Law, Innovation and Technology

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/rlit20

A right to explanation for algorithmic credit decisions in the UK

Alison T. Lui, George Lamb & Lola Durodola

To cite this article: Alison T. Lui, George Lamb & Lola Durodola (27 Feb 2025): A right to explanation for algorithmic credit decisions in the UK, Law, Innovation and Technology, DOI: 10.1080/17579961.2025.2469352

To link to this article: https://doi.org/10.1080/17579961.2025.2469352

Published online: 27 Feb 2025.

LAW, INNOVATION AND TECHNOLOGY
https://doi.org/10.1080/17579961.2025.2469352

Routledge
Taylor & Francis Group

A right to explanation for algorithmic credit decisions in the UK

Alison T. Lui[a], George Lamb[b] and Lola Durodola[c]

[a]Law School, Liverpool John Moores University, Liverpool, UK; [b]Department of Law, University of Salford, Manchester, UK; [c]Centre for Financial and Corporate Integriy, Coventry University, Coventry, UK

ABSTRACT

This article argues for a statutory right to explanation in automated credit decision-making in the UK, as transparency and accountability are central to the rule of law. First, from a moral standpoint, we demonstrate that there is a double level of distrust in financial services and algorithms. Algorithms are unpredictable and can make unreliable decisions. Algorithmic challenges such as bias, discrimination and unfairness are exacerbated by the opacity problem commonly known as the 'black box' phenomenon. The informed consent process in automated credit decision-making is thus incomplete, which requires an ex-post right to explanation for completing the informed consent procedure. Secondly, our doctrinal and comparative legal methodologies reveal that countries such as the USA, Canada, European Union, China and Poland already provide a right to explanation to credit applicants under certain circumstances. We also present new empirical evidence of a public desire to have a right to explanation for unsuccessful credit applications.

ARTICLE HISTORY Received 30 August 2024; Accepted 4 January 2025

KEYWORDS Automated credit lending; algorithmic bias; algorithmic opacity; explainable artificial intelligence; trust; right to explanation

1. Introduction

Algorithmic processing of credit data is widely used in UK banks.[1] 'Algorithmic processing' refers to the processing of personal and non-personal data by automated systems. This includes artificial intelligence (AI) systems such as

CONTACT Alison T. Lui a.lui@ljmu.ac.uk LJMU Law School, Redmonds Building, Brownlow Hill, Liverpool L3 5UZ, UK

[1]Bank of England and Financial Conduct Authority, 'Machine Learning in UK Financial Services' (2019), https://www.fca.org.uk/publication/research/research-note-on-machine-learning-in-uk-financial-services.pdf (accessed 10 July 2024).

machine learning models.[2] The widely used description: 'black box' phenomenon, refers to the challenges of how and why algorithms arrive at a decision, given specific data input.[3] The phrase 'black box' is used because such machine learning algorithms are very complex and constantly adapt to new input through 'deep learning'. 'Deep learning' is a sub-set of machine learning. It allows computers to learn and think independently. In 'deep learning' technology, the decision-making process is often opaque and difficult to explain in a way that can be easily understood by humans. This inscrutability can lead to distrust in algorithms. The stakes are especially high when such algorithms are applied to decisions on access to finance because some of the most inscrutable machine learning models are used.

Indeed, the IIF survey of 2019 shows that financial industry stakeholders such as supervisors, auditors and bank employees are concerned about the 'black box' and explainability problems.[4] They are also fearful of associated problems such as transparency, auditability and interpretability of results. The 'black box' problem is particularly problematic in risk assessment modelling because the algorithms used are too complex to explain or companies will use trade secrets as a defence to explain. The risk is that the algorithms can entrench historic biases, even if protected characteristics such as gender and race are removed from the data input. This means that algorithms might inadvertently create biased outcomes due to, for example, unrepresentative training data. This can lead to unintended harms and outcomes, which include bias or price discrimination against some customers, especially vulnerable groups, in automated credit making decisions.[5] Whilst it is illegal to consider protected characteristics such as race or gender in credit scoring algorithms, algorithms can learn and exploit information such as an applicant's educational level or address to correlate with other demographic information, which can lead to gender or racial biases.

Algorithmic credit decision-making raises the question of whether the process is fair, and its automated nature exacerbates this fairness issue due to its opacity.[6] This opacity makes it almost impossible to find out what

[2]Digital Regulation Co-operation Forum, 'The Benefits and Harms of Algorithms: A Shared Perspective from the Four Digital Regulators' (2022), https://www.gov.uk/government/publications/findings-from-the-drcf-algorithmic-processing-workstream-spring-2022/the-benefits-and-harms-of-algorithms-a-shared-perspective-from-the-four-digital-regulators#fn:29 (accessed 10 July 2024).

[3]A Azzutti, W Ringe and H Stiehl, 'Machine Learning, Market Manilupation, and Collusion on Capital Markets: Why the 'Black Box' Matters' (2021) 43 *University of Pennsylvania Journal of International Law* 79.

[4]Institute of International Finance, 'Machine Learning in Credit Risk' (*Machine Learning in Credit Risk*, 2019) 1, https://www.iif.com/Portals/0/Files/content/Research/iif_mlcr_2nd_8_15_19.pdf (accessed 17 July 2024).

[5]Robert Bartlett and others, 'Consumer-Lending Discrimination in the FinTech Era' (2022) 143 *Journal of Financial Economics* 30, https://www.sciencedirect.com/science/article/pii/S0304405X21002403.

[6]Tal Zarsky, 'The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making' (2016) 41 *Science, Technology, & Human Values* 118.

caused bias or discrimination. The right to an explanation, coupled with transparency on the criteria and weightings used in the credit application, will provide customers with relevant feedback for them to improve their credit scores. In Pavlidis's view, extending the availability of explanations promotes informed decision-making, which promotes transparency and accountability of AI.[7] Internal audits of data sets and processes can also alleviate the problems of bias and discrimination. Concerns regarding transparency, auditability and interpretability of results arising from the 'black box' phenomenon can be partially resolved by developing more maturity with the machine learning techniques and being extra cautious when checking input data.[8]

In 2019, two thirds of the surveyed respondents in the UK financial sector reported that they already use machine learning in financial services. Globally, the IIF survey reveals that 37% of the surveyed 60 international institutions utilise full machine learning algorithms for credit scoring. In the domain of credit analysis, 38% of retail bank respondents use AI enabled credit analytics. In light of this increased trend in the use of AI in finance, it is crucial that we build upon the extant AI and Law literature, particularly to the question of whether a right to explanation to automated credit decision-making is necessary in the UK.

We contribute to this debate by critically evaluating why the use of algorithms in credit processing is problematic. We analyse whether the risks of bias, unfairness and price discrimination are worse when algorithmic processing is used in credit scoring. Secondly, we examine whether current UK legislation is providing sufficient protection to retail banking consumers in relation to algorithmic processing using the doctrinal method. A comparative legal analysis is adopted to compare the positions in the UK, USA, European Union (EU), Poland and China. Finally, our empirical surveys of bank employees and the general public add originality and value to the debate of algorithmic credit scoring. The survey results reveal a desire for a right to explanation, coupled with transparency on the criteria and weightings used in the credit application from banks when credit applications are unsuccessful.

Our article makes four contributions. First, it provides a critical and contemporary review of algorithmic processing in credit scoring and a balanced assessment of the 'black box' phenomenon in the UK financial sector. Second, it conducts a critical evaluation of how UK consumers are protected from the risks of algorithmic processing of credit scoring under existing UK consumer protection legislation, the Data Protection Act 2018. Thirdly, it

[7]G Pavlidis, 'Unlocking the Black Box: Analysing the EU Artificial Intelligence Act's Framework for Explainability in AI' (2024) 16 *Law, Innovation and Technology* 293, https://doi.org/10.1080/17579961.2024.2313795.

[8]Institute of International Finance (n 4).

introduces new empirical evidence of public surveys regarding a desire from the public to have a right to explanation for unsuccessful credit applications. Finally, we argue for a statutory right to meaningful and accessible information to automated decision making, which should include objective criteria and weightings used by banks.

