes the data sets containing inferred information a company's intellectual property.[55]

The application of trade secrets rules to credit scoring in a way that affects the transparency of the financial system is not exactly new: '[t]he trade secrecy surrounding credit scoring risk models, and the misuse of the models coupled with the lack of governmental control concerning their use, contributed to a financial industry wide recession (2007–2008)'.[56]

[47] *Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure* [2016] OJ L 157/1.

[48] 573 U.S. 208 (2014).

[49] Foss-Solbrekk, 'Three Routes to Protecting AI Systems and Their Algorithms under IP Law', 248; Meghan J Ryan, 'Secret Algorithms, IP Rights and the Public Interest' (2020) 21(1) *Nevada Law Journal* 61, 62–63.

[50] Ryan, 'Secret Algorithms', 62–63.

[51] Hiller and Jones, 'Who's Keeping Score?', 83.

[52] Reddix-Smalls, 'Credit Scoring and Trade Secrecy', 117; see also Bartlett et al, 'Consumer-Lending Discrimination in the FinTech Era'.

[53] Gintarė Surblytė-Namavičienė, *Competition and Regulation in the Data Economy: Does Artificial Intelligence Demand a New Balance?* (Cheltenham: Edward Elgar, 2020).

[54] Facebook, 'Submission to the Australian Privacy Act Review Issues Paper' (6 December 2020) 25 <www.ag.gov.au/sites/default/files/2021-02/facebook.PDF>.

[55] Ibid.

[56] Reddix-Smalls, 'Credit Scoring and Trade Secrecy', 89.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

In addition to trade secrets laws, a *sui generis* protection of source code of algorithms is being introduced in international trade law through free trade agreements,[57] which limit governments from mandating access to the source code. The members of the World Trade Organization (WTO) are currently negotiating a new E-commerce trade agreement, which may potentially include a prohibition on government-mandated access to software source code.[58] WTO members, including Canada, the EU, Japan, South Korea, Singapore, Ukraine, and the United States support such a prohibition,[59] which in practice will mean a limited ability for states to adopt laws that would require independent audits of AI and ADM systems.[60] It is argued that adoption of the WTO trade agreement could thwart the adoption of the EU's AI Act,[61] demonstrating how free trade agreements can impose another layer of rules enhancing the opacity of AI and ADM tools.

4.2.3 *'Depersonalising' Information to Avoid Data and Privacy Protection Laws: Anonymisation, De-identification, and Inferences*

Automated Banks' opacity is enabled by the express exclusion of 'anonymised' or 'de-identified' data from the scope of data and privacy protection laws such as the GDPR.[62] In its Recital 26, the GDPR defines anonymised information as not relating to 'an identified or identifiable natural person' or as 'data rendered anonymous in such a manner that the data subject is not or no longer identifiable'. This allows firms to engage in various data practices, which purport to use anonymised data.[63] They argue they do not collect or process 'personal information', thus avoiding the application of the rules, and regulatory enforcement.[64] Also, consumers to whom privacy policies are addressed believe that practices focusing on information that does not directly identify them have no impact on their privacy.[65] This

[57] Kristina Irion, 'Algorithms Off-Limits?' (FAccT'22, 21–24 June 2022, Seoul) 1561 <https://dl.acm.org/doi/pdf/10.1145/3531146.3533212>.

[58] Ibid.

[59] Ibid.

[60] Ibid, 1562.

[61] Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (AI Act) and Amending Certain Union Legislative Acts [2021] OJ COM 206.

[62] *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 (GDPR)* [2016] OJ L 119/1, Recital (26); *Australian Privacy Act 1988* (Cth) s 6.

[63] Katharine Kemp, 'A Rose by Any Other Unique Identifier: Regulating Consumer Data Tracking and Anonymisation Claims' (August 2022) *Competition Policy International TechReg Chronicle* 22.

[64] Ibid.

[65] Ibid, 23.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

in turn may mean privacy policies are misrepresenting data practices to consumers, which could potentially invalidate their consent.[66]

There is an inherent inconsistency between privacy and data protection rules and the uses and benefits that ADM tools using big data analytics promise. Principles of purpose limitation and data minimisation[67] require entities to delimit, quite strictly and in advance, how the data collected are going to be used, and prevent them from collecting and processing more data than necessary for that specific purpose. However, this is not how big data analytics, which fuels ADM and AI models, works.[68] Big data means that 'all data is credit data', incentivising the Automated Banks to collect as much data as possible, for any possible future purpose, potentially not known yet.[69] The exclusion of anonymised or de-identified data from the scope of the protection frameworks opens doors for firms to take advantage of enhanced analytics powered by new technologies. The contentious question is at which point information becomes, or ceases to be, personal information. If firms purchase, collect, and aggregate streams of data, producing inferences allowing them to describe someone in great detail, including their age, preferences, dislikes, size of clothes they wear and health issues they suffer from, their household size and income level,[70] but do not link this profile to the person's name, email, physical address, or IP address – would it be personal information? Such a profile, it could be argued, represents a theoretical, 'model' person or consumer, built for commercial purposes through aggregation of demographic and other information available.[71]

De-identified data may still allow a financial firm to achieve more detailed segmentation and profiling of their clients. There are risks of harms in terms of 'loss of privacy, equality, fairness and due process' even when anonymised data is used.[72] Consumers are left unprotected against profiling harms due to such 'narrow interpretation of the right to privacy as the right to anonymity'.[73]

[66] Ibid, 27–29.

