INTERNATIONAL JOURNAL OF BANKING AND FINANCE
https://e-journal.uum.edu.my/index.php/ijbf

How to cite this article:
Wandmacher, R., Thorhauer, Y., Sturm, C., Driessen, M., Soine, D., & Goerke, T. (2025). A global comparison of credit bureaus based on data utilization in credit scoring. *International Journal of Banking and Finance*, *20*(1), 23-38. https://doi.org/10.32890/ijbf2025.20.2.2

A GLOBAL COMPARISON OF CREDIT BUREAUS BASED ON DATA UTILIZATION IN CREDIT SCORING

[1]Ralf Wandmacher, [2]Yvonne Thorhauer, [3]Christian Sturm, [4]Mareike Driessen, [5]Dominique Soine, [6]Tim Goerke
[1,2,3,4,5&6]accadis Hochschule Bad Homburg, Germany
[1,2&3]accadis Institute of Entrepreneurship, Bad Homburg, Germany

[3]Corresponding author: christian.sturm@accadis.net

Received: 19/12/2023 Revised: 16/2/2024 Accepted: 27/3/2024 Published: 27/10/2024

ABSTRACT

The basic concept of credit scoring is to assess an individual`s payment ability as well as the specific individual`s credit default risk, hence determining an individual`s creditworthiness. Based on the credit score, financial institutions, insurance companies, telecommunication companies and other businesses decide whether consumers are eligible for a mortgage, credit card, auto loan, and other credit products. However, in many countries, potential tenants and insurance applicants also use credit scores extensively for screening. Accordingly, Credit Bureaus (CB) or Consumer Reporting Agencies (CRA) exert an essential gatekeeper function for important economic areas of consumers' everyday life. However, when examining CBs globally, there are considerable differences in the use of data to calculate credit scores. Interestingly, the influence of CBs on credit rating receives little to no attention in academic research. This is particularly evident in the absence of a framework for classifying Credit Bureaus. Therefore, 24 traditional and non-traditional Credit Bureaus operating in 17 different countries are analyzed. First, the study identifies the different data types underlying credit reports and credit scores. Second, CBs are classified and clustered according to the type of information used for credit scoring. Furthermore, promising areas of research, in particular the ethical conflict between data protection and economic participation are highlighted.

Keywords: Credit bureaus, consumer reporting agencies, credit scoring, credit reports, data utilization.

JEL Classification: G20, G24, G28.

INTRODUCTION

In general, credit reports are necessary to evaluate the creditworthiness of consumers to mitigate the potential credit default risk (Kozodoi et al., 2019). In the European Union, Article 8 of the 2008 EU Consumer Credit Directive obliges to assess the creditworthiness of a consumer "where necessary on the basis of a consultation of the relevant database." The updated 2023 EU Consumer Directive replaces the Consumer Credit Directive 2008/48/EC of the European Parliament and the Council. This requires lenders to assess the creditworthiness of customers also for new types of credit, such as "Buy now, Pay later" (Preamble 16 of EU Directive 2023/2225). Especially banks (McIntosh et al., 2013) and similar financial institutions (Markov et al., 2022) rely on credit reports to assess the payment behavior and default risk when issuing loans. However, also insurance companies, telecommunication companies, utility companies, residential real estate companies and other businesses decide – based on the credit report – whether consumers are eligible for health insurance, a mobile phone contract, an auto loan, or a rented home (Asher, 1994; Hartwig & Wilkinson, 2003; Consumer Financial Protection Bureau, 2023).

A credit report is a detailed compilation of financial and personal data that is necessary for assessing the creditworthiness of a private or corporate customer (Miller, 2003; McCann & McIndoe-Calder, 2015). The credit report is based on the analysis of the credit history (Spader, 2010), financial history (Wei et al., 2016), and social-demographic data of the borrower (Siddiqi, 2017). The aggregated data is analyzed based on statistical or mathematical models to create a credit score (Consumer Financial Protection Bureau, 2023); however, there is no common approach or calculation method used by all credit bureaus and consumer reporting agencies (Hertza, 2018). Furthermore, these models often neglect that external conditions such as geographical location or other factors influence the final credit score (Óskarsdóttir et al., 2019).

In addition to credit reports – especially in the USA – also sector-specific, so-called consumer reports provide information for certain industries, particularly regarding insurance, utilities, retail, or gaming (Consumer Financial Protection Bureau, 2023). In addition to financial-related data, a consumer report can compromise additional personal data such as criminal history (Finlay, 2012), driving records (Golden et al., 2016), and employment history (Hayashi, 2019). Alternatively, mobile financial service providers, such as the Santa Monica, California-based FinTech Tala, calculate credit scores based on everyday data from smartphones (Njathi, 2019).

Thus, these mobile financial service providers can also include unbanked customers, who do not have access to traditional financial services (Blechmann, 2016). However, the extent to which credit decisions should be based on non-traditional data remains controversial (Francis et al., 2017). Yet, these alternative credit scores may be the only possibility to access credit, especially in emerging countries (Kumar et al., 2021). Therefore, these alternative credit scores will be increasingly used despite regulatory concerns, as traditional credit scoring is just not possible (Óskarsdóttir et al., 2018). Due to the significant differences in both credit reports and scoring procedures, it is crucial to examine the different types of scoring companies as well as the influence of these scoring companies on the creation of credit scores in more detail (Amat et al., 2017).

LITERATURE REVIEW

Types of Credit Bureaus

In general, credit bureaus are classified based on the ownership structure – public or private (Kusi et al., 2017). Accordingly, legal requirements and regulations – which may differ for private and public credit bureaus – define the purpose of these institutions and set the framework for the scoring system (Ferretti, 2013).

Public credit bureaus – generally referred to as public credit registries (PCRs) – are often operated by central banks or bank superintendencies (Ali, 2022). Partially, PCRs also actively participate in the market for credit information by providing lenders – especially banks – with data on debtors (Girault & Hwang, 2010). The extent to which PCRs actively participate in the credit information market depends primarily on the institutional and legal arrangements that support their activities, as well as the specific characteristics of local credit markets (Girault & Hwang, 2010). However, due to the essential gatekeeping function of PCRs, the inherent risk of financial exclusion arises if the sales and transaction history data used by the PCRs is incomplete or incorrect (Frost et al., 2019). Therefore, public credit registries must meet stricter regulatory requirements, leading to a higher degree of fairness in credit scoring (Girault & Hwang, 2010).

