CODE ENCOUNTERS

Algorithmic risk profiling in housing: a literature review

[Working paper 1]

Alexandra Ciocănel - Dave Beer - Roger Burrows - James Cussens - Alison Wallace

June 2022

Table of Contents

The Nuffield Foundation is an independent charitable trust with a mission to advance social well-being. It funds research that informs social policy, primarily in Education, Welfare, and Justice. It also funds student programmes that provide opportunities for young people to develop skills in quantitative and scientific methods. The Nuffield Foundation is the founder and co-funder of the Nuffield Council on Bioethics, the Ada Lovelace Institute and the Nuffield Family Justice Observatory.

The Foundation has funded this project, but the views expressed are those of the authors and not necessarily the Foundation. Visit www.nuffieldfoundation.org

Introduction

The application of algorithms, enhanced digital data driven technologies, tends to be contentious. The recent expansion of algorithms and the growing importance of data have been met with techno-utopian perspectives celebrating their benefits or techno-dystopian ones drawing attention to their pitfalls, mirroring older debates around the role of technology that have emerged during the 1960s (MacKenzie, 2006, p.26). On the one side, these tools' cost-effective, objective, and innovative aspects are praised, while on the other side, lack of transparency, increased power, control, and surveillance are blamed. Since such deterministic accounts might miss the contingent, constructed, negotiated and contestable aspects of social life, many scholars have also invited more moderate accounts. These aim to assess impacts by considering how such technologies are part of socio-technical systems, i.e., technological products of "human values, desires, and social relations" (Kitchin, 2021) and how they might be context-dependent. Consequently, it is essential to examine how algorithms and data are used in specific domains and societal contexts to fully understand their workings and effects.

In this literature review, we put such a focus by examining the usage of data and algorithms in the case of a risk profiling technology predominantly used in housing decisions: through credit scoring/reporting that impacts access to mortgage finance for homeownership and increasingly tenant referencing for entry to the private rented sector. We review academic and grey literature using as criteria of inclusion a recognition of what are key texts in the field, and a topical framework. The topical framework was derived from Amoore (2020) and involves:

- The wider context of **housing market risk**
- The **attribution of ris**k, the methods, and influences on the calculation of individual risk
- The **processes, mechanics or tools of attribution**, the people, firms, and products that facilitate the processes of attribution
- The **public and professional understanding and exercise of agency** when encountering these technologies, and
- The **impacts** of these regimes of recognition or attribution.

Since our focus is on the societal impact of these technological tools, we prioritised works by sociologists, anthropologists, legal scholars, and human geographers, although some computer science and economics studies are discussed. We employed the University of York's bibliographic databases and Google Scholar to search for the literature by using as keywords: *Algorithmic decision making; Predictive analytics; Automated risk profiling; Credit scoring; Credit rating; Tenant screening; Tenant selection*. Most of the social sciences academic literature on credit scoring is focused on the USA, hence the need to survey the grey literature to gain some knowledge about the UK and which was used only for this context. Given that the academic literature on the social aspects of credit scoring is not so extensive and that digital technologies have been proliferating lately in housing, to develop a greater sense of the context in which risk profiling occurs in housing we combined such literature with key studies on data, algorithms, and PropTech.

The review will inform the research to be undertaken in the project *Code encounters: algorithmic risk profiling in housing* funded by The Nuffield Foundation. The project aims to fill an empirical knowledge gap regarding the usage and impact of housing risk profiling technologies. Despite the importance of housing, less attention has been paid to how data and algorithms are used to make important allocation decisions based on risk profiling in contrast with insurance, health, and finance. This review discusses why this is important and identifies further research questions.

This review's main points are:

1. **Taking a domain-specific approach with a context's specific institutional frame and culture is needed.**

Although critical studies of algorithms are proliferating, there are only a relatively limited number of empirically grounded studies that illuminate the differences and variability in adopting specific technologies and their impact. More specificities are needed in order to more fully do justice to the process of co-constitution (Jasanoff, 2015) at play in the adoption of algorithms.

Further research questions: How specific usages of data and algorithms shape their contexts and are in turn shaped by the contexts in which they are adopted? What is generalisable and what is domain-specific in the study of data and algorithms in housing decisions? Why is this so? How are boundaries drawn between automation and human lead decision making?

2. **The study of algorithms involves the study of data**.

Much of the recent literature on algorithms has been focused on AI/ML techniques and their automated nature. The potential risk here is setting aside their building blocks, data. Given that most of the most critical issues that arise from algorithms are due to the quality or representativity of the training data used, it is extremely important to investigate what data is collected and how it becomes part of algorithmic risk profiling in specific contexts.

Further research questions: What types of data are used in risk assessments in housing decisions? How is their usage justified? How do they impact data subjects? What data frontiers are imagined as being more suitable and by whom?

3. **Empirical evidence regarding credit scoring in the UK and its usage in housing decisions is limited in scope and detail.**

Although there is an extensive literature on credit scoring and its off-label usages (Rona-Tas, 2017), the main focus is on the USA, a country with specific institutional, social, political and economic circumstances. The few existing comparative approaches (Lazarus, 2012; Hohnen, Ulfstjerne and Krabbe, 2021) reveal significant differences between European and North-American contexts that could better illuminate the dual character of algorithmic governance. A much wider engagement with specific empirical cases is needed to address such restrictions in perspective.

Further research questions: How is credit scoring used in housing decisions in the UK and with what effects? What UK country-specific aspects shape the usage and data and algorithm-driven technologies in housing decisions? How does this differ or not from the USA and how such usage is an expression of a global assemblage of actors and technologies?

4. **There is a very limited knowledge of the publics' experiences and understanding of the implications of the use of data and algorithms.**

Currently, further studies that deploy a systematic empirical grounding are needed to understand both the implications of the application of algorithms and how these are understood and perceived by the wider public. A greater sense of these algorithmic subjectivities would be beneficial to expanding knowledge of the housing sector specifically and algorithmic processes more broadly. Public surveys on attitudes fall short in explaining how those attitudes are formed, how they influence behaviour and how they are related to specific understandings.

Further research questions: How do data subjects experience their usage of data in housing decisions? How do they understand and relate to these technologies' opacities, bias and discrimination, privacy and consent issues, and data errors?

***

Approaching data and algorithms from a social science perspective

In broad terms, scholarly studies of algorithms and data might be split between seeing these technologies as embodiments of broader logics requiring a theorization of themes such as rationalization, quantification, discipline and governance (Langley, 2014; Marron, 2009; Langley, 2007) while other studies might be more attuned to the pragmatics and the continuous work of having to create and maintain networks, adopting an Science and Technology Studies (STS) approach focused on the technicality of such tools (Poon, 2008; Seaver, 2017; Hohnen, Ulfstjerne and Krabbe, 2021; Campbell-Verduyn, Goguen and Porter, 2017). The latter have examined especially how the introduction of new calculative tools, the digitalisation of databases and the Internet have changed the conceptions of risk and its management (Poon, 2008; Lauer, 2017; Besedovsky, 2018). Still, it is important not to fall into the traps of technological determinism. As Lauer (2017) argues in the case of the computerisation of credit scoring and reporting, their adoption was not inevitable while the decisions behind it were entirely human. At the same time, power and politics play an important role in shaping their adoption and impact, making some authors argue for the marrying of STS approaches with political economy ones (Fields, 2017).

Given that "authority is increasingly expressed algorithmically" (Pasquale, 2015, p.9), many critical scholars have focused on how numbers, data, and algorithms are shaping various forms of *governance*, understood in the broad Foucauldian sense as the "conduct of conduct" in which the governors are not only state agencies but a multitude of actors, including the self. But rather than relying extensively on Foucault's work on disciplinary societies, Deleuze's related concept, *societies of control* has been proposed as a more apt one for describing contemporary forms of governance and surveillance

(Langley, 2007, 2014). Building on Deleuze's conceptualisation of the control society, Adkins (2018) points to the fact that the role of discipline in ordering a society has been replaced by "processes of disassembly and reassembly" in which the "productive capacities of populations" are enrolled into "movements of flows of money" aimed at making profits for the financial system.

Various studies have drawn attention to how algorithmic governance simultaneously involves instrumental usage and more automated and independent deployment (Hohnen, Ulfstjerne and Krabbe, 2021; Campbell-Verduyn, Goguen and Porter, 2017). Similarly, some scholars speak about the *fetishization of algorithms* either to draw attention to attributing too much agency at the expense of the social relations (Monahan, 2018) or to recognise the important mediation work that they do (Thomas, Nafus and Sherman, 2018). This distinction points to the fact that there might be many variations in the role and effects of algorithms depending on their specific social locations and the networks they are embedded in. Hence, since algorithmic tools have an important role in shaping people's life opportunities (Fourcade and Healy, 2013; Fourcade, 2021a; Citron and Pasquale, 2014), it is vital to understand their workings without over or underestimating their impact or misinterpreting their agency.

Amoore (2020, p.24) argues for a need for more specificity to avoid an "algorithm talk" that conflates the various ways in which algorithms operate by making an analogy with the fact that they are "as varied in their logics and grammars as languages are". Scrutinising the proliferation of "algorithm talk" in public discussions and academic studies from a computer science perspective, Dourish (2016) adopts Niklaus Wirth's formula "Algorithms + Data Structures = Programs" to draw attention to an essentialising trap when an analysis is focused only on algorithms. He notes the computer science usage of the term as referring to "an abstract, formalized description of a computational procedure" (3) and the consequent existence of various types of algorithms (e.g., combinatorial, numerical, probabilistic). Compared with the technical specificity of the term, its usage in social sciences is rather metonymical, referring more broadly to the entire regime of digital automation. Dourish argues that it is important to understand how algorithms operate as part of and not as if they are its entire framework. He draws attention to the important technical and analytical distinctions between algorithm and code, architecture, and their technological materialisation, and he argues that the relationship between algorithms as formalised accounts and as practical tools is not a fixed one. Code platforms, software architecture, storage speeds, network capacities and other technical features matter in how an algorithm operates. Consequently, Dourish invites studies of algorithms to consider more the data and its modes of organisation to make it processable for algorithms, how algorithms keep their identity across different settings, and their temporalities, especially in relation to data streams.

Still, it is important to note that "algorithm" is a term that travels in various arenas and receives different meanings. Burrell (2016) points to how the term stands for a broader polarised debate around who is more prone to errors and who controls a decision in human-technology interactions. Such an interpretative overflow, even if it rests on a synecdoche when the more technical meaning of algorithms is taken into account, suggests that algorithms may be better seen as "manifold consequences of a variety of human practices" rather than "stable objects interacted with from many perspectives" (Seaver, 2017, p.4). Arguing for a sociological understanding of algorithms, Gillespie (2014, p.162) notes that they should not be seen as "abstract, technical achievements" but rather unpacked by examining "the warm human and institutional choices that lie behind these cold mechanisms". This means that studying algorithms should not only aim to unveil their internal workings, but rather examine the political assumptions constitutive of their design, use and legitimacy and the continuous work of changing them concerning their usage.

Machine learning or Big Data algorithms in contrast with rule-based algorithms analyses of data remain unknown in terms of the domains that the data represents even if the formal properties of the data used are known (Dourish, 2016). Some scholars even make a distinction between "statistical models" or "data modelling culture" and "algorithms" or "algorithmic modelling culture", the latter referring to more static and predetermined frameworks to know the world, while the other to more dynamic and less human-controlled ones (Hayles, 1999; Breiman, 2001). The newness in the debate around AI/ML algorithms in contrast with older forms of making up people, or creating certain categories of person, through statistical forms is usually related to the new techniques and larger volumes of data collection that raise concerns around privacy and opting out possibilities; the extension of their application in various domains that have important consequences for life chances; and their opaque workings even for their designers which raises issues of explainability and accountability (Burrell, 2016; Amoore, 2020; Fourcade, 2021a; Langley, 2014).

Data algorithms have also the potential to fully transform the domains in which they operate and consequently their practices and constitutive principles. In the case of the insurance industry, Cevolini and Esposito (2020) argue that machine learning and Big Data and their promise to produce individualised predictions endanger the principle of risk-pooling and mutualisation of risks, and reverse the information asymmetry between insurers and policyholders. It leads to a price personalisation that it is unclear if it is genuinely fairer or will produce new forms of discrimination and negatively affect life chances. But such transformations can also be attributed to "less complex" tools. For example, in the consumer credit industry, credit scoring, based on more classical statistical tools, has similarly reshaped conceptions of risks. It generated a shift from understanding risk as a cost to its assessment as a profitable opportunity that allowed the creation of new financial innovations such as mortgage-backed securities and risk pricing (Poon, 2008; Marron, 2009; Fourcade and Healy, 2013; Adkins, 2018). Some authors argue that credit scoring is one of the most important calculative devices of financialization (Chiapello, 2015; Besedovsky, 2018). More generally, it is the adoption of new and quite computationally unsophisticated mathematical models that has transformed the financial industry in the last five decades rather than the most advanced AI technologies (MacKenzie, 2006). Currently, the UK financial services have started to use AI more like a continuation of existing rule-based models, by augmenting some existing analytical techniques and still leaving a place for human decision making in most cases (The Bank of England and FCA, 2022, p.8). The domains in which ML applications are most used are risk management and compliance; customer engagement; credit; securities sales and trading and general insurance while the greatest perceived benefits are anti-money-laundering, fraud detection and overall efficiency gains (with the associated cost savings) (BoE and FCA, 2019). At the same time, data availability, quality and ethical usage, assuring sufficient explainability, a risk-averse culture, and difficulties in gauging consumer and wider public acceptance seem to be the most important obstacles to innovation from the sector's perspective (CDEI, 2020).

Drawing attention to the newness of these tools, even if in a critical way, mirrors at times the vocabulary around "disruption" in the tech industry. However, historical accounts of automatisation of credit scoring show that debates around their fairness, accountability, explainability and efficiency have been in place since their introduction (Lauer, 2017; Marron, 2009; Krippner, 2017). Hence, it is important to note that they are part of a long tradition of "making up people", i.e. creating categories of persons and associated spaces of action and possibilities, through the usage of statistical thinking (Hacking, 1990). Amoore (2020) calls a similar phenomenon "regimes of recognition" through which various people are sorted, classified and rated in order to facilitate various registres of action. Hence, this raises the question of who is accountable for doing the work of attribution.

Who attributes

Examining who attributes in algorithmic regimes of attribution raises two important questions related to who does the *work* of attributing and what type of work is at play, and who is *accountable* for its effects. Drawing on Beck's work on risk societies, Curran (2018) draws attention to how digital economies reflect a specific configuration of capitalism, state and science in which it is unclear who is responsible for the outcomes of technological innovation that usually escapes substantive democratic evaluation.

Compared to industrial capitalism, digital capitalism conceals labour through the fetishization of AI and automation rather than commodities (Burrell and Fourcade, 2021). As Fields and Rogers (2021) argue in their examination of real estate platform, different forms of labour might be at play in their functioning: unwaged digital labour of users of digital platforms, which might include even inputting personal information that users are willing to share given that it is transformed into commodified data; digitally mediated waged labour associated with the widespread usage of the internet, and; non-human-labour of technological elements and the illusion that it is only it producing the real estate data. But, despite such an distribution of labour, quite a few critical studies of algorithms tend to use a language that invests them with power and agency, being usually the subject of most of their sentences as, for example, in Amoore's "what matters to the algorithm, and what the algorithm makes matter as an actionable output from a set of attributes" (Amoore, 2020, p.4). Such an omission of concrete people and organisations that make and use them might be justified by an understanding of algorithms as having no outside, "an accountable human subject who is the locus of responsibility" (page 5). Amoore, following Foucault, argues that algorithms are implicated in new regimes of verification, that they establish "new patterns of good and bad, new thresholds of normality and abnormality, against which actions are calibrated" (page 6). If the argument that algorithms have a specific way of "seeing" the world by foregrounding some attributes while others might be omitted can hardly be refuted, it is also the case that such a language obfuscates that the agency imbued in algorithms is rather an effect of various socio-technical assemblage. Algorithm providers might as well use automation as a rhetoric to deflate responsibility (Gillespie, 2014). Algorithmic governance involves not only governance *by* algorithms but also *of*, raising questions about how some specific perspectives are enacted through algorithms and *who*, more broadly, does the work, given that such algorithms are designed, deployed, and operated by various people. Consequently, given that they usually bring together different social worlds, their circulation and usage might receive different understandings and responses. Gillespie (2014, p.171) discusses how information must be made *algorithm ready*, arguing that "categorization remains vitally important to database design and management".