[67] See e.g. Art. 5 GDPR.

[68] Tal Zarsky, 'Incompatible: The GDPR in the Age of Big Data' (2017) 4(2) *Seton Hall Law Review* 995, 1004–18.

[69] Ibid, 1010.

[70] Wolfe Christl and Sarah Spiekermann, *Networks of Control: A Report on Corporate Surveillance, Digital Tracking, Big Data & Privacy* (Vienna: Facultas, 2016); Forbrukerrådet (Norwegian Consumer Council), *Out of Control: How Consumers Are Exploited by the Online Advertising Industry* (Report, 14 January 2020) 19–22.

[71] Ibid.

[72] Mireille Hildebrandt, 'Profiling and the Identity of the European Citizen' in Mireille Hildebrandt and Serge Gutwirth (eds), *Profiling the European Citizen: Cross-Disciplinary Perspectives* (New York: Springer, 2008) 305–9; Sandra Wachter, 'Data Protection in the Age of Big Data' (2019) 2 *Nature Electronics* 6, 7.

[73] N Chami et al, 'Data Subjects in the Femtech Matrix: A Feminist Political Economy Analysis of the Global Menstruapps Market' (Issue Paper 6, Feminist Digital Justice, December 2021) 4.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

There is also discussion as to the status of inferences under data and privacy protection laws. Credit scoring processes are often based on inferences, where a model predicts someone's features (and ultimately their riskiness or value as a client) on the basis of other characteristics that they share with others deemed risky by the model.[74] AI models may thus penalise individuals for 'shopping at low-end stores', membership in particular communities or families, and affiliations with certain political, religious, and other groups.[75] While AI-powered predictions about people's characteristics are often claimed to be more accurate than those made by humans,[76] they may also be inaccurate.[77] The question is if such inferences are considered personal information protected by privacy and data laws.

Entities using consumers' data, such as technology companies, are resisting against expressly including inferred information in the scope of data and privacy protections. For example, Facebook openly admitted that '[t]o protect the investment made in generating inferred information and to protect the inferred information from inappropriate interference, inferred information should not be subject to all of the same aspects of the [Australian Privacy Act] as personal information'.[78] The 'inappropriate interference' they mention refers to extending data correction and erasure rights to inferred information.

Second, there is an inherent clash between the operation of privacy and data protection rules and the inference processes AI tools are capable of carrying out. Any information, including sensitive information, may be effectively used by an ADM system, even though it only materialises as an internal encoding of the model and is not recorded in a human understandable way. The lack of explicit inclusion of inferred information, and its use, within the privacy and data protection frameworks provides another layer of opacity shielding financial firms (as well as other entities) from scrutiny of their ADM tools.

When information is 'depersonalised' in some way: de-identified on purpose through the elimination of strictly personal identifiers,[79] through use of anonymous 'demographic' data, through 'pseudonymisation' practices, or because it is inferred from data held (either personal or already de-identified), the result is the same – privacy and data protection rules do not apply. The firms take advantage of that exclusion, sometimes balancing on the thin line between legal and illegal data processing, making their data practices non-transparent to avoid scrutiny by consumers and regulators.

[74] Hurley and Adebayo, 'Credit Scoring in the Era of Big Data', 183.

[75] Ibid.

[76] Wu Youyou, Michal Kosinski, and David Stillwell, 'Computer-Based Personality Judgments Are More Accurate than Those Made by Humans' (Research Paper, *Proceedings of the National Academy of Sciences* 112(4): 201418680, 12 January 2015).

[77] Hurley and Adebayo, 'Credit Scoring in the Era of Big Data', 183.

[78] Facebook, 'Submission to the Australian Privacy Act Review Issues Paper', 25–26.

[79] CM O'Keefe et al, *The De-Identification Decision-Making Framework* (CSIRO Reports EP173122 and EP175702, 18 September 2017), ix.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

As a US judge in a recent ruling put it: '[i]t is well established that there is an undeniable link between race and poverty, and any policy that discriminates based on credit worthiness correspondingly results in a disparate impact on communities of color'.[80] The data used in large-scale AI and ADM models is often de-identified or anonymised, but it inherently mirrors historical inequalities and biases, thus allowing the Automated Banks to claim impartiality and avoid responsibility for the unfairness of data used.