In contrast to public credit registries, private credit bureaus (PCB´s) are for-profit, commercially independent information exchange systems that provide services – i.e., credit scoring of borrowers – to financial institutions and other commercial enterprises to address information asymmetry between lenders and borrowers (Miller, 2003). In general, information about PCB´s has been associated with better prediction of a company's or consumer´s probability of default, making financial institutions more resilient (Kallberg & Udell, 2003). However, as private enterprises, PCBs are only subject to domestic legislation and regulation and are therefore not accountable to public bodies, central banks, or other financial service regulatory authorities (Ferretti, 2015). However, the lower regulatory requirements for PCBs lead to inherent discrimination and opaqueness, as the PCB's internal scoring approach incorporates both public and private information in the credit assessment process (Hiller & Jones, 2022). The frequency of incorrect data is particularly evident. A representative sample of 1,000 credit reports from US consumers found that 26 percent contained at least one data item with a material error (Smith et al., 2013). In addition, scoring methods used by PCBs do not consider local economic conditions, such as regional recessions, which may affect credit history in a specific area but do not affect future repayment behavior (Avery et al., 2000).

In addition, mobile financial service providers are disrupting the market by offering financial services based on alternative scoring approaches to customers who cannot access traditional financial services because of insufficient credit history (Shaikh & Karjaluoto, 2015). As a result of the unavailability of financial data, mobile financial service providers as well as PCBs, are increasingly focusing on alternative data to calculate credit scores (Hiller & Jones, 2022). For this purpose, financial technology companies in particular create a digital personality profile by integrating digital footprints and social media activities into the final credit score (Roa et al., 2021). For this purpose, data on online friendships (Lin et al., 2013), operating systems (Djeundje et al., 2021), time of purchase (Baer et al., 2012), and email correspondence in particular are utilized as input variables for alternative credit scoring (Berg et al., 2020). However, the challenge remains in analyzing various data and integrating the results into the credit scoring process (Kyeong et al., 2021).

Currently, there is no standard classification for organizations assessing creditworthiness, with the result commonly referred to as a credit score (Wong & Dobson, 2019). However, different objectives and divergent data pools are used to create the credit score of the respective organizations (Yu et al., 2015). This circumstance creates a discrepancy in the terminology itself; hence, it is essential to classify the data used and to differentiate more precisely between the term credit score and the associated companies (Consumer Financial Protection Bureau, 2022). For instance, a traditional credit score primarily describes the result of an individual credit-relevant history analysis, which is used to predict the creditworthiness of an individual in the future (Chatterjee et al., 2020). Furthermore, a credit score can compromise – in addition to historical payment behavior- also current payment punctuality when considering, for example, the payment of monthly cellphone bills (Hayashi, 2019).

In contrast, the result of the creditworthiness assessment by financial technology companies – the assessment is mainly based on non-traditional data – is also referred as credit score (Cornelli et al., 2020). However, the literature widely uses the term "credit score" to convey the outcome of an individual's creditworthiness assessment (Wong & Dobson, 2019). Therefore, this situation highlights the need for further differentiation in the general use of the term credit score, which involves grouping the responsible parties and the data used in the respective credit score model.

Classification of Data in Credit Scoring

Regardless of the institution conducting the assessment, credit scoring includes both personal and financial data (Andrus, 2016). In modern credit scoring – as established by Bill Fair and Earl Isaac in 1958 with the introduction of the first scoring system, “Credit Application Scoring Algorithms” (Smith, 2011), primarily financial data is included (Kiviat, 2019). This includes data on credit repayment, credit cards, credit limits, and number of loans (Kiviat, 2019). Since then, the data used has been continuously expanded, and today also includes information on “bill payment records from utilities, cell phone providers, landlords, and cable television companies” (Kiviat, 2019, p. 37).

By contrast, mobile financial service providers mainly use alternative data that relate to an individual's socio-economic status, lifestyle, and habits, such as social media connections, college major, and occupation; cell phone and computer use to calculate credit scores (Kiviat, 2019). This continuous expansion of the data used as well as the application of state-of-the-art technology – especially artificial intelligence – results in increasingly opaque credit rating models (Brainard 2018).

In order to address this increasing complexity – it is becoming increasingly difficult, if not impossible, to understand why a model has produced the result it has (Brainard 2018) – the data usage of credit bureaus in 17 different countries was investigated as part of this research project, and a classification framework for credit bureaus based on data utilization was developed as shown in Table 1.

As a basis for the classification of the data, reference was made to the Regulation (EU 2016/679) of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) and to the Guidelines on Loan Origination and Monitoring EBA/GL/2020/06 of the European Banking Authority.

Table 1

Number of Data Categories per Credit Bureau

Company	Country of Origin	Financial Data Categories	Personal Data Categories
Schufa	Germany	11	5
Creditreform Boniversum	Germany	5	10
Experian	Global	6	9
Credit Kudos “Open-Banking-Technology”	United Kingdom	5	5
TransUnion	USA │Global	5	2
Equifax	USA │Global	7	3
KBIJ	Indonesia	7	0
Acura Labs	Indonesia	2	7
Credit Bureau Singapore	Singapore	8	4
Trusting Social	Vietnam	1	4
CreditVida	India	1	7
Joint Credit Information Center (JCIC)	Taiwan	14	6
Korean Credit Bureau (KCB)	South Korea	8	5
Metropol	Kenya	6	5
Creditinfo	Kenya	3	5
TALA	Kenya	5	26
CRC Credit Bureau Limited	Nigeria	4	5
CreditRegistry	Nigeria	3	6
I-Score	Egypt	5	5
XDS Data Ghana	Ghana	5	4
Círculo de Crédito	Mexico	7	3
Dicom – Equifax	Chile	4	6
Quod	Brazil	8	0
CSR (Credit Information System) Brazil	Brazil	10	1

METHODOLOGY

Current research in credit scoring is predominantly focused on applying new statistical techniques, adopting new technology, and utilizing additional or alternative data for calculating credit scores (Onay & Öztürk, 2018). This illustrates that optimizing predictive accuracy – whether by improving statistical methods or applying state-of-the-art technology – remains the primary goal of academic research. Interestingly, the influence of credit bureaus on credit ratings receives little to no attention in current academic research. In particular, this may be a result of extreme difficulties in obtaining internal data from these companies, as the credit scoring process is an essential part of the business model and is therefore subject to commercial confidentiality (Bundesgerichtshof, 2014). However, the Administrative Court of Wiesbaden has currently submitted a request to the European Court of Justice regarding the interpretation of the EU General Data Protection Regulation with regard to automated individual decision-making (Article 22. para. 1.) in the context of credit scoring (VG Wiesbaden, 2021b). The ECJ's decision is still pending.