Despite the celebration of automation as error and subjectivity free, it is most often the case that it is rarely hands-off and usually adapted to specific social contexts, such as web page ranking algorithms that newer show web pages with a dissident political speech in China or those that promote Nazism in France (Gillespie, 2014). Hence, rather than seeing the advent of digital tools as following an evolutionary path towards more efficiency or as omnipotent agencies, a focus on the different types of work taking place at different levels might be more illuminating in understanding their work. Similarly to information classification systems discussed by Bowker and Star (2000), algorithms have to "satisfy the informational requirements" of various communities creating and using them. They might maintain a sort of constant identity across borders and at the same time they might be being flexible enough to be adopted for more specific usages, acting in this way as *boundary objects*.

At the same time, the adoption of new technologies requires various works of *enrolment* (Callon and Law, 1982). As Marron (2009, p.138)) notes, in the case of credit scoring, their engineers had to convince "Luddite" lender managers to adopt them, and she argues that "systems of risk (…) do not simply disperse through the to their own interminable logic as more rational, more efficient means of governing consumers". Similarly, Lauer (Lauer, 2017, p.231) contends that credit scoring and computers were rather seen in the industry as solutions for the problems of large organisations that did not have enough personnel to access applications and less as a solution for a deficient human risk assessment. Professional judgement was highly respected, while "models" could be seen as suggesting a "lack of reality" (Lauer, 2017, p.189). In addition, banks were highly invested in assuring their clients' privacy, declining initially or reporting limited information to credit bureaus (Lauer, 2017, p.189). The point is rather to fully understand the social worlds in which algorithms circulate and the heterogeneous actants that give algorithms efficiency and persuasiveness. This is especially important since such tools are not only an effect of technoscientific rationality but rather are used in business domains "shaped by the aim of incessantly increasing exchange value as a means of maximizing profit" (Curran, 2018, p.213).

Classification work

The uses of data and algorithms in risk assessments involve various forms of classification work. Risk assessment technologies can be seen as tools to tame future uncertainties by calculating the probability of some event. Risk means measurable uncertainty, i.e., calculations that rely on categorising objects and events into standard categories (Knight, 1921). When applied to humans, risk assessments act as classificatory tools, ranking and sorting people along with various levels of risk. If classification systems are inherent to social life, the growing usage of enumerations and quantified forms of classification bring into focus specific concerns, especially in the contemporary context when ratings and rankings have become ubiquitous (Guyer, 2010; Esposito and Stark, 2019) in what Citron and Pasquale (2014) call the "scoring society". Financial capitalism has been approached as a "new economy of moral judgement" in which algorithmic forms of classifications increasingly use behavioural data to sort and rank consumers/citizens, seemingly rewarding individual acts and choices (Fourcade and Healy, 2017; Fourcade, 2021a).

Interested primarily in the effects of the broad adoption of algorithms for decision making in various domains, Amoore (2020) proposes the concept of *regimes of recognition* in order to approach various forms of classification at play in algorithmic governance. These regimes describe the assemblages of humans, algorithms, and datasets that identify what and who matters by not simply using some predefined attributes but creating in the process the conditions for their recognition. Amoore draws attention to the difference between two types of algorithms, rule-based classifiers that class and recognise objects based on human pre-decided rules of classification and those like machine learning in which recognition is made through input from data clusters[1]. Both operate with establishing weightings, probabilities, and thresholds of recognition, but the human/technological contribution is different. Recognition through input from data clusters is usually "too complex even for the designer of the algorithm to explain the conditional probabilities that are learned" (Amoore, 2020, p.74) and aims toward making algorithms able to adjust themselves through the emergent properties of the data rather than humans adjusting thresholds. In this case, what counts as a design error when such technology fails is hard to establish, given that this is usually due to a misrecognition of input features that they were not trained to. For Amoore, the focus is on regimes of recognition created by machine learning since given their inherent social bias[2] and assumptions, the questioning of the human-technology distinction, and the incomplete knowledge of mechanisms generating a specific output, they raise important political and ethical questions around accountability and responsibility, what she names *cloud ethics*.

However, it is important to note that such questions may be raised on the most fundamental layer of such regimes, the constitution of what counts as data, and their basic processes of measuring, classifying, and aggregating. As Kitchin (2014) argues, data are the building blocks of information and knowledge that are far from being neutral and objective but are framed "technically, economically, ethically, temporally, spatially and philosophically". Such framings are produced in specific social contexts, the regimes' of recognition complexity being one that should not be reduced to their technology, i.e., enabling *recognition via classification rules* or *recognition via input data clusters* (Amoore, 2020). Data and algorithms are corporate products, meaning that they are enacted through specific practices. Gillespie (Lauer, 2017, p.182) notes that an inquiry into algorithms should treat analytically distinct algorithms and databases and start from how data is collected and readied for analysis and by observing what is included and excluded. This might involve moving the focus from the technological workings of an algorithm to its social embeddedness, an approach that we take in the following section

[1] There are different ways of classifying algorithms based on different criteria. Amoore, nonetheless, uses such a binary distinction in order to draw attention to the issue of explainability, easy to trace in rule-based algorithms, less so in certain types of ML algorithms.

[2] Bias in relation to algorithms can have different meanings. Sometimes it can refer to the technical statistical meaning in which it might be reasonable to choose a method that increases bias in order to reduce variance and consequently error. However, in this review we refer to what can be called the moral meaning of bias through which there is a systematic and unfair discrimination of certain individuals or groups in contrast to others. This might be due to a lack of representativity in the data used or the fact that certain variables (e.g., income) might be proxies for societal inequalities that can be mapped along gender, class, ethnicity, race, age.

in which we examine how a regime of recognition around credit scoring has been shaped historically.

***

Credit scoring's regime of recognition

Adoption and development

Computerised credit scoring and reporting have become in the USA a key element of the information economy starting with the 1960s. It pushed the "commodification of personal identity to new limits" through speeding up, centralizing and changing the structure of credit reports from a more narrative and subjective one to a more "cryptic summarization and quantification" (Lauer, 2017, p.182). Credit scoring has been predominantly a statistical tool, involving scorecards in which various variables are given different weights and analysed through discriminatory, regression, multivariate regression, stepwise analysis, or factor analysis (Lauer, 2017, p.182). The combinations of statistics and computers, made correlations between various variables the most important measurement of creditworthiness which has been reduced to "a function of abstract statistical risk" (Lauer, 2017, p.206). Stability became at the time the proxy of character through the newly discovered importance of variables such as homeownership, banking relations, marriage, long term employment to predict a "good" borrower (Lauer, 2017, p.206). Burton calls this a "shift from personal trust to institutional trust" (2017). Lauer (Lauer, 2017) argues that this is a process in which the democratisation of credit seemed to require more constant and centralised surveillance.

During the 1960s, a credit scoring technique was developed by the Fair Isaac Corporation in the USA with the aim of calculating the probability of default, used initially on banks and retailers' customer information and from the 1980s by using credit bureau information (Lauer, 2017). Although various credit scores were developed at the same time, the FICO score version launched in 1989 soon became a market device that managed to coordinate various heterogeneous actors in reducing creditworthiness to a three-digit score mapped along with an interval (300-850) in which thresholds are set for establishing who is creditworthy or at what price (Ritzer, 2001; Poon, 2008). Credit scoring was seen as a response to anti-discriminatory laws, being promoted as a tool that through its focus on credit history went beyond decision making based on protected characteristics (Poon, 2008; Rona-Tas, 2017; Fourcade and Healy, 2017).

The FICO algorithm has been kept secret, but it is known that data such as payment history, credit applications, negative events such as default or bankruptcy, and income play an important role in its calculation by variables allocated different weights (Arya, Eckel and Wichman, 2013). Consumer claims for transparency have pushed Fair Isaac Corporation to make more transparent what is determinant in calculating a consumer score, which has become a commodity to be bought not only by lenders but also by consumers (Marron, 2009). Still, Citron and Pasquale (2014) note that the credit scoring system in the USA is not fully transparent, creating a lot of uncertainties around one's standing while the differences in assessments between the credit bureaus reveal their arbitrary nature. Despite regulations that aim at making credit scores more transparent to consumers, various concessions were still in place, such as immunity from defamation law or reducing the request of disclosure to only four factors that make little light on how a specific score is calculated (Citron and Pasquale, 2014). More recent data privacy laws in the USA and Europe seem nonetheless to give more control to consumers on what data is collected about them and how it is used (Schmeckpeper *et al.*, 2021). At the same time, a centralised credit scoring system might contribute to preventing fraud given its amassing of data from different lenders that make it easier to check a borrowers' debt and credit history (Lauer, 2017).

Credit scoring as a market device in the USA has been developed along with credit bureaus or credit rating agencies who were in the business of collecting information and reselling it. Emerged during the 19th century in order to offer information about the creditworthiness of businesses, such enterprises moved from local coverage to a national one (2017) and from trade to consumers. The computerization of databases during the 1970s and the adoption of computers in business led to a concentration of the credit information market. Most of the local bureaus merged or were replaced by what has become in the last two decades the three most important players in the industry, Experian, Equifax, and TransUnion (Lauer, 2017, p.253), present now also in the UK where they are named credit reference agencies (CRA). CRAs centralise credit data sent by various lenders about consumer loans and credit cards that are used by lenders along with their in-house scoring systems in underwriting mortgages and personal loans. Lauer (2017) notes that their business does not stop at registering defaults and credit history but rather at amassing all sorts of consumer data for risk profiling, consumer segmentation, and direct marketing being in fact *consumer data brokers,* Equifax being the first CRA which at the end of the 1980s created its own marketing division. Along with calculations of credit risk useful for lenders, the CRAs have produced scoring models for gas and electric companies, wireless telephone carriers, and healthcare providers and their credit scoring has been used for employment, insurance, and various forms of lending beyond personal loans. Lauer argues that such a proliferation of statistical thinking based mainly on correlations reduce an individual's lifeworld to "behavioural "facts" devoid of theory or explanation" (Lauer, 2017, p.253).

The very few accounts of credit scoring in the UK point to its introduction during the 1980s. As in the USA, credit scoring had been embraced as a tool able to segment the market of consumers and develop risk pricing and it has also become an important part of mortgage underwriting (Wainwright, 2011). Here the aims seems to be less about fairer access to credit but rather the reduction of informational asymmetries between borrowers and lenders in a context of intense market competition generating chasing for larger market shares (Leyshon and Thrift, 1999). Hence, Leyshon and Thrift (1999) see the growth in the usage of automated computer-based systems of assessment in British retail banking as providing a more cost-effective solution to an old problem by processing larger volumes of information, deskilling the workforce, offering a timely response to economic and social changes, and a more standardized tool that facilitated inter-organizational sharing of knowledge. Nonetheless, the Equal Opportunities Commission supported credit scoring tools as a more "objective" way of granting credit given that this was usually vetted by white middle-class men (Leyshon and Thrift, 1999, p.448). Currently, there is a growing diversification of the data used in credit risk assessment. Deville (2020) notes that the classical three core sources of public information (electoral roll, CCJ data, insolvency information, bankruptcy/IVA data) are being complemented by sources of data such as behavioural scores related to a consumer's credit usage patterns, current account credit turnover, internet and device data, rent, phone and utility payments.

The literature on credit scoring is mostly focused on the Anglo-Saxon countries but as Hohnen et al. (2021) contend in order to fully understand algorithmic governance, inductive comparisons between different contexts might better illuminate its dimensions. The authors compare credit scoring usage in the USA and Denmark and point out that in

Denmark there is a recent growing usage of automated scoring by Danish banks, but these are in-house risk assessment technologies that must obey EU regulations on data protection and credit evaluation and cannot be sold to third parties. Consumers have no access to their scores although they have the right to request explanations for a credit decision and credit scoring. Despite banks' forms of surveillance and profiling of their customers, credit scoring remains rather unnoticed in the public space. The authors propose the concept of *silo-surveillance* to characterise algorithmic governance through credit scoring in the case of Denmark, pointing to how although GDPR rules limit the dissemination of personal information, they do "*not* prevent intensive surveillance and 'credit scoring by approximation'" (Lauer, 2017, p.267). In the case of France, Lazarus (2012) notes a similar usage of credit scores as in-house tools for banks that distinguishes between those accepted and those declined credit while the French data protection agency restricts their sharing between banks and credit institutions. Nonetheless, it seems that irrespective of the regulations regarding privacy in various countries or the financial sector, the way in which the consumer data markets are organised and technically equipped, marks the dominance of "inscrutable surveillance of networks and algorithms" (Lauer, 2017, p.267).

Currently, the credit scoring industry in most of the countries is moving toward new data frontiers that can be used to assess creditworthiness given the Catch-22 dilemma created by the need to have a credit history that excludes people who might be creditworthy (Hohnen, Ulfstjerne and Krabbe, 2021). Since the new mantra of the financial industry is "financial inclusion" and given the rise of Big Data and the consumers' digital footprints, various financial actors have started to explore what other types of data and what other types of attribution tools can be used in order to automate the assessment of creditworthiness and include the "thin file" borrowers (Deville, 2020; Deville and Van der Velden, 2015; Berg et al., 2019; Dunkerley et al., 2021). In 2016, FICO launched in collaboration with Equifax FICO Score XD that includes what are called "alternative data sources" such as mobile phone payments, public records, and property data (Henry and Morris, 2018). The three major credit reference agencies have been collaboratively developing their own score, the VantageScore, which takes into account factors and weights slightly different from the FICO score (DuFault and Schouten, 2020) and score more people by using more of the data from the consumers' credit reports (Henry and Morris, 2018). At the same time, the CRAs act as credit brokers trying to seduce consumers to take loans, pointing to a concern raised by Zarsky (2016) around the marketing purposes of the analysis of personal information for credit scoring (see also Lauer, 2017; Marron, 2009; Adkins, 2018; Fourcade and Healy, 2013).

In addition to the traditional players, fintech start-ups have begun to enter the market, promoting the slogan that "all data is credit data" (Hohnen, Ulfstjerne and Krabbe, 2021; Deville and Van der Velden, 2015). In this way, credit scoring and the personal information it encapsulates have become in themselves commodities circulating between different parties shaping consumers' access to capital goods such as employment and housing (Hohnen, Ulfstjerne and Krabbe, 2021; Langley, 2014). New types of methodologies are tested, the statistical models at the base of credit scoring being problematised and in some cases challenged by new tools such as decision trees and neural network systems. Aggarwal (2021) notes that starting with the mid-2000s in the UK, algorithmic credit scoring has been increasingly adopted especially by fin tech credit lenders such as Zoopla, involving the usage of larger volumes and types of data and more complex ML/DL tools of analysis such as random forests and artificial neural networks.

Although in practice the logistic regression model is dominant given the regulatory requirement to have transparent and auditable models, the literature on the use of machine learning for credit scoring is proliferating and proposes different techniques to rationalise the predictions, i.e., give justifications for decision making (Dastile, Celik and Potsane, 2020; Bücker et al., 2022; Onay and Öztürk, 2018). But as Marron (2009, p.127) argues such methodologies work with the same conception of risk as an objectified quality of the individual looking for the likelihood of default and not its causes, while the occurrence of default is still calculated at the level of groups, the individual level remaining inescapably uncertain. Maroon notes that given that each tool comes with its own technological dilemmas, the only unified position is their contrast with the traditional judgmental decision process, marking their superiority through a discourse of efficiency. Marron contends that logistical regression was considered much more difficult to implement in practice even if it might be more predictive than discriminative analysis due to limited datasets. At the same time, neural networks which are good in working with a smaller number of cases fall short in front of the statutory requirement to offer an intuitive explanation to the consumer as to why they were classed in one way or another.