The reason why privacy and data protection rules lack clear consideration of certain data practices and processes enabled by AI may be due to these tools and processes being relatively new and poorly understood phenomena.[81] This status quo is however very convenient for the companies, who will often raise the argument that 'innovation' will suffer if more stringent regulation is introduced.[82]

4.3 RULES THAT INCENTIVISE THE USE OF ADM TOOLS BY FINANCIAL ENTITIES

In addition to offering direct pathways allowing Automated Banks to evade scrutiny of their AI and ADM models, legal systems and markets in the developed world have also evolved to incentivise the use of automated technology by financial entities. In fact, the use of ADM and AI tools is encouraged, or sometimes even mandated,[83] by legal and regulatory frameworks. After all, the fact that they are *told to* either use the technology, or to achieve outcomes that can effectively only be reached with the application of the tools in question, provides a basis for a very convenient excuse. Though this is mainly an unintended effect of the rules, it should not be ignored.

In this section, we discuss two examples of rules that increase the secrecy of AI or ADM tools used in the context of risk scoring: financial products governance rules and 'open banking' regimes.

[80] Office of the Insurance Commissioner Washington State, *Final Order on Court's Credit Scoring Decision; Kreidler Will Not Appeal* (Media Release, 29 August 2022) <www.insurance.wa.gov/news/final-order-courts-credit-scoring-decision-kreidler-will-not-appeal>.

[81] For example, Prof Sandra Wachter has pointed out the GDPR is based on an outdated concept of a 'nosey neighbour': Sanda Wachter, 'AI's Legal and Ethical Implications' *Twimlai* (Podcast, 23 September 2021) <https://twimlai.com/podcast/twimlai/ais-legal-ethical-implications-sandra-wachter/>.

[82] Microsoft Australia, 'Microsoft Submission to Review of the Privacy Act 1988' (December 2020) 2–3 <www.ag.gov.au/sites/default/files/2021–02/microsoft-australia.PDF>; Facebook, 'Submission to the Australian Privacy Act Review Issues Paper', 25.

[83] See Zofia Bednarz, 'There and Back Again: How Target Market Determination Obligations for Financial Products May Incentivise Consumer Data Profiling' (2022) 36(2) *International Review of Law, Computers & Technology* 138.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

4.3.1 *Financial Products Governance Rules*

Financial firms have always been concerned with collecting and using data about their consumers, to differentiate between more and less valuable customers. For example, insurance firms, even before AI profiling tools were invented (or at least before they were applied at a greater scale) were known to engage in practices referred to as 'cherry-picking' and 'lemon-dropping', setting up firms' offices at higher floors in buildings with no lifts, so that it would be harder for disabled (potential) clients to reach them.[84] There is a risk that the widespread data profiling and use of AI tools may exacerbate issues relating to consumers' access to financial products and services. AI tools may introduce new or replicate historical biases present in data,[85] doing so more efficiently, in a way that is more difficult to discover, and at a greater scale than was possible previously.[86]

An additional disadvantage resulting from opaque risk scoring systems is that consumers may miss out on the opportunity to improve their score (for example, through the provision of counterfactual explanations, or the use of techniques including 'nearby possible worlds').[87] In instances where potential customers who would have no trouble paying back loans are given low risk scores, two key issues arise: first, the bank misses out on valuable customers, and second, there is a risk that these customers' rejections, if used as input data to train the selection algorithm, will reinforce existing biases.[88]

Guaranteeing suitability of financial services is a notoriously complicated task for policymakers and regulators. With disclosure duties alone proving largely unsuccessful in addressing the issue of consumers being offered financial products that are unfit for purpose, policymakers in a number of jurisdictions, such as the EU and its Member States, the United Kingdom, Hong Kong, Australia, and Singapore, have started turning to product governance regimes.[89] An important component of these

[84] Marshall Allen, 'Health Insurers Are Vacuuming Up Details about You: And It Could Raise Your Rates' (17 July 2018) *NPR* <www.npr.org/sections/health-shots/2018/07/17/629441555/healthinsurers-are-vacuuming-up-details-about-you-and-it-could-raise-your-rates>.

[85] Australian Human Rights Commission, *Using Artificial Intelligence to Make Decisions: Addressing the Problem of Algorithmic Bias* (Technical Paper, November 2022) 34–44.

[86] E Martinez and L Kirchner, 'Denied: The Secret Bias Hidden in Mortgage-Approval Algorithms' (25 August 2021) *The Markup*.

[87] Sandra Wachter, Brent Mittelstadt, and Chris Russell, 'Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR' (2018) 31(2) *Harvard Journal of Law & Technology* 841, 848.

[88] European Union Agency for Fundamental Rights, *Bias in Algorithms: Artificial Intelligence and Discrimination* (Report, 2022) 8–9 <https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-bias-in-algorithms_en.pdf>.

