watchit
Expertenforum

3. Veröffentlichung

Algorithmic Credit Scoring in the USA, Europe and China: Between Regulation and 'Petrification'

Udo Milkau

DHBW
Duale Hochschule
Baden-Württemberg
Mosbach

Impressum

DHBW Mosbach
Lohrtalweg 10
74821 Mosbach

www.mosbach.dhbw.de/watchit
www.digital-banking-studieren.de

Algorithmic Credit Scoring in the USA, Europe and China:
Between Regulation and 'Petrification'
Contribution to the Virtual Workshop on AI & FINANCE,
Oct. 29 / Nov. 12, 2021 hosted by Katja Langenbucher
and the associated ZEVEDI investigators,
Goethe University Frankfurt

von Udo Milkau

ISBN 978-3-943656-16-9

Herausgeber:
Jens Saffenreuther

Mosbach, im Dezember 2021

Algorithmic Credit Scoring in the USA, Europe and China: Between Regulation and ‘Petrification’

Udo Milkau, updated version, 13.11.2021

Contribution to the Virtual Workshop on AI & FINANCE, Oct. 29 / Nov. 12, 2021
hosted by Katja Langenbucher and the associated ZEVEDI investigators,
Goethe University Frankfurt

Abstract

The proposal of the European Commission for an ‘Artificial Intelligence Act’ of April 21, 2021, made a categorial classification of any kind of algorithmic credit scoring as ‘high risk system’ - independent whether made by conventional statistical approaches, tradition artificial intelligence or even so-called deep learning. This proposal is a last step in a chain of cross-industry regulation in Europe, which reveal a general scepticism about technology and a paradigm shift from individual responsibility to plain correlations as justification for regulation of digital technologies. While there was never reported discrimination in credit scoring in Europe (of course, errors and incorrect data, but without intension), the USA experienced over fifty years of systemic discrimination in credit scoring and mortgage lending. However, regulators and authorities in the U.S. are more open to possible benefits of advanced technologies and especially financial inclusion as emphasised in 2019 by an 'Interagency statement on the use of alternative data in credit underwriting' and recently with the reported pilot program to use access to bank accounts to issue credit cards to people with insufficient traditional credit scores within the Roundtable for Economic Access and Change project by the Office of the Comptroller of the Currency. A review of the current development of Credit Scoring in Europe, Germany and China allows to evaluate the different environments, but also to derive an answer to the question, whether algorithmic decision-making based on historical data could lead to a ‘petrification’ of historic social development. Whereas the approach in the USA is sector specific, Europa is going towards an omnibus approach for a regulation of algorithms without clear methodology, and in China the regime applies a wait-and-see approach, which allows innovative use of data as far as the communistic party is not challenged. As a conclusion, this paper reveals that (i) the effect of algorithmic credit scoring has to be assessed with an in-depth understanding with the socio-technological nexus of decision-making, while (ii) being based on fundamental understanding about statistical classifiers to avoid misinterpretation.

1. Introduction

For centuries - at least since the Roman 'Bankers of Puteoli' (Jones, 2006) - lenders have been facing the challenge to take the risk of uncertain re-payment by the borrowers in the future for a present premium and, consequently, make a guess about future default of borrowers to calculate this premium. This 'guess' is a statistical estimation based on one's experience about counterparty behaviour, i.e. data from the past, to forecast future development assuming a continuous pattern. Therefore, the use of some kind of 'algorithms' in the sense of an estimation of borrowers' future default after a lender provides own financial resources (i.e. credit scoring) is anything else but new. However, the recent development of analysing tremendous amounts of data with 'non-traditional' statistical classifiers (i.e. 'machine learning) for an 'automated' decision-making (i.e. running pre-programmed computer code instead of manual execution of pre-defined calculation rules) altered the outside-in perception of the credit scoring process.

Vice versa, the current outside-in perception of credit scoring can be applied as a probe for a general understanding of private autonomy and, consequently, regulation of decision-making processes in a market economy. On the one side, 'algorithmic credit scoring' is an integrated part of the overall debate of regulation of digitalisation in the 21st century, while on the other side a comparison between (i) the perspectives in the USA, Germany (representing Europe) and China reveals different approaches to regulation and institutional design in general and (ii) to sector-specific vs. cross-industry approaches across the three regions. Finally, many decision-making processes in the 21st century are encountering challenges including - unfortunately - enduring discrimination, never ended antisemitism (see: Longerich, 2021), or denial of scientific knowledge by groups claiming their own 'truth'.

Therefore, algorithmic credit scoring will be assessed from three different points of view: first, as archetype of decision-making between manual instructions and automated[1] statistical classifiers; second, as benchmark for the different concepts in the USA, Germany/Europe, and China; and third, as example for intertemporal relations, because statistical estimations of future behaviour have an impact on the social

[1] Especially in the context of artificial intelligence, there is a blurring line between three different terms: automated, autonomous, and autarch, which will be discussed later in this paper. While some 'autonomous driving systems' may be able to solve complicated technological (sic!) tasks according to their programming and without human invention during real-time control automation, no current system is even near to an autonomy in the sense of an 'own' decision with a free will independent from the pre-programmed intension of the designer. Contemporary artificial intelligence is [quote] '*able to fit a function to a collection of historical data points*' [Pearl, 2018].

context and a potential for perpetuating or 'petrification' of existing social structures into the future.

The recent public debate about 'algorithms' with the archetype of algorithmic credit reveals a multilayer socio-technological context. One recent example is the proposal of the European Commission (2021) for a new 'Artificial Intelligence Act' (AIA), which - *inter alia* - defines any[2] type of algorithmic credit scoring as a 'high-risk system[3]' [quote from recital 37]:

> *'In particular, AI systems used to evaluate the credit score or creditworthiness of natural persons should be classified as high-risk AI systems, since they determine those persons' access to financial resources or essential services such as housing, electricity, and telecommunication services. AI systems used for this purpose may lead to discrimination of persons or groups and perpetuate historical patterns of discrimination, for example based on racial or ethnic origins, disabilities, age, sexual orientation, or create new forms of discriminatory impacts.'*

This recital reveals a structural shift of paradigm. While credit risk has always been the risk of the lender due to information asymmetry (in favor of borrowers) that the borrower will not be able to repay in future, the AIA takes a new perspective with a (high) risk for the borrower concerning '*access to financial resources'*, which are still the financial resources of the lender (aka 'other people's money'). This shift of paradigm seems to be part of a trend, which can be illustrated by a quote taken from Steffen Mau (2020) [quote with a reference to (Eubanks, 2017) in the original]:

> *'On the basis of an ever-growing body of data, algorithms and high-tech tools are now 'automating inequality', creating forms of exclusion or assigning status (Eubanks, 2017), be it on the credit market, at the labour market, in the welfare system, for insurers, for product pricing or for targeted advertisement.'*

---

[2] As discussed later in detail, the proposed AIA defines 'artificial intelligence' in annex I in a never seen and extremely overarching way including 'statistical approaches' and 'Bayesian estimations' - or in other words all traditional statistical methods - and, consequently, regulates any (sic!) algorithmic credit scoring including existing statistical concepts.

[3] There is a long list of requirements, with which 'high-risk systems' shall comply. Unfortunately, the requirements are phrased in a blended terminology taken from different perspectives. For example, the data sets used for training et cetera shall be not only representative, but [quote from AIA, article 10] *'free of errors and complete'*. While it is standard in statistics that data has to be 'representative' (concerning the specific context), it remains unclear what *'free of errors and complete'* should mean, as any data set taken from a 'measurement' of the real world will be limited and will usually have (statistical and systematical) errors.

Concerning the possibility to '*perpetuate historical patterns of discrimination*', the AIA seems to pick up a line of argumentation already given by the German Datenethikkommission (2019; German Commission on Ethics in Data) [example 10; quote; translation by the author near to the original]:

> *'Within the scope of the estimation of credit worthiness, the household income is used as information". This [household income] is different in average for the genders in Germany. Consequently, an algorithmic system, which applies household income, can generate different distributions for the estimations of the credit worthiness of men and women.'*

Trivially, two different distributions (depending on gender) as input will lead to different distributions (depending on gender) as output of a statistical estimation. A - simplified - credit scoring <u>only</u> based on household income does not use gender as a parameter and is demographically blind. Even more, such an algorithmic system will produce <u>exactly the same score value for a woman as for a man with the same household income</u> - i.e. the scoring is fully gender agnostic. The credit scoring algorithm cannot change social structures of average income distribution, which are far beyond its scope and not changeable by a decision-making about a loan approval without any causal interaction with salary. However, as the input (distribution of household income) determines the output (distribution of credit approval), there will be an underlying 'perpetuation' of social pattern but without any causal link to the decision about credit approval (equal income = equal approval).

This example underscores that the statistical and causal concepts of (algorithmic) decision-making processes fades from the spotlight, and the debate centers about 'how' algorithms interact with the context, 'how' we want them to impact the society, and 'how' they should be regulated to do what we expect them to do. The present debate about algorithmic decision-making and artificial intelligence does not focus on the regulation of a technology (e.g. concerning objective safety criteria) but on socio-technical nexus and the outside-in perception of the outcome.

This paper will start with an assessment of decision-making in the social context and one illustrative recent example (in this case: health care) to set the scene. Then, it will review the (social) reality, the current regulation and on-going discussion about regulation of algorithmic decision-making exemplified by credit scoring: sector-specific in the USA, cross-industry in Europe, and functional for the regime in China.

As conclusion, this paper discusses the possibilities and limitations of sector-specific versus cross-industrial design of regulation and the interaction of regulation with the social context.

2. Decision-Making, Algorithms, and the Social Context

Any decision-making requires some structured approach (including rule of thumb or ‘heuristics’) and is in principle ‘algorithmic’: whether these ‘instructions’ are followed by humans or computer programs. The debate about ‘algorithmic’ decision-making focus on the role of technology and use of data and reveals an underlying fear that ‘machines’ could decide about the fate of a human being. In reality, decision-making typically follows statistical classification: Either we know what the future will bring (statistically speaking: a perfect prediction), and we do not need to ‘decide’ but follow a deterministic way without any individual responsibility. Or we have a free will but limited information about the uncertain future and take our experience to estimate a best guess.

Typically, we take our ‘experience’ about past events, assume continuation and apply a statistical classifier to estimate whether a new situation belongs to ‘blue’ or ‘orange’ class of events given a binary decision problem. This process of decision-making does not depend on the technical media: whether done by humans with an instruction manual, a pencil and a piece of paper, or whether this procedure is programmed in software and executed on a computer. Remarkably, hundred years ago a ‘computer’ was a term for a human being doing computations.

Likewise, it is a technical feature, but no fundamental difference, whether the classification of recorded events was done by drawing a line by hand to separate ‘blue’ and ‘orange’ events (or even only done in mind, and maybe unconsciously), by calculation a regression formula (with pencil or computer), or by so-called machine learning (for high-dimensional statistical classifying with many parameters and tremendous amounts of events). All those algorithms are mathematical methods to find an ‘optimal’ classifier to group ‘blue’ and ‘orange’ with minimal overlap based on a measured control parameter - in our case ‘household income’.