The social impact of credit scoring

The adoption of new risk technologies involves reshaping market actors' commercial practices (Poon, 2008, 2016), regulatory responses (McFall and Moor, 2018; Weiss, 2016) and consumers reactions, generating various political struggles and debates (Weiss, 2016; Krippner, 2017; Marron, 2009). The capacity of credit scoring to discriminate between different classes of people and assign individual risk has meant a reinforcement of structural inequalities, despite the advocates' claim of technological neutrality, making their capacity to include or exclude people and impact on life chances one of the most discussed aspects (Fourcade and Healy, 2013, 2017; Schmeckpeper et al., 2021; Langley, 2014, 2009; Marron, 2009). Importantly, credit scoring is not only a tool to decide who gets and who gets not credit but rather one through which different categories of consumers are priced and sold different financial products, leading to risk-based pricing or what Poon (2008) calls a shift from *risk minimising* to *risk management* strategies in lenders' models. Some scholars have argued that credit scoring are a continuation of older classificatory practices in the consumer credit history (Krippner, 2017; Hohnen, Ulfstjerne and Krabbe, 2021) while others point to how such technologies are "an emergent and controlling form of power" marking a shift from disciplinary to control societies (Langley, 2014; Adkins, 2018). As Fourcade and Healy (2017, p.23) put it: "the old classifier was outside, looking in. …The new classifier is inside, looking around".

Reshaping social positions

Firstly, citizenship is digitalised. This continues a long tradition in which data practices constitute populations as governable entities and a distinct form of peoplehood (Cakici, Ruppert and Scheel, 2020). But it is not only the state who controls this process. Pasquale (2015, p.9) notes that the actors deciding the distribution of a society's benefits and burdens are rather part of a network in which "the distinction between state and market is fading". Similarly, Fourcade (2021a) proposes the concept of *ordinal citizenship* to capture how in contemporary societies the figure of the citizen is morphed with that of the consumer, through public and private actuarial and quantitative tactics that redefine membership in a society and the individualization of biopolitical enrolment. Reshaping citizenship takes place not only through algorithmic tools to provide public services (Dencik et al., 2019; Eubanks, 2018) but also through the broader dynamics at play in movements such as digital and financial inclusion reconfigure social rights and obligations. Krippner (2017) examines how credit risk technologies can reshape *economic citizenship*, the capacity to claim control over economic resources, showing how different consumers movements in the USA have tried to counteract the gender or race discrimination embodied by such tools, moving from redlining to greenlining.

Fourcade (2021a, p.158) argues that the financial and tech industries work on an idea of citizenship "that bypasses the state, and even transcends national boundaries", "financial inclusion" and "digital incorporation" promising equal status and the freedom to craft oneself to anyone. Importantly, in different settings and economic circumstances, financial inclusion can take different forms, promoting not only the use of new data for assessing individual creditworthiness but also using "social collateral", i.e., capitalizing on an individual's social capital through guarantors (Muriel, 2021; Lofranco, 2021). Gabor and Brooks (2017), examining development interventions' ethos of promoting digital-based financial inclusion, point to how poor populations are connected to algorithms through the digital footprints of their mobile phone usage or even by being invited to play online games generating behavioural data. The authors argue that such interventions "obscure the desire and momentum of financial capital to access high-risk/high return markets" (Gabor and Brooks, 2017, p.433) while the monitoring of digital footprints and the importance attributed to "nudging techniques" in financial education programmes aimed at making the poor more capable financial subjects reveal the disciplinary traits of financialization.

Accominotti (2021) argues that citizens are not only subjects but also observers of the ratings and rankings generated by algorithmic tools of classification and moves the focus from their social usage as an index of moral character to their consideration as *aesthetic objects*. Through this lens, classifications such as credit scores appear neat and definitive despite the messy sources of information feeding them, creating an artificial orderliness with rankings that can easily pass as legitimate given that they are rooted in an individual's doing. Such single, crisp and clean-cut figures make observers less inclined to question them, Accominotti (Fazi, 2021; Amoore, 2020; Burrell, 2016) suggesting that "they cultivate a hierarchical gaze in the eyes of those experiencing the world through the sorting that they create" that undercuts egalitarian understandings of society. The persuasiveness of such an aesthetic can be encountered also in how what is technologically easy to measure - steps taken, number of Facebook friends, money spent - become "superficial stand-ins for much more complex concepts, such as overall health and wellbeing, the strength of social relationships, and the value of an employee" (Whitson, 2014, p.352). Trying to account for the fact that credit ratings have been used in the USA starting with the 19 century although without being reliable assessments of risks, Carruthers (2013) makes a similar argument pointing to how their categorical and quantitative information was easy to incorporate into rule-based decision processes, favoured by the business sector and regulators. Such techniques enacted *procedural rationality*, or what Porter (1995) calls "mechanical objectivity" allowing a broad coordination between heterogeneous actors and being a good enough tool in the face of uncertainty.

Secondly, given that risk assessment tools in housing decisions mediate access to an important good, they not only create subjects to be governed as citizens but also delineate between different class positions. The growing importance of credit in most developed societies has drawn attention to how economic extraction from productive labour might be supplemented by debt relations, pointing to a need to reconceptualise class beyond relations to the means of production (Graeber, 2014). Fourcade and Healy (2013) argue that markets' "classification situations" intensify existing class inequalities given how credit scoring leads to differential access and pricing of credit. At the same time, scholars have drawn attention to how more generally data-driven algorithmic governance challenges equal liberal citizenship given their unequal distributional effects (O'Neil, 2016; Eubanks, 2018; Fourcade and Healy, 2013, 2017). In the "algorithm society", the ownership of digital means of production has created the coding elite with extensive cultural, political and economic claims (Dencik et al., 2019; Deville, 2020; Leyshon and Thrift, 1999; Rona-Tas and Guseva, 2018; Zarsky, 2016). Moreover, risks are unevenly distributed, some being able to benefit from it, others being more protected, while others might be even more disadvantaged (Curran, 2015). In a context in which risks are increasingly individualised, how one responds might be judged through a moral lens specific to class distinctions such as in judgements of debt through moral pronouncements about profligacy and irresponsibility delineating between middle class and underclass (Davey, 2020).

Another related strand of research points to how class positions are being reshaped in the context of financialization by looking at wealth accumulation strategies in relation to housing. Forrest and Hirayama (2018), comparing the UK and Japan, both countries in which the expansion of homeownership has been stalled in the last decade, argue that housing wealth and intergenerational transfers have become highly important in determining life chances and consumption patterns given that housing price inflation and stagnating wages make access to the housing market hard to reach by increasing levels of populations. Similarly, Lisa Adkins et al. (2021, 2020) argue that housing wealth and tenure reconfigure the social space even more than productive labour. The authors propose an asset-based class scheme distinguishing between investors, outright homeowners, homeowners with a mortgage, renters and the homeless as a more adequate one for describing social inequalities in conditions of financial capitalism. Hence, in this context, it becomes even more important to understand the socio-technical assemblages shaping access to housing especially since such economic divisions usually overlap with racial and ethnic ones.

Thirdly, with the advent of digitalisation and datafication the understanding of identity in terms of classical sociological categories like gender, race or class is reshaped giving birth to new forms of identity such as "algorithmic identity" (Cheney-Lippold, 2011), "the digital person" (Solove, 2004), or "data doubles" (Nichols, 2004; Lupton, 2019). Such identities are built from what is most legible to an algorithm from the digital traces of a person, emphasising some attributes and excluding others. It is not anymore labour power and the physical body that makes humans valuable to capital but rather "the millions of bits of information about who they are, deep down, and how they relate to one another" (Burrell and Fourcade, 2021, p.220). Built on tools aimed at sorting people based on the minimum necessary information, digital identities might not be only gross oversimplifications of complex persons but also powerful categories for self-categorization and important public metrics (Gillespie, 2014). In 2018, Experian UK released an advert in which viewers were explicitly invited to 'Get to know your data self, with Experian', showing how a person was followed all the time by his "data double", a character that comprised from his financial history and that fully mirrored him. Douglas-Jones (2021, p.162) draws attention to how the imagery and rhetoric around attributing agentic personhood to our "data doubles" in popular culture reflect the "relocation of the metaphor of self-knowledge outwards" in which an external unknown archive is to be known rather than the interior depths under the surface of one's body.

Data ethics and challenges

Such a proliferation of actors and risk assessment tools have also given rise to various debates around their ethics and politics, raising concerns around their opacity (Pasquale, 2015; Citron and Pasquale, 2014; Amoore, 2020; Burrell, 2016; Fazi, 2021; Esposito and Stark, 2019); privacy and consent (Ferreri and Sanyal, 2021; Gabor and Brooks, 2017; Marron, 2009; Lauer, 2017; Whitson, 2014); bias and discrimination (Boeing et al., 2021; Citron and Pasquale, 2014; Decker, 2021; Dencik et al., 2019; Hassani, 2021; Rosen, Garboden and Cossyleon, 2021); and data errors with important effects (Dencik et al., 2019; Deville, 2020; Leyshon and Thrift, 1999; Rona-Tas and Guseva, 2018; Zarsky, 2016).

Opacity

The growing production of digital footprints and shadows (Fazi, 2021; Amoore, 2020; Burrell, 2016) raises various questions regarding who has access to what data, for how long, and how such data is analysed to produce a specific output. Pasquale (2015) uses the concept of "black-boxed society" to draw attention to the continuous scrutiny of consumers/citizens and a prevailing opacity surrounding how data is used by businesses or governments. He contends that asking for more transparency is not a solution since this might add new layers of complexity and in fact hinder an attempt to more intelligibility. Still, he argues for more regulation in order to make algorithms more scrutable even if this might be made by a third party, a "trusted auditor" that might assure privacy and work in the interest of the public. If Pasquale takes a legal approach to black-boxing as secrecy constituted by social structures and mechanisms, other scholars treat opacity as a *technical condition*, especially when the focus is on AI/Machine Learning, and are less inclined to see audits as a solution (Fazi, 2021; Amoore, 2020; Burrell, 2016). For Amoore (2020) this means that transparency is an impossible achievement while the boundaries between human and technology agency are hard to draw especially in the case of AI/Ml systems.

As Citron and Pasquale (2014) note in the case of credit scoring, opacity is maintained without necessarily this being an effect of technical complexity consumers being unable to understand what constitutes an optimal credit behaviour. Even if CRAs announce the relative weight of some of the variables they are vague and non-indicative about what one should do. As Gillespie (2014) argues, knowledge around algorithms' working is unevenly distributed, making backstage access a form of power. Usually protected as a trade secret, some stakeholders nonetheless might have more knowledge, such as commercial actors or third party developers that use APIs in contrast with their users. Nonetheless, arguments for more transparency have been made built on the assumption that more transparency can improve efficiency and accuracy. But, Zarsky (2016) notes that at least in the case of credit scoring it is rarely the case that consumers review and correct their credit files while if consumers learn what is considered an indicator of credit risk might try to "game the system" avoiding some behaviours but without necessarily acting in a less risky way. Still, it is hard to see what "gaming the system" might mean in this context, since credit scoring data points are built on registering specific actions. Hence, if someone takes a credit card and performs specific usages only because they think that this will improve their credit score, this could hardly be called "gaming the system" but rather a form of powerful disciplinary force.

On the other hand, the explainability of ML models is an important aim in their development in the financial sector given regulatory requirements to explain decision making (BoE and FCA, 2019). Hence, in the case of "black-box" ML deployment, firms use a variety of model validation techniques in order to better understand how the model makes its predictions (e.g. black-box testing, explainability tools) and implement a series of safeguards to address risks (e.g. alert systems, so-called 'human-in-the-loop'). In the case of credit scoring where explainability is a priority, the main technique used is logistic regression with some ML elements, even if some firms might use "more opaque 'challenger' models alongside these" (CDEI, 2020).

Bias and discrimination

Many scholars point out that discrimination is constitutive for algorithms and statistical models at least in the sense of creating a meaningful difference between various items (Amoore, 2020; Lauer, 2017) but also by institutionalising social values through code (O'Neil, 2016; Rodima-Taylor, 2022). Hence, these differences are far from benign and usually reflect larger societal biases and forms of discrimination. Credit scoring has long been recognised as reflecting "people's prior history of financial privilege or disadvantage" (Foohey and Greene, 2021, p.3) most of its variables being proxies for racial, ethnic, and economic disparities (Hinnant-Bernard and Crull, 2004; Krippner, 2017; Hassani, 2021; Herrine, 2016; Pager and Shepherd, 2008). The ability of credit scoring as a good measure of trustworthiness has been questioned given that it more likely reflects precarious forms of employment and housing conditions and a lack of social safety nets rather than an unwillingness to pay one's debt (Foohey and Greene, 2021;

Marron, 2009; Kirwan, 2021). This is because “credit scoring systems are ill-equipped to discriminate according to situational and contextual factors” (Burton, 2012), a low score due to problems created by unemployment not being differentiated from one due to excessive spending. Consequently, even if credit scoring data does not directly take into account protected characteristics such as gender, race, and ethnicity, it is more likely that people with lower income have lower scores. This reflects in countries such as the USA not only economic but also racial divides, credit scoring and reporting being less a case of discrimination through differential treatment but rather through disparate impact. Their racial discrimination has been recognised at times, some federal state laws in the USA banning, for example, credit scoring usage in the case of employment and insurance (Foohey and Greene, 2021). Beyond proxy variables, the UK financial industry seems to be aware that their databases are skewed towards specific socio-demographics, few training datasets contain minority communities data (CDEI, 2020).

In other sectors, such as lending and housing, credit scoring was rather legitimised as a less biased tool to make decisions (Lauer, 2017; Marron, 2009; Rona-Tas, 2017) some studies pointing that in contrast to decisions taken based on individual discretion, automated underwriting increased the approval rate for minority and low-income clients (Pager and Shepherd, 2008; Bartlett *et al.*, 2019). But if automation might increase acceptance rates for such groups this is also correlated with higher interest rates (Fuster et al., 2022), being able to enhance distributional fairness through larger access but also undermine it through price differentials (Aggarwal, 2021). Issues have been rather raised about the disproportionate role played by credit history in its calculation given that significant parts of the population do not have one, generating in various national contexts, including the UK, debates around what other alternative data could be considered. However, as Foohey and Greene (2021) argue, “financial inclusion” talk distracts attention from the fact that whatever data points are taken into account, the “credit scoring game” is one in which those on the lower side of the socioeconomic ladder are always more likely to fail. And in countries where economic disparities overlap with racial and ethnic discrimination, such tools will perpetuate these inequalities. As argued by Tilly (1998), “durable inequality” is mainly a product of organisational forms which create categories shaping access to valuable resources. Consequently, it remains an open question if such tools should dominate important life decisions even if they might be less biased than decisions based solely on individual discretion. And, as a policy review

(CDEI, 2020) of bias in algorithmic decision making argues, it is more important to establish how such a decision might be overall biased rather than only focusing on the algorithm.

Privacy and consent

Gabor and Brooks (2017) note that in the context of the "digital revolution" rights to privacy might be suspended especially for the most vulnerable citizens. This is not only because digital footprints and shadows might be under continuous surveillance without consumers being aware of it (Gillespie, 2014; Kitchin, 2021) but also because in order to have access to important goods such as credit or housing they might willingly share whatever they are asked (Marron, 2009). As Whitson (2014) points out personal data has become the "price of participation" in various settings in a digitalised society. But such a price is differently paid, Pasquale (2015) notes that there is a blatant asymmetry between how the law protects privacy in commerce versus that of persons.

This is especially worrying in the case of credit scoring given that since the late 1970s, the advent of digital databases has brought centre stage new concerns related to individual privacy given that electronic information made possible various forms of merging and matching data from different sources (Lauer, 2017). Although since then, various regulations have been adopted, the circulation of consumer information between credit bureaus, financial institutions and other commercial actors is still largely unfettered due to various provisions or the fact that they belong to the same "corporate family". At the same time, current regulations might be more inclined to encourage data sharing even if only with consumer consent. EU's Revised Payment Service Directive (PSD2), enforced also in the UK, makes possible the sharing of consumers' account and financial payment data through open Application Programming Interface of banks. If this might be a challenge to the monopolisation of data by specific banks or represent new opportunities for FinTechs (Onay and Öztürk, 2018) still it is unclear if consumers fully understand what is asked and consent out of will and not out of pressure.