This article is structured as follows. Following this introductory Section 1, in Section 2 we critically evaluate the algorithmic processes utilised in credit decision-making. Section 3 evaluates the legal reasons for a right to explanation. In Section 4, we examine the right to explanation in other jurisdictions including the European Union, US, Canada, China and Poland. Section 5 is a critical evaluation of the current UK legislative framework in relation to a right to explanation. Section 6 provides the methodology for the surveys and Section 7 is a detailed analysis of the data obtained from these surveys. In Section 8, we provide a discussion and recommendations for the UK government as to why an ex-post statutory right to explanation is needed in automated credit decision-making. Section 9 analyses the chance of our proposed right of explanation under the Labour Government of 2024. Section 10 concludes this article.

2. The use of algorithmic process in credit decision-making

We begin our evaluation of AI-driven credit decision-making processes by clarifying the different algorithmic models used in these processes. Algorithms, used in machine learning, can be very complex. Within the spectrum of machine learning, Gadzinski and Castello state that rules-based systems such as discriminant analysis, decision tree and linear regression analysis are more explainable and interpretable than neural networks which are not rules-based.[9] The challenge posed by neural networks is that we do not know which rules are used to handle unforeseen information.[10] Rules-based systems are complicated but transparent.[11] Rules-based models explore causal relationships and there is no room for machine interpretation. Thus, they are easier to explain than models not based on rules. Discriminant models are objective methods of finding the differences between good and bad customers, based on historical data of past applicants.[12] Mis-classification errors occur because for example, a customer with little credit

[9]G Gadzinski and A Castello, 'Combining White Box Models, Black Box Machines and Human Interventions for Interpretable Decision Strategies' (2022) 17 *Judgement and Decision-Making* 598.
[10]J Zerilli and others, 'Transparency in Algorithmic and Human Decision-Making: Is There a Double Standard?' (2019) 32 *Philosophy and Technology* 661.
[11]R Williams, 'Rethinking Administrative Law for Algorithmic Decision Making' (2022) 4 *Oxford Journal of Legal Studies* 468.
[12]R Fisher, 'The Use of Multiple Measurements in Taxonomic Problems' (1936) 7 *Annals of Human Genetics* 179.

history ('thin files')[13] can be incorrectly classified as a high risk and therefore a bad customer. One way to reduce mis-classification is to permit a wider range of evidence so that customers can prove that they have good repayment records with for example, utility companies.[14] This would provide more objectivity in discriminant models.

Banks use linear regression models to establish a threshold credit score. The score relies on the linear relationship between past applicant data and associated weights attached to each variable.[15] Credit repayment history is the highest weighted variable in both models.[16] For example, credit scoring services company Fairfax, Isaac and Company attach 35% to repayment history; 30% to amounts owed; 15% to length of credit history; 10% to new credit and 10% to credit mix.[17] Applicants who score above the threshold will be accepted and those who score below will be rejected. However, there will be instances where this model will deny creditworthy customers or accept customers who may be prone to default. Hence, linear regression models do not generally perform well with outliers and are not sufficiently accurate in credit scoring.[18] Decision tree models support non-linearity and are very useful to predict the probability of an event, i.e. the probability of a customer defaulting.

By mimicking the human brain, neural networks operate by mapping independent variables onto a set of corresponding dependent variables. Due to the multiple layers of neurons involved within such networks, scholars have criticised neural networks for being opaque.[19] The stakes are high in areas such as access to finance and education because the least interpretable machine learning models are used.[20] Moreover, many neural network-based machine learning models are deemed to be 'black boxes'.[21] Whilst they are often accurate, many scholars opine that the models are too difficult for humans to comprehend, let alone explain,

[13]F Ostman and C Dorobantu, 'AI in Financial Services. The Alan Turing Institute' (2021), https://www.turing.ac.uk/sites/default/files/2021-06/ati_ai_in_financial_services_lores.pdf (accessed 12 July 2024).
[14]Aire, 'Written Evidence by Aire (ALG0066)' (2017), https://committees.parliament.uk/writtenevidence/80004/html/.
[15]Terry Harris, 'Quantitative Credit Risk Assessment Using Support Vector Machines: Broad versus Narrow Default Definitions' (2013) 40 *Expert Systems with Applications* 4404, https://www.sciencedirect.com/science/article/pii/S0957417413000754.
[16]Nikita Aggarwal, 'The Norms of Algorithmic Credit Scoring' (2021) 80 *The Cambridge Law Journal* 42, https://www.cambridge.org/core/product/identifier/S0008197321000015/type/journal_article (accessed 23 July 2024).
[17]Lorena Rodriguez, 'All Data Is Not Credit Data: Closing the Gap between the Fair Housing Act and Algorithmic Decision Making in the Lending Industry' (2020) 120 *Columbia Law Review* 1843, https://www.proquest.com/scholarly-journals/all-data-is-not-credit-closing-gap-between-fair/docview/2468394719/se-2?accountid=12118.
[18]D West, 'Neural Network Credit Scoring Models' (2000) 27 *Computers and Operations Research* 1131.
[19]Williams (n 11); Y Wang, S Wang and K Lai, 'A New Fuzzy Support Vector Machine to Evaluate Credit Risk' (2005) 13 *IEEE Transactions on Fuzzy Systems* 820.
[20]J Waa and others, 'Interpretable Confidence Measures for Decision Support Systems' (2020) 144 *International Journal of Human Computer Studies* 1.
[21]Gadzinski and Castello (n 9).

even by subject experts.[22] This can lead to credit finance decision outcomes which are biased, discriminatory and unfair for certain customers.

In the next section, we focus on the right to explanation for automated decision-making. We commence by stating the legal foundations for the right to explanation in the UK and EU jurisdictions as being Article 22 GDPR 2016 and the academic debate surrounding this. We then critically evaluate the recent Advocate General's (AG) opinion of *SCHUFA Holding and Others (Scoring)* before the Court of Justice for the European Union (CJEU),[23] (hereinafter, SCHUFA case) which features useful advice from Advocate-General Pikamae as to what constitutes 'meaningful' information and explanation to customers. This is followed, in section 4, by a critical examination of the statutory right to explanation in the US, Canada, China and Poland.

3. The right to explanation in automated decision-making

Post-Brexit, the UK retains the GDPR under the UK Data Protection Act 2018. Thus, Article 22 remains applicable and protects data subjects when automated individual decision-making and profiling take place. A data subject should have 'the right not to be subject to a decision based solely on automated processing … which produces legal effects concerning him or her or similarly significantly affects him or her'.[24] Article 22 only applies to automated decision-making with 'legal effects' or 'similarly significant effects'. Recital 71 names two examples of such effects: automatic refusal of an online credit application and e-recruiting practices. The wording of Recital 71 has created a debate as to whether it gives individuals the right to an explanation. Recital 71 states that

> such processing should be subject to suitable safeguards, which should include … the right to obtain human intervention, to express his or her point of view, **to obtain an explanation of the decision reached after such assessment and to challenge the decision**. (emphasis added)

The phrase highlighted in bold found in the non-binding Recital 71 (but missing from Article 22 GDPR) has triggered an ongoing global debate on whether a customer (data subject) should be given a right to explanation of automated decision-making within the GDPR framework – witness, for example, contributions by Goodman and Flaxman;[25] Wachter

[22] E Pinetlas, I Livieris and P Pintelas, 'A Grey-Box Ensemble Model Exploiting Black-Box Accuracy and White-Box Intrinsic Interpretability' (2020) 13 *Algorithms* 17.
[23] Judgment of 7 December 2023, Case C-634/21 *SCHUFA Holding and Others (Scoring)*.
[24] Article 22, UK GDPR.
[25] B Goodman and S Flaxman, 'EU Regulations on Algorithmic Decision-Making and a 'Right to Explanation'' (*arXiv.org*, 2016), https://arxiv.org/abs/1606.08813 (accessed 18 September 2023).

et al;[26] Selbst and Powles.[27] These contributions, inter alia, discuss the legal basis of the right to explanation; as well as the alleged practical and mechanical challenges of explaining algorithms to the public. Scholars' views on Article 22 UK GDPR range from the narrow approach adopted by Wachter et al[28] requiring only prior explanations of how the system functions and not an explanation of reasons for decisions obtained through the functions of such systems to Casey, Farhangi and Vogl's[29] opinion that there is a right to explanation under Article 22 GDPR. Adopting a holistic approach, Casey et al posit that although Article 22 may not provide full transparency to data subjects seeking a complete explanation, the European data protection agencies see Article 22 not just as a remedy, but as a broader algorithmic design mechanism. Casey et al's submission is that Article 22 is more powerful when combined with the 'data protection by design' principles in Article 25 of the GDPR. These principles promote compliance in automated systems during the product's whole cycle to reduce algorithmic bias.