Data Collection and Analysis

Precisely for this reason, the underlying data was collected from various sources – especially from the websites of the credit bureaus, but also from academic journals, current legislation, governmental organizations, and related websites. In particular, information on relevant data for the calculation of each credit bureau´s credit score was collected. The identified data was subsequently subdivided into financial data – which include, for example, data on bank accounts, credit cards, installment loans, or mobile leasing – and personal data – which include, among others, data on demographics (name, age, gender, etc.), criminal record information, social media data, telecommunication data, and geographical data.

Classification of Data

Subsequently, the data – both financial and personal – were categorized into "data necessary for the calculation and exact personal assignment of a credit score "and" additional data for the calculation and exact personal assignment of a credit score." Separating the data into "data necessary for the calculation and exact personal assignment of a credit score" and "additional data for the calculation and exact personal assignment of a credit score" was difficult, as there is no standardized method for calculating credit scores in the literature or in practice (Hertza, 2018). Therefore, the current legislation and regulatory requirements of the European Union were taken as a reference to derive the minimum requirements regarding the data usage in credit score calculation.

Accordingly, financial data is categorized into two groups: "data necessary for the calculation and precise personal assignment of a credit score" and "additional data for the calculation and precise personal assignment of a credit score." This is in accordance with the existing guideline (EBA/GL/2020/06) on Loan Origination and Monitoring issued by the European Banking Authority (EBA). Chapter 5.2 "Assessment of Borrower's Creditworthiness," of the EBA Guideline defines the general criteria of a creditworthiness assessment in paragraph 98, which should be included in a creditworthiness assessment by banks. These are “servicing obligations, their remaining duration, their interest rates and the outstanding amounts, and repayment behavior, e.g., evidence of any missed payments and their circumstances, as well as directly relevant taxes and insurance” (European Banking Authority, 2020, p. 35). The EBA Guideline classifies all data mentioned as "data necessary for the calculation and exact personal assignment of a credit score" (abbreviated as "Necessary Data" in the Framework). Hence, data exceeding the requirements of the EBA are classified as "additional data for the calculation and exact personal assignment of a credit score" (abbreviated as "Additional Data" in the framework), as these are not necessary – at least for regulatory reasons.

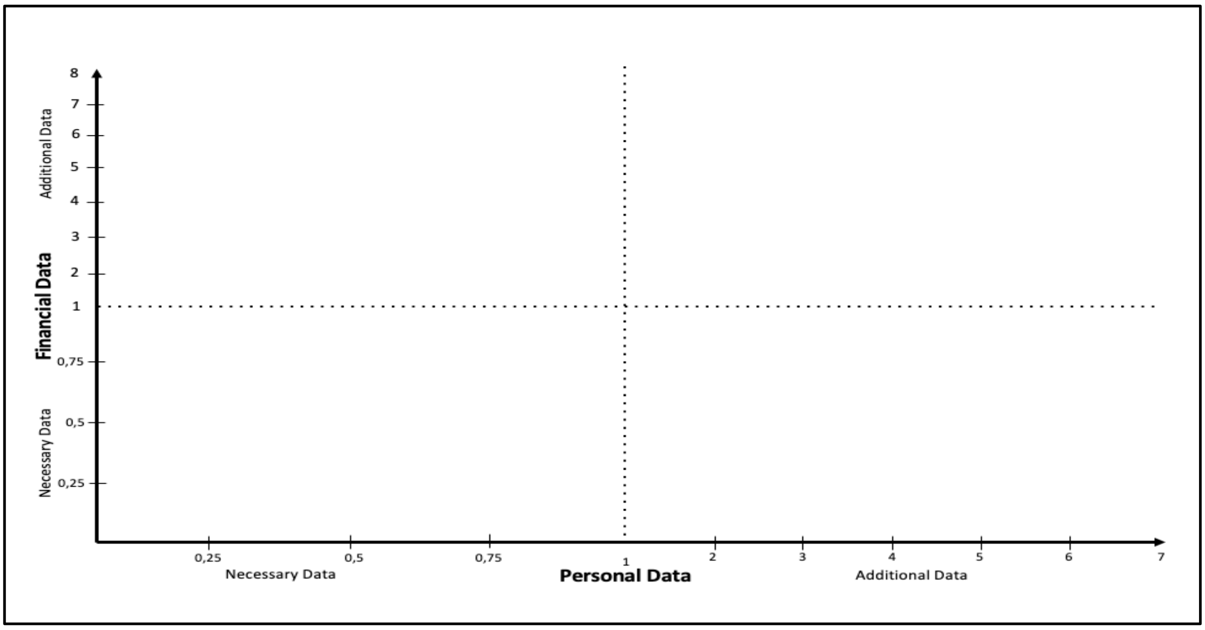
Conversely, the Regulation (EU 2016/679) of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and on the repeal of Directive 95/46/EC (General Data Protection Regulation) establishes the distinction between "data necessary for the calculation and exact personal assignment of a credit score" and "additional data for the calculation and exact personal assignment of a credit score." Article 9, paragraph 1 of the regulation lists particularly sensitive data, including “data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation” (European Union, 2016, p. 38). Therefore, the framework classifies all the data mentioned in the regulation as "additional data for the calculation and accurate personal assignment of a credit score" (abbreviated as "additional data"). Consequently, the Framework characterizes any other personal data that pertains to an identified or identifiable natural person as "data necessary for the calculation and exact personal assignment of a credit score" (abbreviated as "Necessary Data").

Development of the Framework

Hence, the position of each credit bureau indicates the relative use of “Additional Data” compared to “Necessary Data”. The general framework displays the ratio of "additional" to "necessary" personal data on the X-axis and the ratio of "additional" to "necessary" financial data on the Y-axis. The intervals of the axes – for necessary data in 0.25 intervals and for additional data in intervals of 1 – are chosen to represent the ratio in the most accurate manner.

Figure 1

General Framework for the Classification of Credit Bureaus based on Data Utilization



The center of the axis is 1, as the ratio is illustrated here. Accordingly, the lower left quadrant illustrates the ratios < 1 of both financial and personal data. Thus, more necessary than additional data – both for financial and personal data – is utilized in the calculation of credit scores. Accordingly, the lower right quadrant illustrates the ratios < 1 of financial data and > 1 of personal data. The upper left quadrant illustrates the ratio > 1 of financial data and the ratio < 1 of personal data. The upper right quadrant illustrates ratios > 1 for both financial and personal data.