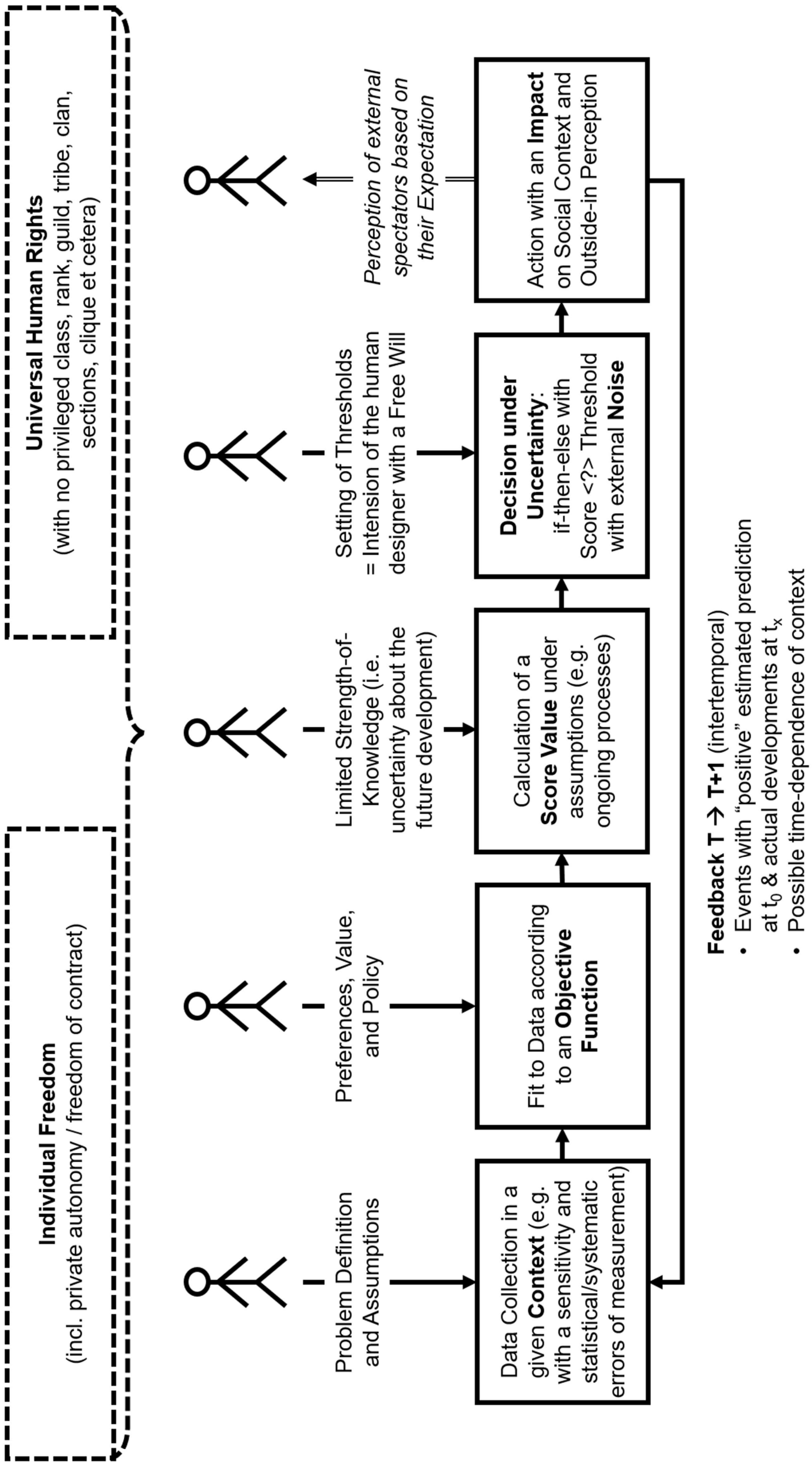

Figure 1: Schematic six-step process for algorithmic decision-making (see text for details). It has to be noted that the problem of 'fitting' pictures (aka pattern recognition with machine learning methods) with 'hidden' details in the input data will not be discussed in this paper except of footnote 10 at the end.

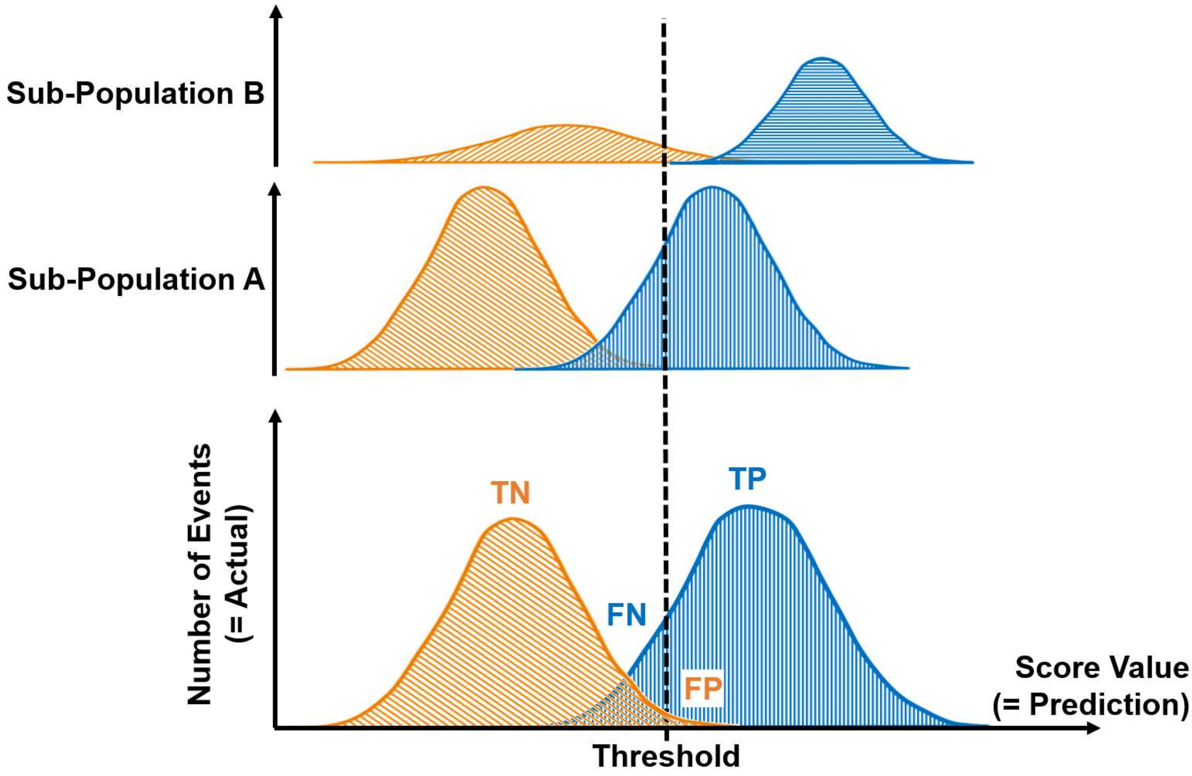

Figure 2: Statistical estimation of a score value and if-then-else decision with a threshold in case of two overlapping sub-groups (TP, TN, FP, and FN stand for True Positive, True Negative, False Positive and False Negative). The lower distribution is 'demographically blind' for the sub-populations, whereas the upper distributions are separated by an additional parameter with values of 'A' and 'B' for two sub-populations. As discussed in the current debate about the so-called identity policy, one can postulate many more sub-sub-…-groups by all possible combinations of sensitive parameters, which will end up - in the limit $n \rightarrow \infty$ - with micro-populations of few people or even one person, which is the upside-down of the central limit theorem rendering all statistical properties meaningless.

Already this very stylised example reveals a six-step process for algorithmic decision-making (see Figure 1):

1. Collection of statistical data (i.e. an ensemble with a given context) with measurable parameters (and always with measurement errors) and correlations based on the problem formulation and assumptions.
2. Human decision about an economic objective function (or problem formulation) to be optimized and fit of the function to the historic data based on preferences, value,
   and policy of the decision-maker.
3. Calculation of a score value with this function for a new event with the assumption that this fit to data from the past also applies for future predictions (process without time dependence or 'repeated game').
4. Human decision for the setting of a threshold parameter x for the if-then-else decision for allocation of a scarce resource:
   if 'economic' score value (of new event) > x, then do positive action, else do not
   (depending on the risk-appetite and contingency for uncertainty due to noise etc.).
5. Execution of the action, but always with the social responsibility for the impact (e.g. affordability of loans and avoiding indebtedness of consumers) and with an outside-in perception by external spectators.
6. Feedback from current situation to next interval (T → T+1) with an uncertainty about time-dependence and question how wide the scope of context should be seen (i.e. question what is a 'representative' data set when conditions are changing).

As long as the 'blue' and 'orange' populations are either fully separated (enabling a 'perfect prediction') or fully identical (i.e. with so-called 'equal base rate'), the situation is trivial. However, in reality (i) there are overlapping distributions for the *ex-post* results compared to *ex-ante* scoring and (ii) different potential subgroups will have different distributions as illustrated in Figure 2.

Comparing a threshold with such overlapping distributions result in correct classifications (True Positive 'TP' and True Negative 'TN') but also misclassifications (False Positive 'FP' and False Negative 'FN'), if *ex-ante* classified events are re-assessed *ex-post* with the actual result. While in medical studies, especially FN could be detected if all patents are monitored until the end of a therapy study, in decision-making FN are excluded from the whole process typically, which could introduce a bias and a shift in the distribution of the portfolio (if not corrected for in later analysis). In medical tests, FP could be severe problem, as healthy patients will receive a wrong result as "having a disease" and the negative (psychological) impact has to be balanced against the benefits of detection (the well-known problem of sensitivity vs. specificity, see also Pearl, 2018, for the example of breast cancer surveillance).

Such tests - whether health care or lending - always come with a trade-off, which require a modification of the 'optimal choice' of threshold according to the objective function e.g. to balance:

a. possible margin (as market-price for loans of TP) of a credit portfolio vs.
b. risk provision (coverage of expected risk and economic capital for unexpected risk of FP) vs.
c. excluded profit of misclassified 'positive' customers as FN (which could have re-paid).

The situation is even more challenging if the population, which is the base for the scoring, has a sub-structure based on sensitive data (as defined in the context of anti-discrimination regulation, see below). In the example above, the household income has a relation to a sensitive characteristic 'gender'. If a separation based on the sensitive characteristic is allowed by regulation, the problem can be solved with two different thresholds and individual setting. If the people with the sensitive characteristic must not be 'discriminated' by this parameter, classification requires demographic blindness (i.e. one threshold for all people with the same 'household income' independent of the sensitive characteristic). Given the precondition that the sensitive characteristic may be applied to choose such a setting, it depends on a - not two-sided - objective function, how to optimize the threshold. If this precondition is not given, the decision-maker cannot make an adjusted setting with respects to people with sensitive characteristics.

Especially in the context of decision-making with machine learning, there is an ongoing discussion about so-called ‘algorithmic fairness’ to make these decisions with same ‘fairness parameters’ (see e.g. Mitchell et al, 2021) for the sub-structure. Such ‘fairness parameters’ are the standard statistical measures of performance for (binary) classification test and usually defined as Sensitivity := TP/(TP+FN), Specificity := TN/(TN+FP), Precision := TP/(TP+FP), FNR = 1 - Sensitivity, FPR = 1 - Specifity et cetera. However, it follows from the definition of these ratios of type A/(A+B) and was pointed out by Kleinberg et al. (2017) that it is impossible that three ‘fairness parameters’ should be the same for two sub-structures (except for special cases of equant base rate of perfect classifiers). It always depends on the (economic) objective function how choose a threshold for a scoring value, but always as a trade-off. A ‘fairness’ in allocating a scarce resource can be achieved only in cases such as a birthday cake to be distributed in a family - and even here it remains open, whether the birthday child gets a bigger piece?