In contrast, Wachter et al adopt a very narrow interpretation of the phrase in Article 22 GDPR 'meaningful information about the logic involved', as 'a limited right to explanation of the functionality of automated decision-making systems' due to the missing phrase highlighted in bold in Recital 71. They argue for a narrow right to explanation to specific, post decision-making process because they believe that explaining specific decisions is more onerous than system functionality allows, i.e. details about the models, decision-trees, criteria.[30]

Referring to the first point, Selbst and Powles posit that whilst Recital 71 is not binding, it supports data subjects' rights in Articles 13–15 GDPR.[31] Further, they argue that a holistic and purposive interpretation of the GDPR is necessary to give effect to its goals. Articles 13(2)(f), 14(2)(g), and 15(1)(h) require companies to provide subjects with information regarding 'the existence of automated decision-making, including profiling, referred to in Article 22 … and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject'. Moreover, Recital 71 of the GDPR, which is not binding but can provide useful guidance, provides 'the right to express his or her point of view, to obtain an explanation of the decision reached after such assessment and to challenge the decision'.

[26]S Wachter, B Mittelstadt and L Floridi, 'Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation' (2017) 7 *International Data Privacy Law* 76.
[27]A Selbst and J Powles, 'Meaningful Information and the Right to Explanation' (2017) 7 *International Data Privacy Law* 233.
[28]Wachter, Mittelstadt and Floridi (n 26).
[29]B Casey, A Farhangi and R Vogl, 'Rethinking Explainable Machines: The GDPR's "Right to Explanation" Debate and the Rise of Algorithmic Audits in Enterprise' (2019) 34 *Berkeley Technology Law Journal* 143.
[30]Wachter, Mittelstadt and Floridi (n 26).
[31]Selbst and Powles (n 27).

Referring to the second point, Mendoza and Bygrave[32] refute Wachter et al's submission that it is not possible to provide explanations to specific individual decisions due to system complexity. Some machine learning algorithms are deterministic and once there is data input and the same model is used, the outputs are the same. There are no 'case specific decision rules'. Mendoza and Bygrave[33] advocate that this right to explanation should be interpreted flexibly and functionally which support data subjects' human rights. Do their human rights extend to a right to explanation? Article 10 of the European Convention on Human Rights (ECHR) states that: 'Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers.' A broad interpretation of 'receiving information' would support Articles 13(2)(f), 14(2)(g), and 15(1)(h) of the GDPR, which require companies to give data subjects 'meaningful information about the logic involved' in the decision-making. Recital 71 of the GDPR further endorses the right to express views, supporting the freedom of expression under Article 10 of the ECHR.

Kaminski[34] submits that the right to explanation is a prerequisite to other rights of the applicants. It is only after obtaining an explanation that an applicant can express their view and challenge the decision if they wish. Her views are broadly consistent with Selbst and Powles[35] as well as Mendoza and Bygrave.[36] The right to challenge a decision aligns with the principles of 'lawfulness, fairness and transparency' in Article 5(1)(a) of the GDPR. It is only right that unsuccessful credit applicants can challenge a decision if for example there is either substantive unfairness (incorrect information being used) or procedural unfairness (protection and non-disclosure of algorithms as trade secrets).

Nevertheless, the right to explanation is not an absolute right. It requires careful balance of all stakeholders' legitimate interests. Particularly tricky is the balance of lenders' algorithms as trade secrets against data subjects' right to an explanation. Algorithms give financial institutions a competitive business advantage. They are often protected as intellectual property assets in the form of trade secrets or patents, which are of immense value to the financial institutions. However, trade secrets should not shield financial institutions from transparency and accountability. The AG's opinion of *SCHUFA Holding and Others (Scoring)* (Case C-634/21 in 2023) from the CJEU rebuts Wachter et al's submission that their proposed narrow right

[32]I Mendoza and L Bygrave, 'The Right Not to Be Subject to Automated Decisions Based on Profiling', *EU Internet Law: Regulation and Enforcement* (2017).
[33]ibid.
[34]M Kaminski, 'The Right to Explanation, Explained' (2018) 34 *Berkeley Technology Law Journal* 189.
[35]Selbst and Powles (n 27).
[36]Mendoza and Bygrave (n 32).

to explanation 'could be heavily curtailed to protect the controller's interests (i.e. trade secrets, intellectual property)'.[37] In expressing this, Wachter et al relied on first, Recitals 47 and 63 of the GDPR, which recognise that data controllers may have relevant rights and interests and they also expressly provide for the overriding rights of data controllers. Secondly, there is also the German legal commentary and jurisprudence such as the German judgment of SCHUFA Holdings and Others. The German SCHUFA judgements[38] reveal that data subjects do not have the right to investigate how accurate the credit making algorithm process is. This is because trade secrets law protects the underlying formulae used in algorithms. Since their article in 2017 however, the case has been referred to the CJEU. The latest CJEU philosophy is to facilitate 'meaningful information' and explanation to data subjects using appropriate communication methods. On the issue of trade secrets, AG Priit Pikaemae said (in para 56):

> It follows that, although the protection of trade secrets or intellectual property constitutes, in principle, a legitimate reason for a commercial information company to refuse to disclose the algorithm used to calculate the data subject's score, it cannot, on the other hand, justify an absolute refusal of information. Moreover [this is so] when there are appropriate means of communication, which facilitate understanding, while ensuring a certain degree of confidentiality.

Whilst not binding, AG Priit Pikamae's opinion provides useful guidance on what is considered 'meaningful information' under Article 15 of the GDPR. Under Art. 15(1)(h) GDPR, the natural person has the right to obtain, based on access right, amongst others, information on Automated Decisions taken in his/her respect (if any), including the existence of the Automated Decision; meaningful information on the logic involved in this case, and meaningful information on the significance and the envisaged consequences of such an Automated Decision. That said, the Institution will not be obliged to provide the aforesaid information under Art. 15 paragraph 1 letter (h) GPDR on the Automated Decision taken by the Credit Agency, because the Institution does not hold this information.

AG Pikamae submitted that this information shall include sufficient detailed explanation on the method used for the calculation of credit scoring and the causes which led to a specific result. AG Pikamae also stated that the details shall comprise general information on the factors taken into consideration for the Automated Decision and the importance of each of such factors in an aggregated manner, such information being useful to the natural person for contesting the Automated Decision. AG

[37] Wachter, Mittelstadt and Floridi (n 26).

[38] Judgment of the German Federal Court Bundesgerichtshof 28 January 2014 – VI ZR 156/13. Also LG Gießen 6 March 2013 – 1 S 301/12. Also, AG Gießen 11 October 2014 – 47 C 206/12.

Pikamae, however, does not believe that there is a need to apply the algorithm used considering its complexity. Wachter et al's submission that the right to explanation could be heavily restricted due to the data controller's right to protect its trade secrets is weakened in light of AG Pikamae's opinion.

The EU's Artificial Intelligence Act[39] (EU AI Act) 2024 provides some compromise between the right to explanation and protection of trade secrets. The EU AI Act is a landmark piece of legislation since it is the first in the world to regulate AI. The Act, which came into force on 1st August 2024, adopts a risk-based approach, so the higher the risk of harm to society an AI use case poses, the more onerous the rules that are triggered. AI used in credit scoring is according to Article 6(2) and section 5(3) Annex III a high-risk AI system. These systems must assess and reduce risks, maintain use logs, be transparent and accurate, and ensure human oversight. Under Article 11 of Annex 4, high-risk AI product developers must provide technical documents to national authorities and bodies to comply with the principles of explainability and transparency. Citizens will have a right to submit complaints about AI systems and receive explanations about decisions based on high-risk AI systems that affect their rights under Article 13 of Annex 4.

Article 13 specifies that the information provided by organisations such as banks to users should be 'concise, complete, correct and clear information that is relevant, accessible and comprehensible to users'. This is useful in terms of the language used, but different people in diverse situations require tailored explanations. There are three main categories of people that require explainability in credit scoring. First, loan officers that are said to prefer local sample-based explanations. This involves comparing the unsuccessful applicant's profile against other similar profiles. Second, rejected loan applicants that are said to prefer local feature-based explanations. This involves providing individual case specific reasons. Third, regulators or data scientists that are said to prefer global model explanations.[40] This is because by having a global picture of logic and reasoning used by the model, it enables regulators to ensure that the model is fair and consistent in making decisions. In section 7, we confirm that through our surveys of both the banking sector and the general public, most of the surveyed participants would like clear, concise accessible information of the reasons for rejecting a

[39]Regulation (EU) 2024 of the European Parliament and of the Council of laying down harmonised rules on Artificial Intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act).