[89] Hannah Cassidy et al, 'Product Intervention Powers and Design and Distribution Obligations: A Cross-Border Financial Services Perspective' (Guide, Herbert Smith Freehills, 11 June 2019) <www.herbertsmithfreehills.com/latest-thinking/product-intervention-powers-and-design-and-distribution-obligations-in-fs>.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

financial product governance regimes is an obligation placed on financial firms, which issue and distribute financial products, to ensure their products are fitness-for-purpose and to adopt a consumer-centric approach in design and distribution of the products. In particular, a number of jurisdictions require financial firms to delimit the target market for their financial products directed at retail customers, and ensure the distribution of the products within this target market. Such target market is a group of consumers of a certain financial product who are defined by some general characteristics.[90]

Guides issued by regulators, such as the European Securities and Markets Authority[91] and the Australian Securities and Investment Commission,[92] indicate which consumers' characteristics are to be taken into account by financial firms. The consumers for whom the product is intended are to be identified according to their 'likely objectives, financial situation, and needs',[93] or five 'categories': the type of client, their knowledge and experience, financial situation, risk tolerance, and objective and needs.[94] For issuers or manufacturers of financial products these considerations are mostly theoretical: as they might not have direct contact with clients, they need to prepare a *potential* target market, aiming at *theoretical* consumers and their *likely* needs and characteristics.[95] Both issuers and distributors need to take reasonable steps to ensure that products are distributed within the target market, which then translates to the identification of real consumers with specific needs and characteristics that should be compatible with the potential target markets identified. Distributors have to hold sufficient information about their end clients to be able to assess if they can be included in the target market,[96] including:

– indicators about the likely circumstances of the consumer or a class of consumers (e.g. concession card status, income, employment status);

– reasonable inferences about the likely circumstances of the consumer or a class of consumers (e.g. for insurance, information inferred from the postcode of the consumer's residential address); or

[90] Martin Hobza and Aneta Vondrackova, 'Target Market under MiFID II: the Distributor's Perspective' (2019) 14 *Capital Markets Law Journal* 518, 529.

[91] European Securities and Markets Authority (ESMA), 'Guidelines on MiFID II Product Governance Requirements' (ESMA35–43-620, 5 February 2018).

[92] Australian Securities and Investment Commission (ASIC), 'Regulatory Guide 274: Product Design and Distribution Obligations' (December 2020).

[93] ASIC, 'Regulatory Guide 274', para 274.6.

[94] ESMA, 'Guidelines on MiFID II Product Governance Requirements', 34–35.

[95] ESMA, 'Final Report: Guidelines on MiFID II Product Governance Requirements' (ESMA35–43-620, 2 June 2017) 34, para 17.

[96] 'The MiFID II Review – Product Governance: How to Assess Target Market' *Ashurst* (Financial Regulation Briefing, 3 October 2016) <www.ashurst.com/en/news-and-insights/legal-updates/mifid-12-mifid-ii-product-governance-how-to-assess-target-market/#:~:text=Regular%20review%20by%20the%20manufacturer,how%20to%20get%20that%20information>.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

> – data that the distributor may already hold about the consumer or similar consumers, or results derived from analyses of that data (e.g. analysis undertaken by the distributor of common characteristics of consumers who have purchased a product).[97]

Financial products governance frameworks invite financial firms to collect data on consumers' vulnerabilities. For example in Australia, financial firms need to consider vulnerabilities consumers may have, such as those resulting from 'personal or social characteristics that can affect a person's ability to manage financial interactions',[98] as well as those brought about by 'specific life events or temporary difficulties',[99] in addition to vulnerabilities stemming from the product design or market actions.

The rationale of product governance rules is to protect financial consumers, including vulnerable consumers,[100] yet the same vulnerable consumers may be disproportionately affected by data profiling, thus inhibiting their access to financial products. Financial law is actively asking firms to collect even more data about their current, prospective, and past customers, as well as the general public. It provides more than a convenient excuse to carry out digital profiling and collect data for even more precise risk scoring – it actually *mandates* this.

4.3.2 *How 'Open Banking' Increases Opacity*

Use of AI and ADM tools, together with ever-increasing data collection feeding the data hungry models,[101] is promoted as beneficial to consumers and markets, and endorsed by companies and governments. Data collection is thus held out as a necessary component of fostering AI innovation. Companies boast how AI insights allow them to offer personalised services, 'tailored' to individual consumer's needs. McKinsey consulting firm hails 'harnessing the power of external data' noting how

[97] ASIC, 'Regulatory Guide 274', para. 277.180.

[98] ASIC's RG para. 274.47 provides examples of such personal and social characteristics: 'speaking a language other than English, having different cultural assumptions or attitudes about money, or experiencing cognitive or behavioural impairments due to intellectual disability, mental illness, chronic health problems or age'.

[99] ASIC, 'Regulatory Guide 274' para. 274.47: 'an accident or sudden illness, family violence, job loss, having a baby, or the death of a family member'.

[100] For example, Indigenous Australians, whose lack of financial literacy historically made them an easy target for mis-selling of inadequate products: Commonwealth of Australia, *Royal Commission into Misconduct in the Banking, Superannuation and Financial Services Industry* (Interim Report Vol. 2, 2018) 452–57.