Finally, it should be notes that a new direction in the debate appeared with ‘algorithmic recourse’ and ‘fairness of recourse’. As it is beyond the scope of this paper to elaborate on this debate, only two issue should be mentioned. First, Karimi et al. (2021) provided a comprehensive survey of ‘algorithmic recourse’ and introduced this concept with the example of an individual applying for a loan and receiving a refusal, who askes two questions: 'Why was I rejected the loan?' and 'What can I do in order to get the loan in the future?' One the one side, this leads to an approach of ‘gaming the system’, which contradicts the objective of a statistical estimation of future defaults. On the other side, this approach ignores the responsibility of a lender to avoid borrowers’ indebtedness. Some feedback with a link to public assistance or debt advice service would be reasonable alternatives. Second, Kügelgen et al. (2021) discussed ‘fairness of recourse’ and elaborated on an example of credit card approval for two sub-groups with distributions with the same mean, but different variance. They elaborated that the ‘cost of recourse’ (i.e. average distance from the ‘not approved’ part of the distribution to the decision boundary) is much larger for individuals in group with larger variance. While this is statistically true (and trivial), it ignores that any decision-making is a non-symmetric function (see Figure 3) and any constructed sub-groups with distinct variances will show this difference.

3. An Illustrative Example

Although the next example is about algorithmic prediction in the U.S. health care system - taken from a recent study of Obermeyer et al. (2019) - it illustrates the challenges of objective functions, implicit assumptions, data versus proxies, and the danger of naïve use of ‘available’ data without proper understanding of the context. It is also - unfortunately - one more example for systemic discrimination in the USA based on ‘widely used’ algorithms [quote]:

> *“Health systems rely on commercial prediction algorithms to identify and help patients with complex health needs. We show that a widely used algorithm, typical of this industry-wide approach and affecting millions of patients, exhibits significant racial bias: At a given risk score, Black patients are considerably sicker than White patients, as evidenced by signs of uncontrolled illnesses.”*

This result reveals much about such programs as cornerstone of health management in the USA. The evaluated prediction algorithms produced <u>different prediction scores for patients with the same health condition depending on being White or Black</u> (while not using the sensitive characteristic parameter ‘race’) and amplifies discrimination better score values lead to more future allocation of health care resources to those people with already better conditions. According to this study, most health systems in the USA use similar programs as the cornerstone of population health management efforts.

Strangely, this specific program might be ‘correct’ in the sense of achieved result compared to the commercial (sic!) objective function, which is based on the assumption that patients with <u>higher health care cost in the past</u> may require provision of additional resources in the future <u>to reduce future costs</u> in the long run. This ‘widely used algorithm’ does not measure health condition directly but applies a proxy for ‘sickness’ based on cost recorded in the past for medical treatments. For the commercial objective to reduce costs (and not to improve health as medical objective!) the program seam to work and - in some disturbing way - works in a manner that all patients with similar costs in the past will receive similar support in the future.

However, this approach is wrong from a statistical perspective and wrong from the perspective of anti-discrimination. The - primarily - objective or ‘problem formulation’ of a health care system is, trivially, health care and support of individual health condition, but not - primarily - cost cutting.

The proxy parameter ‘health care cost in the past’ does not depend on the health of the patient only, but additionally on the patients’ access to the health care system as a statistical confounder (see especially Pearl, 2018, for particularly good explanation of the statistical back-ground of ‘cofounders’). In simple words, Black patients do not have the same access to the health care systems as White patients and generate lower costs, which in turn do not qualify Black people in average for the same and better prophylactic treatment as White people (to avoid future costs for the system).

Ignoring this confounder, which is of significant importance in the context of health care, causes a wrong statistical approach and a wrong choice of the proxy for the statistical prediction. If a program applies statistics, it has to apply statistics correctly! Although this proxy ‘costs in the past’ does not refer to ‘race’ directly, the confounder ‘accesses to the system’ (or economic possibility to access to system as needed) depends on ‘race’ significantly, which could be formalized:

Race → Access → Cost in the Past ← Health Condition (no reference to Race)

↓

Special Support in the Future

Consequently, the outcome is not only correlated with ‘race’ but has a causal relationship with the parameter ‘race’ via the parameter ‘access’, although the parameter ‘race’ is not applied in the algorithm. It is rather concerning that the medical industry in the USA is highly regulated, but such ‘commercial’ algorithms seam to flight under the radar of sector-specific regulation or supervision.

This example is embedded in the specific U.S. context but, nevertheless, is illustrative for the following questions concerning regulation:

- What do we - as society - want ‘regulation of algorithms’ to achieve, and how should the institutional design of regulation and supervision be designed?
- How much expertise of (advanced) statistics and (algorithmic) decision-making process is required?
- And how can ‘successes’ of regulation been assessed or measured?

Different approaches to regulation of algorithms will be review for the case of credit scoring next, and be compared with the developments in the USA, in Germany (representing Europe) and in China following on.

4. Design of Regulation of Algorithms

Like health care, banking and financial services are highly regulated industries. Since George Stigler's essay fifty years ago, the normative foundation of regulation has been discussed. While regulation, de-regulation, and re-regulation came in waves - oscillating between believe in the market economy and claims of 'market failure' (or, respectively, economic theories unable to describe actual developments of markets). The basic question in scope is how regulation and supervision could be designed to achieve reasonable results in a social context: sector-specific, cross-industry, or case-based.

The case of Algorithmic Decision-making and Credit Scoring is a special one, as it links - at least - three different aspects of regulation:

1. Individual decision-making as part of private autonomy and freedom of contract (with a need for information about the past but with the responsibility for the future outcome in the context of the economic objective) as core component of any market economy.
2. Centuries old religious concerns about 'usury' of unethical lending that enrich the lender in an 'unfair' manner (nowadays pointing to information asymmetries between the lender and the borrower) up to recent questions in the flagship magazine of the International Monetary Fund about a 'New Morality of Debt' (Aggarwal, 2021).
3. New worries about 'robots' or 'algorithms' taking control to subjugate mankind (from Dennis Feltham Jones' novel 'Colossus' in 1966 to 'Skynet' in the Terminator movies).

Although lending is a fundamental function in economy, the outside-in perception of the decision-maker's intension depends on the social context and varies between the assumption of *objective / discriminative / opaque* criteria of a lender. In its economic function to provide liquidity to the market and unknown potential borrowers, a lender always takes the risk (with a responsibility to avoid indebtedness of the lender, if the lender is financial institution but no credit shark).

Nonetheless, there is a growing discussion that the lender has a - non-causal - obligation to redress for historical discrimination in the society independent of any individual responsibility of the lender or legal claim of the borrower. The following quotes may illustrate different points of view in the debate without claiming completeness of contents.

Aggarwal (2021) wrote in the magazine of the International Monetary Fund [quote, underlining by the author]: '*Notably, the <u>datafication of consumer lending</u> has amplified moral concerns about <u>harm to individual privacy, autonomy, identity, and dignity</u>. ... These practices diminish consumers' ability to craft their own identity as they become increasingly chained to their "data self," or algorithmic identity. ... Is it moral for lenders to use highly intimate health and relationship data - for example, captured from social media and dating apps - to determine consumer creditworthiness? ... Yet the datafication of consumer lending could also uphold the morality of debt, by improving other dimensions of <u>distributional fairness in consumer credit markets</u>. ... Firms should be subject to more rigorous obligations to justify the processing of personal data under the paradigm of datafied lending.*'

In 2019, five U.S. agencies[4] (CFPB, 2019) published an 'Interagency statement on the use of alternative data in credit underwriting' [quote, underlining by the author]: '*The agencies recognize that use of <u>alternative data may improve the speed and accuracy of credit decisions</u> and may help firms evaluate the creditworthiness of <u>consumers who currently may not obtain credit</u> in the mainstream credit system. ... In addition, the agencies are aware that the <u>use of certain alternative data may present no greater risks</u> than data traditionally used in the credit evaluation process. For example, the agencies are aware that some firms are automating the <u>use of cash flow data to better evaluate borrowers' ability to repay loans</u>. … Consumers can expressly permission access to their cash flow data, ... and disclosed to the borrower, as may be required under the Equal Credit Opportunity Act and the Fair Credit Reporting Act.*'

On May 13, 2021 the Wall Street Journal (Rudegeair and Andriotis, 2021) reported that that JPMorgan Chase, Wells Fargo, U.S. Bancorp and other large U.S. banks are going to start a pilot program, under which they would have common

[4] Statement from the Consumer Financial Protection Bureau (CFPB), the Federal Reserve Board (Federal Reserve), the Federal Deposit Insurance Corporation (FDIC), the National Credit Union Administration (NCUA), and the Office of the Comptroller of the Currency (OCC).

access to customers' checking or deposit accounts to increase their chances of being approved for credit cards. This pilot would be part of Project REACh (Roundtable for Economic Access and Change) launched last summer by the Office of the Comptroller of the Currency. Accouding to sources, pilot program is expected to launch this year.

A recent working paper of the Bank of International Settlement 'BIS' (Gambacorta et al., 2020) discussed a possible substitution of collateral by data for corporate loans [quote, underlining by the author]: '*The use of <u>massive amounts of data by large technology firms (big techs) to assess firms' creditworthiness</u> could <u>reduce the need for collateral in solving asymmetric information problems</u> in credit markets. Using a unique dataset of more than 2 million Chinese firms that received credit from both an important big tech firm (Ant Group) and traditional commercial banks, ... We find that big tech credit ... reacts strongly to changes in firm characteristics, such as transaction volumes ... a greater use of big tech credit ... could reduce the importance of collateral in credit markets* ...'.

These quotes illustrate that '*data*' is the core concern, but with antagonistic positions. While Aggarwal (2021) focussed on <u>theoretically possible harm</u> for consumers and point to '*distributional fairness in consumer credit markets'* (disregarding freedom of contract and private autonomy of the lenders), CFPB (2019) did <u>not necessarily see greater risks</u>, but the benefits of data like payment transaction history, when provided according to data protection and anti-discrimination regulation. Finally, Gambacorta et al. (2020) pointed out the <u>measurable advantages</u> of the use of data and improvements for the lender with asymmetrical information about borrowers' future behaviour.

Undoubtedly, we live in the age of digitalization and proliferation of the use of data. These three opinions correlate with different design principles for regulation and supervision:

- A general (vertical) and more prohibitive regulation with prescriptive consumer protection / data protection and anti-discrimination focussed on the outcome of decision-making and the impact on the social context.
- A sector-specific (horizontal) regulation with regard to the business requirements with a balanced approach to 'data', which acknowledges possible benefits including more financial inclusion and possible risks.

- An innovation friendly regulation to avoid tangible misuse or harm embedded in the framework of a resilient market economy (see: Markus K. Brunnermeier, 2021).

Especially in cases like decision-making, a cross-industry regulation without any specific background of actual process and business requirement can *per se* be prohibitive only. In a perfect utopia - with the old question about human freedom at a Pareto optimum - decision-making would be deterministic and without problems, but the reality shows how 'imperfect' the society is. Consequently, decision-making in the real world - uncertain, imperfect, ambivalent, path-dependent - has to be scrutinised whether it supports perpetuating or 'petrification' of the existing social situation with all the deficits or support the overall search process of a market economy for dynamic optimization.