[40]L Demajo, V Vella and A Dingli, 'An Explanation Framework For Interpretable Credit Scoring' (2021) 12 *International Journal of Artificial Intelligence and Applications* 19.

credit application. Thus, the local feature-based explanations are of the most use to them.

The EU also stipulates additional duties for financial intermediaries when relying on third-party IT and software systems via the EU Digital Operational Resilience Act (EU DORA) 2023,[41] which came into force on 16th January 2023. The reliance on third-party credit scores could qualify under Article 2 of the EU DORA Act 2023. By reinforcing their operational resilience against cyber-security threats, the EU DORA Act 2023 should instil trust and enhance data privacy amongst financial customers. It does not however, do much to solve the issue of algorithmic bias. Nevertheless, the EU legislative landscape provides consumer protection to financial customers in respect of explainability and privacy. We now turn to the legislative framework of the USA, Canada, Poland and China where these countries provide financial consumers a right to explanation against AI related harms to some degree.

4. The statutory right to explanation in different jurisdictions

Regulators in the USA have already called for adverse action codes in legislation for credit denial in section 1002.9 (a)(2) of Equal Credit Opportunity Act 1974 ('creditor must provide the applicant with the principal reason for the action taken') and section 15 US Code 161 g of the Fair Credit Reporting Act 1970

> key factors that adversely affected the credit score of the consumer in the model used, the total number of which shall not exceed 4 … The term 'key factors' means all relevant elements or reasons adversely affecting the credit score for the particular individual, listed in the order of their importance based on their effect on the credit score.

The White House Blueprint for an AI Bill of Rights[42] is a guide for society setting out the key principles when using automated systems. Central to this is the principle that developers and designers of AI should give 'clear, timely and accessible' explanations of outcomes to the public when automation is used. The phrase 'clear, timely and accessible' resembles Article 13 Annex 4 of the EU AI Act 2024.

Around the world, various countries provide citizens with a statutory right to explanation in relation to automated algorithmic decisions. The Canadian C-27 Digital Charter Implementation Act ('the Canadian C-27

[41] Regulation (EU) 2022/2554 Of The European Parliament And Of The Council of 14 December 2022 on digital operational resilience for the financial sector and amending Regulations (EC) No 1060/2009, (EU) No 648/2012, (EU) No 600/2014, (EU) No 909/2014 and (EU) 2016/1011.

[42] The White House Office of Science and Technology Policy, 'Blueprint for an AI Bill of Rights Making Automated Systems Work For The American People' (*Blueprint for an AI Bill of Rights*, 2022), https://www.whitehouse.gov/ostp/ai-bill-of-rights/ (accessed 22 July 2024).

Act')[43] requires companies which use automated decision-making systems for the purposes of prediction, recommendation or decision which would have a *significant impact* on them, to provide data subjects with an explanation. The explanation should indicate 'the type of personal information that was used to make the prediction, recommendation or decision, the source of the information and the reasons or principal factors that led to the prediction, recommendation or decision'.[44] Arguably, the choice of 'significant impact' in the Canadian C-27 Act is wider than 'legal effects concerning him or her or similarly significantly affects him or her' found in Article 22 of the EU GDPR.

Poland passed legislation in 2019 which amended Article 70a of its Banking Act 1997. Under this provision, customers now have a statutory right to explanation when a bank makes a credit decision. The bank must then provide the list of criteria it used, no matter whether the decision was made automatically (using algorithms), or by a bank employee. However, the Polish Financial Supervision Authority ('the OPFSA') has found that Polish banks have provided only general data categories to customers without providing the applicant's data or their financial circumstances.[45] The OPFSA recommended that Polish banks need to give applicants personalised and detailed information, including on what applicants can do to improve their creditworthiness. The OPFSA also noted that Polish banks did not give a deadline by which they would provide the explanation for an unsuccessful application. Whilst an indefinite period of asking and responding to explanations has some benefits, the OPFSA opined that it would be more practical to have a deadline by which banks have to respond to a request for explanation when turning down an application.

China specifically carved out algorithmic governance as an important pillar of its AI regulatory framework. The 2021 Recommendation Algorithm Provision; 2022 Deep Synthesis Provisions and 2023 Generative AI Measures all provide some degree of protection to users by demanding companies not to discriminate based on protected characteristics and file algorithmic information with the Cybersecurity Administration of China[46] It can be argued that China is ahead of other countries in protecting consumers from

[43]Digital Charter Implementation Act 2022. An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts.

[44]ibid.

[45]Office of the Polish Financial Supervision Authority, 'Announcement of the Office of the Polish Financial Supervision Authority' (2020), file:///C:/Users/aliso/Downloads/Komunikat_UKNF_ws_prawa_do_uzyskania_wyjasnien_nt_oceny_zdolnosci_kredytowej_wersja_szczegolowa_70332 (1).pdf.

[46]J Gong, H Qu and H Dorwart, 'AI Governance in China: Strategies, Initiatives, and Key Considerations' (*AI Governance in China: Strategies, Initiatives, and Key Considerations*, 2024), https://www.twobirds.com/en/insights/2024/china/ai-governance-in-china-strategies-initiatives-and-key-considerations#:~:text=Consequently%2Cthestatutoryresponsibilitiesof,obligations%2Cprovisionsrelatedtoautomated (accessed 22 July 2024).

algorithmic harms because first, it has established an Algorithmic Registry, where security assessments of registered algorithms are performed. It is unclear however, to what extent the Registry can perform meaningful insight into the 'black box' phenomenon.[47] Secondly, China adopted the Personal Information Protection Law (PIPL) in August 2021. The PIPL gives users the right to ask the handler to explain its decision-making and can prohibit the handler from making decisions based solely on its use. When handlers use automated decision making, they must undertake impact assessments to ensure that first, the purpose and method used by the handlers are lawful, legitimate and necessary. Second, they must also have regard to the influence on individual's rights; and finally, to consider whether protective measures are legal, effective and proportionate to the risk.[48]

5. Current UK legislative framework for a right to explanation

In contrast with the different jurisdictions noted above, the UK currently does not have equivalent legislation applicable to customers who have been denied credit. However, the UK government acknowledges in its public consultation report: 'Data: a new direction' of 2021[49] that there will be a substantial increase in the use of automated decision-making in the next few years. It also recognised that AI should be a force for good and does not inadvertently harm customers. The UK government, after noting considerable confusion amongst the public regarding Article 22 GDPR, is pondering how to reform this law. It is leaning towards providing data subjects with a right to specific safeguards, rather than a general ban on solely automated decision-making.[50]

The EU AI Act 2024 applies to businesses in the UK that develop or deploy an AI system that is used in the EU; they will need to comply with the EU AI Act. Some UK citizens are thus protected to a degree if the AI system is used in the EU. Under Article 86(1) EU AI Act 2024, they are protected if a high-risk AI system produces legal effects or similarly effects that adversely affect someone's health, safety or fundamental rights. In this case, the affected users can ask for clear and meaningful explanations of the role of AI in the decision-making process and the principal factors taken into account. Thus, rejected credit applicants, where credit scoring is a 'high risk', will benefit from the EU AI Act. Nevertheless, not all UK citizens will benefit from this right to explanation.

[47]A Kachra, 'Making Sense of China's AI Regulations' (*Making Sense of China's AI Regulations*, 2024), https://www.holisticai.com/blog/china-ai-regulation (accessed 22 July 2024).
[48]ibid.
[49]UK Government, 'Consultation Outcome Data: A New Direction – Government Response to Consultation' (*Data: A New Direction*, 2021) 1, https://www.gov.uk/government/consultations/data-a-new-direction/outcome/data-a-new-direction-government-response-to-consultation.
[50]ibid.