Data errors

Issues have been raised also around quality problems of data used for risk assessments. In the USA governmental studies reveal that a significant percentage of people had important errors in their files, some leading to credit denials of larger costs, and it is usually Black and Latino households that are the most affected (Foohey and Greene, 2021). Rona-Tas (2017), discussing various survey studies of credit scoring in the USA aimed at evaluating their accuracy, shows how given that the data is provided by lenders, they have no incentive to use the standard format, while "broken records" can be quite frequent. This means that the same person can have her information spread across different records as if she was two or more persons or pieces of information that belong to different persons are put together as if they are one. She argues that although credit bureaus might use complex algorithms to match information, there are still around five to ten per cent broken records. Unsurprisingly, those that are the most affected by bad data are the young, the poor, minorities or person with lower scores and thin files, given that they have less resources to correct them or are not properly registered in databases. Such errors might have important consequences for consumers who are unable to control them. Tims (2017) and the Responsible Finance report (Pughe, 2018) document how in the UK lenders or credit reference agencies' failures to properly document consumers' information lead to them being refused credit while attempts to correct them are highly bureaucratic and take a lot of time. At the same time, consumers do not receive any default notice being unaware that they might have problems with their credit score, although these might be due to falsely registered events, such as a mobile phone account that they never had, or recorded debts that they were unaware of, since even a £1 debt default can tarnish a record for six years. Still, Zarsky (2016) asks what are the alternatives to automated algorithmic processes in making credit decisions and points out that their errors are less significant than those from manual credit scoring techniques.

Artificial Intelligence Public-Private Forum final report commissioned by the Bank of England and FCA (2022) draws attention to the fact that "AI begins with data". Hence data quality, a full understanding of its provenance, completeness, and representativeness, as well as its ongoing transformations have to be taken into account in AI inclusion in financial services. Such issues are made even more salient in the usage of alternative data which are usually coming from third-party providers and present "additional challenges of quality, provenance, and sometimes, legality" (The Bank of England and FCA, 2022, p.8).

Beyond errors with data, others have questioned if indeed such calculative tools do what they claim they do. In the case of credit-scoring systems, the distribution and representativeness of the data raise important questions about the statistical validity of these models since their databases are highly unbalanced and skewed (Onay and Öztürk, 2018; Wainwright, 2011; Leyshon and Thrift, 1999; Aggarwal, 2021). Given that minority and low-income populations are systematically underrepresented in such databases (Gates, Perry and Zorn, 2002) questions have been raised about their predictive capacities. At the same time, because the rise of algorithmic credit scoring is mostly a post- 2008 financial crisis phenomenon, the training data might reflect a benign macro-economic environment, being less prepared to anticipate borrowers' behaviour in moments of crisis (Aggarwal, 2021). But predictability may be a more general issue rather than one attributed to deficient population samples since statistical tools address only averages and general trends (Mackenzie, 2015; Lazarus, 2012). Given that the uniqueness of each applicant is also recognised in the credit industry this has led to the development of various statistical approaches (Van Dijk and Garga, 2006) and constant monitoring of the predictive ability of scorecards (Wainwright, 2011). It might also mean that approaches like predictive analytics with their promise of producing individualised predictions might soon be broadly adopted.

***

Credit information as a tool of attribution in housing decision making

Credit scoring and mortgage markets

Mortgage markets are considered by economists as "rationed credit markets" in which access to credit is limited based on income, wealth, and credit quality (Acolin, Goodman and Wachter, 2019). Hence, access to mortgage credit might be denied to households with lower incomes, wealth constraints and poor credit scores, even if it might be more optimal for such households to own rather than rent. In most developed countries, access to mortgage credit has been enlarged due to the way in which the financial industry has generated an extension of loanable capital given a broader integration in global financial markets and the development of new statistical tools for risk assessment and mortgage securitization (Aalbers, 2017a; Langley, 2008). The path taken by financialization has taken different shapes in different countries and political economists refer to this as *variegated financialization* (Aalbers, 2017b; Lapavitsas and Powell, 2013). Mortgage-backed securities originating in the USA are, for example, increasingly adopted in European countries such as the UK, Spain, the Netherlands and Italy, but less so in Germany, France and Portugal (Aalbers, 2008) or Eastern European Countries (Pellandini-Simányi and Vargha, 2018; Mikuš and Rodik, 2021).

Such scholarly accounts usually focus on the structural causes and effects of financialization of housing, privileging a macro level of analysis, leaving behind the mezzo level of organisational change or the micro-level of everyday experiences that such a phenomenon might generate. In these analyses, the financialization of housing is usually seen in contrast with previous Fordist regimes of housing in which housing finance was considered too socially important to be set only by markets. Housing in the post-Fordist, neoliberal and financialized regime is seen as an important driver of financial markets coupling households and financial markets risks. Aalbers (2008) argues that the securitisation of mortgage loans, credit scoring and risk pricing have been important tools of financialization leading not only to an increase in the number of homeowners but also in housing prices, amplifying the levels of social inequality. At the same time, Micozzi (2020) shows how in specific settings, credit scoring is an important technology for financial institutions, property developers and corporate landlords to exclude the urban poor and target rather the middle- and upper-income sections of society. In the case of South Africa that he describes, given the important role attributed to income levels, only 20% of the population qualifies for a mortgage, reinforcing existing inequalities along race and class dimensions. Marxist approaches to financialization have long seen it as a phenomenon involving extracting value from households' income streams (Lapavitsas, 2014) although some would argue that industry innovations such as risk pricing and mortgage-backed securities made calculations on current incomes less relevant than their probability in the future (Adkins, 2018). However, the usage of credit scoring for accepting/rejecting an application or as a tool for risk pricing seems to take different shapes in different countries reflecting local regulations and understandings of risks.

Credit scoring and automated underwriting tools' usage on the mortgage market was initiated in the USA during the 1990s by the government-sponsored enterprise (GSE) Freddie Mac and Fannie Mae aiming to make the mortgage underwriting process less affected by brokers' manipulations, to be able to better monitor credit standards and create a common denominator for describing mortgage-backed securities (MBS) (Poon, 2008). If initially AI solutions were tested they raised serious problems of accuracy given a lack of sufficient data on mortgage default to be properly quantified and the hard to formalise intuitive aspects of mortgage underwriting (Foote, Loewenstein and Willen, 2019). Despite Fair Isaac's plans to construct a mortgage specific score, the FICO® consumer lending score was adopted. It quantified creditworthiness by setting the 660 scores as a delineator between prime and subprime markets based on statistical analyses that revealed that those under 660 were more likely to default. The FICO® score was incorporated in a mortgage score including mortgage lending data (e.g., loan-to-value ratios, property type) and Freddie Mac created the statistically-based automated underwriting system Loan Prospector and Fannie Mae the Desktop Underwriter (Gates, Perry and Zorn, 2002). Lenders were initially reluctant to see credit scoring relevant in predicting default on mortgage loans since in contrast with consumer credit it was predominantly thought at the time that the equity held in collateral makes borrowers less likely to default (Foote, Loewenstein and Willen, 2019). Given the GSEs important role in the mortgage market in the USA, credit scoring and computerised models of mortgage underwriting were soon adopted by most lenders. This has involved a relaxation of debt-to-income ratios, which statistically proved to be less relevant than credit scoring in predicting defaults, lenders approving more loans at the lower end of the income distribution (Foote, Loewenstein and Willen, 2019). However, along with the technological changes, during the 1990s various governmental measures have also been put in place in order to enlarge access to credit for lower-income households (Foote, Loewenstein and Willen, 2019). But the credit expansion facilitated has resulted in having proportionally the younger and better-educated households being better represented rather than low-income, less-educated households (Foote, Loewenstein and Willen, 2019). Consequently, the crisis of default following the financial crisis has been predominantly a middle class one. Stout (2019) designates American debtors' struggles with foreclosures as *post-middle-class-life projects,* to draw attention that in the contemporary USA the middle class as a subject position becomes increasingly unreachable given the financialization of housing.

Outside the GSEs market, the FICO® score adoption as a central underwriting tool was pushed by Standard & Poor, the main risk rating agency for MBS, managing to unite around it a series of heterogenous actors who otherwise had different information systems, risk assessments and products. Rating agencies such as Standard & Poor used the same category system as for corporate bonds (e.g., AAA) for the sub-prime-mortgage-based CDOs, lending in this way legitimacy to what proved much riskier financial products (Krippner, 2017; Marron, 2009). Martha Poon (2008) argues that the broad adoption of credit scoring marks a shift from *credit by screening* to *credit by risk*. Previous lenders' rule-based assessments distinguished between who gets and who not credit and were used as *risk minimising strategy*. The new statistical calculations of risk through credit scoring created a spectrum of borrowers who can be offered different products and prices, reflecting a *risk management strategy*. Proposed as a neutral market device that does not consider any protected characteristics, the FICO® score has allowed the financial industry to put a price on risk and enlarge their pool of "creditworthy" mortgage borrowers and change the traditional forms of financing debts through their securitisation.

The distinctions created along income scales have pointed also to the discriminatory practices of lending institutions in the USA. These have long been a point of contention given their racial component, credit scoring and reports augmenting rather than preventing them. Various scholars point to redlining, the denial of lending to African American and Hispanic applicants, and also yellow lining, the approval of mortgages with higher interest rates, or reverse redlining through which the same victims have to confront forms of predatory lending, or just being given less time and information (Pager and Shepherd, 2008; Hinnant-Bernard and Crull, 2004; Aalbers, 2011). This discrimination has been met with various consumer movements and regulations (Krippner, 2017; Marron, 2009) while Freddie Mac's automated writing tool, Tool Inspector, has been changed in order to assure more inclusion of underserved populations (Gates, Perry and Zorn, 2002). But the logic of credit scoring remains fundamentally one in which the legacy of economic, racial, and ethnic inequalities is still reflected (Foohey and Greene, 2021).

The rise of FinTechs in mortgage lending that usually digitalise their whole mortgage lending process and automate decisions is seen as a possible remedy given increased competition, reduced cost of information, and fewer personal interactions (Rodima-Taylor, 2022; Fuster et al., 2019). Current evidence suggests that although FinTech algorithms still discriminate, they do it 40% less than face to face lenders (Bartlett *et al.*, 2019). If studies confirm that algorithms might reduce discrimination in comparison to face to face lending and cater for lending to underserved populations (Bartlett et al., 2019; Gates, Perry and Zorn, 2002), it is important to note that algorithms still have *consistent* biases that are applied on a large scale especially given biases inherent in the data sources themselves (Schmeckpeper et al., 2021; Courchane and Ross, 2019).

UK

In the UK, the automatization of mortgage underwriting and the usage of credit scoring, either developed in-house or through credit reference agencies, have started to be used more broadly in the 2000s especially by large and medium-size lenders, although very few used a fully automated process at the time (Van Dijk and Garga, 2006) while currently, most banks use logistic regression models rather more advanced ML algorithms (CDEI, 2020). Credit scoring models were seen as useful time-efficient tools and as a way of reaching new market segments, both qualities highly desirable for acting in a very competitive market, and also as tools able to facilitate more consistent and transparent decisions and better control, reporting, and governance (Van Dijk and Garga, 2006). They were part of broader underwriting decisions that included policy rules (e.g., minimum/maximum age of applicant, maximum LTV), affordability assessments, and other checks such as fraud and money laundering. The distinction between prime and subprime lending played a role in the extension of automatization. Most of the lenders operating in the subprime market used a manual assessment given that sub-prime applicants have profiles less amenable to automated policy rules and they were also more likely to use a credit score given the need for a more detailed assessment.

Credit scores were used for boundary and within-market classifications (Fourcade and Healy, 2013) of applicants: a threshold was used for accepting/rejecting them while those accepted might receive different loan amounts, interest rates, and periods depending on the value of their credit score. In mortgage lending, creditors might use a combination of data provided by CRAs (currently, TransUnion, Experian, and Equifax), their own, and other third parties. The scores created by CRAs are based on three main sources: 1) public information that includes electoral roll, CCJ data, insolvency information, bankruptcy/IVA data; 2) credit search and inquiry information; 3) credit reporting data submitted by lenders according to SCOR's Principle of Reciprocity (2021). Given FCA's regulations, credit risk assessments include calculations of "credit risk" to the lender and "affordability risk" to the borrowers but the latter seems to raise some challenges given unreliable declarations of income and expenditure in applications, lack of cross-referencing across the industry, and the incapacity of Current Account Turnover (CATO) data to fully capture the complexity of an individuals' financial arrangements or to convince every consumer to consent (Deville, 2020, p.11).

More automated decision-making tools tend to be used more by fintech companies which have more "flexibility and space" to develop ML tools, while the banking sector tends to be more conservative, to has stricter regulatory requirements, and constraints due to legacy systems (Marron, 2009; Davey, 2020). However, big data credit scoring and analytics seem to be the future that all actors will embrace (Deville, 2020).

Social collateral

In all countries, mortgage contracts and the usage of credit scores seem at a first glance to favour the individual, reinforcing the idea that capitalism is centred on individuals, while in practice it is households or extended families who take risks (Zaloom, 2019). Most of the critical studies on credit scoring decry their moralisation of *individual* borrowers (Marron, 2009; Davey, 2020), missing the importance of the fact that banks ask for co-debtors and guarantors to the loan agreement, especially in the case of low-incomes or more precarious forms of employment. This suggests that credit-debt relations are far from being interpersonal, involving rather wider inter-domestic networks. Muriel's (2021) study of mortgages in Spain points out how at different stages of mortgage borrowing, inter-domestic networks have to be mobilised, arguing that neither individuals nor households are good units of analysis in order to fully understand mortgage indebtedness and its risks. Firstly, the decision to take a mortgage is usually encouraged and rewarded by a broader social circle, various relatives financially contributing to the down payment or acting as guarantors of the mortgage contract. Secondly, various life circumstances might create difficulties in mortgage repayment, financial or in-kind kin support being the first strategy that people can retort to. This can be unconditional but also fraught with tensions and feelings of shame. Thirdly, in the case of repossessions, it is usually the extended family that offers a substitute for shelter. Moreover, relatives who agreed to act as guarantors might have their incomes seized by creditors for any outstanding debts, the risk of repossession being spread beyond individual borrowers. Muriel argues that predatory lending practices on the Spanish mortgage market can be partially explained by the fact that creditors saw a borrower's social networks as a reliable repayment insurance or collateral in case of default. In this way, a debtor's social capital has been monetized and transformed into "a vehicle of extended financial expropriation" (Muriel, 2021, p.50). Such a study draws attention to the fact that parents might not be only important actors for long-term accumulation strategies focused on housing (Adkins, Cooper and Konings, 2020; Forrest and Hirayama, 2018) but also be exposed to financial hardship when their children's mortgage arrangements fail. Examining the types of intergenerational support for access to homeownership in the UK, Suh (2019) notes that several banks have already adapted to the important role played by parents, creating new first-time buyers mortgages that are directly linked to their parents' wealth. She also draws attention to the fact that the median wealth of people aged between 55-64 was £21,000 in 2014/2015 while the average deposit was £33,000, suggesting that parents might go into debt to help their children.

Digital technologies and credit information in rental markets

Apps and digital technologies are increasingly being adopted on rental markets to select tenants and manage properties. Various scholars have drawn attention to the importance of studying how digital technology is reshaping the real estate industry given the growing pace of innovation targeting this industry in the last decade, a domain called "PropTech" (Fields and Rogers, 2021; Porter et al., 2019; Ferreri and Sanyal, 2021; Shaw, 2020; Langley and Leyshon, 2021; Maalsen et al., 2021; Short et al., 2006). Maalsen et al. (2021, p.16) argue that rental technologies are becoming a key new mechanism of discrimination on the rental market, noting that "the types of information recorded and digitised, the ways it is algorithmically transformed into data products, and who has access to those products, are crucial considerations for discrimination-focussed research and policy". McElroy and Vergerio (2022) note that the industry has been developing since the 1980s through the adoption of technologies such as the internet, data aggregators, phone apps, APIs, machine learning and technology. Such technologies have been imagined as fully disrupting the social organisation of the rental market, in a more general understanding of electronic markets as agents of disintermediation (Schmeckpeper et al., 2021). But their adoption seems more patchy in practice, the roles of intermediaries such as letting agents being still important and involving continuity with older practices (Dunning *et al.*, 2019). According to Unissu, a prop-tech listing service, 8,521 companies were operating globally, out of which 5,305 were on the residential market (636 in the UK). Still, McElroy and Vergerio (2022) draw attention to the inadequacy of corporate data given that their case study, although operating as a PropTech was listed as an information technology, manufacturing, and security corporation.