Credit scoring (as a testing probe for algorithmic decision-making in a specific industry) is a 'neutral' statistical estimation credit defaults: 'blind' to demographic structure but based on 'objective' economic parameters such as disposable household income. Nevertheless, any 'scoring' is always sensitive to the social context beyond the technical parameters and mirrors access to (financial) services and (economic) capabilities of the potential borrowers. In a speech in September 2021 Joachim Wuermeling (2021), Member of the Executive Board of the Deutsche Bundesbank, elaborated on the mismatch of (horizontal) banking regulation and the (vertical) proposal of the European Commission on Artificial Intelligence [quote, emphasise by the author]:

> *Let me start with an example of two frameworks that don't yet complement each other well. As banking supervisors, we are particularly interested in processes and applications that have a bearing on risk management, such as artificial intelligence in credit assessments, liquidity planning, or portfolio management.* <u>*The use of artificial intelligence is supervised under the existing banking regulation*</u>*. I am therefore rather critical about introducing special authorisation requirements, such as those proposed by the European Commission for creditworthiness checks. Banks should continue to be supervised in a technology-neutral manner - without duplicating any regulation, and without duplicating supervisory processes.*

As regulation is always an intervention into a market economy, it is a ‘second best’ choice on the thin line between benefits vs. bureaucracy. From this pragmatic perspective, a comparison of existing regulations of (algorithmic) credit scoring can provide some insight into regulatory design and the results.

5. USA

In the USA, sector-specific anti-discrimination legislation for housing and financial services started already in the late 1960s/early 1970s, but could not prevent ongoing - and systemic[5,6] - discrimination today. On the one side there are: Civil Rights Act incl. Fair Housing Act (FHA) and Consumer Credit Protection Act (CCPA) of 1968, Fair Credit Reporting Act (FCRA) of 1970, and Equal Credit Opportunity Act (ECOA) of 1974 - which either states that it is illegal to discriminate based on race, sex, age, national origin, or marital status, or because one receives public assistance et cetera, or protects consumers data incl. the right for correct data and transparency about the use of data. In addition, the Clinton administration introduced significant regulatory changes to the Community Reinvestment Act (CRA), which ‘encouraged’ the lending industry to extend to low- and moderate-income neighborhoods (see: Clinton, 1995). A new regulatory approach can be seen in the cross-industry California Consumer Privacy Act (CCPA) in force since mid-2020, which provides consumers control over personal data collected by businesses (without any link to the respective industry).

On the other side, there is a long list of deficits, which cast doubt on the efficiency of regulation for credit scoring. An alarming problem was reported in the Annual Economic Report 2020 of the Bank of International Settlement (BIS, 2020) that "*nearly half of Black and Hispanic US households are unbanked or underbanked*" (approx. 15% unbanked and an additional 30% underbanked).

Although this has nothing to do with credit decision-making or algorithmic scoring directly, any type of credit decisions in a society with minorities being dramatically underrepresented - if not to say: discriminated - can never achieve any ‘fairness’ at all.

---

[5] The term ‘systemic’ is used to indicate the overall problems of the financial services industry (along the society), which are no assemblage of separated cases but an industry-wide (and society-wide) phenomenon. However, the term ‘systemic’ is not meant to construct special groups in the industry or society, which are ‘good’ or ‘bad’ in principle due some characteristic but without looking on the individuals, the circumstances, and the social context.

[6] Additionally, there is intrinsic dynamics in the systemic discrimination as e.g. Dang et al. (2020) showed as an effect of the Covid-19 pandemic [quote]: ‘*Further, the pandemic has exacerbated anti-Asian xenophobia and racism, which have historically acted as barriers to equity. ... Asian Americans have been the target of a unique rise in racist rhetoric and discrimination.*’

A recent meta study by Quillian et al. (2020) suggested that [quote]: '*racial gaps in loan denial have declined only slightly, and racial gaps in mortgage cost have not declined at all,*' in the U.S mortgage market. It is worth to quote the whole abstract [quote, underlining by the author]:

> *We examine trends in racial and ethnic discrimination in U.S. housing and mortgage lending markets through a quantitative review of studies. We code and analyze as a time series results from 16 field experiments of housing discrimination and 19 observational studies of mortgage lending discrimination. Consistent with prior research, we find evidence of a decline in housing discrimination from the late 1970s to the present. Our results show that this trend holds in both the national audits sponsored by the U.S. Department of Housing and Urban Development (HUD) and in non-HUD studies. The <u>decline in discrimination is strongest for discrimination that involves direct denial of housing availability</u>, for which discrimination has declined to low levels. The downward trend in discrimination is weaker for measures reflecting the number of units recommended and inspected, and significant discrimination remains for these outcomes. <u>In the mortgage market, we find that racial gaps in loan denial have declined only slightly, and racial gaps in mortgage cost have not declined at all, suggesting persistent racial discrimination</u>. We discuss the implications of these trends for housing inequality, racial segregation, and racial disparities in household wealth.*

It is a positive result that direct discrimination in the U.S. housing market declined - more than fifty years after the Fair Housing Act. However, one can only agree to this study, that it is a '*disturbing finding*' to see not much decline [quote]: '*in loan denial and cost has not declined much over the previous 30 - 40 years'*. The different development between direct discrimination (using sensitive data) and more subtle forms of discrimination indicates a complex development, especially in a society with an underlying systemic discrimination. Even if a decision is neither direct discrimination (using sensitive characteristics), not indirect discrimination (having a different outcome for people with same economic conditions but different sensitive characteristics), a delicate disparity in the economic conditions (simplified: similar household income, but different credit file history) seems to be beyond the reach of current regulation.

The historic discrimination of ‘redlining’ certain areas deemed to be ‘high-risk’ for mortgages, is reflected in the current exposure of this areas to the effect of global warming. A recent study of the National Academies of Sciences, Engineering, and Medicine (NSA, 2021) on local climate effects pointed out [quote]: ‘*Today, historically redlined urban areas are on average about 4.5 degrees Fahrenheit hotter than their green counterparts, with some cities seeing discrepancies as high as 20 degrees*.’

Different to Europe, where credit agencies started historically recording customers’ defaults as protection for banks (and other type of lenders) in a situation of information asymmetry with ‘unknown’ consumers, the U.S. financial system has the approach to assume ‘thick file’ credit history, i.e. the regular use of many and different loans (from student and auto loans to credit card/consumer finance and mortgages) provides a high score to get more loans. Consequently, any credit decision-making based on (existing) credit files cannot produce equal results even in the case of comparable actual economic conditions, if people have no or no significant credit history: the problem of so-called ‘thin files’. A new and not yet peer-reviewed study by Blattner and Nelson (2021) showed that credit scores are statistically ‘noisier’ indicators of default risk for historically under-served people: Credit scores have less explanatory power for consumers with ‘thin credit files’ and, consequently, these people have a limited chance to receive a loan compared with ‘thick file’ consumers. According to the study - based on one specific credit score but for a long-term development of consumers' credit report data from 2009 to 2017 - the ‘noise’ or limited explanatory power resulted from features of credit report data and especially data sparsity (of ‘thin files’). The study used two innovative approaches. The long-term credit report history allowed a monitoring of rejected mortgage applicants on other loans to achieve a proxy for default risk and analyse the prediction power (prediction of the credit score vs. proxy for future credit performance, given the mortgage would have be approved). Additionally, two populations with economic disadvantage could be identified, i.e. low income and racial or ethnic minority, by a standard method of Bayesian Improved Surname Geocoding ('BISG') to estimate income and race or ethnicity from a combination of name and geographic information, which was included in the credit files but not used for the credit scoring.

The study concluded [quote]:

"… approximately half of the gap in credit scores' overall explanatory power, and at least half of data bias, can be explained by observable features of credit report data, such as data sparsity".

An additional issue is the possibility of misconduct 'at the fringe' of the regulated perimeter. For example, the U.S. Department of Justice (2012) filed a settlement [quote]: '*to resolve allegations that Wells Fargo Bank, the largest residential home mortgage originator in the United States, engaged in a pattern or practice of discrimination against qualified African-American and Hispanic borrowers in its mortgage lending from 2004 through 2009*'. While in one part, Wells Fargo discriminated by charging higher fees and rates [quote]: '*because of their race or national origin rather than the borrowers' credit worthiness or other objective criteria related to borrower risk'*, in another part the discrimination was conducted by [quote]: '*steering approximately 4,000 African-American and Hispanic …, into subprime mortgages when non-Hispanic white borrowers with similar credit profiles received prime loans'.* Therefore, discrimination could result from a subtle 'pre-selection of' or 'steering to' unfavorable products compared to people with other characteristics. Such misconduct does not require an intension to discriminate, but has many similarities to the well-known 'fraud triangle' of Cressey (1953) of opportunity (typically: inadequate and/or consciously bypassed internal control systems), rationalization (e.g. the long-lasting governmental support for mortgage lending to 'sub-prime' borrowers without sufficient financial resources before the sub-prime crisis), and motivation (e.g. due to incentive structures to sell as much as possible and with highest possible margin).

At this point it is important to make a warning against too much anecdotical evidence, as not every allegation proves true, and not every scoring in financial services is prone to discrimination. One - not representative, again - example was the so-called 'Apple Card debacle'. On March 23, 2021, the New York State Department of Financial Services (DFS, 2021) issued a report summarizing the findings after investigating consumer complaints about the Apple Card and concluded [quote]: *'No Fair Lending Violations Found*' and elaborated [quote]: '*..., consumers voiced the belief that if they shared credit cards with spouses, even if only as authorized users, they were entitled to the same credit terms as spouses. In reality, however, underwriters are not required to treat authorized users the same as account holders, and may* *consider many other fac-tors. In terms of gender, the Department found, based on its data analysis, that Apple Card applications from women and men with similar credit characteristics generally had similar outcomes. For all consumers who reported concerns about their Apple Card credit application outcomes to the Department, evidence showed that those decisions were explainable, lawful, and consistent with the Bank's credit policy*.' While the individual claims of consumers, who were not able to understand the background of the scoring algorithms, are understandable and should trigger better communication by the lenders and more financial literacy in general[7], the tendency to presume anecdotical correlation for causal evidence is critical.

Despite over fifty years of sector-specific anti-discrimination and data protection regulation, there is no comprehensive insight into regulatory efficiency especially in the credit card business, which amounts to 2,383 $bn in 2020 compared to 1,856 $bn mortgage lending volume (see KPMG, 2021). For the housing and mortgage market, the most comprehensive and recent meta study by Quillian et al. (2020) conclude [quote]:

> '*To the extent that anti-discrimination enforcement is one factor accounting for the decline of explicit forms of discrimination, the Fair Housing Act has been successful. However, white applicants are still given more options and overall better treatment in the housing search process, and their advantages in mortgage pricing and availability have not decreased. In sum, the results suggest that anti-discrimination enforcement in the housing and mortgage markets should continue, and efforts should be increased to ensure that all home seekers receive equal treatment regardless of their race or ethnic background.*'

[7] See also the FTC (2021) Consumer Information 'A Special Note To Women' concerning the understanding of credit scoring for so-called 'partner card'.