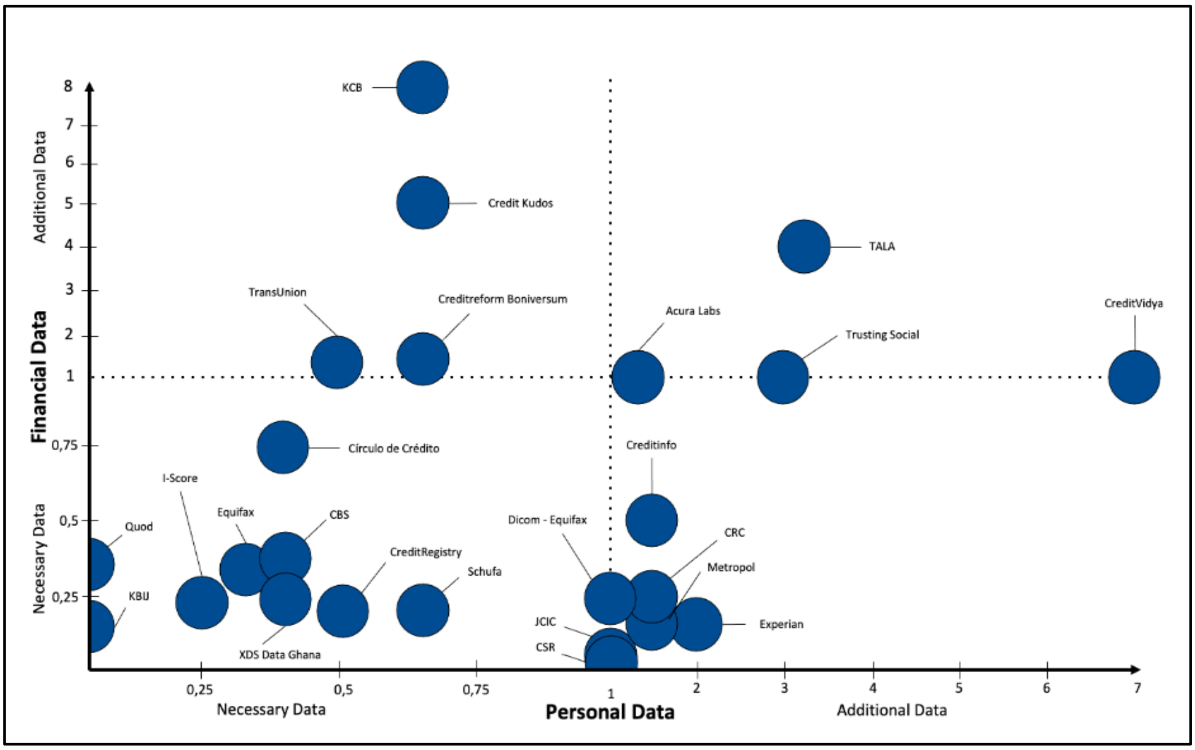
As the classification is based on the relative distribution of “Necessary Data” and “Additional Data”, a value = of 1 – this applies to both financial and personal data – describes a balanced distribution of “Necessary Data” and “Additional Data.” A value of < 1 indicates a smaller distribution of “Additional Data” compared to “Necessary Data.” Similarly, a value > 1 indicates a greater distribution of “Additional Data” compared to “Necessary Data.”

RESULTS AND DISCUSSION

As being shown in Figure 2, it clearly demonstrates the significant variations in data utilization across various credit bureaus. However, two agglomerations are noticeable: firstly, the large number of credit bureaus relying primarily on necessary data – both, financial and personal data – to calculate the credit score, and secondly, an agglomeration of credit bureaus increasingly utilizing additional personal data. In addition, some credit bureaus also use additional financial data or – such as Tala – mainly additional personal and financial data. Interestingly, no clear patterns are apparent in terms of data utilization, irrespective of the type of credit bureau or geographical location of operation.

Figure 2

Classification of Credit Bureaus based on Data Utilization



As the analysis illustrates, comparing the credit bureaus is – particularly due to the differences in data utilization – difficult, if not impossible, as the fundamentals of the assessment of creditworthiness vary significantly. As a result, the comparison of credit scores is also limited. Hence, it is reasonable to introduce explicit terminology to differentiate between credit scores with different data emphases. Accordingly, the categorization – and therefore the explicit terminology – of the credit score depends on the data utilization of the credit bureau. Consequently, it is possible to differentiate between four different types of credit scores:

- Credit Scores on Financial Data (CSF) are defined by the predominant utilization of additional financial data. Accordingly, the ratio of additional to necessary financial data is > 1, whereas the ratio of additional to necessary personal data is < 1.
- Credit Scores on Personal Data (CSP) are defined by the predominant utilization of additional personal data. Accordingly, the ratio of additional to necessary personal data is > 1, whereas the ratio of additional to necessary financial data is < 1.
- Credit Scores on Balanced Data (CSB) are defined by the predominant utilization of necessary financial and personal data. Accordingly, the ratios of additional to necessary for financial and personal data are < 1.
- Credit Scores on Mass Data (CSM) are defined by the utilization of mass data. Accordingly, the ratios of additional to necessary for financial and personal data are > 1.

However, due to the inaccuracy of traditional credit scores in reflecting a borrower's ability to repay, researchers are increasingly exploring the use of real-time financial data (Toh, 2023). Therefore, it might be useful to prefix the different credit scores to indicate the type of financial data – historical or real-time – utilized for calculating the credit score. Accordingly, credit scores – regardless of whether CSF, CSP, CSB, or CSM – should be prefixed with HT for historical data or RT for real-time data, depending on the focus of financial data utilization. This further differentiation becomes particularly necessary, once the utilization of real-time financial data becomes the norm. Accordingly, the following distinctions are possible.