Trying to establish a research agenda, Fields and Rogers (2021) propose an examination of PropTech or what they call platform real estate by taking into account their *platform logic*, *digital labour* and *financialization*. They classify the current existing services as 1) *trading platforms*, digital services and products aimed at the buying and selling of real estate; 2) *operational platforms* that automate aspects of the rental and property management process; and 3) *data platforms* which provide comprehensive real estate data and AI and big data analyses. For Fields and Roger (2021, p.77), taking into account such technologies' platform logic means examining "matters of control, corporate dominance, and the profit driven platform's drive for capital accumulation". Such a focus on power and politics, Fields and Rogers argue, could also be orientated towards the various forms of labour at play in the working of these technologies: the unwaged digital labour that users might engage with when producing data that is commodified; the digitally mediated wage labour or the perception of the impact of non-human labour. Finally, the authors propose financialization as a useful analytical term to approach these technologies inviting researchers to examine how new connections between financial instruments and real estate are forged through them.

Among scholars focused on operational platforms, Boeing et al. (2021) analysis of rental housing platforms, draws attention to the fact that the data used might increase existing socio-spatial inequalities by serving the specific business interest and without being fully transparent to users. The authors find more problematic short-term rental rather than long-term rental platforms since they not only amplify existing inequalities but also create new ones given the low stock of accommodation available for low-income renters. Still, long term rental can be affected in multiple ways. McElroy and Vergerio (2022) show how in New York post-9/11 era tenant screening and surveillance have proliferated as a landlord technology justified as a way of "keeping buildings safe" while the post-2008 financial crisis period witnessed an extension of corporate landlordism and digital property management platforms. Their focus is on tech companies that produce building-access and surveillance systems, and they argue that such technologies produce a "carceral spatiality" by making the space of the home one of surveillance by using new technologies such as facial recognition to "automate gentrification" by policing lease violations in order to evict poorer tenants and attract richer ones.

The growing use of digital technologies in the management of tenants and properties in the rental market has been described by Fields (2022) as giving rise to the phenomenon of the *automated landlord*. Her study of large investors in the single-family homes to rent in the USA shows how such technologies support the assestisation of housing and permit control at a distance through acquisition algorithms that make the evaluation and purchase of housing more efficient while tenant facing technologies offer functionalities such as search and application for the property and the management of rent payments and maintenance requests. She argues that rather than fully replacing landlords' capacities, such technologies augment them since, given the idiosyncrasies of properties and people, a full automation process is hard to reach. Still, following Fourcade and Healy (2020, p.9) she points to how such an amassing of data or what Fourcade and Healy call *information dragnet,* creates new ways of categorising people that allow the industry various forms of profit extractions. Fields (2017) notes how the datafication of tenants makes them to pay two rents, one under the form of money, the other under the form of data.

Beyond management, digital technology has an important role in shaping access to rental market. This can start with online property listing and communication, Hogan and Berry's (2017) finding that in Toronto, email inquiries to landlords showed a significant level of discrimination against Muslim/Arabic men and a modest one against Asian men, Blacks, and Muslim/Arabic women. Their study shows an important affordance enabled by digitalisation which given its asymmetrical communication facilities an "opportunity denying" discrimination (exclusion through nonresponse). Hence, as Maalsen et al. (2017) argue digital technologies on the rental market can have a powerful discriminatory effects through opportunity denying exacerbated in the context of digital means of communication, differences in access to and the ability to use internet, and an increase in the informal ways of renting in which gender, race, age, disability play a covert role of selection. Another way of excluding/including applicants is through the development of tenants digital databases. For Australia, Short et al. (2006) note how they might enable property managers to exclude "high risk" tenants but such a constructions of risk further marginalise vulnerable persons.

Similarly, credit scoring and reports can also be used to mediate access to rental properties, being increasingly adopted on the rental markets in various countries as part of an automated process of tenant screening. Nonetheless, the empirical literature on the automatization of tenant screening is scant and predominantly focused on the USA. If retracing a potential tenant's financial history and behaviours is seen as a good index for assessing the future behaviour of a tenant and avoiding financial losses, in practice this might be a new logic of social sorting that usually reinforces existing social segregation (Migozzi, 2020) by being only a "crude proxy for poverty" (Rosen, Garboden and Cossyleon, 2021) in the case of low-income households. In the USA, automated tenant screening based on what are allegedly clear, objective criteria are adopted in order to comply with fair housing regulations and allow better management of properties when scale is important (Reosti, 2020; Decker, 2021; Rosen, Garboden and Cossyleon, 2021). Hence, there is still an important difference in its adoption given the different types of landlords and the market niches in which they operate, Rosen et al (2021) distinguish between *algorithmic screening* and *gut proxies* as the main tenants' selection procedures. The former refers to the adoption of algorithms that automate selection based on criteria such as income, credit report, criminal history, and residential history and is used by more professionalised and resourced landlords. The latter refers to the more subjective evaluation of tenants by smaller landlords. However, both end up enacting various forms of racial discrimination, the algorithms through the usage of variables such as credit score or eviction history that are proxies for marginalisation based on race. An important point made by Rosen et al. is that the usage of automation might be under revision given a business's objectives, showing how some criteria might be changed to accommodate the specific characteristics of their target market (e.g., what counts as valid proof of income).

Given the split of the rental market between larger, professional landlords and smaller amateur ones, subjective evaluation still seems to play an important role for the latter. Reosti's (2020) case study of how antidiscrimination law is applied to the rental market in Seattle shows how despite the fact that tenant screening tools might be adopted as a more uniform way of evaluation, the role played by subjective assessment is more important in the case of small-scale rental housing providers. So long as they were not explicitly discriminatory, such subjective evaluations were in some cases to the benefit of disadvantaged renters, going beyond a black and white assessment. However, they were also less transparent, increasing the search costs for renters with a more discrediting background. A similar point is made by Decker (2021) who shows how in USA automated tenant screening is used more by larger portfolio landlords who due to scale tend to routinize their property management and tenant selection. Although smaller-scale landlords might use similar screening, asking for credit checks and employer references, a subjective assessment based on in-person interaction occupied an essential role. In this way, criteria for becoming and remaining a tenant were more explicit in the case of large scale, professional landlords who preferred to work with more objective, clear rules not only for ease of work but also to protect themselves against accusations of discrimination. After controlling for tenants' incomes and property and neighbourhood characteristics, Decker argues that such differences in tenant selection account also for the fact that large scale landlords, who usually operate low-cost rental housing, are more likely to rent to someone who will later miss rent payments. But such larger landlords might also have more restrictive policies, e.g., longer criminal lookback periods or more frequent consideration of all credit factors, and be less likely to consider mitigating factors, as Smith and Byrne (2021) show in their analysis of subsidized housing providers in the USA who used tenant screening companies.

UK

In a review of the literature on mechanisms and forms of exclusion in the contemporary UK housing market, Preece and Bimpson (2019) note that Big Data, algorithms, and social media are increasingly being adopted by landlords to mitigate against risks, leading to a *direct exclusion* of some tenants based on individual characteristics and risk factors and to an *indirect exclusion* through advertising some housing services only to some specific groups. The authors comment that there is very limited evidence on the adoption of credit scoring within the rented sector in the UK or more generally of automated tenant screening processes and it is unclear what their impact is despite their problematic aspects related to the enactment of larger structural inequalities under a veneer of technological neutrality and to the breach of privacy. Similarly, a Big Brother Watch (2021) report examining the use of algorithms in welfare based on Freedom of Information requests, points to how Risk-Based Verification risk scoring used for housing benefits allocation, developed in the UK by private companies, one of them TransUnion, although not explicitly using protected characteristics, might nonetheless have important biases. The uses of geodemographics from the Office of National Statistics' 2011 Area Classification for Output Areas, of data points related to children, statutory sick pay, income, capital, employment, bereavement or carer status have, according to them, the potential to incur discrimination against disadvantaged groups.

As mentioned in the case of automated tenant screening, there is a difference between the various types of landlords and the extent to which they adopt digital technologies or their rationales. Pleace's (2005) case study on digital services on the social housing market in the UK, notes that larger social landlords tended to use such services aiming to increase efficiency and decrease the administrative burden of some routine tasks such as rent collections and repair requests. Savings from digitalisation were reused in programmes aimed at helping tenants to improve their digital skills or to maintain non-digital routes of interaction. At the time, smaller landlords were less inclined to use such digital services due to their costs. Interestingly, both types of landlords saw full automation less possible in the case of more complex transactions such as the assessment process for a household seeking access to social housing or establishing that a need for assistance was 'legitimate'. Pointing to the fact that digitalisation is shaped by the contexts in which it is introduced, Pleace argues that contrary to the US model of the of "citizen consumer", British social landlords were rather promoting a more redistributive approach through reinvesting savings from digitalisation into improving access to services and the digitalisation of their tenants but this was prior to the recession of 2008 and subsequent austerity measures.

But the impact of more recent welfare reforms and the overall path to digitalisation taken in the UK, might make such a difference less relevant currently. Hickman et al.'s (2018) study of the impact of welfare reforms on housing associations note that reforms such as Universal Credit have accelerated the digitalisation of their services. Digitalisation was seen as a measure to reduce costs, as a way of making services more accessible and responsive, although it was acknowledged that some tenants are less prone to using digital services. At the same time, Preece et al. (2020) note the increased use of affordability assessments in the case of social housing given the impact of the latest welfare reforms that are based on various digital technologies. In their study, some housing associations mentioned that they were working with CRAs and were trying to develop models for tenant risk that were including data from credit reports. Preece et al. argue that the implications of such usages were unclear, as the process remained opaque and tenants had limited ability to challenge them.

Looking at the private market, Ferreri and Sanyal (2021) note how in the UK, a service such as Tenant Assured asked tenants to give access to their social media accounts or search for red flags such as "high risk language alerts" and "new to country alert" signalling potential privacy issues and a bias against immigrants. Hence, tenant digital services seem to have proliferated in the last decade, having different usages on the private rental market and social housing, but most of the existing studies have a limited empirical reach, pointing to a need of more comprehensive studies that look at how digital technologies and models of risks are used in order to mediate access to housing on rental markets.

Regulations

That algorithms govern but are also being governed is reflected by the fact that shifting networks of governance exist in various national contexts that determine what type of data can be used and how (Wainwright, 2011). Still, this should not be taken as an unilinear process in which regulations have the power to control the detrimental effects of technologies since sociotechnical assemblages might involve various political struggles in order to gain a dominant position, reflecting various interest groups (2009). Ferreri and Sanyal (2021) draw, for example, attention to the state's desires to encourage technological innovation, allowing companies to exploit the ambiguity of rules. At the same time, Graeber (2014) notes that the state through institutions such as bailiffs, prison, police protect economic extraction from debt relations, opening questions around whose "public interests" is best represented through regulations.

The history of credit scoring reveals that its accelerated and widespread adoption was facilitated by a need for businesses to legally comply with regulations related to data protection, discrimination, and financial risk protection. The literature on the USA reveals how fair access to housing and credit and protection of privacy has shaped the credit information industry and the public image of credit scoring. Here, the Equal Credit Opportunity Act of 1974, amended in 1976, outlawed discrimination in credit based on gender, marital status, race, national origin, religion, age, or income source. Marron (2009) and Lauer (2017) argue that the way in which credit scoring has been encoded in this regulation recognised it as an objective tool able to eliminate "subjective" discrimination in lending markets, even if the usage of "secondary" variables such as type of employment or tenure was still correlated with structural inequalities in terms of gender or race. Given debates around privacy due to the centralisation of credit information facilitated by computer databases, Federal Credit Reporting Act was adopted in 1970, restricting the uses of credit files, making them slightly more transparent, and giving more power to consumers to amend them. But as Lauer (2017) notes it had a series of loopholes that made it, in the end, less a guarantee of privacy and it left open the issue of what counts as *relevant* information for credit granting. Given its shortcomings, various privacy regulations have started to be adopted at the state level asking companies to be more transparent about how and what data they collect on their consumers (Schmeckpeper *et al.*, 2021).

For the UK, various regulations shape the design and usage of credit scoring, including consumer credit and data protection regulations, anti-discrimination, competition, intellectual property, and general consumer protection laws, while the main regulatory institutions are the Financial Conduct Authority, Information Commissioner's Office, Financial Ombudsman Service, and Equality and Human Rights Commission (Aggarwal, 2021). The financial sector's supervisors are considered "technologically neutral", i.e. they do not intervene in the technologies used, but they do monitor the potential risks created (BoE and FCA, 2019). There seems to be strong public support for better regulation of data and algorithms, but it is unclear what the public sees as the most important and what they find as acceptable trade-offs (Ada Lovelace Institute, 2022).

Aggarwal (2021), focused more on the implications of algorithmic credit scoring and consequently on consumer credit and data protection regulations, argues that *allocative efficiency, distributional fairness*, and *consumer privacy* (as autonomy) are the three main norms underpinning its regulation. Allocative efficiency is meant to assure that capital is allocated to borrowers who can repay according to the contractual conditions and is reflected primarily in the requirement of affordability assessment in defining creditworthiness. It is also meant to assure market integrations which in terms of data involves the free movement of personal data. Consequently, such an efficient allocation would have implications also for the distributional fairness principle which aims at avoiding allocating unaffordable credit to more vulnerable consumers. The regulatory regime of data recognises the importance of consumer privacy as autonomy given the growing importance of "digital selves" in contemporary societies, going beyond privacy as non-interference, informational secrecy, or confidentiality through various forms of control and limitations of processing of data (e.g., stricter control on the usage of personal sensitive data) and a recognition of the right of self-determination. Aggarwal notes that current empirical studies, few and focused only on the USA, support the thesis that algorithmic credit scoring might enhance allocative efficiency and distributional fairness through better risk assessment analyses, but this is highly dependent on the quality of training data used. As for consumer privacy, this is highly reduced; the constant surveillance of all sorts of data involves a need for new solutions and regulations.

Currently, UK's regulatory choices are between the EU model of "trustworthy AI" which places a greater emphasis on consumer protection and ethics and the USA's or China's models, which have a lower bar in terms of personal data protection and collection (Ada Lovelace Institute). A Bank of England and FCA (2019) report on the usage of ML in the financial sector notes that the industry does not see regulation as an "unjustified barrier" to ML deployment, which has as more important constraints legacy IT systems and data limitations. According to a Centre for Data Ethics and Innovation (2020) review report, the collection of data on protected characteristics or proxies for them is generally avoided given that it is unlawful, a "fairness through unawareness" being more generally adopted in algorithmic decision making. But this might create challenges for assessing biased outcomes, and as a response some banks use protected characteristics data not in their models for decision making but in testing their bias. Risk around algorithmic bias is more an organisational concern, given that existing standards bodies do not define specific bias thresholds. Still, issues raised by commercial sensitivity and competition make such models opaque and hard to assess if there is an issue of algorithmic bias or larger structural inequalities.

Table 1: UK Regulatory Landscape

<table>
<tr><td>Body</td><td>Financial Conduct Authority (FCA):</td></tr>
<tr><td colspan="2">Relevant Regulations:
• CONC rules, FCA consumer credit regulations: https://www.handbook.fca.org.uk/handbook/CONC/1/?view=chapter

• FCA 2018 Rules for assessing creditworthiness
Creditworthiness is understood as credit risk and affordability. Affordability assessment based only on a customer's income raised concerns that this might lead to financial exclusion arguing for taking into account a household's income where appropriate.