	Observation	Regulation	Benefits of new approaches (Data & ML)
Inclusion to financial services in general (access)	Underlying systemic discrimination in the society and - at least - ambivalent experiences of 'positive actions' such as the CR	Task for government / promotional banks, but with clear link of decision, risk, and responsibility	Possibly better inclusion of the FN domain by a two-step Rooney Rule like approach (see text)
Indirect discrimination (or disparate impact: by outcome without business needs, without intension)	See e.g. Michaelson et al. (2020) for CFPB enforcement actions	Disputed issue, as not literally covered in ECOA, but developed as a doctrine*	No, but the doctrine could be used to extend the scope of a legal test to (non-sensitive) demographic parameters in a non-traditional) data-set with social disparity
Subtle discrimination e.g. by pricing a home as collateral according to the neighborhood, but not by a 'stand-alone' value	Possibly a gap in current regulation: more research necessary	Not covered by regulation as prices would be calculated according to business needs / industry standards	Possible shift from collateral to data (see chapter about China for more details)
Marketing of high price products to specific groups in the society	Specific cases (e.g. Wells Fargo)	Only by court decisions (to be included in future regulations)	No, but potentially ML could help governments to identify dubious cases
Direct discrimination	Effectively suppressed by current sector-specific regulation	Positive effect of sector-specific regulation	No

Table 1: Summary of current sector-specific regulations of (algorithmic) decision-making in the U.S. financial service industry compared to selected results as discussed in the text.

**) For the discussion see the interpretation in the 'Federal Fair Lending Regulations and Statutes' of the FED (2017), the historical perspective in Taylor (2018), an expectation about current developments in* Michaelson et al. (2020), and Marshall (2020) about the ruling of the U.S. Supreme Court in 2015 'Inclusive Communities', which narrowed 'disparate impact' to cases where direct causality can be conclusively shown.

More insight would be required especially into the roots of the '*subtle forms of discrimination*' as pointed out by Quillian et al. (2020): e.g. with targeted marketing of high-cost mortgage products in minority communities or due neighborhood contingency of race discrimination with different calculated values of comparable homes in different 'affluent' neighborhood as a modern variant of 'redlining'. Nonetheless, this is an issue of human responsibility and the human choice of an objective function, but not of the calculation of a programmed algorithm (whether as manual for human work, as rule-based software or as statistical classifier with some artificial neural network technology).

Of course, it is out of question that credit scoring should be risk-based (concerning the risk taken by the lender), but vice versa it is a responsibility of the lender to calculate the affordability of loans and to decline loans, if there is a risk for indebtedness of consumers. If the society wants to promote the financial situation of these people not qualified for a loan, either governmental promotional banks have a mandate to improving economic, social, and other living conditions, or governments can redistribute via taxation. While economic objective functions are within the business of banks as agents in the market economy, governments and promotional banks follow such social objectives.

Recently a new question has raised, whether advanced methods such as 'machine learning' (see below) and usage of non-traditional data could provide redress for perpetuated bias in traditional credit scoring and credit data, while being in-scope of commercial lenders.

As said in the already mentioned 'Interagency statement on the use of alternative data in credit underwriting' (CFPB, 2019) the use of alternative data may help consumers currently not qualified to obtain credit in the traditional credit system. The possible pilot program of JPMorgan and others lang U.S. bank to issue credit cards to people with no credit scores' (as reported recently by Rudegeair and Andriotis, 2021, in the Wall Street Journal) would be a crucial test for this concept to analyse cash flow data/payment transaction data to 'upgrade' an insufficient credit score.

However, Fuster et al. (2020) found mixed results in a simulation with detailed administrative data on US mortgages [quote, underlining by the author]:

> *'Machine learning models slightly increase credit provision overall, but increase rate disparity between and within groups; effects mainly arise from flexibility to uncover structural relationships between default and observables, rather than from triangulation of excluded characteristics [i.e. 'proxies']. We predict that <u>Black and Hispanic borrowers are disproportionately less likely to gain from new technology</u>.'*

Another study by Wang and Perkins (2019) with consumer loan data from the largest FinTech lender of personal loans in the USA (Lending Club) simulated the effect of machine learning (ML) and found [quote, underlining by the author]:

> *'This suggests that unconventional data can help, but are most likely to bring about materially more accurate credit ratings only for consumers with little or no credit history, as such data substitute for the absence of the more informative credit variables." … "It finds that the ML methods produce more favorable ratings for different groups of consumers, although <u>those already deemed less risky seem to benefit more on balance</u>.'*

Both simulations - based on real lending data, but without data from 'social media' - reveal that such approaches and the use of non-traditional data and/or machine learning as kind of advanced 'statistical classifiers' could improve inclusion. While non-traditional approaches to credit decision-making / credit scoring could support better inclusion, this benefit seems to be focused on not-/under-banked and/or people with 'thin' credit files, for people with similar economic conditions and similar products offered, but not on impaired economic conditions in a society with historical discrimination.

These simulations are consistent with empirical evidence from LendingClub data (Jagtiani and Lemieux, 2019) that alternative data could help borrowers to get better score value [quote]:

> *'The use of alternative data has allowed some borrowers who would have been classified as subprime by traditional criteria to be slotted into "better" loan grades, which allowed them to get lower-priced credit.'*

It remains unclear whether alternative data only support an ‘upgrade’ for some potential borrowers or could improve lending conditions for all borrowers[8].

Summarizing recent results about the status-quo, a fragmented picture as shown in Table 1 can be deduced. While it would require more research to disentangle the different interaction between (sector-specific) regulation and efficiency in the real world, a preliminary hypothesis for the specific situation in the USA could be discussed with three issues:

1. Sector-specific regulation of credit scoring / decision-making seems to be effective concerning direct discrimination and indirect discrimination with a well-defined context and enforceability: Especially the scope of the business objectives have to be transparent, lenders have to calculate their expected future risk and take responsibility according to these calculations. However, this approach opens backdoors for circumvention and misconduct, which renders the overall impact on systemic discrimination in the society with limited efficiency.
2. The long-term effect of positive (or affirmative) actions such as the CRA is very ambiguous. As the subprime crisis revealed, the general problem is the separation between decision and responsibility (e.g. by a combination of a governmental ‘request’ and a fragmented value chain with a decoupling of underwriting - accounting - final risk taking). While there has been a lot of debate about the contribution of the CRA to the financial crisis, it should be clear that such a fundamental crisis has no mono-causal explanation. As well-known in history, there is always a blend of over-ambitious governments, greedy investors, inflexible central banks, ‘risk-free’ governmental agencies (up to the problem of Alt-A mortgages, i.e. so-called ‘Liar Loans’ with insufficient documentation), silo-thinking bankers and other market players interacted to exploit the original regulation, which was designed (good-minded) to provide inclusion for low- and moderate-income neighborhoods but disregarding the principles of risk.

[8] A recent working paper of the Bank of England (Eccles et al., 2021) simulated market competition between ‘traditional’ and ‘innovative’ lenders with three classes of borrowers: ‘good’ and ‘risky’ with ‘risky’ split in ‘high risk’ and ‘low risk’, which could be distinguished only by the ‘innovative’ lenders. Depending on the borrower mix and the cost for the corresponding ‘innovative’ scoring systems different market dynamics will develop and different classes of borrowers could benefit from more selective scoring systems.

3. As far as today, non-traditional data could provide more granular statistical prediction of future re-payments but limited to (i) near-traditional financial information linked to expected re-payment and (ii) offering benefits for already better-off people in the 'False Negative' category, but not for excluded people due to underlying economic problems.

With this hypothesis of efficient but limited sector-specific regulation, well-intended but often ambivalent positive actions, and promising but constrained benefits of advanced methods, it could help to take a step aside and apply 'non-mathematical' approach such as the Rooney Rule (see especially: Kleinberg and Raghavan, 2018) together with a simple but transparent anti-discrimination regulation. The so-called 'Rooney Rule' was a protocol of the U.S. National Football League (NFL) in 2002 to improve the (historically low) representation of African-Americans in head coaching positions and requires that - instead of a fixed quota - at least one of the finalists for an open position should be African-American (or from an underrepresented minority in general). Especially high-tech firms implemented this protocol when hiring executive managers with at least one candidate in the final round of interviews has to be a member of a minority (especially women or other underrepresented groups).

Of course, hiring (from a group of finalists) is a different process compared to credit approval (with always one applicant to be scored), but the idea of the Rooney Rule could be transferred to credit decision-making: Applicant with a negative score value in the range with a high probability for False Negative could be given a 'second chance' (a chance, but no approval) by asking for access to his/her bank account with payment transaction history, account at most used e-commerce merchant with similar history of payments, or other considerable registers of cash flow history. After receiving the customer's mandate to access these data, the lender could make a re-calculation including this non-traditional data to - possibly - enhance FN to TP (but not the other way, i.e. no reduction of standardized score value).

Such a 'second chance' could avoid the problems of quotas (typically conflicting with risk-based statistical prediction based on economic data). Such a 'second chance' could help to avoid or at least reduce the 'petrification' of historic social development with this additional opportunity.

Of course, ‘second chance’ should not be regarded in a negative sense but stand for a modification of standard statistical estimation (based on the ‘credit file history’) with a possible enhancement from the FN domain by using additional non-traditional data in a second decision cycle. Additionally, such a two-step scoring (traditional + enhancement possibility for the FN domain) would avoid - by design - any discussion about theoretically possible bias and/or proxy for sensitive characteristics within the non-traditional.

Whereas problems due to bias or proxies could always be constructed hypothetically (but without any credible evidence in real applications until today), a two-step approach with an additional non-traditional scoring as unidirectional chance for improvement of score results would avoid any negative impact of any possible bias from the beginning.

However, in mid-2020, then-Senator Harris, Senator Warren and Senator Sherrod Brown wrote a letter to the CFPB relating to fair use of educational attainment data in credit decisions [quote]: *'The risk of discrimination arises because the lender is not evaluating the applicant based on their own characteristics, but instead based on the characteristics of other students at their school ..*.' While it can be discussed how much predictive value for future re-payment capabilities a school or basic university degree has (see Langenbucher and Corcoran, 2021), it is clear that any population with similar economic circumstances will have similar income and will be able to re-pay in a similar pattern. And it is the task of statistical estimation of credit scores to estimate future re-payment pattern from a population - that’s what statistics does! Any statistical prediction is - principally - an approach based on forecasting collective development under the assumption of continuation. This misunderstanding is disturbing because no lender - as economic agent trying to make statistical estimation about future re-payment probability (i.e. opposite to forecast of default) based on collective data - can change the structure of society, and it is the task of the administration to provide remedy e.g. with social benefits or access to governmental promotional banks’ programs.