The recent wide interpretation of Article 15(1)(h) GDPR, the right to receive 'meaningful information about the logic involved' in automated decision-making (Article 15(1)(h) GDPR), to encompass both 'sufficiently detailed explanations of the method used to calculate the [output] and the reasons for a certain result' by AG Pikamae is a persuasive ground for UK citizens to rely on for a right to explanation. AG Pikamae's interpretation indicates that the pendulum has swung to reading a right to explanation into the EU GDPR. Further, his opinion is that decisions based solely on automated processing within the meaning of Article 22(1) GDPR should be understood to also include a fully automated part of the final decision if such a part plays a determining role in the decision (para. 48). The impact of this is that AG Pikamae's opinion has opened up Article 22 GDPR to partially automated decisions, when previously, Article 22 GDPR was thought to apply only to solely automated decisions.

AG Pikamae's opinion is not binding though, which means there is no certainty for UK citizens. Does this mean that there should be an ex-post right to explanation in algorithmic processing in the UK? Before answering this question, the UK Consumer Protection Act 1974 provides some protection to credit applicants. They can ask credit providers to check the applicants' personal details and credit history. If denied credit unfairly, applicants can ask their credit providers which credit reference agency or agencies they used. Upon payment, credit applicants can receive a copy of their credit files. If the information on the files is inaccurate or out-of-date, applicants can ask for that information to be amended under Consumer Credit (Credit Reference Agency) Regulations 2000.

However, the current consumer protection laws do not provide a right to explanation as to why a credit decision is unsuccessful or what criteria were used. We believe that the statutory rights enshrined in the EU AI Act 2024, the rights for the technical documentation (Article 11), system logs (Article 12) and instructions for use (Article 13) of a high-risk AI system, are useful but not necessarily accessible. Rather than relying on Article 86(1) EU AI Act 2024 where the applicant needs to show a high-risk AI system was used and adverse impact on their rights, it seems easier and quicker if unsuccessful credit applicants petition to their Courts under national law for information such as the technical documentation (Article 11) or system logs (Article 12). However, the average consumer is unlikely to fully understand the technical language or details of the technical documentation, system logs or instructions for use. A right to explanation will make information more accessible to consumers, because it is an opportunity to ask why a credit decision is unsuccessful or what criteria were used. This is important, because when deciding the fairness of algorithms, there is limited attention to the *subjective* fairness perceptions of the recipient of the decision that the algorithm

made.[51] Even with counterfactual AI models making AI predictions more accessible and interpretable, the decisions are not necessarily fair. According to proponents of the Fairness Theory,[52] recipients of AI decisions focus on three questions. First, would a different decision by the decision-maker result in a better outcome? Secondly, could the decision-maker have made a different decision? Finally, should the decision-maker have made a different decision? Some recipients focus more on the outcome; others focus more on the process. Shaw et al's research shows that when decision makers explain the decisions to the recipients, the latter group is more likely to accept a poor outcome, because they think the decision-making process is fair.[53] A right to explanation thus injects subjective fairness into the decision-making process.

A further reason why the UK needs a right to explanation is because explainable artificial intelligence (XAI) is more aimed at experts and focuses on technical points. XAI explains how algorithms make decisions to human *experts* and it facilitates understanding to decision-makers.[54] XAI increases confidence in businesses by encouraging compliance to regulations such as the GDPR; allowing developers to detect flaws in the models and builds customers' trust in automated decisions. AI approved loans to unpromising customers will seriously diminish trust in AI systems, much more so than human experts.[55] Janssen et al suggest that 'XAI embraces the social right to explain decisions to the public'. Explanations are often only necessary when the recipients of the outcomes are impacted negatively. XAI facilitates the right to an explanation, but XAI per se is insufficient. As explained earlier, a right to explanation incorporates a recipient's subjective fairness into the algorithmic decision-making process. This helps with the recipient, who is often less powerful than businesses and decision makers, in accepting the outcome, even if it is a poor one. To a certain extent, we also believe that the right to explanation addresses the power imbalance between decision makers and recipients of outcomes.

If a statutory right to an explanation were introduced in the UK for automated credit decision-making, it could form the basis for legal challenges in cases where applicants feel their rights have been infringed. This statutory right would provide a basis for judicial review, complaints to data protection

[51]D De Cremer, 'What Does Building a Fair AI Really Entail?' (2020) *Harvard Business Review*, https://hbr.org/2020/09/what-does-building-a-fair-ai-really-entail.

[52]R Folger and R Cropazano, 'Fairness Theory: Justice as Accountability', *Advances in organization justice* (Stanford University Press 2001).

[53]J. Shaw, E Wild and J. Colquitt, 'To Justify or Excuse? A Meta-Analytic Review of the Effects of Explanations' (2003) 88 *Journal of Applied Psychology* 444.

[54]Marijn Janssen and others, 'Will Algorithms Blind People? The Effect of Explainable AI and Decision-Makers' Experience on AI-Supported Decision-Making in Government' (2022) 40 *Social Science Computer Review* 478.

[55]Swati Sachan and others, 'An Explainable AI Decision-Support-System to Automate Loan Underwriting' (2020) 144 *Expert Systems with Applications* 113100.

authorities, and potential civil claims, particularly in scenarios where credit applicants are denied access to finance and where decisions are made in a way that is opaque or seems biased. It would likely serve as a robust foundation for challenging automated credit decisions on grounds of transparency, fairness, and accountability. By creating new avenues for appeal and redress, such a right would empower applicants to hold financial institutions accountable for opaque decision-making processes, helping to promote procedural fairness. However, implementing and balancing this right with financial institutions' interests in maintaining intellectual property protections would require careful legal guidance and oversight. This raises questions about how such a statutory right as a basis for challenges could work, how it might be operationalised –

5.1. Grounds for challenging algorithmic decisions

A statutory right to explanation would empower individuals to challenge automated credit decisions on the following grounds:

- Transparency and Fairness: Applicants could argue that they were not given sufficient or meaningful information regarding why their credit applications were denied, contrary to their statutory right to an explanation. This argument could focus on the adequacy of the explanation provided, particularly if it fails to clarify the key criteria or weightings that influenced the decision. Drawing from Selbst and Powles,[56] meaningful information must be sufficiently detailed to help applicants understand and, if necessary, contest the decision.
- Bias and Discrimination: If applicants believed the credit scoring algorithm involved bias or discrimination, whether directly, through protected characteristics, or indirectly, through proxies, they could use the statutory right to obtain explanations to identify and argue potential discriminatory patterns. As demonstrated by Mendoza and Bygrave,[57] the right to explanation allows individuals to probe algorithmic outcomes for fairness and assess whether data processing aligns with principles of anti-discrimination.
- Accuracy and Relevance: Applicants could challenge credit denials by questioning the accuracy of the data used in the decision or the relevance of the criteria applied. Under this statutory right, they would have a basis to request explanations of the data sources, input variables, and logic applied by the algorithm, making it easier to identify if incorrect or outdated information led to a flawed outcome. This aligns with the principles

[56]Selbst and Powles (n 27).
[57]Mendoza and Bygrave (n 32).

of lawfulness and accuracy under Article 5(1) of the UK Data Protection Act 2018.

5.2. Legal mechanisms and processes for challenges

If a statutory right to explanation were given, it could enable applicants to pursue challenges through various legal avenues:

- Administrative Appeals and Judicial Review: Individuals denied credit based on automated decisions could challenge the adequacy of explanations through judicial review. For instance, applicants could argue that the explanation failed to meet statutory standards or that the decision-making process breached principles of transparency and fairness under data protection law. Courts could then review whether the explanations provided were consistent with the statutory requirements and whether they offered enough insight into the criteria that led to the adverse decision.
- Complaints to Data Protection Authorities: The UK's Information Commissioner's Office (ICO) would play a key role in overseeing compliance with this statutory right. Individuals could lodge complaints with the ICO if they believe that financial institutions have failed to provide adequate explanations to them. The ICO would have the authority to investigate these complaints and, if warranted, impose fines or require remedial actions from the institution. Precedents in data protection law, for instance under the GDPR, suggest that regulators can compel organisations to provide detailed information on automated decisions affecting individuals.[58]
- Civil Claims for Damages: In cases where individuals suffer demonstrable harm from the decision-making process, such as reputational damage, financial loss, or stress, they could pursue civil claims for damages under the UK Data Protection Act 2018. AG Pikamae's interpretation in the SCHUFA case indicated that individuals should receive 'meaningful information about the logic' of decisions, suggesting that courts could support individuals seeking redress when this right is violated.[59]

5.3. Balancing trade secrets with the right to explanation

One key legal challenge would involve balancing applicants' right to explanation with financial institutions' protection of intellectual property (IP)

[58]Casey, Farhangi and Vogl (n 29).
[59]Judgment of 7 December 2023, Case C-634/21 *SCHUFA Holding and Others (Scoring)*.

and trade secrets in their proprietary algorithms. Courts and regulatory bodies might need to assess whether the institution's explanation sufficiently respects both transparency and confidentiality. As we have already noted, the SCHUFA case highlights that while trade secrets are legitimate interests, they cannot fully override individuals' rights to meaningful explanations.[60] Thus, if any financial institution claims that explaining a decision would compromise trade secrets, it may still need to disclose specific, objective criteria in a way that respects both transparency and proprietary concerns. Conversely, financial institutions may mount legal challenges themselves, which could add to operating costs that may impact them, including leading to reduced investments in improving the AI systems for the effective operationalising of the statutory right to an explanation.