The policy statement notes that FCA rules are neutral in terms of automation without aiming to discourage its usage. But: "where processes are automated, we expect firms to have appropriate policies and procedures to ensure they can adequately manage any risks associated with those processes. The same applies if the firm relies to a significant extent on data or information from CRAs or other third parties" (5)
It takes a principle-based rather than a prescriptive approach that stipulates that "lenders must use "sufficient information" but without mentioning what it should comprise, and what types and sources of information can be used.

HMRC tax, council tax, rental payment information, and Open Banking - have an ambiguous status and are not currently included in the regulation
• PSD2 ((Revised) Payment Services Directive) - EU legislation, enforced by FCA; subject to customer consent, it aims to promote data-sharing to third parties-providers (TPP) to share payment account information
• Consumer Credit Act 1974</td></tr>
</table>

Consumers have the right to access and correct errors in their credit reports; credit providers must notify applicant from what CRA they took the data; CRAs must disclose credit files to the consumers. Note that FCA is now undertaking a review of the credit information market that might involve changes to current legislation.	
Body	**Competition and Markets Authority**
Relevant Regulations: ● Competition and Markets Authority's Retail Banking Market Investigation Order 2017 - Open Banking; requires the nine largest banks in the UK to share current account information in a standardised format (application programming interface, API)	
Body	**The Information Commissioner's Office (ICO)**
Relevant Regulations: ● Data Protection Act 2018 (currently under review) If a "significant decision" (i.e., has adverse legal effects or it significantly affects the data subject) is fully automated, the controller has to notify the data subject about the decision's nature. The data subject has to right to request the controller to reconsider the decision or take a new one that is not based only on automatisation. The controller has to comply with the request and announce the data subject to the steps taken and outcomes. But there is no clear demarcation of what a decision based only on automation means, allowing for insignificant human interventions to change the status of such decisions (Big Brother Watch, 2021). Overarching data protection principles: data minimisation and purpose limitations; lawful data processing (consent or/and "legitimate interests"); "data protection by design and default"; "data protection impact assessment" for situations of high risk. Consumers do not have to give consent for data sharing if there is a fair reason. In the case of data errors in CRAs files, this makes consumers powerless in having control over what and how data is shared between them, lenders, and other actors such as landlords (Tims, 2017).	
Body	**Steering Committee on Reciprocity**
Relevant Regulations: ● SCOR 2021 Principles of Reciprocity - industry forum responsible for "administration and development of the data sharing rules known as the Principles of Reciprocity" Usages of credit report data allowed: employment vetting process for "certain government departments"; "qualifying police service"; Members of Closed User Groups; CRAs; CIFAS; Social Housing Sector (based on the reciprocity principle social landlords share social housing rental data that can be accessed by credit performance data providers); Private Housing Sector (only by landlords or agencies with more than 100 properties);	

Horizons

The future of the credit information industry seems to be orientated towards new data frontiers and way of collecting them. "Financial inclusion" has become an important strategy for the financial sector to enlarge access to credit for what are considered to be "thin files", i.e., people who do not have a credit history. This has reverberated on the credit information market, where talk about "alternative data" has become dominant. Attention is given to how various other regular payments, digital footprints, or psychometric and behavioural characteristics can be included in a credit score (Foohey and Greene, 2021; Equifax, 2021; Aggarwal, 2021; Berg et al., 2019; Anon; Campbell-Verduyn, Goguen and Porter, 2017; Bernards, 2019).

Deville (2020) notes that big data credit scoring is one of the latest developments of the credit risk assessment practices but at the moment it is only tentatively embraced by the credit industry. The usage of alternative data seems to be more readily embraced in short term and subprime lending to enhance the risk assessment of "thin" credit files. Such examples include Sunny which combines traditional credit scoring with data from a user's device and browser behaviour and Aire which uses machine learning to derive information about borrowers from all sorts of sources. According to Deville, the following data represents the credit industry's new experiments with data, but their adoption is rather restricted either to some organisations or to specific goals:

Table 2: Newer data sources being used in UK credit risk assessment adapted from Deville (2020, p.9)

Data type	**Description**
Behaviour scores	Predictive scores derived from a borrower's credit usage patterns with a particular creditor. Potentially combines data relating to financial behaviour – such as repayment history, or changes in the credit limit to credit balance ratio – with non-financial information – changes in address or marital status, for instance. Particularly used in credit card lending.
Current account credit turnover (CATO)	Reported by retail banks, drawing on income and expenditure information from users' current accounts. Uses algorithmic modelling to estimate borrower incomes. This is a challenging task, according to a number of interviewees, giving the complexities of many individuals' financial arrangements (see section 2.3.1). Access to data varied according to the POR, but even those lenders not reporting CATO data can employ services offered by CRAs to undertake income verification. Initially provided just to Callcredit (in 2008, according to one interviewee), but now used by all UK CRAs.
Internet and device	Digital information released by a user accessing an app or webpage. Includes IP address, which is used to estimate a user's location, and information about the type of device (e.g. make, model, operating system) being used. Used by at least one credit reference agency in their fraud prevention work.
Omni channel	Using data derived from communications with borrowers across various 'channels' – for example phone, email, SMS. Can allow a more dynamic assessment of risk, in which risk is determined by patterns in how a user responds to a prompt via a particular channel. Reported to be being used by

	some creditors in the UK and more widely overseas.
Phone & broadband	Data concerning payment history with communications providers. A relatively new addition to CRA data.
Social housing rental	Data reported to CRAs by social housing providers regarding tenant payment histories and listed in credit reports. A new addition to certain CRA data.
Private rental	Data reported to a CRA by either an individual tenant or a property management organisation. Currently restricted to one CRA, via Experian's recently launched 'The Rental Exchange'. This which allows individual private tenants and large property management organisations (letting more than 500 properties) to have private rental payment data recorded on an individual's credit file. For most private renters it remains an 'opt-in' possibility
Utilities	Data concerning payment history with utilities providers. A relatively new addition to CRA data.
Vehicle finance	Data concerning payment history with vehicle finance providers. A relatively new addition to CRA data.
Added by the authors: (Mentioned in the literature, not known yet how widely they have been adopted, they are more likely to be used for consumer loans from FinTechs)	
Transactional data	Broad access through Open Banking In use: FusionScore/Financial Health Index from Account Score and Equifax. Key metrics: "Behaviours of the debit and credit transactions observed in a customer's current accounts"; "Attitudes to spending and saving"; "Early warning to changes in lifestyle and financial status".
Social media activity	There seems to be a low adoption in the financial sector and a high scepticism around its relevance. It was used by a CRA and a major bank a few years ago but without making their models more accurate to justify further usage (CDEI, 2020). Tenant Assured offered by the firm Score Assured in 2016 asked for access to social media accounts for tenants applications but it seems to have been withdrawn (Ferreri and Sanyal, 2021).
Mobile phone usage	Call duration, time calls are made, numbers frequently called, who initiates calls, grammar and punctuation in text messaging are used as variables in credit scoring by FinTechs but mostly in emerging markets (Berg et al 2019).
Payments on "subprime" loans	Such as payday loans (Foohey and Greene, 2021)

Deville notes that Open Banking creates new opportunities for credit risk assessments based on "real-time analysis of transactional data". The industry seems already to have moved forwards and create such products (e.g. Bud, Zopa). An Equifax report (Equifax, 2021) praises Open Banking as an important and "disruptive" tool in credit risk analysis allowing for more "accurate and holistic" assessments of customers data and enabling automation at scale. Nonetheless, a broader adoption of such new sources of data does not seem to be always a straightforward process. In the case of rental data, the Creditworthiness Assessment Bill which was aimed at forcing lenders to include rental data in credit risk assessments failed to gain support and did not pass the parliament[3]. But, Experian is making use of such data and is being supported by PropTech companies and agents such as Canopy, CreditLadder, and Howsy (Lewis, 2019). At the same time, governmental bodies seem open to new developments around Big Data and AI, the UK National AI Strategy in 2021 revealing the aims of a ten-year plan to make the UK an "AI superpower". And the financial sector perceives that a broader adoption of ML and their improvement will lead to new forms of personalisation of products, new analytical insights, and improved services in the next three years (BoE and FCA, 2019)

Given the new opportunities and challenges of the credit information market, FCA has commissioned in 2019 a broad study, one of its first reports being a scenario-thinking approach to assess its possible futures. In the *Future of Credit Information Market* (Marron, 2009; Langley, 2014; DuFault and Schouten, 2020) four scenarios are envisioned based on the fact that demographics and economy, technology, consumers trends and preferences, credit information and lending-markets trends and regulatory tools and initiatives shape the information market. The most "preferable" scenarios seem those in which Big Data and AI will play a much more important role in shaping the market, favouring an imaginary of growth led by technological changes.

***

The lived experience and understandings of "data subjects"

Much has been said about the capacity of risk-assessment technologies to make up people, but less is known about how ordinary people experience and understand them. Amoore (2020, p.16) argues that algorithms are aperture instruments or tools of perception that extract, reduce, and condense data environments showing what is to be perceived. This raises questions related to how one identifies with their data doubles or what Gillespie (2009, p.114) calls making ourselves *algorithmically recognizable*. Gillespie (2014, p.184) argues that a focus should be put not on the "effects" or "impact" of algorithms on people but rather on the "multidimensional "entanglement" between algorithms put into practice and the social tactics of users who take them up". This is because there is rather a "recursive loop" between what algorithms and people do, which instils changes not only

[3] UK Parliamentary Bills: https://bills.parliament.uk/bills/2005/news (Accessed 8th of June 2022)

in people but also in algorithms.

Most influential studies of credit scoring have focused on how financial devices envision a specific type of subject such as the data entrepreneur (Marron, 2009; Langley, 2014; DuFault and Schouten, 2020) while others might point to the creation of the "lumpenscoretariat" and digitally invisibles (Fourcade and Healy, 2017; Fourcade, 2021a). Marron (2009, p.114) notes that through credit scoring, consumers are "objects of surveillance and self-surveillance in their exercise of choice". Langley (2014, p.452) contends that the commodification of consumer scoring might involve various directions of subjectification, creating not only subjects who meet outstanding obligations but also who "have an entrepreneurial disposition to credit and to their credit score". But empirical studies of how ordinary people understand and experience credit scoring are rather scarce, inviting a more empirically based phenomenology of ordinality (Fourcade, 2021b).

Credit scores generated by the credit reference agencies bear striking similarities with self-tracking devices, their marketing strategies employing gamification strategies aimed at alluring their clients to "boost" their score or find out how they compare with neighbours, while incentive platforms (e.g., Waypoints in the USA) are designed to gamify rent extraction (Fields, 2017). The credit score and nudging advices become techniques through which consumers are invited to make behavioural changes, encouraging positive steps towards building an "excellent" score and penalising wrongdoings. Similarly, gamification in all sorts of self-tracking apps uses the same feedback loops and performance metrics to govern behaviour rather than offering experiences of "fun" (Whitson, 2014). In this way, consumers become not only subjects of markets' classification situations (Fourcade and Healy, 2013) but also subjects that have to govern, regulate, and optimize themselves. Focused mostly on the USA, the literature on credit scoring as experienced by common people reveals a strong awareness and significance of it in the public consciousness, with people actively engaging in its construction and using it as a marker of identity (Hohnen, Ulfstjerne and Krabbe, 2021; Gusterson, 2019). In a more methodological oriented study, Ziewitz and Singh (2021) note that there is rather a spectrum of attitudes and experiences shaped by different levels of awareness, a need for ongoing inquiry and an open sense-making process when one is forced to encounter them. Hence, at the one extreme, there might be people who are completely ignorant about their credit scoring, given a lack of events requiring them to use it even though they might be data subjects of credit reference agencies, while at the other extreme, there might be people actively engaged in monitoring and transforming their credit scoring.

Analysing material from online forums in the USA through the lens of quantified self-ontology, DuFault and Schouten (2020) note the emergence of the *datapreneur consumer identity*. Such posts reveal a high level of transparency by sharing highly personal financial data, an optimisation mindset, and an engagement with practices of feedback loops and score-hacking. Among these users, there seems to be a broad acceptance of credit scoring as a marker of creditworthiness, internalising the financial lingo and engaging in various financial practices to boost their credit score. At the same time, a bad credit score raises feelings of shame, fear, and self-recrimination, reflecting the moral framework of such tools. Nonetheless, constant monitoring and acting upon it through the consumption of financial products make such users feel as if agentive and able to control how they are seen despite realising that the scoring rules keep on changing. Given such broad acceptance of the authority of the market-assigned score and their willingness to enter the game, DuFault and Schouten propose the notion of *datapreneurial identity* to articulate consumer practices that involve an iterative process of becoming aware and improving one's commercial data identity marking new ways of making more responsible consumers in neoliberal markets.

Based on an ethnographic study, Kear (2017) proposes an understanding of credit scoring as a *key legal technology of arbitration*. This reflects the structural role of credit score in economic life with important consequences in societies where credit plays an important role in sustaining livelihoods, making "playing the credit scoring game" obligatory even for people who would rather not. Such games entail various technologies of surveillance but also digital self-representation strategies, while the credit score has an arbitration function by mediating interactions between market participants through taming a future uncertainty of default. He points to how his USA informants' usage and understanding of credit scoring revealed a mismatch between the quantification at the base of a credit score such as MyFICO and their personal sense of creditworthiness and financial realities. Nonetheless, they engaged in various activities of building credit even if these were in contrast with their values (e.g., antipathy toward debt but using a credit card) given the importance of credit in allocating resources. Kear notes that if in the past credit reporting and scoring involved a "restriction of borrower' ability to knowingly affect their scores, the industry considering this as a form of "manipulation", currently borrowers are more able to "consciously and strategically affect their score" (Kear, 2017, p.358). In other words, subjects learn to tell creditworthiness stories that are legible to algorithms even if these might conflict with their understandings and recognise the importance of data. Consequently, if previous credit social movements involved debates around rights to credit through an anti-discrimination discourse, current debates shaped around financial inclusion are played in terms of positive data, i.e., what other data could be assessed to get more inclusion. This, even if it recognises existing inequalities nonetheless reinforces the hegemony of credit score and its distributional role of wealth and income.

UK

In the case of the UK, the existing studies, mostly grey literature, are more focused on attitudes rather than experiences, being usually based on survey studies, focus groups, or experiments (CDEI, 2021; Miller et al., 2020; Hartman et al., 2020; Ada Lovelace Institute, 2022; Waind, 2020; Samson, Gibbon and Scott, 2019; BritainThinks, 2021). According to these, there seems to be strong public support for a broad usage of data and algorithms, but only as long as there is sufficient regulation to secure consumers' rights and specific technologies are deployed in a manner consistent with public benefit. For example, if technology like facial recognition has strong support if used by police in criminal investigations or to unlock phones, it has less so if used in supermarkets to track shoppers' behaviour or in hiring processes, consumers being more reluctant to share data only for commercial benefits (Ada Lovelace Institute, 2022, p.15). Consequently, the nature of the data shared, the context of sharing, and who accesses and uses that data influences the public's comfort with data sharing.

Public attitudes surveys reveal a "data trust deficit" in the UK, with people being more willing to share their data with organisations that they trust to use it in the public interest. NHS, universities and banks are the most trustworthy organisations, while the less trustworthy are social media companies and big tech (CDEI, 2021). When it comes to the popularity of using algorithmic decision-making tools, a Centre for Data Ethics and Innovation report (2020) notes that in the UK financial services are the most popular. Experimental attempts to assess the perception of fairness of algorithm-based lending decisions revealed that research participants tended to drop a bank's services if they were informed that they use proxy data for protected characteristics or social media data, especially those who felt more at risk of being discriminated against. Still, the usage of proxy data such as salary was more ambiguous since, although it might speak about structural discrimination, it was considered a legitimate variable for assessing affordability.

More experiential studies reveal a more ambiguous attitude towards such technologies and a more shifting usage. Kirwan's (2021) ethnographic study of clients of a debt advice service in the UK points to how the "governmentality of the credit file" might have different receptions according to one's circumstances and hopes for the family's security. If some of his interlocutors refused to go into bankruptcy in order not to affect their credit scoring and hence the possibility to get credit in a context of "deficit budgets", others preferred to be "enforcement free" and protect their family from the experience of eviction, bailiffs, and committal proceedings, even this badly affected their credit score. Kirwan's study shows the different relevance that credit scoring can have and how they are embedded in broader social relations, imaginaries of the future, and specific affective states. Similarly, Davey (2020) notes how some of his indebted informants from a council estate refused debt advice services given that some of them, e.g., insolvency and debt management plans, might reduce their credit scores and consequently their ability to access credit.