6. Germany and Europe

In Europe - and following in Germany - the situation is somehow complementary to the USA with a cross-industry regulation of data protection and anti-discrimination, but a strange ‘mishmash’ in the currently proposed ‘Artificial Intelligence Act’ (AIA):

- The General Data Protection Regulation (GDPR) of May 5, 2016 is focused on data protection, but only for ‘data processing’ in the sense of electronic data processing, while the same algorithmic procedure calculated by humans according to a ‘manual’ would be out of scope. Additionally, the GDPR regulates (i) decision-making based on automated processing of personal data - with the same problem that decision based on manual processing of the same data are out-of-scope - and (ii) anti-discrimination, but with the same limitation (see below).
- The two pillars of European anti-discrimination regulation: the EU Race Equality Directive of June 29, 2000 (2000/43/EC) and the Framework Employment Directive of Nov. 27, 2000 (2000/78/EC) regulate specific issues (i.e. race equality and hiring) independent of the type of processing (manually and electronically), but are limited in scope and, consequently, only the Race Equality Directive would apply to manual credit scoping, but not the other anti-discrimination issues covered in the Framework Employment Directive. Additionally, both directives define ‘indirect discrimination’ (see below), which is not defined in the GDPR.
- The proposed Artificial Intelligence Act (AIA) of April 21, 2021, is a ‘mishmash’ of many disjunct topics and only motivated by a technology-focussed regulation - violating the claim of technology agnostic regulations by the European Commission. Based on an exceptional and far-reaching definition of ‘Artificial Intelligence’[9], the AIA blends three approach: 1. amendments to existing product-specific regulations (concerning the use if ‘Artificial Intelligence’ within those products, which would typically be no regulation, but an amendment to the specific text of a product regulation how a product should comply to so-called ‘CE marking’), 2. prohibition of specific ‘public’ applications of AI (which

[9] The AIA applies a completely new definition of ‘artificial intelligence’, which has no historical foundation by defining in annex I everything from machine learning (i.e. artificial neural networks et cetera) via ‘logic-based’ approaches (which should mean so-called ‘symbolic’ approaches with rule-based concepts) to statistical approaches (which have never been included in a definition of AI before).

could be regarded as the core objective of the AIA), and 3. an unstructured list of so-called high-risk systems (but in a strange mixture of appliance of AI in public services such as justice, social benefits, or asylum with some ‘stand-alone’ private sector implementations in (safety) infrastructures, employment (overlapping with the Framework Employment Directive) and credit scoring (as a solitary issue).

- And finally, the proposal for a new Directive on consumer credits (CCD) of June 30, 2021, included non-discrimination in Art. 6 (referring to in Article 21 of the Charter of Fundamental Rights of the European Union), the concept of automated processing (taken from GDPR), and the AIA concept of ‘*risk to* a *persons’ access to financial resources’* in recital 48 [quote]: “*establishes that AI systems used to evaluate the credit score or creditworthiness of natural persons should be classified as high-risk AI systems, since they determine those persons’ access to financial resources*”. However, the proposed new CCD did not change the ‘*Obligation to assess the creditworthiness of the consumer*’ in Article 18 [quote]: ' *... creditor ... makes a thorough assessment of the consumer’s creditworthiness. That assessment shall be done in the interest of the consumer, to prevent irresponsible lending practices and over-indebtedness, ...*’.

The AIA is a proposal of the European Commission for the time being and changes can be expected during the trialogue process with the European Council and European Parliament. One indication for the overall problem can be find in the Explanatory Memorandum of the AIA [quote, page 3]: ‘*The proposal lays down a solid risk methodology to define “high-risk” AI systems that pose significant risks to the health and safety or fundamental rights of persons*.’

One the one side, the ‘*solid risk methodology’* is a very inhomogeneous lists in the annex for products with AI components or, respectively, services based on AI (but also including all traditional statistical approaches as defined in the proposal). On the other side, this mix is repeated in the different perspectives of ‘*health and safety or fundamental rights of persons’* with the mishmash of specific scopes with product safety versus potential harm to fundamental rights by public services like police or justice using AI for the prediction of behavioral pattern.

Roberto Viola (2021), Director-General of DG CONNECT of the European Commission, mentioned the importance of ‘A Human-Centric Approach to AI’ and especially the importance of ‘fundamental rights risk’ of AI at a recent European conference on ‘A Framework for Trustworthy AI’ - but explained this with the risk of AI in credit scoring. Over the decades, there was a shift is the concept of fundamental human rights from defense rights against oppressive governmental actions to include second-generation rights to subsistence and third-generation solidarity rights. However, there was never a ‘right to get a credit’. There are undisputable legal and ethical questions concerning e.g. biometric recognition and behavioral detection in public spaces (see e.g. Wendehorst and Duller, 2021), and a new report of the United Nations High Commissioner for Human Rights (UNHCR, 2021) summarized [quote]: ‘*[AI] ... affects people’s right to privacy and other rights, including the rights to health, education, freedom of movement, freedom of peaceful assembly and association, and freedom of expression*.’ However, it is doubtful how the same approach could fit with credit-scoring with a <u>risk taken by lenders</u> for future default of borrowers. Of course, credit-scoring has to be compliant with existing laws including non-discrimination and data protection - but lending is always a free decision of the lender in a market economy.

Fully incoherent is the nearly hidden problem that any type of statistical classification for credit scoring would be defined as ‘high-risk’ including all existing practices, which are well established for decades. With that said, it may be a minor - but characteristic - inconsistency that AIA Art. 10 (3) states that [quote]: ‘… *data sets shell be relevant, representative, free of errors and complete’*. While representative data is a fundamental in statistical analysis and - for example - medical studies of new drugs or treatments, ‘*free of errors and complete’* is rather unscientific, as all measurements have errors (statistical and systematical) and no data set is complete (but only representative due to the limitation of finite data).

Nonetheless, one can attempt a synopsis of the regulations [quotes, underlining by the author]:

1. [GDPR recital 71]: ‘*The data subject should have the right not to be subject to a decision, … which is based solely on automated processing and which produces legal effects concerning him or her or similarly significantly affects him or her, such as automatic refusal of an online credit application … without any*

human intervention. ... However, decision-making based on such processing, including profiling, should be allowed where … or necessary for the entering or performance of a contract between the data subject and a controller, or when the data subject has given his or her explicit consent. In any case, such processing should be subject to suitable safeguards, which should include …, to obtain an explanation of the decision reached after such assessment and to challenge the decision. ... In order to ensure fair and transparent processing in respect of the data subject, …, the <u>controller should use appropriate mathematical or statistical procedures for the profiling, …and that prevents, inter alia, discriminatory effects</u> on natural persons on the basis of racial or ethnic origin, political opinion, religion or beliefs, trade union membership, genetic or health status or sexual orientation, <u>or that result in measures having such an effect</u>. Automated decision-making and profiling based on special categories of personal data should be allowed only under specific conditions.'

2. [EU Race Equality Directive and Framework Employment Directive (both) Art. 2, 2.(b]): *'… indirect discrimination shall be taken to occur where an apparently neutral provision, criterion or practice would put persons of a … particular [characteristic] at a particular disadvantage compared with other persons, <u>unless that provision, criterion or practice is objectively justified by a legitimate aim</u> and the means of achieving that aim are appropriate and necessary*.' and Recital (15): *'indirect discrimination to be established by any means including on the basis of statistical evidence*.'

3. [(proposed) AIA, recital 37] concerning credit scoring: *'In particular, AI systems used to evaluate the credit score or creditworthiness of natural persons should be classified as high-risk AI systems, since they determine those persons' access to financial resources or essential services such as housing, electricity, and telecommunication services. AI systems used for this purpose <u>may lead to discrimination of persons or groups and perpetuate historical patterns</u> of discrimination, for example based on racial or ethnic origins, disabilities, age, sexual orientation, or create new forms of discriminatory impacts.'*

It remains unclear, whether the European Commission has been intending a consistent and integrated approach[10]. However, from an isolated perspective of credit scoring, the framework - and ignoring inconsistencies on lower levels - could be disentangled into two main approaches.

While there is no holistic anti-discrimination framework, one can assume that European values prohibit any kind of direct discrimination: whether concerning race /ethnic origins or due to any other sensitive characteristics (either mentioned in the texts or to be assumed for reasons of consistency). On the other side, wordings like '*result in measures having such an effect*' or '*may lead to ... perpetuate historical patterns*' seem to constitute the case of indirect discrimination, although indirect discrimination is defined only in the case of sensitive characteristics of race or ethical origin (and to employment). This implicit extension of the aspect of ‘indirect discrimination’ raises the principal question, whether there is any responsibility without casualty but based on correlation only?

While there is - unfortunately - existing antisemitism in Europe and in Germany, the public discussion in Germany about algorithmic decision-making is centered around the two following issues:

- The so-called ‘equal payment gap’ between men and women and impact of - even simple and rule-based - credit scoring using household income as key indicator (as illustrated above with the example of the Datenethikkommission).
- Use of ‘automated processing’ in general and especially of non-tradition data (e.g. for a ‘second chance’ credit scoring using the regulation of third party access to payment accounts with a customer’s mandate according to PSD2 regulation).

[10] According to the AIA: ‘*For AI systems …by financial institutions regulated …, the market surveillance authority for the purposes of this Regulation shall be the relevant authority responsible for the financial supervision of those institutions under that legislation*.’ This splits the original approach of a cross-industry regulation again into industry-specific supervision. Additionally, the proposed Digital Operational Resilience Act (DORA) of 2020 already includes the regulation of ICT Risk Management and ICT Third-party Risk, which would be a redundancy of regulation of (i) AI in financial services and (ii) ICT risk and ICT outsourcing in financial services.

The public discussion about a possible discrimination of women due to (algorithmic) credit scoring mirrors the fifty years old struggle in the USA for equal rights and especially for equal rights for women applying for a loan. The practice at this time to discriminate women in credit decision was one root for the Equal Credit Opportunity Act of 1974. Looking to Germany (and Europe) today, three issue should be discussed separate, but are sometimes blended.

First, there is the so-called ‘gender pay gap’ in Germany, as women have 18% lower income compared to men in average (i.e. calculated for the whole population without adjustment) according to the German statistics authority ‘Statistisches Bundesamt’ (destatis, 2021), which ranks Germany in those countries with at the largest gap in Europe. One the one side, this average value has to be adjusted in the case of <u>comparable</u> conditions and equivalent qualification, and an adjusted gap was only 6% in 2018 (for ‘statistical twins’). According to Statistisches Bundesamt (destatis, 2021) this value still contains a dependency on occupational biography, which was not adjusted due to missing data. However, it is disturbing that also ‘clean’ data for public university professors of one faculty show that the ‘additional salary’ to the base renumeration for female professors in mechanical engineering is nearly 1800.-€ less compared to male professors (see: Kortendiek, 2021). This very specific analysis underlines that there is a ‘gender pay gap’ in Germany, but more on a specific level and not so much for average statistical twins in the whole population.

Second, the related example of the German Datenethikkommission (2019) that [quote]: ‘*an algorithmic system, …, can generate different distributions for … men and women’* is correct, but not even near discrimination. Of course, if an income distribution with an *ex-ante* imbalance between men and women is used as input for an algorithm, this will produce different outcome results (except for trivial cases). Nonetheless, a differently distributed input for a decision-making system is generated by the reality of the social situation in a country but is not causally linked to any impact. <u>Only if individual cases of decision-making with similar economic condition (and for the same product offered) of applicants - whether men or women - are compared, potential indirect discrimination could be deduced</u>.

Additionally, one standard example for ‘indirect discrimination’ does not fit here: The example of ‘blanket no beards policy’ in food industry, which would exclude e.g. people with Sikh believe (and could be remedied by a policy of wearing hair nets) is an ex-ante rule without any individual evaluation of the (economic) situation of the applicant. Vice versa, any (algorithmic) credit scoring is always using individual economic data about the financial condition of the individual applicant as input for a statistical estimation of predictable future re-payment problems and/or default.

Third, there is a worrying misunderstanding of the fundamentals of statistics, and one non-representative example might be a blog on the Goethe University Frankfurt - SAFE Finance Blog website (Bauer, 2020), which explains [quote]:

> ‘*For example, data with a disproportionately high percentage of women who were unable to repay a consumer loan could lead to an ML model systematically predicting lower repayment probability for women. In an automated system of credit allocation, women would then receive a loan less frequently. This kind of an AI system could therefore increase social inequality and significantly reduce the economic welfare of women*.’