5.4. Potential outcomes and remedies from a challenge

The outcomes of a successful challenge could vary depending on the nature of the claims brought and the relief sought:

- Improved Explanations: Courts or regulators may require financial institutions to improve the transparency of their automated credit decision processes. This could involve mandating that explanations include clear descriptions of the criteria and weightings that significantly influenced the decision, as well as guidance on how individuals might improve their credit standings to prepare for future applications.
- Rectification and Reassessment: If an explanation reveals that inaccurate or irrelevant data influenced the decision, applicants might be entitled to a reassessment of their credit application. Under the UK Data Protection Act, individuals have rights to rectification and correction of inaccurate personal data, which could be invoked if the automated decision relied on incorrect information. The rectification may include restoring an applicant's credit rating to a position they were in, prior to the inclusion of inaccurate data, to make good.
- Damages for Procedural or Substantive Unfairness: In cases of significant financial or psychological harm caused by inadequate explanations or biased algorithms, courts could award damages. Such claims would likely require evidence of specific harm, such as denial of financial access due to discriminatory criteria, aligning with case law around algorithmic accountability and procedural fairness.[61]

[60] Pavlidis (n 7).
[61] Zarsky (n 6).

5.5. Challenges in implementing and enforcing the statutory right to an explanation

While the right to explanation could be the basis for a challenge, practical challenges in enforcing it would remain. Financial institutions may struggle to provide explanations for highly complex models, like deep learning algorithms, which are inherently difficult to interpret and explain.[62] Furthermore, balancing between granular explanations and protecting trade secrets might create variability in the explanations provided, leading to inconsistencies across institutions and even confusion for applicants, and potential grounds for even further legal challenges.

In the next section, we evaluate the UK bank sector and general public surveys as to what the surveyed participants want regarding the right to explanation in unsuccessful automated credit decisions.

6. Surveys of public and bank employee attitudes to automated credit decision-making

To add originality to this paper, this study has also conducted an opinion survey of both the general public and bank employees on their attitudes to the use of AI for credit finance decisions. We obtained university ethical approval for the surveys (reference 21/LAW/004). The results of these public and employee opinion surveys buttress the call for a right of explanation to be introduced within the British banking and credit finance sectors. Moreover, the use of mixed methods in this article provides greater robustness for the policy arguments expressed in the previous sections of this article. Our empirical evidence reveals two very important results. First, most of both public and banking employee participants would like to see a statutory right to explanation for unsuccessful automated lending decisions. Secondly, the explanation should include the criteria used by banks in making the decision and concerns that the bank has regarding the application.

6.1. Survey inception and design process

Public opinion on automated credit decision-making in the UK is still emerging, which is why we believe that empirical research in this area is necessary and overdue. A cross-sectional design was employed to generate data about the opinions of the public and bank employees on AI and its use in banking, specifically credit decisions. Data was collected via online surveys created via

[62]Goodman and Flaxman (n 25).

the Joint Information Systems Committee Online (JISC) Surveys. For both surveys, all participants confirmed that they were over 18 years of age. The participant information sheet also stressed that it was particularly interested in hearing the views of under-represented groups in AI such as ethnic minorities and women.

In the general public survey, participants were invited via email and social media advertisement to participate in this study, which was conducted online via the JISC website. 62 participants completed at least part of the survey. For the bank sector employee survey, participants were invited to participate in the survey administered through questionnaires via email, which had a link attached they could click on to gain access to the survey. 37 bank employees completed at least part of the survey. Participants were recruited from seven banks in the UK: HSBC Holdings, Lloyds Banking Group, Royal Bank of Scotland Group, Barclays, Santander, Standard Chartered and Halifax. Whilst all these banks were contacted, only employees from Barclays and HSBC participated in the survey. Invitations for participants were disseminated via email to them, with reminders sent midway into the research timeframe. Despite our best attempts, we are aware that the sample size of 99 participants is not the largest. Nevertheless, established authority[63] shows that a sample size of a minimum of 100 participants is recommended for meaningful results. Falling one short of 100, we believe that the quality of our data is still valid and can produce meaningful results. Besides, we use a triangulation method of corroborating our data with other sources such as doctrinal and comparative legal methodologies.

The seven banks approached were chosen primarily due to their performance in the Bank of England's most recent stress test table in 2021. The Bank of England carries out stress testing to ensure banks and insurance companies are robust enough to withstand another financial crisis. The banks contacted performed best and have performed consistently well on stress tests since the Bank of England introduced stress testing in 2008, following the global financial crisis. Another reason these banks were contacted was because these banks have invested in similar types of AI that have been used in chatbots, online banking, credit decisions, and bank robots. This is evidenced by the same third party FinTech companies several UK banks have used for delivering AI solutions, products and services. For these reasons, we believe that these banks represent an accurate cross-section of banking employees' attitudes towards the use of AI in banking and credit decisions.

[63]R Gorsuch, *Factor Analysis* (Second, Erlbaum 1983).

6.2. Survey design

The public survey consisted of 25 questions: seven concerning personal and demographical information and 18 concerning attitudes and views of artificial intelligence. The bank employee survey consisted of 20 questions; seven concerning personal and demographical information and 13 concerning attitudes and views of artificial intelligence. Most questions in both surveys were multiple choice questions (consisting of sliding scales and multiple answer questions) and short answer questions.

In both surveys, Likert scales were also used to gauge participants' views. Likert scales are used in studies concerning human-robot interaction.[64] They typically provide five possible answers to a statement or question that allows respondents to indicate their positive-to-negative strength of agreement or strength of feeling regarding the question or statement. Matell and Jacoby's research[65] reveals that the number of rating categories is independent of reliability and validity. Higher number of rating categories does not lead to more reliable or valid results. The number of rating categories is very much dependent on the context. We believe that a five to seven Likert rating in the context of AI, banking and credit is sufficient. This is because the use of AI in this area is still growing, and it is unlikely that the surveyed participants will have accumulated a huge amount of experience of AI and credit making decisions. Further, our surveys are detailed and providing too many choices will increase the cognitive load and fatigue of participants.[66] Thus, a five to seven point rating scale in our view is reliable and valid.

7. Data analysis

7.1. General public survey results

98.4% of participants said they thought customers are entitled to an explanation when refused a bank loan. Of participants who thought customers are entitled to an explanation, most said they thought that a customer's right to an explanation should be a right enshrined in law. 13.3%, however, did not think this should be the case. 65.6% thought it was very important that customers had a legal right to an explanation. Human involvement in credit applications is key to many of them. 90.2% of participants thought AI and humans should work together to decide credit applications. Only 4.9% of participants thought humans or AI should work alone. Most

[64]M Schrum and others, 'Concerning Trends in Likert Scale Usage in Human-Robot Interaction: Towards Improving Best Practices' (2023) 12 *ACM Transactions on Human-Robot Interaction* 1.
[65]Michael S Matell and J Jacob, 'Is There an Optimal Number of Alternatives for Likert Scale Items? Study I: Reliability and Validity' (1971) 31 *Educational and psychological measurement* 657.
[66]Schrum and others (n 64).

participants (77%) thought transparency was very important in the decision-making process.

Participants said the most important piece of information that should be included in an explanation for an unsuccessful credit application is a full list of criteria that was used in deciding the application (83.3%). A large proportion of participants (81.7%) also said they thought a list of concerns of the bank should be included. Over half thought an explanation should include who decided the application (53.3%). A smaller proportion (43.3%) thought an explanation of the role AI played in the decision should be included and a small proportion (31.7%) thought information about who made the decision should be included.[67] When asked what other information should be included, one participant thought advice as to what the applicant could do to improve their chances to get approved next time round should be included.