[101] Machine Learning in particular has been described as 'very data hungry' in the World Economic Forum and Deloitte; WEF and Deloitte, *The New Physics of Financial Services: Understanding How Artificial Intelligence Is Transforming the Financial Ecosystem* (Report, August 2018) <www.weforum.org/reports/the-new-physics-of-financial-services-how-artificial-intelligence-is-transforming-the-financial-ecosystem/>.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

'few organizations take full advantage of data generated outside their walls. A well-structured plan for using external data can provide a competitive edge'.[102]

Policymakers use the same rhetoric of promoting 'innovation' and encourage data collection through schemes such as open banking.[103] The aim of open banking is to give consumers the ability to direct companies that hold financial data about themselves to make it available to financial (or other) companies of the consumer's choice. Thus, it makes it possible for organisations to get access to consumers' information they could never get from a consumer directly, such as for example their transaction data for the past ten years.

Jurisdictions such as the EU, United Kingdom, Australia, and Hong Kong have recently adopted regulation promoting open banking, or 'open finance' more generally.[104] The frameworks are praised by the industry as 'encourag[ing] the development of innovative products and services that help consumers better engage with their finances, make empowered decisions and access tailored products and services'.[105]

While open banking is making it possible for financial firms to develop new products for consumers, the jury is still out as to the scheme's universally positive implications for consumers and markets.[106] One thing that is clear, however, is that because of its very nature, open banking contributes to information and power asymmetry between consumers and Automated Banks.

Traditionally, in order to receive a financial product, such as a loan or an insurance product, consumers would have to actively provide relevant data, answering questions or prompts, in relation to their income, spending, age, history of loan repayments, and so on. Open banking – or open finance more broadly – means that consumers can access financial products without answering any questions. But these questions provided a level of transparency to consumers: they knew what they were being asked, and were likely to understand why they were being asked such questions. But when an individual shares their 'bulk' data, such as their banking transaction history, through the open banking scheme, do they really know what a financial firm is looking for and how it is being used? At the same time, in such a setting, consumers are deprived of control over which data to share (for

[102] Mohammed Aaser and Doug McElhaney, 'Harnessing the Power of External Data' (Article, 3 February 2021) *McKinsey Digital*.

[103] Nydia Remolina, 'Open Banking: Regulatory Challenges for a New Forum of Financial Intermediation in a Data-Driven World' (SMU Centre for AI & Data Governance Research Paper No 2019/05, 28 October 2019).

[104] EMEA Center for Regulatory Strategy, 'Open Banking around the World' *Deloitte* (Blog Post) <www.deloitte.com/global/en/Industries/financial-services/perspectives/open-banking-around-the-world.html>.

[105] UK Finance, 'Exploring Open Finance' (Report, 2022) <www.ukfinance.org.uk/system/files/2022–05/Exploring%20open%20finance_0.pdf>.

[106] Joshua Macey and Dan Awrey, 'The Promise and Perils of Open Finance' *Harvard Law School Forum on Corporate Governance* (Forum Post, 4 April 2022) <https://corpgov.law.harvard.edu/2022/04/04/the-promise-and-perils-of-open-finance/>.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

example, they cannot just hide transaction data on payments they made to merchants such as liquor stores or pharmacies). The transparency for financial firms when data is shared is therefore significantly higher than in 'traditional' settings – but for consumers the process becomes more opaque.[107]

4.4 CAN CORPORATE SECRECY COEXIST WITH CONSUMER RIGHTS? POSSIBLE REGULATORY SOLUTIONS

ADM tools contribute to maintaining corporate secrecy of Automated Banks, and as we argue in this chapter, legal systems perpetuate, encourage, and feed the opacity further. The opacity then increases the risk of consumer harm, such as discrimination, which is more difficult to observe, and more challenging to prove.

In this section we provide a brief outline of potential interventions that may protect against AI-facilitated harms, particularly if applied synchronously. This discussion does not aim to be exhaustive, but rather aims to show something can be done to combat the opacity and resulting harms.

Interventions described in academic and grey literature can be divided into three broad categories: (1) regulations that prevent businesses from using harmful AI systems in financial markets, (2) regulations that aid consumers to understand when ADM systems are used in financial markets, and (3) regulations that facilitate regulator monitoring and enforcement against AI-driven harms in financial markets. Approaches to design (including Transparency by Design[108]) are not included in this list, and while they may contribute to improved consumer outcomes, they are beyond the scope of this chapter.

The somewhat provocative title of this section asks if corporate secrecy is the real source of the AI-related harms in the described context. The interventions outlined below focus on preventing harms, but can the harms really be prevented if the opacity of corporate practices and processes is not addressed first? Corporate secrecy is the major challenge to accountability and scrutiny, and consumer rights, including right to non-discrimination, cannot be guaranteed in an environment as opaque as it currently is. We submit that the regulatory interventions urgently needed are the ones that prevent secrecy first and foremost. AI and ADM tools will continue to evolve, and technology as such is not a good regulatory target[109] – the focus must be on harm prevention. Harms can only be prevented if the practices of financial firms, such as credit scoring discussed in this chapter, are transparent and easily monitored both by regulators and consumers.

[107] Bednarz et al, 'Insurance Underwriting in an Open Data Era'.