Given the <u>assumption</u> is correct and women had a higher probability for default in the past, then it is trivial - with the statistical hypothesis of continuously on-going processes - that women will receive less approvals for future loans. In this case the credit scoring - whether machine learning (ML) or any straightforward tool based on household income - will simply make an economic estimation for the lender, which has to take the risk. Consequently, any statistical prediction would produce an output <u>distribution</u> of estimated repayment probabilities in accordance to the <u>given</u> input distribution In a free market economy, it is not a task of the lender to change the social reality i.e. different wealth distributions, which exists <u>as assumed</u> above. A lender can only contribute to the economic welfare of an individual applicant (with an unknown gender), if this applicant has a reasonable probability for re-payment. This example is even more strange, as e.g. a quantitative meta-study about “Women and Repayment in Microfinance” (D’Espallier et al., 2011) revealed [quote]: *“… a higher percentage of female clients in MFIs [microfinance institutions] is associated with lower portfolio risk, fewer write-offs, and fewer provisions, all else being equal*.”

This discussion in Germany shares a connotation with the AIA due to an implicit or explicit assumption that credit scoring [quote] '*determine those persons' access to financial resources'*, while in reality credit scoring is a tool of the lender to estimate whether the money of the lender would be re-paid after a credit is approved. In other words: There is no 'right to borrowing' and lending is within the private autonomy (and freedom of contract) of the lender.

However, there is a subtle shift in the public debate from 'casualty' to 'correlation' when it comes to algorithmic decision-making. This shift can be exemplified by a study about 'Discrimination, artificial intelligence, and algorithmic decision-making' [Zuiderveen Borgesius, 2018] published by the Council of Europe [quotes, underlining by the author]:

> *'Most non-discrimination statutes apply only to discrimination on the basis of protected characteristics, such as skin colour. Such statutes do not apply <u>if an AI system invents new classes</u>, which do not correlate with protected characteristics, to differentiate between people. Such differentiation could still be unfair, however, for instance when it reinforces social inequality.'*
>
> *'Suppose, for instance, that poorer people rarely live in the city centre and must travel further to their work than other employees. Therefore, poorer people are late for work more often than others because of traffic jams or problems with public transport. The company could choose "<u>rarely being late often</u>" as a class label to assess whether an employee is "good". But if people with an immigrant background are, on average, poorer and live further from their work, that <u>choice of class label would put people with an immigrant background at a disadvantage</u>, [...]'*

Of course, no existing AI system would be able to '*invent new classes*' (and the labelling of training data is done by human beings), and '*being late often*' is a simple and direct measure for the correct performance of an employment contract with an obligation of the employee. Nonetheless, the understanding in a report published by the Council of Europe (sic!) reveal that a longer construction of

"*Suppose, ... Therefore, ... could ... But if ... would* ..."

is used to show how some algorithm would generate an outcome (i.e. 'being late often' = <u>individual responsibility</u> for non-performance of an employment contract) with a hypothetically correlation to a sub-group.

An anecdotical evidence - but exactly about a ‘second chance’ credit scoring based on payment transaction information with the mandate of the consumer - was the debacle of the ‘Check Now’ field test of Germany's leading credit bureau ‘SCHUFA’, which tried to test such a ‘second chance’ approach in Germany very much the same as recently reported about a pilot program of large bank in the U.S. (Rudegeair and Andriotis, 2021).

In this test, the consumer was asked for two consents: first a consent according to PSD2 for the access to the consumer’s bank account (incl. processing of the data) and a second consent for storing and analyzing the data to collect a long-term data-set. Both consents were independent as (i) the PSD2 consent allows only the processing of the account information without storing and (ii) according to GDPR any consent has to be very specific and decoupled. While such a double consent (with two boxes to be checked) may be a difficult way to understand for consumers, it would be exactly according to the existing regulations GDPR and PSD2.

However, only few days after the press release about the new service and start of a field test with a mobile telco provider with some 100 customers in Nov. 2020, a report in public television (Bognanni et al., 2020) discussed hypothetically possible risks and disadvantages for the consumers. After more negative comments in different media, public television reported (Busch et al., 2021) in March 2021 that the ‘Check Now’ was terminated [quote] ‘*obviously*’ due to the original report.

According to press reports (Börsen-Zeitung, 2021), a subsidiary of Schufa will provide a simple PSD2-compliant access-to-account history service „Girocheck“ with consumer’s consent, but without any correlation to a population. While the acceptance of this limited service - only rule-based inspection of payment history instead of statistical estimation of credit re-payment - has to be monitored, it can be questioned whether it is beneficial for consumers with low credit scoring to be excluded from a possible ‘second chance’.

However, the public awareness in Germany concerning data processing, automated credit scoring, and algorithmic decision-making is prejudiced with constructed scenarios. This assessment is somehow disturbing, because many e-business merchants are using (i) the payment history of existing customers (with their consent) to decide about payment options offered, (ii) partly alternative data such as consumers’ digital footprint (e.g. device type / operation system and e-mail-address provider) to estimate spending patterns and re-payment probabilities (see e.g.: Berg et al., 2019) and (iii) the correlations in the existing data to decide the payment options for new customers (including measures against identity fraud as significant issue with new customers in general). While such fully automated procedures are state-of-the-art in e-business to avoid fraud, similar approaches in automated credit scoring are used to exemplify the negative capabilities of algorithms.

Especially in this case, which is anecdotical but archetypical, the European approach of regulation is quite complicated and difficult to communicate, as there is:

- Cross-sectoral regulation of GDPR, but …
- … with industry specific examples (as '*online credit application'* in recital 71)
- Sectoral regulation of PSD2 with defined 'consent' to access account, but …
- … different concepts of 'consent' in PSD2 vs. GDPR according to a guideline of the European Data Protection Board (EDPB, 2020)
- Erratic approach within the AIA, but
- … with the industry specific example of credit scoring completely independent from any use of AI due to the definition that 'statistical approaches' would be enough for credit scoring to qualify as 'high-risk system'.

These approaches leave wide room for interpretations and does not provide an innovation-friendly framework for digitalization in financial services. It will require future research to monitor how the European regulatory framework will impact the development of financial services in the global competition.

7. China

China is an autocratic regime and command economy, which emulates a market-like development in a try-and-search approach (to fulfil the promises of the regime for economic benefits) until economic agents get near to rival the communist party (including the challenge of those benefits compared to governmental distribution). Consequently, law and regulation do not provide an *ex-ante* framework for economic development but are applied *ex-post* to 're-align' the economy to the regime.

Under these preconditions, one can analyze the recent development in China concerning algorithmic decision-making and credit scoring vs. government regulation.

As developments in China are not always accessible in English language, a series of working papers of the BIS provide insight into advanced credit scoring methods in China: Gambacorta et al. (2020) discussed ‘Data vs collateral’ as already mentioned, Frost et al. (2019) provided insight into the opportunities of small merchant on online-platforms to access credit (from the non-financial platform providers) in China and Latin America (on the Mercado Libre / Mercado Pago platform), and Gambacorta et al. (2019) discussed [quote]: ‘*How do machine learning and non-traditional data affect credit scoring? New evidence from a Chinese fintech firm’*. In the last case, traditional information (credit card information) and non-traditional information (usage of mobile apps and e-commerce) was compared [quote]:

i. *The fintech’s machine learning-based credit scoring models outperform traditional empirical models […] in predicting borrowers’ losses and defaults.*

ii. *Non-traditional information improves the predictive power of the model.*

iii. *While the models perform similarly well in normal times, the model based on machine learning is better able to predict losses and defaults following a negative shock […]*

iv. *The predictive power of all the models improves when the length of the relationship between bank and customer increases. However, the comparative advantage of the model that uses the fintech credit scoring technique based on machine learning tends to decline when the length of the relationship increases.*

Ant Financial, as partial subsidiary of Alibaba, started in 2015 with the ‘Sesame Score’ (Ant Financial, 2015) based on [quote, underlining by the author]:

a. *‘Credit History reflects a <u>user's past payment history and indebtedness</u>, for example credit card repayment and utility bill payments.*

b. *Behavior and Preference reveals a user's <u>online behavior on the websites they visit, the product categories they shop</u>, etc.*

c. *Fulfillment Capacity shows a user's ability to fulfill his/her contract obligations. Indicators include use of financial products and services and <u>Alipay account balances</u>.*

d. *Personal Characteristics examine the extent and accuracy of personal information, for example home address and length of time of residence, mobile phone numbers, etc.*

e. *Interpersonal Relationships reflect the online characteristics of a user's friends and the interactions between the user and his/her friends.'*

The first and the third element resembles financial scores as discussed above, the second and fourth are typical for online merchants (the origin of Alibaba, but unique in the combination of financial and shopping history), and the last element (behaviour in social media) seems questionable. Nevertheless, the other major payment system in China, Tencent's WeChatPay, recently announced an own competing credit score system (Gill, 2020), which should be based on based on consumers' personal and credit records, but also 'habits' of players of online games - one of the traditional business lines of Tencent.

Nonetheless, China strengthened financial services regulation. The Chinese Financial Supervision required the dominating payment platforms AliPay and Tencent/WeChatPay to connect to a central clearing via the Peoples Bank of China in 2017, designed AliPay the status of a 'financial conglomerate' in 2020 by a special legislation applied only to AliPay until now (as far as reported in English language), and announced more regulation in late 2020 (Kharpal, 2020). Legislation and regulation in China is never an *ex-ante* framework, but typically an *ex-post* alignment of actual development to the concepts of the communist party. In 2021, Beijing require Ant Financial to separate the two lending units Huabei (similar to credit card lending) and Jiebei (small consumer loans) into 'independent' companies. According to Financial Times (Yu and McMorrow, 2021), this development is part of a governmental plan to require AliPay to hand over all the user data into a separate credit scoring 'joint-venture', in which state-owned companies would have a majority. In parallel, the central bank (People's Bank of China) in developing a Central Bank Digital Currency (CBDC, in China as Digital Currency/Electronic Payments "DC/EP"), which should be issued by the central bank, operated by commercial banks and other companies, but with a so-called 'backup' of the record of all consumers' payment transactions at a central, governmental entity (Zhou, 2020; Gao, 2020).

In this context, a concerning observation should be noted. As Gerd Gigerenzer (2021) mentioned recently, a survey in western countries revealed a development that people are going to ‘accept’ such surveillance systems: While in 2018 only 10% of the respondents approved this, in 2019 the number doubled, and people explained that they ‘have nothing to hide’.

The actual development towards a ‘communization’ of personal data has to be compared with the formal situation. China included protection of personal data in the Civil Code of the People's Republic of China in 2017 (Winkler, 2021) and continued this development in the current version of the civil code of mid 2020 (Ding et al., 2020). Although direct translations from Mandarin could be misleading due to the specific cultural and language background - data for example: 数据 - shùjù means basically ‘[something] concerning numbers’, which is a much more ‘digital’ approach compared to Western languages - the work of Ding et al. (2020) included an in-depth understanding of Chinese language and legal tradition and has to be highly appreciated.

The civil rights code included:

- General civil rights: §111 protection of personal information, §127 data and asset protection on the internet, §128 protection of special people (children, elderly persons, disabled persons, women, consumers)
- Chapter 6 / right of privacy and protection of personal information: §1032 protection of private space (which is more specific than ‘anti-discrimination’ as it protects every person / persons’ information independent of any definition of sensitive characteristics),
  §1034 protection of personal information, §1035 requirement of legal processing of personal information (including requirement for consent, i.e. similar to GDPR, but in few lines) plus: §1037 rights of natural persons against information processors, $1038 obligation of information processors,

This approach is more ‘general’ compared to European GDPR or even U.S. sector-specific regulation and does not discuss hypothetical disadvantages for consumers. On the other side, the reality in China is not the written legislation, but the actual handling of legislation by the regime and (government-depending) courts. In this sense, Chinese regulation may help consumers on a micro-social level but has always an overall governmental goal.