In terms of communicating the explanation, 88.1% of participants thought an explanation should be provided by email, with 49.2% of participants saying it should be given by a human via telephone conversation. Just 6.8% of participants said they thought a chatbot should give an explanation. It could be argued that responses to this question indicate a distrust amongst participants about AI's ability to give a detailed and adequate explanation.

The level of trust amongst participants for banks appears to be high. 54.1% of participants said they thought banks were somewhat trustworthy or trustworthy. Only 14.7% thought banks were not trustworthy. 31.1%, however, were indifferent. 75.4% of participants said AI did not affect their trust in banks. 14.7% said it increased their trust in banks, with just 9.8% saying it made them trust banks less. The survey results show that AI is a force for good if humans remain in the loop and the decision-making process is transparent.

7.2. Bank sector survey results

At the time of completing the survey, 83.8% of participants worked in a bank, thrift (which includes associations) or credit union. Of the remaining participants, the majority preferred not to say where they worked, two said they worked in another type of lending institution and, the participant who said 'other', said they worked in 'retail banking'.

70.3% of those surveyed indicated that they were aware that AI was used in their institutions. Interestingly, eight participants (21.6%) said they were not sure if AI was used. Of the fifteen participants who provided a response to this question, seven (18.9%) participants of the thirty-seven displayed an

[67] To avoid any misunderstanding, the 31.7% who ticked this response were indicating that they thought that information should be included about the particular individual who made the decision.

awareness of the type of AI used in their banking institution. This is a very low level of awareness.

63.9% of the participants said they believed AI should be used in banks and other lending institutions. Eight (22.2%) of participants said they were not sure if AI should be used. Interestingly, just two (5.6%) participants said they thought AI should not be used. This is a very low number. The responses here indicate high support for the use of AI in banking.

Participants were marginally more supportive of AI's use in the early stages of the lending process (59.4% supported customer enquiries for quotes and applications) than the decision-making stage with lending acceptances and rejections. 57.1% of those surveyed said they thought bank employees should be involved in the credit (lending) process. This contrasts with 90.2% of the public sector participants who thought AI and humans should work together to decide credit applications. Bank survey participants opined that humans should be employed predominantly in the later stages of lending, i.e. lending offer and rejection stages (70% and 55% respectively).

86.1% of participants supported AI and humans working together in the credit (lending) process. No participant thought humans should work alone, but interestingly, two (5.6%) said they thought AI should work alone, although this number is so small in comparison to most participants who said AI and humans should work together, it has no implications for the consensus. When asked the roles of AI and human working together, the highest amount of support (35.5%) was for humans only intervening when there was a rejected application. There was an almost equal amount of support, however, for humans and AI working through every step of the process (32.3%) and AI processing the application and humans reviewing it (29%).

91.4% said that if AI works alone in automated credit decisions, they either agreed or strongly agreed that customers have a right to know how the lending decision was made. This is a most encouraging sign from the banking sector survey for our submission of a right to explanation to unsuccessful automated credit decisions. 88.9% of surveyed participants indicated that there should be transparency when AI is used to make any part of a credit decision. 86.5% agreed that customers should have a mandatory right to an explanation when a customer's credit application is rejected in the automated decision-making process.

When it came to the statement of 'Banks are transparent enough, and do not need to be mandated by UK law to explain automated credit decisions to customers', the banking sector participants did not have consensus. 29.7% disagreed that banks are transparent enough; 35.1% said they neither disagreed nor agreed that banks are transparent enough and 34.1% agreed that banks are transparent enough. This is not surprising, given that banks are concerned about protecting their algorithms as trade secrets. Our

survey shows that 51.3% believe that customers' rights to an explanation should supersede banks' need for protection of trade secrets. We note however, 29.7% of participants neither disagreed nor agreed. The authors feel that the percentage of 51.3 supporting the right to an explanation should prevail over protection of trade secrets is high, given the narrative of the high commercial value of algorithms and trade secrets.

The banking sector survey participants believe that the right to explanation should include the criteria that banks used in making the decision; detailed comments as to why the application was rejected and what type of AI was used and at what stage. The least popular option is providing a list of concerns from banks and explanations to customers. The criteria are voted by both the public and banking sector as important. Divergence can be seen with regards to providing a list of concerns from banks, which was viewed by the public sector survey participants as very important.

8. Discussion and recommendations

Our empirical evidence reveals two very important results. First, most of both public and banking employee participants would like to see a statutory right to explanation for unsuccessful automated lending decisions. Secondly, the explanation should include the criteria used by banks in making the decision and concerns that the bank has regarding the application. Such explanation requirements are already incorporated in the US, EU, Canada, China and Poland.

For UK citizens, the Consumer Protection Act 1974 provides some protection to inaccurate information and right to obtain information on credit files. However, it does not protect them from algorithmic harms such as algorithmic biases, price discrimination and opacity. That said, linear regression is most used in banks' credit scoring models. However, decision trees and neural networks can also be used as inputs to linear regression models. More advanced machine learning algorithms are used in credit scoring for larger loans. Retail credit decision-making in the UK is not yet fully automated, nor can it be regulated under Article 49 Data Protection Act 2018. AI is only an input in the overall decision-making process of credit lending. Therefore, the risk of AI harms exists but it is not at a large scale at present, unless the situation changes. Some protection is thus necessary against algorithmic harms for UK citizens.

Article 86(1) EU AI Act 2024 only gives protection to applicable users when a high-risk AI system is used and there is an adverse impact on their health, safety or fundamental rights. Although algorithmic credit scoring falls under a 'high-risk' AI system, citizens still need to prove the second limb of adverse impact. Arguably, it is easier to rely on Articles 11–13 of the EU AI Act 2024 and ask for technical documentation, system

logs and instructions for use in the national courts. However, this still requires claimants to issue their claims before national Courts.

Both the bank employees and general public survey participants are not asking for comprehensive explanations of the algorithmic innards. They would like to know what criteria and weightings are used as well as how they can improve their creditworthiness. Therefore, we do not ask for a higher standard of explanation or transparency of algorithms as compared with humans. Nor are we suggesting that the standard of explanation should be higher for neural networks compared to linear models such as discriminant models because of the more complex nature of neural networks. Our empirical results do not call for either. We respond to the survey results by proposing a statutory right to accessible, meaningful explanation, which includes the objective criteria used in the decision-making process and what type of AI is used in the credit decision making process. This will be a standalone right, similar to say Article 11 of the EU AI Act 2024, giving EU citizens the right to technical documentation. We do not think that the right to *technical* documentation will help the general public. Instead, as stated in section 4, customers would like local feature-based explanations that are tailored to them. Information to customers must be meaningful and accessible to be effective.

Meanwhile, given the prevailing CJEU philosophy in the SCHUFA case, the right to explanation in China, USA, Canada, Poland, market forces may drive organisations such as financial institutions to some form of right of explanation in algorithmic credit decision making processes. Lessig's New Chicago School approach to algorithmic bias argues that behaviour is partly shaped by market forces. The UK government acknowledges in its public consultation report: 'Data: A New Direction' of 2021 that there will be a substantial increase in the use of automated decision-making in the next few years.[68] Capgemini Research Institute conducted a survey 4,400 consumers in six countries on ethical AI. Their results reveal that 62% of consumers have more trust in companies which they perceive as practising ethical AI.[69] Any business would be financially motivated to maximise profits under the shareholder maximisation theory. Nevertheless, consumer demand for businesses practising ethical AI may shift the pendulum to a more stakeholder approach to operating businesses. The impact may be that in the future, more companies will practise ethical AI and perhaps voluntarily provide information about how the use of high-risk AI systems impacts upon customers.

[68]UK Government (n 49).

[69]Capgemini Research Institute, 'AI and the Ethical Conundrum How Organizations Can Build Ethically Robust AI Systems and Gain Trust' (2019) 1, https://www.capgemini.com/wp-content/uploads/2021/02/AI-and-the-Ethical-Conundrum-Report-1-40-1.pdf (accessed 23 July 2024).

9. A right to explanation under the Labour Government of 2024

The election of a new Labour government for the first time in 14 years could have significant consequences for the prospects of a statutory right to explanation in automated credit decision making in the UK. Labour's approach diverges noticeably from the approach of the previous Conservative government under Rishi Sunak as it seeks to take a tougher stance on the regulation of AI.