The Ada Lovelace Institute review report (2022) on public attitudes to data notes important gaps in our knowledge that should be investigated in future research avenues:

1) More specificity related to what the public sees as a beneficial regulation of data-driven technologies;
2) More detailed analyses of what the public considers "public benefit" and responsible data use given the inherent tensions of such technologies that span the private/public boundaries;
3) How the public expects for transparency to look in practice and how concretely transparency and awareness link. If most of the existing studies measure awareness, given the preponderance of public attitudes studies, it is important to enquire more about understandings, what people concretely know and how this shapes their demands for transparency.
4) More explicit exploration of the connections between data understanding, awareness, and attitudes.

***

Conclusion

The increased use of data and algorithms in housing decisions is just beginning to receive scholarly attention despite the proliferation of such technologies in the industry. In this review, we have put a focus on credit scoring and reporting, showing how its adoption involves specific social organisations and has an impact on changing conceptions and management of risk and consequently shaping who has access or not or at what price on the housing markets. The usage of credit scoring and reporting can be placed in the broader socio-technical ecosystem built around digital technologies aimed at managing and shaping access to housing either through buying with a mortgage or renting in the private or social sectors. Such an ecosystem is characterised by a proliferation of actors, data circulating between apps and websites developers, CRAs, various intermediaries, public authorities, and financial institutions or landlords. Its environment is mainly digital, and despite regulations of data processing, there might still be various overflows that raise concerns around its opacity, biases, privacy and informed consents, or accountability for data errors. Such an ecosystem creates data subjects that have a reduced capacity to grasp such processes, not to mention to challenge them, pointing to a need of more in-depth examinations of how such technologies are adopted and with what effects.

Below we briefly summarise some of the most important points raised by this review:

Credit scoring and reporting has reshaped the lending history through new conceptions and management of risks that enabled the development of risk pricing. Despite their initial development for consumer lending, their usage has been adopted in other sectors, making their measure of creditworthiness a general one for moral character although in fact it reflects socio-economic structurally un/privileged positions.

The statistical tools for credit scoring and the data available through reports have facilitated the automation of mortgage underwriting and tenant screening. They have been promoted as less biased tools for housing decisions. However, the current available data is thin in measuring their social impact, mostly focused on the USA, and reveals new possible forms of discrimination due to variables that are proxies for larger societal inequalities.

The usage of electronic means for credit scoring has enabled the centralisation of data leading to a disappearance of local credit bureaus and the creation of three major credit reference agencies in the USA that currently operate globally offering not only credit reference services but also consumer data analytics. Although there have been important data breaches and there seems to be a thin line between the usage of data for credit references vs. marketing purposes, little is known about how data is used by CRAs and how consumers understand and experience such a collection of data about them.

There is an increased use of data and digital technologies on the real estate market, various platforms and products being created to trade and operate on the market. The rental market is being reconfigured through new technologies developed to risk assess tenants and manage properties, making data a new type of rent that tenant are constrained to pay either when applying for a rent or by having property under a full surveillance.

Future prospects of the ecosystem built around credit information data seem to recognize its current limit, as having a credit history, for example, is limited for various kinds of people that might be otherwise creditworthy. This has generated a search for new alternative data that could prove more useful but despite a rhetoric built around the slogan of financial inclusion there is still no open discussion around the opacity of these tools, how they are able to protect privacy and get informed consent, or who is accountable for data errors. At the same time, the usage of alternative data will still involve an analysis that prioritises corelation and not causality and that is blind to contextual and structural factors at base of a specific individual's life circumstances.

In several countries, many consumers have started to "play the credit score game" as they are aware that such a technical artifact mediates their access to important goods. Still, the current literature suggest, that there are various experiences and attitudes towards such tools and more generally the collection of data by the industry or government inviting for more specificity in trying to understand differences between situations and groups of people.

References

Aalbers, M. B. (2008). The Financialization of Home and the Mortgage Market Crisis. *Competition & Change*, 12 (2), pp.148–166.

Aalbers, M. B. (2011). *Place, Exclusion and Mortgage Markets*. Malden: Wiley-Blackwell.

Aalbers, M. B. (2017a). *The Financialization of Housing: A political economy approach*. London and New York: Routledge.

Aalbers, M. B. (2017b). The variegated financialization of housing. *International Journal of Urban and Regional Research*, 41 (4), pp.542–554.

Accominotti, F. (2021). The aesthetics of hierarchy. *The British Journal of Sociology*, 72 (2), pp.196–202.

Acolin, A., Goodman, L. and Wachter, S. M. (2019). Accessing Homeownership With Credit Constraints. *Housing Policy Debate*, 29 (1), pp.108–125.

Ada Lovelace Institute. (2022). *Who cares what the public think? Public Attitudes to UK Data*. London: Ada Lovelace Institute.

Ada Lovelace Institute. *Expert explainer: The EU AI Act Proposal*. [Online]. Available at: https://www.adalovelaceinstitute.org/resource/eu-ai-act-explainer/ [Accessed 9 April 2022].

Adkins, L. (2018). *The time of money*. Stanford: Stanford University Press.

Adkins, L., Cooper, M. and Konings, M. (2020). *The Asset Economy*. Cambridge: Wiley.

Adkins, L., Cooper, M. and Konings, M. (2021). Class in the 21st century: Asset inflation and the new logic of inequality. *Environment and Planning A: Economy and Space*, 53 (3), pp.548–572.

Aggarwal, N. (2021). The Norms of Algorithmic Credit Scoring. *The Cambridge Law Journal*, 80 (1), pp.42–73.

Amoore, L. (2020). *Cloud Ethics: Algorithms and the Attributes of Ourselves and Others*. Durham: Duke University Press.

Arya, S., Eckel, C. and Wichman, C. (2013). Anatomy of the credit score. *Journal of Economic Behavior & Organization*, 95, pp.175–185.

Bartlett, R. et al. (2019). *Consumer-Lending Discrimination in the FinTech Era*. [Online]. Available at: https://www.nber.org/system/files/working_papers/w25943/w25943.pdf. [Accessed 20 April 2022]

Berg, T. et al. (2019). On the Rise of FinTechs: Credit Scoring Using Digital Footprints. *The Review of Financial Studies*, 33 (7), pp. 2845–2897.

Bernards, N. (2019). The poverty of fintech? Psychometrics, credit infrastructures, and the limits of financialization. *Review of International Political Economy*, 26 (5), pp.815–838.

Besedovsky, N. (2018). Financialization as calculative practice: the rise of structured finance and the cultural and calculative transformation of credit rating agencies. *Socio-Economic Review*, 16 (1), pp.61–84.

Big Brother Watch. (2021). *Poverty Panopticon: The hidden algorithms shaping Britain's welfare state*. London: Big Brother Watch.

BoE and FCA. (2019). *Machine learning in UK financial services*. Bank of England and FCA . [Online]. Available at: https://www.fca.org.uk/publication/research/research-note-on-machine-learning-in-uk-financial-services.pdf. [Accessed 15 February 2022].

Boeing, G. et al. (2021). Tilted platforms: rental housing technology and the rise of urban big data oligopolies. *Urban Transformations*, 3 (1), p.6.

Bowker, G. C. and Star, S. L. (2000). *Sorting things out: Classification and its consequences*. Cambridge: The MIT Press.

Breiman, L. (2001). Statistical Modeling: The Two Cultures. *Schweizerische Monatsschrift fur Zahnheilkunde = Revue mensuelle suisse d'odonto-stomatologie / SSO*, 16 (3), pp.199–231. [Accessed 30 March 2022].

BritainThinks. (2021). *How can the public sector be meaningfully transparent about algorithmic decision making?* CDEI. [Online]. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/995014/Complete_transparency__complete_simplicity_-_Accessible.pdf. [Accessed 15 February 2022].

Bücker, M. et al. (2022). Transparency, auditability, and explainability of machine learning models in credit scoring. *The Journal of the Operational Research Society*, 73 (1), pp.70–90.

Burrell, J. (2016). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, 3 (1), p.1-15.

Burrell, J. and Fourcade, M. (2021). The Society of Algorithms. *Annual Review of Sociology*, 47 (1), pp.213–237.

Burton, D. (2007). *Credit and Consumer Society*. London: Routledge.

Burton, D. (2012). Credit Scoring, Risk, and Consumer Lendingscapes in Emerging Markets. *Environment & Planning A*, 44 (1), pp.111–124.

Cakici, B., Ruppert, E. and Scheel, S. (2020). Peopling Europe through Data Practices: Introduction to the Special Issue. *Science, Technology & Human Values*, 45 (2), pp.199–211.

Callon, M. and Law, J. (1982). On Interests and Their Transformation: Enrolment and Counter-Enrolment. *Social Studies of Science*, 12 (4), pp.615–625.

Campbell-Verduyn, M., Goguen, M. and Porter, T. (2017). Big Data and algorithmic governance: the case of financial practices. *New Political Economy*, 22 (2), pp.219–236.

Carruthers, B. G. (2013). From uncertainty toward risk: the case of credit ratings. *Socio-Economic Review*, 11 (3), pp.525–551.

CDEI. (2020). *Review into bias in algorithmic decision-making*. [Online]. Available at: https://www.gov.uk/government/publications/cdei-publishes-review-into-bias-in-algorithmic-decision-making/main-report-cdei-review-into-bias-in-algorithmic-decision-making [Accessed 20 April 2022].

CDEI. (2021). *Public Attitudes to Data and AI Tracker Survey. Wave 1*. [Online]. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1064525/Public_attitudes_to_data_and_AI_-_Tracker_survey.pdf. [Accessed 15 February 2022].

Cevolini, A. and Esposito, E. (2020). From pool to profile: Social consequences of algorithmic prediction in insurance. *Big Data & Society*, 7 (2), pp.1–11.

Cheney-Lippold, J. (2011). A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control. *Theory, Culture & Society*, 28 (6), pp.164–181.

Chiapello, E. (2015). Financialisation of Valuation. *Human Studies*, 38 (1), pp.13–35.

Citron, D. K. and Pasquale, F. (2014). The scored society: Due process for automated predictions. *Washington Law Review* , 89, p.1.

Courchane, M. J. and Ross, S. L. (2019). Evidence and Actions on Mortgage Market Disparities: Research, Fair Lending Enforcement, and Consumer Protection. *Housing Policy Debate*, 29 (5), pp.769–794.

Curran, D. (2015). Risk illusion and organized irresponsibility in contemporary finance: rethinking class and risk society. *Economy and Society*, 44 (3), pp.392–417.

Curran, D. (2018). Risk, innovation, and democracy in the digital economy. *European Journal of Social Theory*, 21 (2), pp.207–226.

Dastile, X., Celik, T. and Potsane, M. (2020). Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing*, 91, p.106263.

Davey, R. (2020). Retail Finance and the Moral Dimension of Class. Debt Advice on an English Housing Estate. In: Hann, C. and Kalb, D. (Eds). *Financialization. Relational Approaches*. New York and Oxford: Berghahn. pp.220–240.

Decker, N. (2021). How Landlords of Small Rental Properties Decide Who Gets Housed and Who Gets Evicted. *Urban Affairs Review* , p.1-17.

Dencik, L. et al. (2019). The 'golden view': data-driven governance in the scoring society. *Internet policy review*, 8 (2). [Online]. Available at: doi:10.14763/2019.2.1413. [Accessed 17 February 2022].

Deville, J. (2020). *Futures of credit risk assessment in the UK*. [Online]. registry-trust.org.uk. Available at: https://www.registry-trust.org.uk/documents/46/Futures_of_Credit_Risk_Assessment_-_March_2020.pdf [Accessed 1 February 2022].

Deville, J. and Van der Velden, L. (2015). Seeing the invisible algorithm: The practical politics of tracking the credit trackers. In: *Algorithmic Life*. London: Routledge. pp.101–120.

Douglas-Jones, R. (2021). Bodies of data: doubles, composites, and aggregates. *The journal of the Royal Anthropological Institute*, 27 (S1), pp.159–170.

Dourish, P. (2016). Algorithms and their others: Algorithmic culture in context. *Big Data & Society*, 3 (2), pp.1–11.

DuFault, B. L. and Schouten, J. W. (2020). Self-quantification and the datapreneurial consumer identity. *Consumption Markets & Culture*, 23 (3), pp.290–316.

Dunkerley, F. et al. (2021). *Future of the credit information market*. Centre for Future and Foresight Studies and FCA. [Online]. Available at: https://www.fca.org.uk/publication/research/future-credit-information-market-report.pdf. [Accessed 15 February 2022].

Dunning, R. et al. (2019). Technological change and estate agents' practices in the changing nature of housing transactions. *Housing Studies*, 34 (5), pp.849–867.

Equifax. (2021). *The digital future of Finance and Credit Risk: The 2021 Open Banking Research Report*. London: Equifax.

Esposito, E. and Stark, D. (2019). What's Observed in a Rating? Rankings as Orientation in the Face of Uncertainty. *Theory, Culture & Society*, 36 (4), pp.3–26.

Eubanks, V. (2018). *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's Publishing Group.

Fazi, M. B. (2021). Beyond Human: Deep Learning, Explainability and Representation. *Theory, Culture & Society*, 38 (7–8), pp.55–77.

Ferreri, M. and Sanyal, R. (2021). Digital informalisation: rental housing, platforms, and the management of risk. *Housing Studies*, pp.1–19.

Fields, D. (2017). *Rent, datafication, and the automated landlord*. Shaw, J. and Graham, M. (Eds). Oxford: Meatspace Press., pp.16–19.

Fields, D. (2022). Automated landlord: Digital technologies and post-crisis financial accumulation. *Environment & Planning A*, 54 (1), pp.160–181.

Fields, D. and Rogers, D. (2021). Towards a Critical Housing Studies Research Agenda on Platform Real Estate. *Housing, Theory and Society*, 38 (1), pp.72–94.

Foohey, P. and Greene, S. S. (2021). *Credit Scoring Duality*. Law and Contemporary Problems, Forthcoming 2022. [Online]. Available at: doi:10.2139/ssrn.3992749 [Accessed 13 April 2022].

Foote, C., Loewenstein, L. and Willen, P. (2019). *Technological innovation in mortgage underwriting and the growth in credit: 1985–2015*. Federal Reserve Bank of Boston. [Online]. Available at: doi:10.29412/res.wp.2019.11.

Forrest, R. and Hirayama, Y. (2018). Late home ownership and social re-stratification. *Economy and Society*, 47 (2), pp.257–279.

Fourcade, M. (2021a). Ordinal citizenship. *The British Journal of Sociology*, 72 (2), pp.154–173.

Fourcade, M. (2021b). The unbearable rightness of being ranked. *The British Journal of Sociology*, 72 (2), pp.203–206.

Fourcade, M. and Healy, K. (2013). Classification situations: Life-chances in the neoliberal era. *Accounting, Organizations and Society*, 38 (8), pp.559–572.

Fourcade, M. and Healy, K. (2017). Seeing like a market. *Socio-Economic Review*, 15 (1), pp.9–29.

Fuster, A. et al. (2019). The Role of Technology in Mortgage Lending. *The Review of Financial Studies*, 32 (5), pp.1854–1899.

Fuster, A. et al. (2022). Predictably Unequal? The Effects of Machine Learning on Credit Markets. *Predictably Unequal? The Effects of Machine Learning on Credit Markets*, LXXVII (1), pp.5–47.

Gabor, D. and Brooks, S. (2017). The digital revolution in financial inclusion: international development in the fintech era. *New Political Economy*, 22 (4), pp.423–436.

Gates, S. W., Perry, V. G. and Zorn, P. M. (2002). Automated underwriting in mortgage lending: Good news for the underserved? *Housing Policy Debate*, 13 (2), pp.369–391.

Gillespie, T. (2014). The relevance of algorithms. *Media technologies: Essays on communication, materiality, and society*, 167 (2014), p.167.