Table 2

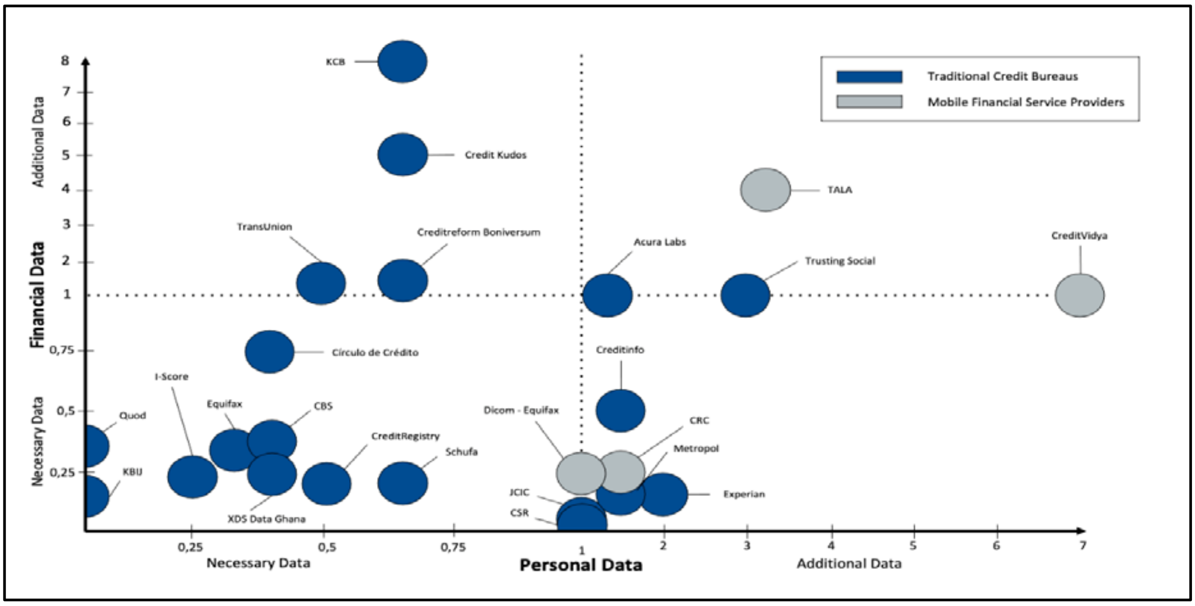
Prefix to Distinguish Credit Scores based on Type of Financial Data

Credit Scores on Historic Financial Data	Credit Scores on Real-Time Financial Data
$_{HT}$CSF	$_{RT}$CSF
$_{HT}$CSP	$_{RT}$CSP
$_{HT}$CSB	$_{RT}$CSB
$_{HT}$CSM	$_{RT}$CSM

Based on the differences of the four different types of credit scores, relevant challenges exist for borrowers and lenders. In particular, when comparing the three major credit bureaus in the US – Experian, Equifax, and TransUnion – it demonstrates that one consumer´s creditworthiness is reported differently. Thus, based on the proposed terminology, Experian reports a Credit Score on Personal Data (CSP), Equifax reports a Credit Score on Balanced Data (CSB), and TransUnion reports a Credit Score on Financial Data (CSF). Therefore, it is not surprising that the actual credit scores of the US credit bureaus differ from each other – sometimes significantly.

Figure 3

Classification of Credit Bureaus based on Data Utilization by Type

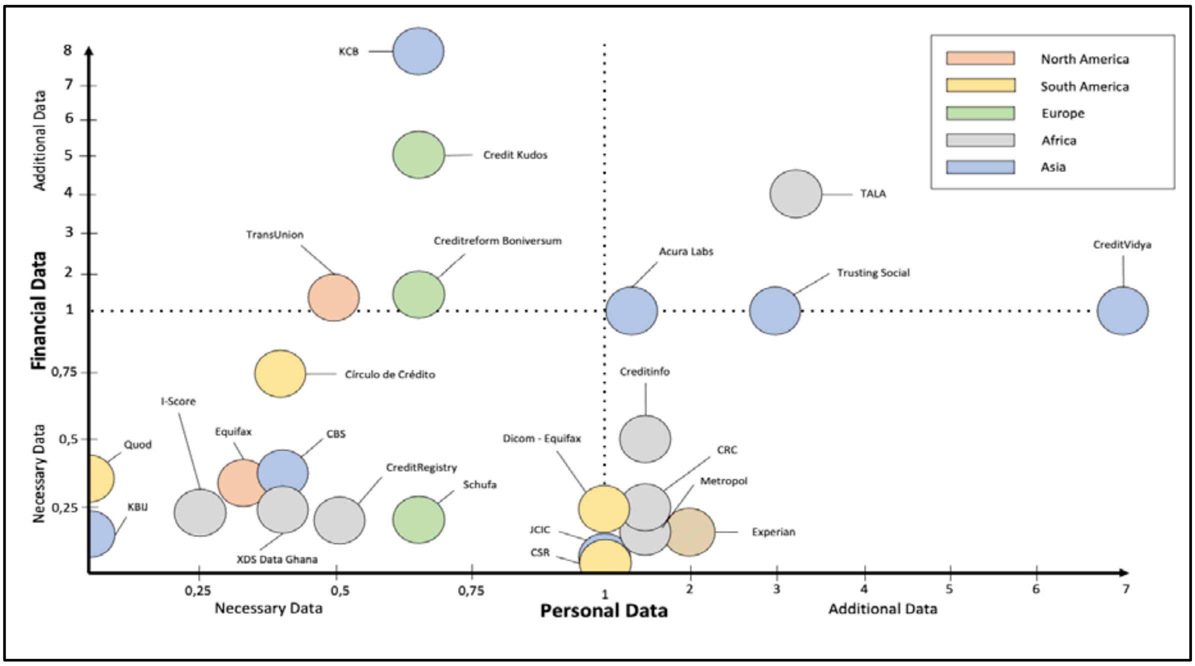


Although it is evident that mobile financial service providers MFSPs considerably rely on additional personal and financial data as shown in Figure 3, this pattern of data utilization is not exclusive to MFSPs. Additionally, traditional credit bureaus are increasingly utilizing more personal data. Therefore, a clear distinction – at least based on data utilization – from traditional credit bureaus (CB) is not possible. As shown in Figure 4, when considering the geographic location of operation of the CBs, noticeably European, North, and South American institutions rely significantly less on additional personal data – except Experian – compared to institutions operating in Asia and Africa. One reason could be the stricter data protection guidelines resulting from the General Data Protection Regulation of 2016 in the European Union and the California Consumer Privacy Act of 2018 in the USA, which significantly impact the processing and utilization of personal data.

Furthermore, the availability of data – especially financial data – in less industrialized regions presents a challenge, as both the financial infrastructure is insufficiently developed, and large parts of the population have no access to the traditional financial markets. While FinTechs have improved access to financial markets for customers who do not have access to traditional financial services, this has included replacing financial data with non-traditional – primarily additional personal – data in credit scoring.