Little was mentioned in Labour's 2024 manifesto regarding AI, other than a commitment to supporting the development of the AI sector through its industrial strategy, by removing planning barriers to new data centres.[70] It also stated that it would 'ensure the safe development and use of AI models by introducing binding regulation on the handful of companies developing the most powerful AI models'.[71]

More details about the Labour government's approach to AI were, however, laid out in more detail in the King's speech at the opening of the first session of the new parliament in July 2024.[72] The King's speech stated that Labour would seek to 'establish the appropriate legislation to place requirements on those working to develop the most powerful artificial intelligence models' and that it would seek to 'harness the power of artificial intelligence as we look to strengthen safety frameworks'.[73] With the new Labour government being so early into its new term in government, the finer details of this legislation are yet to be outlined. This statement in the King's speech, coupled with its manifesto pledge, however, indicates that Labour will seek to specifically regulate developers of the most powerful AI tools and models. Whilst this does not speak to a right to an explanation of how an AI credit decision-making algorithm came to its decision, it could have implications for improving the efficacy of the algorithm itself.

Further indications about the direction of Labour's proposed new AI legislation can also be gleaned from what has been said by individuals within Prime Minister Starmer's new cabinet. The new technology and science secretary, Peter Kyle, stated in February 2024 that if a new Labour government was formed, it would seek to implement a 'statutory code' under which AI companies would be legally required to share testing data with the Government.[74] Kyle also stated that AI companies would have to

[70]Labour Manifesto (2024), https://labour.org.uk/change/ (accessed 2 November 2024).
[71]Labour Manifesto (2024) 33, https://labour.org.uk/wp-content/uploads/2024/06/Change-Labour-Party-Manifesto-2024-large-print.pdf (accessed 2 November 24).
[72]The King's Speech (2024), https://www.gov.uk/government/speeches/the-kings-speech-2024 (accessed 2 November 2024).
[73]The King's Speech (2024): Briefing Notes, https://www.gov.uk/government/publications/kings-speech-2024-background-briefing-notes (accessed 8 November 2024).
[74]D Milmo, 'Labour would force AI Firms to share their Technology's Test Data' (*The Guardian*, 4th February 2024), https://www.theguardian.com/technology/2024/feb/04/labour-force-ai-firms-share-technology-test-data (accessed 8 November 2024).

inform the Government whether they were planning to develop AI systems over a certain level of capability and would need to conduct safety tests with independent oversight, placing significant requirements on developers of AI.[75] Again, although this does not speak specifically to the introduction to a right to explanation in credit decisions, it does indicate that Labour will seek to make those tools and algorithms safer and more transparent in how they are developed.

Further detail about Labour's plans for AI was provided by the former Economic Secretary to the Treasury and City Minister, Tulip Siddiq.[76] Siddiq used a keynote speech at an industry-focused FinTech conference in Westminster in October, to heap praise on the sector's achievements in areas such as innovation and job creation whilst highlighting topic areas that the new Labour government sees as most significant. Siddiq stated that the fields of data and AI are the UK government's two priority areas when it comes to financial services. Siddiq also stated that the UK is 'in a global race to develop a "smart data" economy' and she believed that if governments and regulators worked effectively in partnership with the FinTech industry in harnessing data, that this would have huge potential for the UK economy. Importantly, Siddiq stated that, under Labour, lawmakers would aim to pass an 'AI Bill' as soon as possible, to lay the foundations for a long-term regulatory framework for what she described as 'open banking' and 'open finance'. Although the details of the proposed Bill are yet to be released, let alone passed, this again indicates a clear intention by Labour to take a tougher stance on AI when it comes to regulation.

It is too early to judge Labour's record against the previous Conservative government when it comes to AI, however, given the limited existing literature on its approach to AI, it seems Labour is seeking to strengthen AI regulation beyond what the last government was able to achieve. The previous Sunak-led Government took an agile, non-binding approach to regulating AI. The previous government did acknowledge in its AI Government Response paper, that they would consider introducing binding measures on the developers of the most capable general purpose AI systems if certain circumstances arise, but they had no plans in the short-term to do so.[77] Contrasting this with what the new Labour Government has said to date, it is clear that Labour is making AI regulation a priority as it seeks to establish AI legislation within its first parliamentary term. Whilst it cannot

[75] ibid.

[76] I Hall, 'Data and AI are 'greatest focus', UK City Minister tells Fintech Conference' (Global Government Fintech, 11th October 2024), https://www.globalgovernmentfintech.com/data-ai-focus-tulip-siddiq-fintech/ (accessed 8 November 2024).

[77] UK Government (2024). *A pro-innovation approach to AI regulation: government response*, https://www.gov.uk/government/consultations/ai-regulation-a-pro-innovation-approach-policy-proposals/outcome/a-pro-innovation-approach-to-ai-regulation-government-response (accessed 9 November 2024).

be assumed that this will include a statutory right to an explanation for algorithmic credit decisions, it could be argued that it would seem to indicate a greater likelihood that this could happen – or at least more likely to happen than under the previous Conservative government.

10. Conclusion

From a moral and ethical standpoint, we believe that there is a strong case to submit an ex-post right to explanation regarding automated credit decision-making due to the double level of distrust in financial services and algorithms. Our survey results show that customers want a right of explanation, even if they do not need detailed comments or reasons as to why their applications were unsuccessful. Algorithmic bias continues to exist in automated credit lending. Thus, we need explainable AI to safeguard data subjects' interests. Currently, there is some protection under EU legislation from algorithmic related harms. The EU AI Act 2024 and the wide interpretation of AG Pikamae in the SCHUFA case means that where Article 86(1) of the EU AI Act 2024 applies, UK citizens can benefit from a degree of right to explanation in partially automated credit finance decisions if they fulfil the relevant conditions. In reality, it is easier to petition to their national Courts and assert their statutory rights under Articles 11–13 EU AI Act. We have however, identified the right to technical documentation is unlikely to be accessible or meaningful to the public. We therefore call upon the UK government to introduce a statutory right to accessible and meaningful local feature-based information for credit finance applicants such as objective criteria and weightings used, so that data subjects in this category can be given sufficient information to improve their creditworthiness and make informed decisions.

The legal arguments and empirical results presented in this article make several important contributions to the debate over transparency and accountability of AI use in credit decision-making processes. First, we enrich the scholarship and literature on algorithmic opacity. Secondly, we contribute to the extensive scholarship[78] on the debate in the interpretation of Article 22 GDPR and the right to explanation in automated decision-making. More significantly, we demonstrated that there is a double level of distrust in both financial services and algorithms when it comes to automated credit decision-making. Therefore, there is a strong argument for an ex-post right to explanation from moral and ethical grounds.

Our study adopted a doctrinal legal analysis alongside empirical methods involving survey data analysis to draw insights from the multi-disciplinary nature of the Law and AI relationship, thereby enhancing the justification

[78]Wachter, Mittelstadt and Floridi (n 26); Selbst and Powles (n 27); Kaminski (n 34).

for law to hold AI-related decision-making processes to account through the right of explanation for credit applicants. Our empirical data analysis fills a big gap in the views of the public and banking sector on general AI issues; AI in credit lending and the right to explanation in automated credit decisions. It reveals that both the general public and even bank employees value the legal provision of such a right to explanation.

Finally, we also critically reviewed the comparative legal status of the right to explanation within different national jurisdictions such as the European Union, United States of America, Canada, China and Poland. Balancing their views and current legislation, we believe a statutory right to meaningful and accessible explanation in algorithmic credit finance decisions is practical. A statutory right to explanation will not solve all the 'black box' related challenges, but it will provide some protection to the most vulnerable groups of consumers in the UK.

Acknowledgements

We thank Professor David Ong; Professor Claire Hannibal; Professor Rosalind Jones; Professor Rebecca Parry and Dr Magnus Strand for their insightful comments of earlier drafts.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by a LJMU Qualitative Research grant.

Notes on contributors

Dr Alison T Lui is a Reader in Corporate and Financial Law at Liverpool John Moores University and the Associate Dean (Global Engagement) for the Faculty of Business and Law.

Dr George Lamb is a Sessional Lecturer in Law at the University of Salford.

Dr Lola Durodola is from Coventry University's Centre for Financial and Corporate Integrity. A senior leader at a Global Financial institution, she specialises in Financial Crime, AML and ABC. She teaches AI Law, IP Law and Law of Business Enterprises at Roehampton University.