[108] Heike Felzmann et al, 'Towards Transparency by Design for Artificial Intelligence' (2020) 26 (6) *Science and Engineering Ethics* 3333, 3343–53.

[109] Lyria Bennett Moses, *How to Think about Law, Regulation and Technology: Problems with 'Technology' as a Regulatory Target* (SSRN Scholarly Paper No ID 2464750, Social Science Research Network, 2013) 18–19.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

4.4.1 *Preventing Automated Banks from Designing Harmful AI Systems*

International and national bodies in multiple jurisdictions have recently adopted, or are currently debating, various measures with an overarching aim of protecting consumers from harm. For example, the US Federal Trade Commission has provided guidance to businesses using AI, explaining that discriminatory outcomes resulting from the use of AI would contravene federal law.[110] The most comprehensive approach to limiting the use of particular AI tools can be found in the EU's proposed *Artificial Intelligence Act*. Its Recital 37 specifically recommends that 'AI systems used to evaluate the credit score or creditworthiness of natural persons should be classified as high-risk AI systems'. This proposal is a step towards overcoming some opaque practices, through the provision of 'clear and adequate information to the user' along with other protections that enable authorities to scrutinise elements of ADM tools in high-risk contexts.[111] Early criticisms of the proposed Act note that while a regulatory approach informed by the context in which ADM is used has some merit, it does not cover potentially harmful practices such as emotion recognition and remote biometric identification,[112] which could be used across a range of contexts, generating data sets that may later be used in other markets such as financial services.

An alternative approach to regulating AI systems before they are used in markets is to limit the sources of information that can be used by ADM tools, or restrict the ways in which information can be processed. In addition to privacy protections, some jurisdictions have placed limitations on the kinds of information that can be used to calculate a credit score. For example, in Denmark, the financial services sector can use consumers' social media data for marketing purposes but is explicitly prohibited from using this information to determine creditworthiness.[113] Similarly, the EU is considering a Directive preventing the use of personal social media and health data (including cancer data) in the determination of creditworthiness.[114] Such prohibitions are, however, a rather tricky solution: it may be difficult for the regulation to keep up with a growing list of data that should be excluded from

[110] Elisa Jilson, 'Aiming for Truth, Fairness and Equity in Your Company's Use of AI' *US Federal Trade Commission* (Business Blog Post, 19 April 2021) <www.ftc.gov/business-guidance/blog/2021/04/aiming-truth-fairness-equity-your-companys-use-ai>.

[111] European Commission, 'Regulatory Framework Proposal on Artificial Intelligence' *European Commission* (Web Page) <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai#:~:text=encourages%20dangerous%20behaviour.-,High%20risk,life%20(e.g.%20scoring%20of%20exams>.

[112] Daniel Leufer, 'EU Parliament's Draft of AI Act: Predictive Policing Is Banned, but Work Remains to Protect People's Rights' (4 May 2022) *Access Now* <www.accessnow.org/ai-act-predictive-policing/>.

[113] Hohnen et al, 'Assessing Creditworthiness'.

[114] Proposal for a Directive of the European Parliament and of the Council on consumer credits [2021] OJ COM 347 (47).

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

analysis.[115] One way of overcoming this challenge would be to avoid focusing on restricted data sources, and instead create a list of acceptable data sources, which is a solution applied for example in some types of health insurance.[116]

Imposing limits on how long scores can be kept and/or relied on by Automated Banks is another important consideration. In Australia, credit providers are bound by limits that stipulate the length of time that different pieces of information are held on a consumer's file: credit providers may only keep financial hardship information for twelve months from the date the monthly payment was made under a financial hardship arrangement, whereas court judgements may be kept on record for five years after the date of the decision.[117] In Denmark, where the credit reporting system operates as a 'blacklist' of people deemed more likely to default, a negative record (for instance, an unpaid debt) is deleted after five years, regardless of whether or not the debt has been paid.[118] A challenge with these approaches is that the amount of time particular categories of data may be kept may not account for proxy data, purchased data sets, and/or proprietary scoring and profiling systems that group consumers according to complex predictions that are impossible to decode.

4.4.2 *Aiding Consumers Understand When ADM Systems Are Used in Financial Services*

Despite the development of many principles-based regulatory initiatives by governments, corporates, and think tanks,[119] few jurisdictions have legislated protections that require consumers to be notified if and when they have been assessed by an automated system.[120] In instances where consumers are notified, they may be unable to receive an understandable explanation of the decision-making process, or to seek redress through timely and accessible avenues.

Consumers face a number of challenges in navigating financial markets, such as understanding credit card repayment requirements[121] and failing to accurately

[115] For examples of such potentially harmful data sources see: Pasquale, *The Black Box Society*, 21, 31; Hurley and Adebayo, 'Credit Scoring in the Era of Big Data', 151–52, 158; Hiller and Jones, 'Who's Keeping Score?'.

[116] E.g., health insurers in the United States under the US Public Health Service Act, 42 USC § 300gg(a)(1)(A) may only base their underwriting decisions on four factors: individual or family coverage; location; age; and smoking history.