Figure 4

Classification of Credit Bureaus based on Data Utilization by Geographic Location of Operation



The application of additional personal data in credit scoring, however, remains highly controversial and involves various risks. First, the analysis of the various data and the integration of the results into the credit assessment process still pose a significant challenge (Kyeong et al., 2021). In this case, it must be emphasized that the creditworthiness may therefore be based on possibly random correlations derived from an accumulation of vast amounts of data. Hence, many FinTechs – such as Tala and CreditVidya – rely on artificial intelligence to analyze and process the vast amounts of data. Problematically, the high complexity of artificial intelligence only allows for a limited comprehension of its reasoning (Hamon et al., 2020). Therefore, a considerable risk of incorporating prejudices or deriving correlations containing discriminatory tendencies, exists.

FUTURE AVENUES OF RESEARCH

The analysis of data utilization in credit scoring and the resulting classification framework of credit bureaus highlights several avenues for future research. Specifically, the categorization of credit scores according to data usage and the implementation of innovative technologies such as artificial intelligence are generating additional inquiries. These inquiries primarily pertain to ethical and legal concerns, as well as the methods used and the comparability of ongoing research. Furthermore, it is important to consider the future significance of credit scoring, particularly in light of ongoing technology advancements and the growing use of real-time data for evaluating creditworthiness.

The analysis of the data utilization in credit scoring has demonstrated that significant differences exist in the data employed to calculate credit scores between the credit bureaus. Of course, these differences raise legitimate concerns regarding the comparability of credit scores from various credit bureaus. Are credit scores with different databases comparable? How significant are the differences in terms of the credit rating's informative value? These possible differences are particularly important to consider between traditional credit scores – which are mainly calculated on the basis of financial data – and alternative credit scores – primarily calculated based on alternative and personal data.

The general necessity of credit scoring must also be questioned in light of the current technological advancements. The European Payment Services Directive 2 (PSD2) regulates account access for third-party providers, establishing a common method for credit assessment in many EU and US countries (Saia et al., 2021). This method of credit assessment – also referred to as Open Banking – is a veritable alternative to traditional credit scoring.

The different data bases of the considered credit bureaus indicate a potential problem regarding the comparability of credit scores. The differences in the data basis – in particular when comparing credit scores of traditional credit bureaus with credit scores of mobile financial service providers – are not the only significant difference. Also, the type of financial data – historical or real-time – may possibly make a difference. Considering new regulatory approaches and the increasingly advancing technology for the use of real-time data, it is necessary to analyze any potential variations in detail.

Ethical Considerations

The analysis demonstrates different approaches to calculating credit scores, especially concerning the underlying data. However, applying personal and financial data to determine an individual's creditworthiness raises several ethical considerations and expands the scope of research. First, simply collecting a large amount of data is misleading because not all sets of information correlate with a person's creditworthiness. CBs should limit their activities to collecting only the kind of data that is related to an individual’s financial strength or payment behavior. Consequently, the correlation between data and creditworthiness is the subject of research by academics and commercial banks. For example, is a person's sexual orientation a pertinent indicator?

Second, provided this is the case, is utilizing such information in the societal interest? Finding a broad societal consensus on a blacklist of data that credit scores should not refer to is necessary. Similarly, there is information that should not be collected in the first place. There is a need for academic research on which personal or financial data should be utilized for credit scoring or not due to ethical concerns. It is important to examine to what extent the credit assessment justifies an intrusion into the most personal areas of an individual.

Third, the employment of artificial intelligence must be viewed critically. On the other hand it allows the combination and bundling of relevant information. Conversely, the more sophisticated an algorithm is, the less transparent the compilation of the credit score becomes~~.~~ Because an algorithm is only as good as the data it receives, prejudices and manipulations can easily find their way into the scores. Moreover, a major ethical issue concerning responsibility arises when a person with a faulty score cannot blame anyone anymore (Thorhauer, 2020). Such claims fizzle out because a hypercomplex algorithm is in charge, not a real person.

Fourth, mobile financial service providers can be viewed critically because of the type of data they use. However, in emerging and developing countries, alternative credit scores are the only possibility for currently unbanked consumers to access credit (Kumar et al., 2021). Still, providers are not interested in changing the current state – e.g., the lack of access to traditional financial services for consumers in emerging and developing countries – as their business model is based on the status quo. Consumers must weigh the protection of their personal rights against the potential for economic participation, given the scarcity of credit scoring alternatives.

Legal Considerations

The analysis demonstrates different approaches to calculating credit scores, especially concerning the underlying data. Particularly, the stricter data protection regulations of the European Union (GDPR) and the United States (CCPA), along with the growing use of novel technologies like AI, necessitate further consideration regarding legal aspects.

First, the consequences – especially after court rulings on the interpretation of data protection laws – must be examined with respect to credit scoring. As previously stated, the Administrative Court of Wiesbaden has submitted a request to the European Court of Justice regarding the interpretation of the EU General Data Protection Regulation relating to automated individual decision-making (Article 22. para. 1) in the context of credit scoring (VG Wiesbaden, 2021a).

On December 7th, 2023, the EU's highest court initially ruled that SCHUFA's scoring violates the European General Data Protection Regulation if banks consider the credit score to be essential for contractual decisions. The GDPR prohibits making decisions solely based on automatically processed data. However, the ECJ clarified that exceptional cases may permit this practice. For instance, the national legislator could implement an exceptional clause. The Federal Data Protection Act has such a provision in Germany.

The ECJ judgement has not concluded the legal proceedings. The Administrative Court of Wiesbaden must now consider whether the exceptional provision in the Federal Data Protection Act is lawful at all. Should the Administrative Court conclude that the exceptional provision violates European Law, credit scoring in Germany would be unlawful if companies base contractual decisions solely on credit scores.

In addition, the Administrative Court of Wiesbaden has also submitted a request to the ECJ concerning the interpretation of Articles 77 and 78 of the General Data Protection Regulation of the EU concerning the permissibility of the storage of personal data from a public register by credit bureaus without any reason (VG Wiesbaden, 2021a). This underscores the need for a more detailed examination of the implications of legal risks on credit scoring.