Graeber, D. (2014). Anthropology and the rise of the professional-managerial class. *HAU: Journal of Ethnographic Theory*, 4 (3), pp.73–88.

Gusterson, H. (2019). Introduction: Robohumans. In: Bestheman, C. and Gusterson, H. (Eds). *Life by algorithms: How roboprocesses are remaking our world*. Chicago: The University of Chicago Press. pp.1–19.

Guyer, J. I. (2010). The eruption of tradition? *Anthropological Theory*, 10 (1–2), pp.123–131.

Hacking, I. (1990). *The Taming of Chance*. Cambridge: Cambridge University Press.

Hartman, T. et al. (2020). Public perceptions of good data management: Findings from a UK-based survey. *Big Data & Society*, 7 (1), pp.1–16.

Hassani, B. K. (2021). Societal bias reinforcement through machine learning: a credit scoring perspective. *AI and Ethics*, 1 (3), pp.239–247.

Hayles, K. (1999). *How We Became Posthuman. Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago: The University of Chicago Press.

Henry, N. and Morris, J. (2018). Scaling Up Affordable Lending: Inclusive Credit Scoring. *Coventry: Responsible Finance*. [Online]. Available at: https://www.researchgate.net/profile/John-Morris-20/publication/342010616_Scaling_Up_Affordable_Lending_Inclusive_Credit_Scoring/links/5ede54c692851cf138697fa2/Scaling-Up-Affordable-Lending-Inclusive-Credit-Scoring.pdf. [Accessed 15 February 2022].

Herrine, L. (2016). Credit Reporting's Vicious Cycles. *NYU Rev. L. & Soc. Change*, 40, pp.305–350.

Hickman, P., Pattison, B. and Preece, J. (2018). *The impact of welfare reforms on housing associations: A scoping study*. [Online]. Available at: http://shura.shu.ac.uk/22302/1/cache-impact-welfare-reforms-housing-associations.pdf. [Accessed 15 February 2022].

Hinnant-Bernard, T. and Crull, S. R. (2004). Subprime Lending and Reverse Redlining. *Housing and Society*, 31 (2), pp.169–186.

Hogan, B. and Berry, B. (2011). Racial and Ethnic Biases in Rental Housing: An Audit Study of Online Apartment Listings. *City & Community*, 10 (4), pp.351–372.

Hohnen, P., Ulfstjerne, M. A. and Krabbe, M. S. (2021). Assessing Creditworthiness in the Age of Big Data: A Comparative Study of Credit Score Systems in Denmark and the US. *Journal of Extreme Anthropology*, 5 (1), pp.29–55.

Jasanoff, S. (2015). Future imperfect: Science, technology, and the imaginations of modernity. In: Jasanoff, S. and Kim, S.-H. (Eds). *Dreamscapes of modernity: Sociotechnical imaginaries and the fabrication of power*. Chicago and London: The University of Chicago Press. pp.1–33.

Kear, M. (2017). Playing the credit score game: algorithms, 'positive' data and the personification of financial objects. *Economy and Society*, 46 (3–4), pp.346–368.

Kirwan, S. (2021). Between a knock at the door and a knock to your score: re-thinking 'governing through debt' through the hopeful 'imaginaries' of UK debtors. *Journal of Cultural Economy*, 14 (2), pp.159–175.

Kitchin, R. (2014). *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. London: SAGE Publications.

Kitchin, R. (2021). *Data Lives: How Data Are Made and Shape Our World*. Bristol: Bristol University Press.

Knight, F. H. (1921). *Risk, uncertainty and profit*. Boston: Houghton Mifflin.

Krippner, G. R. (2017). Democracy of Credit: Ownership and the Politics of Credit Access in Late Twentieth-Century America. *The American Journal of Sociology*, 123 (1), pp.1–47.

Langley, P. (2007). Uncertain Subjects of Anglo-American Financialization. *Cultural Critique*, (65), pp.67–91.

Langley, P. (2008). Sub-prime mortgage lending: a cultural economy. *Economy and society*, 37 (4), pp.469–494.

Langley, P. (2009). Debt, Discipline, and Government: Foreclosure and Forbearance in the Subprime Mortgage Crisis. *Environment & Planning A*, 41 (6), pp.1404–1419.

Langley, P. (2014). Equipping entrepreneurs: consuming credit and credit scores. *Consumption Markets & Culture*, 17 (5), pp.448–467.

Langley, P. and Leyshon, A. (2021). The Platform Political Economy of FinTech: Reintermediation, Consolidation and Capitalisation. *New Political Economy*, 26 (3), pp.376–388.

Lapavitsas, C. (2014). *Profiting Without Producing: How Finance Exploits Us All*. London and New York: Verso Books.

Lapavitsas, C. and Powell, J. (2013). Financialisation varied: a comparative analysis of advanced economies. *Cambridge Journal of Regions, Economy and Society*, 6 (3), pp.359–379.

Lauer, J. (2017). *Creditworthy: A History of Consumer Surveillance and Financial Identity in America*. Columbia University Press.

Lazarus, J. (2012). The Ambition of Credit Scoring: Forecasting Credit Failure. *Raisons politiques*, 48 (4), pp.103–118.

Lewis, N. (2019). *Legislation to help renters get on property ladder dies a quiet death in parliament*. [Online]. Available at: https://thenegotiator.co.uk/creditworthiness-assessment-dies/ [Accessed 12 April 2022].

Leyshon, A. and Thrift, N. (1999). Lists come alive: electronic systems of knowledge and the rise of credit-scoring in retail banking. *Economy and society*, 28 (3), pp.434–466.

Lofranco, Z. T. (2021). “I would not be a loan guarantor even for my blood brother”: Financialization of inter-household solidarity and the privatization of risk in Bosnia and Herzegovina. In: *Households and Financialization in Europe*. Routledge. pp.56–78.

Lupton, D. (2019). *Data Selves: More-than-Human Perspectives*. Newark: Polity Press. [Accessed 8 June 2022].

Maalsen, S. et al. (2021). *Understanding discrimination effects in private rental housing*. Australian Housing and Urban Research Institute Limited. [Online]. Available at: doi:10.18408/ahuri7324501.

Mackenzie, A. (2015). The production of prediction: What does machine learning want? *European Journal of Cultural Studies*, 18 (4–5), pp.429–445.

MacKenzie, D. (2006). *An engine, not a camera: How financial models shape markets*, Inside Technology. London, England: MIT Press.

Marron, D. (2009). *Consumer Credit in the United States: A Sociological Perspective from the 19th Century to the Present*. New York : Palgrave Macmillan.

McElroy, E. and Vergerio, M. (2022). Automating gentrification: Landlord technologies and housing justice organizing in New York City homes. *Environment and planning. D, Society & space*, pp.1–20.

McFall, L. and Moor, L. (2018). Who, or what, is insurtech personalizing?: persons, prices and the historical classifications of risk. *Distinktion: Journal of Social Theory*, 19 (2), pp.193–213.

Migozzi, J. (2020). Selecting Spaces, Classifying People: The Financialization of Housing in the South African City. *Housing Policy Debate*, 30 (4), pp.640–660.

Mikuš, M. and Rodik, P. (2021). Introduction: Households and peripheral financialization in Eastern and Southern Europe. In: Mikuš, M. and Rodik, P. (Eds). *Households and Financialization in Europe*. New York: Routledge. pp.1–34.

Miller, C. et al. (2020). *People, Power and Technology: The 2020 Digital Attitudes Report*. Doteveryone. [Online]. Available at: https://doteveryone.org.uk/wp-content/uploads/2020/05/PPT-2020_Soft-Copy.pdf. [Accessed 15 February 2022].

Monahan, T. (2018). Algorithmic Fetishism. *Surveillance and Society*, 16 (1), pp.1–5.

Muriel, I. S. (2021). Within and beyond the household: Spanish mortgage debtors' tactics to cope with financialization in a semi-periphery. In: Mikuš, M. and Rodik, P. (Eds). *Households and Financialization in Europe*. New York: Routledge. pp.37–55.

Nichols, J. (2004). Data doubles: Surveillance of subjects without substance. *CTheory*, Online, pp.2–17.

Onay, C. and Öztürk, E. (2018). A review of credit scoring research in the age of Big Data. *Journal of Financial Regulation and Compliance*, 26 (3), pp.382–405.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Penguin Books.

Pager, D. and Shepherd, H. (2008). The Sociology of Discrimination: Racial Discrimination in Employment, Housing, Credit, and Consumer Markets. *Annual Review of Sociology*, 34, pp.181–209.

Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge and London: Harvard University Press.

Pellandini-Simányi, L. and Vargha, Z. (2018). Spatializing the future: financial expectations, EU convergence and the Eastern European Forex mortgage crisis. *Economy and society*, 47 (2), pp.280–312.

Pleace, N. (2005). The shaping of electronic service delivery. *Information, Communication and Society*, 8 (4), pp.524–541.

Poon, M. (2016). Corporate Capitalism and the Growing Power of Big Data: Review Essay. *Science, technology & human values*, 41 (6), pp.1088–1108.

Poon, M. A. (2008). From New Deal Institutions to Capital Markets: Commercial Consumer Risk Scores and the Making of Subprime Mortgage Finance. *https://papers.ssrn.com › papershttps://papers.ssrn.com › papers*. [Online]. Available at: https://papers.ssrn.com/abstract=1458545 [Accessed 13 January 2022].

Porter, L. et al. (2019). Planning, Land and Housing in the Digital Data Revolution/The Politics of Digital Transformations of Housing/Digital Innovations, PropTech and Housing – the View from Melbourne/Digital Housing and Renters: Disrupting the Australian Rental Bond System and Tenant Advocacy/Prospects for an Intelligent Planning System/What are the Prospects for a Politically Intelligent Planning System? *Planning Theory & Practice*, 20 (4), pp.575–603.

Porter, T. M. (1995). *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton: Princeton University Press.

Preece, J. and Bimpson, E. (2019). *Forms and mechanisms of exclusion in contemporary housing systems*. UK Collaborative Centre for Housing Evidence. [Online]. Available at: https://thinkhouse.org.uk/site/assets/files/1469/cache0419b.pdf. [Accessed 25 April 2022]

Preece, J., Hickman, P. and Pattison, B. (2020). The affordability of "affordable" housing in England: conditionality and exclusion in a context of welfare reform. *Housing Studies*, 35 (7), pp.1214–1238.

Pughe, E. (2018). *Responsible Finance calls for credit reference agencies to do more to support financial inclusion*. [Online]. Available at: https://responsiblefinance.org.uk/2018/02/responsible-finance-calls-for-credit-reference-agencies-to-do-more-to-support-financial-inclusion/ [Accessed 1 February 2022].

Reosti, A. (2020). "We Go Totally Subjective": Discretion, Discrimination, and Tenant Screening in a Landlord's Market. *Law & social inquiry: journal of the American Bar Foundation*, 45 (3), pp.618–657.

Ritzer, G. (2001). *Explorations in the Sociology of Consumption: Fast Food, Credit Cards and Casinos*. London, Thousand Oaks, New Delhi : SAGE Publications.

Rodima-Taylor, D. (2022). Introduction. In: Rodima-Taylor, D. and Shipton, P. (Eds). *Land and the Mortgage: History, Culture, Belonging*. Berghahn Books. pp.1–34.

Rona-Tas, A. (2017). The Off-Label Use of Consumer Credit Ratings. *Historical Social Research / Historische Sozialforschung*, 42 (1 (159)), pp.52–76.

Rona-Tas, A. and Guseva, A. (2018). Consumer Credit in Comparative Perspective. *Annual Review of Sociology*, (44), pp.55–75.

Rosen, E., Garboden, P. M. E. and Cossyleon, J. E. (2021). Racial Discrimination in Housing: How Landlords Use Algorithms and Home Visits to Screen Tenants. *American Sociological Review*, 86 (5), pp.787–822.

Samson, R., Gibbon, K. and Scott, A. (2019). *About Data, About Us*. Open Data Institute and Royal Society for the encouragement of Arts, Manufactures and Commerce. [Online]. Available at: https://www.thersa.org/globalassets/pdfs/reports/data-about-us-final-report.pdf. [Accessed 13 January 2022]

Schmeckpeper, K. et al. (2021). Algorithm transparency through the Fair Credit Reporting Act (FCRA). *Journal of Science Policy & Governance*, 18 (04). [Online]. Available at: doi:10.38126/jspg180415.

Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society*, 4 (2), pp.1–12.

Shaw, J. (2020). Platform Real Estate: theory and practice of new urban real estate markets. *Urban geography*, 41 (8), pp.1037–1064.

Short, P. et al. (2006). Tenancy Databases, Professional Practices and Housing Access among Low-Income Tenants in the Private Rental Sector in Australia. *International Journal of Urban and Regional Research*, 30 (4), pp.930–943.

Smith, M. C. and Byrne, T. H. (2021). Locked Out: The Systematic Exclusion of Poor Renters From Federally Subsidized Housing. *Housing policy debate*, Ahead-of-Print, pp.1–19.

Solove, D. J. (2004). *The Digital Person: Technology and Privacy in the Information Age*. New York: New York University Press.

Stout, N. (2019). *Dispossessed: How Predatory Bureaucracy Foreclosed on the American Middle Class*. Oakland: University of California Press.

Suh, E. (2019). *Young British Adults' Homeownership Circumstances*. [Online]. Available at: https://www.housing-studies-association.org/articles/2-young-british-adults-homeownership-circumstances [Accessed 18 March 2022].

The Bank of England and FCA. (2022). *Artificial Intelligence Public-Private Forum*. [Online]. Available at: https://www.bankofengland.co.uk/-/media/boe/files/fintech/ai-public-private-forum-final-report.pdf?la=en&hash=F432B83794DDF3F580AC5A454F7DFF433D091AA5.

Thomas, S. L., Nafus, D. and Sherman, J. (2018). Algorithms as fetish: Faith and possibility in algorithmic work. *Big Data & Society*, 5 (1), pp.1–11.

Tilly, C. (1998). *Durable Inequality*. Berkley: University of California Press.

Tims, A. (2017). *How credit score agencies have the power to make or break lives*. [Accessed 10 May 2022].

Van Dijk, R. and Garga, S. (2006). *UK mortgage underwriting*. The Council of Mortgage Lenders. [Online]. Available at: https://www.oxera.com/wp-content/uploads/2018/03/UK-mortgage-underwriting-3.pdf. [Accessed 23 February 2022]

Waind, E. (2020). *Trust, Security and Public Interest: Striking the Balance*. Administrative Data Research UK and Economic & Social Research Council. [Online]. Available at: https://www.adruk.org/fileadmin/uploads/adruk/Trust_Security_and_Public_Interest-_Striking_the_Balance-_ADR_UK_2020.pdf. [Accessed 13 January 2022]

Wainwright, T. (2011). Elite Knowledges: Framing Risk and the Geographies of Credit. *Environment & planning A*, 43 (3), pp.650–665.

Weiss, H. (2016). *Creditworthiness and the consumer perspective: on credit scoring in Israel*. [Online]. Available at: http://pubman.mpdl.mpg.de/pubman/item/escidoc:2350698/component/escidoc:2350697/mpi-eth-working-paper-0173.pdf. [Accessed 17 January 2022]

Whitson, J. R. (2014). Foucault's fitbit: Governance and gamification. In: Waltz, S. and Deterding, S. (Eds). *The Gameful World*. Boston : The MIT Press Cambridge. pp.339–358.

Wigand, R. T. (2020). Whatever happened to disintermediation? *Electronic Markets*, 30 (1), pp.39–47.

Zaloom, C. (2019). *Indebted: How Families Make College Work at Any Cost*. Princeton: Princeton University Press.

Zarsky, T. (2016). The Trouble with Algorithmic Decisions: An Analytic Road Map to Examine Efficiency and Fairness in Automated and Opaque Decision Making. *Science, technology & human values*, 41 (1), pp.118–132.

Ziewitz, M. and Singh, R. (2021). Critical companionship: Some sensibilities for studying the lived experience of data subjects. *Big Data & Society*, 8 (2), pp.1–13.

The Rental Exchange. [Online]. Available at: https://www.experian.co.uk/business/customer-insights/risk-analysis/rental-exchange [Accessed 10 February 2022].