[117] 'Your Credit Report', *Financial Rights Legal Centre* (Web Page, 6 February 2017) <https://financialrights.org.au/>.

[118] Hohnen et al, 'Assessing Creditworthiness', 40.

[119] Anna Jobin, Marcello Ienca, and Effy Vayena, 'The Global Landscape of AI Ethics Guidelines' (2019) 1(9) *Nature Machine Intelligence* 389, 2–5.

[120] See e.g. Art. 22 GDPR.

[121] Jack B Soll, Ralph L Keeney, and Richard P Larrick, 'Consumer Misunderstanding of Credit Card Use, Payments, and Debt: Causes and Solutions' (2013) 32(1) *Journal of Public Policy & Marketing* 66, 77–80.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

assess their credit.[122] For individuals, it is crucial to understand how they are being scored, as this will make it possible for them to be able to identify inaccuracies,[123] and question decisions made about them. Credit scoring is notoriously opaque and difficult to understand, so consumers are likely to benefit from requirements for agencies to simplify and harmonise how scores are presented.[124]An example of a single scoring system can be found in Sri Lanka, where credit ratings, or 'CRIB Scores' are provided by the Central Information Bureau of Sri Lanka, a public-private partnership between the nation's Central Bank and a number of financial institutions that hold equity in the Bureau. The Bureau issues CRIB Score reports to consumers in a consistent manner, utilising an algorithm to produce a three-digit number ranging from 250 to 900.[125] In Sri Lanka's case, consumers are provided with a singular rating from a central agency, and although this rating is subject to change over time, there is no possibility of consumers receiving two different credit scores from separate providers.

Providing consumers with the opportunity to access their credit scores is another (and in many ways complementary) regulatory intervention. A number of jurisdictions provide consumers with the option to check their credit report and/or credit score online. For example, consumers in Canada[126] and Australia[127] are able to access free copies of their credit reports by requesting this information directly from major credit bureaus. In Australia, consumers are able to receive a free copy of their credit report once every three months.[128]

However, such approaches have important limitations. Credit ratings are just one of many automated processes within the financial services industry. Automated Banks, with access to enough data, can create their own tools going outside the well-

[122] Marsha Courchane, Adam Gailey, and Peter Zorn, 'Consumer Credit Literacy: What Price Perception?' (2008) 60(1) *Journal of Economics and Business* 125, 127–38.

[123] Beth Freeborn and Julie Miller, *Report to Congress under Section 219 of the Fair and Accurate Credit Transactions Act of 2003* (Report, January 2015) i <www.ftc.gov/system/files/documents/reports/section-319-fair-accurate-credit-transactions-act-2003-sixth-interim-final-report-federal-trade/150121factareport.pdf>. In one study of 1001 US consumers, 26 per cent found inaccuracies in their credit reports.

[124] Heather Cotching and Chiara Varazzani, *Richer Veins for Behavioural Insight: An Exploration of the Opportunities to Apply Behavioural Insights in Public Policy* (Behavioural Economics Team of the Australian Government, Commonwealth of Australia, Department of the Prime Minister and Cabinet, 2019) 1, 14. Studies have shown simplifying and standardising information in consumer markets aids comprehension and assists consumers in making choices that result in better outcomes.

[125] Credit Information Bureau of Sri Lanka, 'CRIB Score Report Reference Guide' (Guide) <www.crib.lk/images/pdfs/crib-score-reference-guide.pdf>.

[126] 'Getting Your Credit Report and Credit Score' *Government of Canada* (Web Page) <www.canada.ca/en/financial-consumer-agency/services/credit-reports-score/order-credit-report.html>.

[127] 'Access Your Credit Report' *Office of the Australian Information Commissioner* (Web Page) <www.oaic.gov.au/privacy/credit-reporting/access-your-credit-report>.

[128] Ibid.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

established credit rating systems. Also, it is consumers who are forced to carry the burden of correcting inaccurate information which is used to make consequential decisions about them, while often being required to pay for the opportunity to do so.[129]

In addition, explainability challenges are faced in every sector that uses AI, and there is considerable investigation ahead to determine the most effective ways of explaining automated decisions in financial markets. It has been suggested that a good explanation is provided when the receiver 'can no longer keep asking why'.[130] The recent EU Digital Services Act[131] emphasises such approach by noting that recipients of online advertisements should have access to 'meaningful explanations of the logic used' for 'determining that specific advertisement is to be displayed to them'.[132]

Consumer experience of an AI system will depend on a number of parameters, including format of explanations (visual, rule-based, or highlighted key features), their complexity and specificity, application context, and variations suiting users' cognitive styles (for example, providing some users with more complex information, and others with less).[133] The development of consumer-facing explainable AI tools is an emerging area of research and practice.[134]

A requirement of providing meaningful feedback to consumers, for example, through counterfactual demonstrations,[135] would make it possible for individuals to understand what factors they might need to change to receive a different decision. It would also be an incentive for Automated Banks to be more transparent.