CONCLUSION

The primary contribution of this paper is first, the classification framework for credit bureaus based on data utilization and second, the suggested terminology to differentiate between credit scores with different data emphases. Current credit scoring research has failed to recognize and consider the significant differences in data utilization, as demonstrated by comparing credit scores of traditional credit bureaus to those of mobile financial service providers. This paper aims to take a first step towards eliminating this deficiency by suggesting a terminology to differentiate between credit scores in the first place. Further, it outlines future research fields related to using data for credit rating.

Furthermore, the analysis of existing literature on this topic has demonstrated that the influence of CBs on credit score calculations has received almost no attention so far. CBs have a considerable influence in this regard, both through the choice of calculation method – in many cases, they are developed in-house –and through the choice of data used to calculate the credit score. Despite this significant influence, research still views credit bureaus as a "black box" and pays little attention to them. This proposed differentiation of credit scores – as well as of credit bureaus – provides new avenues of research.

Upcoming studies should first clarify whether credit scores with different databases are comparable. In this context, the influence of CBs on credit scoring has to be a central part of future research. Furthermore, credit scores that are different depending on how data is used, and the use of new technologies (such as artificial intelligence) raises further questions, particularly relating to ethical and legal considerations, as well as with respect to methodological approaches and the comparability of current research. Moreover, the future relevance of credit scoring has to be discussed, especially to about current technological developments and the increasing application of real-time data to assess creditworthiness. Furthermore, are CBs still necessary in the 21st century? The fundamental task of research is to increase the quality of knowledge for the good of society. Hence, academia may take the lead in actively shaping the future of credit scoring, considering social, legal, and technological developments.

ACKNOWLEDGEMENT

This research did not receive any specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

Ali, M. S. B. (2022). Credit bureaus, corruption and banking stability. *Economic Systems*, *46*(3), 1–8

Amat, O., Manini, R., & Renart, M. A. (2017). Credit concession through credit scoring: Analysis and application proposal. *Intangible Capital*, *13*(1), 51–70.

Andrus, M. T. (2016). The right to be forgotten in America: Have search engines inadvertently become consumer reporting agencies? *Business Law Today*, *1*(1), 1.

Asher, J. (1994). Look what credit scoring can do now. American Bankers Association. *ABA Banking Journal*, *86*(5), 51.

Avery, R. B., Bostic, R. W., Calem, P. S., & Canner, G.B. (2000), Credit scoring: Statistical issues and evidence from credit-bureau files. *Real Estate Economics*, *28*, 523-547.

Baer, T., Goland, T., & Schiff, R. (2012). *New credit-risk models for the unbanked*. McKinsey & Company.

Berg, T., Burg, V., Gombovic, A., & Puri, M. (2020). On the rise of Fintech: Credit scoring using digital footprints. *The Review of Financial Studies*, *33*(7), 2845–2897.

Blechmann, J. G. (2016). Mobile credit in Kenya and Tanzania: Emerging regulatory challenges in consumer protection, credit reporting and use of customer transactional data. *The African Journal of Information and Communication*, *17*, 61–88.

Brainard, L. (2018). *What are we learning about artificial intelligence in financial services?* A speech at Fintech and the new financial landscape. Philadelphia, Pennsylvania (No. 1021). Board of Governors of the Federal Reserve System (US).

Bundesgerichtshof. (2014). *VI ZR 156/13.* http://juris.bundesgerichtshof.de/cgi-bin/rechtsprechung/document.py?Gericht=bgh&Art=pm&sid=99f39ef8e8587ba87abfaf9e6104 0eb8&nr=66910&linked=urt&Blank=1&file=dokument.pdf

Chatterjee, S., Corbae, D., Dempsey, K. P., & Ríos-Rull, J. V. (2020). A quantitative theory of the credit score. *National Bureau of Economic Research*, *91*(5), 1803-1840

Consumer Financial Protection Bureau. (2022). *What is a credit score?* Federal Reserve, Consumer Financial Protection Bureau.

Consumer Financial Protection Bureau. (2023). *List of consumer reporting agencies.* Federal Reserve, Consumer Financial Protection Bureau.

Cornelli, G., Frost, J., Gambacorta, L., Rau, P. R., Wardrop, R., & Ziegler, T. (2020). *Fintech and big tech credit: A new database*. https://papers.ssrn.com/sol3/papers.cfm? abstract_id=3707437

Djeundje, V. B., Crook, J., Calabrese, R., & Hamid, M. (2021). Enhancing credit scoring with alternative data. *Expert Systems with Applications*, *163*, 113766.

European Banking Authority. (2020). *Guidelines on loan origination and monitoring*. European Banking Authority.

European Union. (2016). Regulation 2016/679. *Regulation (EU 2016/679) of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).* https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

Ferretti, F. (2013). The legal framework of consumer credit bureaus and credit scoring in the European Union: pitfalls and challenges over indebtedness, responsible lending, market integration, and fundamental rights. *Suffolk University Law Review*, *46*(3), 791–828.

Ferretti, F. (2015). Credit bureaus between risk-management, creditworthiness assessment and prudential supervision. *EUI Department of Law Research Paper* (2015/20).

Finlay, S. (2012). Credit scoring, response modeling, and insurance rating: A practical guide to forecasting consumer behavior (2nd ed.). Palgrave Macmillan UK.

Francis, E., Blumenstock, J., & Robinson, J. (2017). *Digital credit: A snapshot of the current landscape and open research questions.* UC Berkeley: Center for Effective Global Action. https://escholarship.org/uc/item/88r1j7sz

Frost, J., Gambacorta, L., Huang, Y., Shin, H. S., & Zbinden, P. (2019). BigTech and the changing structure of financial intermediation. *Economic Policy*, *34*(100), 761-799.

Girault, M. G., & Hwang, J. C. (2010). Public credit registries as a tool for bank regulation and supervision. *World Bank Policy Research Working Paper* (5489).

Golden, L. L., Brockett, P. L., Ai, J., & Kellison, B. (2016). Empirical evidence on the use of credit scoring for predicting insurance losses with psycho-social and biochemical explanations. *North American Actuarial Journal*, *20*(3), 233–251.

Hamon, R., Junklewitz, H., & Sanchez, I. (2020). Robustness and *explainability of artificial intelligenc*e. *From technical to policy solutions*. JRC Technical Report. https://www.google.de/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjHpsSMuMqAAxVf2QIHHYEtDrEQFnoECBMQAQ&url=https%3A%2F%2Fpublications.jrc.ec.europa.eu%2Frepository%2Fbitstream%2FJRC119336%2Fdpad_report.pdf&usg=AOvVaw1JlIaGHrx9Tl2KGtptWFWK&opi=89978449

Hartwig, R. P., & Wilkinson, C. (2003). The use of credit information in personal lines insurance underwriting. *Insurance Issues Series*, *1*(2), 1–20

Hayashi, Y. (2019, August 9). *Banking & finance: New credit criteria raise fairness issue*. Wall Street Journal, 2019. https://www.proquest.com/newspapers/banking-amp-finance-new-credit-criteria-raise/docview/2269852016/se-2

Hertza, V. A. (2018). Fighting unfair classifications in credit reporting: Should the United States adopt GDPR-inspired rights in regulating consumer credit. *New York University Law Review*, *93*(1), 1707–1743.

Hiller, J. S., & Jones, L. S. (2022). Who's keeping score? Oversight of changing consumer credit infrastructure. *American Business Law Journal*, *59*(1), 61–121.

Kozodoi, N., Lessmann, S., Papakonstantinou, K., Gatsoulis, Y., & Baesens, B. (2019). A multi-objective approach for profit-driven feature selection in credit scoring. *Decision Support Systems*, *120*, 106–117.

Kallberg, J. G., & Udell, G. F. (2003). The value of private sector business credit information sharing: The US case. *Journal of Banking & Finance*, *27*(3), 449–469.

Kiviat, B. (2019). Credit scoring in the United States. *Economic sociology the European Electronic Newsletter*, *21*(1), 33-42.

Kumar, A., Sharma, S., & Mahdavi, M. (2021). Machine learning (ML) technologies for digital credit scoring in rural finance: A literature review. *Risks*, *9*(11), 192.

Kusi, B. A., Agbloyor, E. K., Ansah-Adu, K., & Gyeke-Dako, A. (2017). Bank credit risk and credit information sharing in Africa: Does credit information sharing institutions and context matter? *Research in International Business and Finance*, *42*, 1123–1136.

Kyeong, S., Kim, D., & Shin, J. (2021). Can system log data enhance the performance of credit scoring?—Evidence from an internet bank in Korea. *Sustainability*, *14*(1), 130.

Lin, M., Prabhala, N. R., & Viswanathan, S. (2013). Judging borrowers by the company they keep: Friendship networks and information asymmetry in online peer-to-peer lending. *Management Science*, *59*(1), 17–35.

Markov, A., Seleznyova, Z., & Lapshin, V. (2022). Credit scoring methods: Latest trends and points to consider. *Journal of Finance and Data Science*, *8*, 180–201.

McCann, F., & McIndoe-Calder, T. (2015). Firm size, credit scoring accuracy and banks' production of soft information. *Applied Economics*, *47*(33), 3594–3611.

McIntosh, C., Sadoulet, E., Buck, S., & Rosada, T. (2013). Reputation in a public goods game: Taking the design of credit bureaus to the lab. *Journal of Economic Behavior & Organization*, *95*, 270–285.

Miller, M. J. (2003). *Credit reporting systems and the international economy*. The MIT Press.

Njathi, A. (2019). The effects of mobile lending apps on bottom of the pyramid consumers in Kenya: A case of Tala FinTech Company. *Journal of African Interdisciplinary Studies*, *3*(8), 112–126.

Óskarsdóttir, M., Bravo, C [Cristián], Sarraute, C., Baesens, B., & Vanthienen, J. (2018). *Credit scoring for good: Enhancing financial inclusion with smartphone-based microlending*. Thirty Ninth International Conference on Information Systems.

Óskarsdóttir, M., Bravo, C [Cristián], Sarraute, C., Vanthienen, J., & Baesens, B. (2019). The value of big data for credit scoring: Enhancing financial inclusion using mobile phone data and social network analytics. *Applied Soft Computing*, *74*, 26–39.

Roa, L., Correa-Bahnsen, A., Suarez, G., Cortés-Tejada, F., Luque, M. A., & Bravo, C [C.] (2021). Super-app behavioral patterns in credit risk models: Financial, statistical and regulatory implications. *Expert Systems with Applications*, *169*, 114486.

Saia, R., Giuliani, A., Pompianu, L., & Carta, S. (2021). *From payment services directive 2 (PSD2) to credit scoring: A case study on an Italian Banking Institution*. In KDIR (pp. 164-171). https://www.scitepress.org/PublishedPapers/2021/106532/106532.pdf

Shaikh, A. A., & Karjaluoto, H. (2015). Mobile banking adoption: A literature review. *Telematics and Informatics*, *32*(1), 129–142.

Siddiqi, N. (2017). *Intelligent credit scoring: Building and implementing better credit risk scorecards* (2nd ed.). John Wiley & Sons.

Smith, B. C. (2011). Stability in consumer credit scores: Level and direction of FICO score drift as a precursor to mortgage default and prepayment. *Journal of Housing Economics*, *20*(4), 285-298.

Smith, L. D., Staten, M., Eyssell, T., Karig, M., Freeborn, B. A., & Golden, A. (2013). Accuracy of information maintained by US credit bureaus: Frequency of errors and effects on consumers' credit scores. *Journal of Consumer Affairs*, *47*, 588-601.

Spader, J. S. (2010). Beyond disparate impact: Risk-based pricing and disparity in consumer credit history scores. *The Review of Black Political Economy*, *37*(2), 61–78.

Thorhauer, Y. (2020). Enlightenment in the digital age – Implications for applied ethics. In Y. Thorhauer & C. A. Kexel (Eds.), *Facetten der digitalisierung: Chancen und herausforderungen für mensch und management* (pp 7-29). Springer Gabler.

Toh, Y. L. (2023). Addressing traditional credit scores as a barrier to accessing affordable credit. *Economic Review*, *108*(3), 1-22.

VG Wiesbaden. (2021a). *6 K 788/20.WI.* Court Wiesbaden.

VG Wiesbaden. (2021b). *6 K 441/21.WI.* Court Wiesbaden

Wei, Y., Yildrim, P., van den Bulte, C., & Dellarocas, C. (2016). Credit scoring with social network data. *Marketing Science*, *35*(2), 234–258.

Wong, K. L. X., & Dobson, A. S. (2019). We're just data: Exploring China's social credit system in relation to digital platform ratings cultures in Westernised democracies. *Global Media and China*, *4*(2), 220–232.

Yu, L., Li, X., Tang, L., Zhang, Z., & Kou, G. (2015). Social credit: A comprehensive literature review. *Financial Innovation*, *1*(1), 1–18.