4.4.3 *Facilitating Regulator Monitoring and Enforcement of ADM Harms in Financial Services*

The third category of potential measures relies on empowering regulators, thus shifting the burden away from consumers. For example, regulators need to be able

[129] Some consumers discovered that their reports 'featured inconsistent or misleading claims descriptions and statuses, included personal information unrelated to insurance at all, and no explanation of the terms used to assist in comprehensibility'. See Roger Clarke and Nigel Waters, *Privacy Practices in the General Insurance Industry* (Financial Rights Legal Centre Report, April 2022) vii <https://financialrights.org.au/wp-content/uploads/2022/04/2204_PrivacyGIReport_FINAL.pdf>.

[130] Leilani Gilpin et al, 'Explaining Explanations: An Overview of Interpretability of Machine Learning' (2019) v3 *arXiv*, 2 <https://arxiv.org/abs/1806.00069>.

[131] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services and amending Directive 2000/31/EC (Digital Services Act) [2022] OJ L 277/1, para 27.10.2022.

[132] Ibid, para 52.

[133] Yanou Ramon et al, 'Understanding Consumer Preferences for Explanations Generated by XAI Algorithms' (2021) *arXiv*, 9–14 <http://arxiv.org/abs/2107.02624>.

[134] Jessica Morley et al, 'From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices' (2020) 26(4) *Science and Engineering Ethics* 2141.

[135] Rory Mc Grath et al, 'Interpretable Credit Application Predictions with Counterfactual Explanations' (2018) v2 *arXiv*, 4–7 <https://arxiv.org/abs/1811.05245>.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

to 'look under the hood' of any ADM tools, including these of proprietary character.[136] This could be in a form of using explainable AI tools, access to raw code, or ability to use dummy data to test the model. A certification scheme, such as quality standards, is another option, the problem however is the risk of 'set and forget approach'. Another approach to providing regulators insight into industry practices is the establishment of regulatory sandboxes, which nevertheless have limitations.[137]

Financial institutions could also be required to prove a causal link between the data that they use to generate consumer scores, and likely risk. Such approach would likely reduce the use of certain categories of data, where correlations between data points would not be supported by a valid causal relationship. For example, Android phone users are reportedly safer drivers than iPhone users,[138] but such rule would prevent insurers from taking this into account when offering a quote on car insurance (while we do not suggest they are currently doing so, in many legal systems they could). In practice, some regulators are looking at this solution. For example, while not going as far as requiring direct causal link, the New York State financial regulator requires a 'valid explanation or rationale' for underwriting of life insurance, where external data or external predictive models are used.[139] However, such approach could result in encouraging financial services providers to collect more data, just to be able to prove the causal link,[140] which may again further disadvantage consumers and introduce more, not less, opacity.

4.5 CONCLUSIONS

Far from being unique to credit scoring, the secrecy of ADM tools is a problem affecting multiple sectors and industries.[141] Human decisions are also unexplainable and opaque, and ADM tools are often made out to be a potential, fairer and more transparent, alternative. But the problem is secrecy increases, not decreases, with automation.[142]

There are many reasons for this, including purely technological barriers to explainability. But also, it is obviously cheaper and easier not to design and use

[136] Ada Lovelace Institute, *Technical Methods for the Regulatory Inspection of Algorithmic Systems in Social Media Platforms* (December 2021) <www.adalovelaceinstitute.org/wp-content/uploads/2021/12/ADA_Technical-methods-regulatory-inspection_report.pdf>.

[137] Sophie Farthing et al, *Human Rights and Technology* (Australian Human Rights Commission, 1 March 2021) 1, 95–97.

[138] Henry Hoenig, 'Sorry iPhone Fans, Android Users Are Safer Drivers' *Jerry* (Blog Post, 20 April 2023) <https://getjerry.com/studies/sorry-iphone-fans-android-users-are-safer-drivers>.

[139] New York State Department of Financial Services Circular Letter No 1 (2019), 18 January 2019, 'RE: Use of External Consumer Data and Information Sources in Underwriting for Life Insurance'.

[140] Gert Meyers and Ine Van Hoyweghen, '"Happy Failures": Experimentation with Behaviour-Based Personalisation in Car Insurance' (2020) 7(1) *Big Data and Society* 1, 4.

[141] See for example Chapters 8, 10 and 11 in this book.

[142] Pasquale, *The Black Box Society*.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press

transparent systems. As we argue in this chapter, opacity is a choice made by organisations, often on purpose, as it allows them to evade scrutiny and hide their practices from the public and regulators. Opacity of ADM and AI tools used is a logical consequence of the secrecy of corporate practices.

Despite many harms caused by opacity, the legal systems and market practice have evolved to enable or even promote that secrecy surrounding AI and ADM tools, as we have discussed using examples of rules applying to Automated Banks. However, the opacity and harms could be prevented with some of the potential solutions which we have discussed in this chapter. The question is whether there is sufficient motivation to achieve positive social impact with automated tools, without just focusing on optimisation and profits.

https://doi.org/10.1017/9781009334297.007 Published online by Cambridge University Press