China claims financial inclusion (see especially: Mu, 2021) for non- or underbanked in a developing country with high efficiency - but always under the precondition of the supremacy of the regime of the communistic party.

All-in-all, China develops a rather political approach to regulation - always from the point of view of a command economy under supervision of the communistic party. In the past, a trail-and-error approach gave much room to the development of innovative payment platforms (i.e. AliPay and WeChatPay) and their development into credit scoring connected to consumer and merchant finance. However, the current development with the 'full-stop' on the development of AliPay - including suspension of the planned IPO and separation of the profitable lending business - makes clear that the regime in China does not support independent legislation, but legislation is part of the communistic party's autocratic regime.

8. Discussion: Regulation versus Innovation

The approaches in the USA, Germany/Europe, and China reveal that any regulation of (algorithmic) credit scoring - whether traditional statistical or advanced AI approaches, although always statistical classifiers - is no isolated 'technical' issue but always an issue of the socio-technical context and the political situation. With much simplification, the three situations could be characterized by:

USA: long-term systemic discrimination, fifty year of <u>sector-specific regulation with limited improvements</u> in the lending sector (unfortunately), but openness to 'use of alternative data in credit underwriting' (CFPB, 2019)

Germany/Europe: no systemic discrimination concerning lending (but continuing anti-semitism not relevant for credit scoring) but <u>public debated with constructed cases for hypothetical (indirect) discrimination, cross-industry regulation</u> with the exception of the current proposal for AIA, and a current tendency to a subtle 'indirect discrimination' regulation not based on causality, which could render even pure statistical credit scoring as 'high-risk systems' without any evidence (and without any justification for such an additional regulation)

China: formally more case-based regulation, but in general a trail-and-error approach of an <u>autocratic regime with benefits for non-/underbanked consumers</u> during the last years based with financial inclusion on the one side but increasing control in a command economy on the other side.

Consequently, these different approaches correlate to different perspectives on the schematic six-step process for algorithmic decision-making as illustrated in Figure 1:

- Sector-specific regulation (like in the USA) allows supervisors to discuss the ‘use of alternative data’ with respect to specific steps in decision-making. Especially in step 3, alternative data such as payment transaction history can create possible improvements due to an enhancement of the ‘strength-of-knowledge’ with additional insight especially for those potential borrowers with thin credit files as a ‘second chance’ (i.e. non-sufficient information to calculate a statistical estimation to estimate the lender’s risk). However, recent studies revealed that this can help some people but cannot remedy systemic discrimination in a society.
- Cross-industry approaches cannot take into account specific ‘internal’ aspects of decision-making *per se*. This is reasonable for anti-discrimination regulation (against direct discrimination, i.e. usage of sensitive characteristics) based on individual cases, while any statistical approach is limited to correlations between input (step 1) versus output (step 5) without any insight into causal relations. As illustrated in a simplifies way in Figure 3, two selected groups within the population with different distribution of household income will - in the simplified model - lead to different distributions of approved loans after credit scoring with a threshold on household income, i.e. the decision-making causes a trivial non-linear transformation (cut-off) but without any casualty. It can be defined as ‘*perpetuate historical patterns*’ that this transformation (i) has no impact on the income distribution due missing causal interaction and (ii) the outcome will differ between these two groups as already the input differed. However, a focus of the public perception on such correlation without casualty is rather characteristic for Germany, and the Schufa case is an illustrative - while anecdotical - example for the construction of hypothetical danger for the consumers.
- The specific approach in China with the extreme ambivalence between global leadership in the use of non-conventional data for credit scoring (with benefits especially for those consumers with weak banking history or underbanked population) and an ‘social scoring’ (see e.g. Chen, 2021) of an autocratic regime raises strongest concerns. Nonetheless, the work of Gambacorta et al.

(2019) revealed that non- or underbanked people at the beginning of a credit relationship (either with an e-business merchant or with a bank) benefit from innovative machine learning approach for credit decision-making and financial inclusion.

Especially the debate in Europe about a possible ‘discrimination’ in credit scoring due to the gender pay gap resembles with well-known example of the so-called Berkeley admission paradox (see Pearl, 2018, for more details, and Figure 4 with a causal diagram). In 1973, the university of California found that 44% of the men who applied at Berkeley graduate school were accepted, compared to 35% of the women. Was there a direct effect, i.e. did the university discriminate women?

An analysis found that on the level of departments there was no discrimination (even some departments accepted more women compared to men). But a higher proportion of women applied to the humanities and social sciences with a higher number of applicants at all and a smaller number of places. In total, fewer women were accepted at university level, as women preferred to apply to departments with a lower probability for acceptance at all (independent of gender), whereas men preferred engineering et cetera. The causal relation was straightforward as ‘higher proportion of women’ * ‘lower probability at all’ = lower acceptance rate. Of course, there was (and still is) a bias in the society as women prefer an education in ‘soft’ sciences, but with many potential reasons (i.e. unobservable so-called mediators) for an individual decision for those ‘soft’ sciences, which are out-of-scope to be controlled by the university. If we as society want to get women to study ‘hard’ sciences (‘MINT’), the society can set-up campaigns, do marketing for those departments, or consider a ‘Rooney Rule’ to encourage more women in the final selection step (without any fixed quota). Nonetheless, the university has no responsibility of for any <u>individual</u> decision! Likewise, no lender is responsible for the household income of a potential borrower (but has a responsibility to decline all people, who would run into over-indebtedness).

9. A Final Remark Concerning ‘Artificial Intelligence’

As a final remark concerning ‘artificial intelligence’ should be made: Did the development of the recent years in artificial intelligence, machine learning and so-called deep learning change the understanding of decision-making in general and credit scoring in particular? The shortened answer is simply: no!

Nearly all contemporary AI tools (with few exemptions, which have no application in credit scoring up to now) are - still - statistical classifiers or are in the words of Pearl (2018) ‘*able to fit a function to a collection of historical data points*’. Although the fit function of an artificial neural network is more complicated compared to traditional classifiers (see e.g. Milkau, 2021) and leading-edge ‘generative pre-trained transformer’ (GPT) as currently *en vogue* in language processing may have multi-billions of parameters to be ‘fitted’, the general principle of AI is illustrated in Figure 5. While in Figure 2 a one-dimensional distribution with ‘blue’ and ‘orange’ events was discussed, Figure 5 shows a two-dimensional distribution of two types of ‘blue’ events with ‘+’ and ‘orange’ events with ‘-’, which are used as input data set to train some traditional AI approaches to achieve a classification.

Without much theory, it is clear that only within the original scope of input data (within an envelop containing all ‘blue’ and ‘orange’ events) any classification of a future event as statistical estimation is reasonable, or vice versa the input data set has to be ‘representative’ for the future scope. However, any real measurement of real-world data will include some errors (noise in the input or simple measurement problems) as indicated with the ‘orange’ ‘-’ events at the bottom left. Consequently, any requirement for ‘error free data’ is not applicable to any real-world data set. Likewise, it is clear that for any event outside the original scope (marked as ‘?’) not consistent classification is feasible.

Additionally, the graph illustrates that within the scope different methods (i.e. different types of multidimensional ‘fit functions’) provide different statistical qualities (e.g. better sensitivity vs. better specificity). It depends on the objective function, which method or ‘fit function’ will deliver the best results, but there is no ‘best’ ‘fit functions’ at all. Without (i) a defined scope and (ii) defined objective function, no selection of a suitable method would be possible.

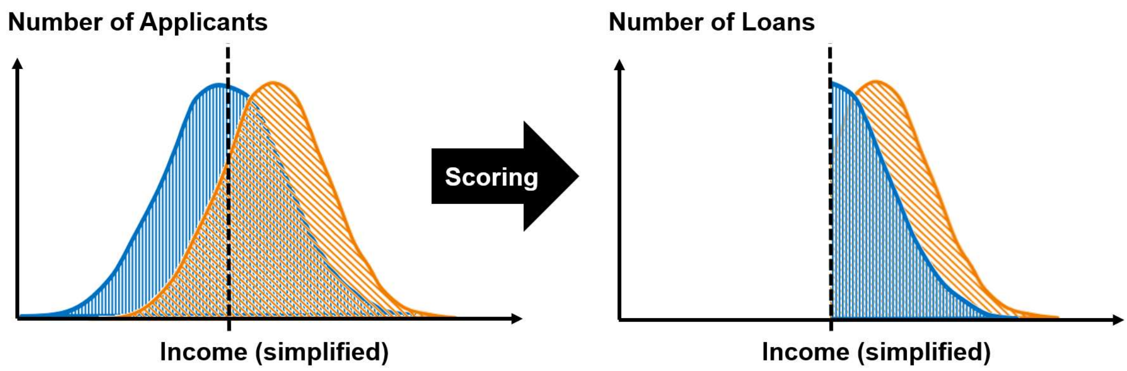

Figure 3: Correlation between Input (step 1) and Output (step 5) for credit decision-making based on the (very) simplified model with household income as only parameter in the case of two groups defined by some selected characteristic, which has no causal relationship to the credit scoring but with different income distributions of these groups. In this (very) simplified illustration the 'orange' group will have 2/3 positive approvals and the 'blue' group will have only 1/2 positive approvals, because the 'blue' group has a income distribution with a lower average value

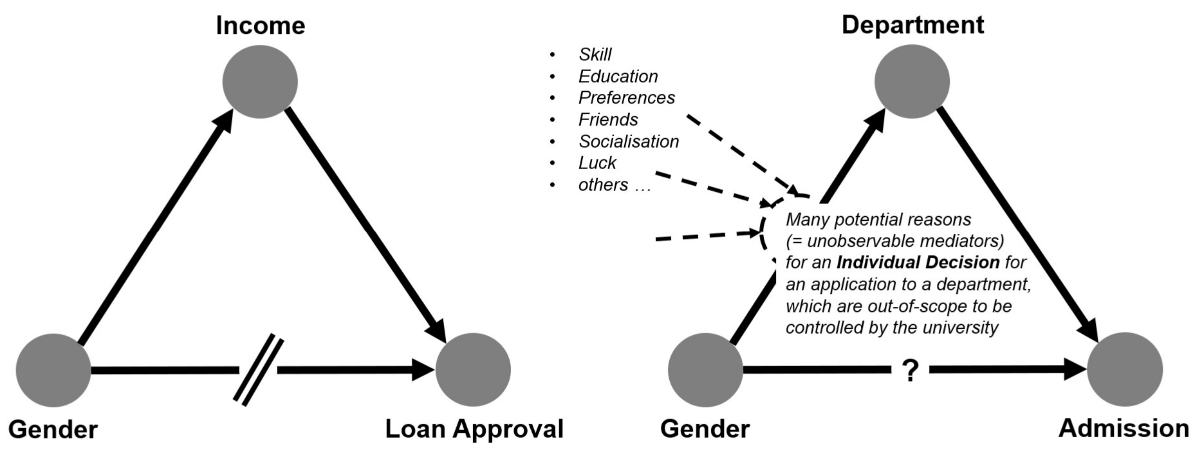

Figure 4: Comparison between the (non-existent) effect of the gender pay gap on credit scoring (left) with the Berkeley admission paradox (right) in a causal diagram (details see text)

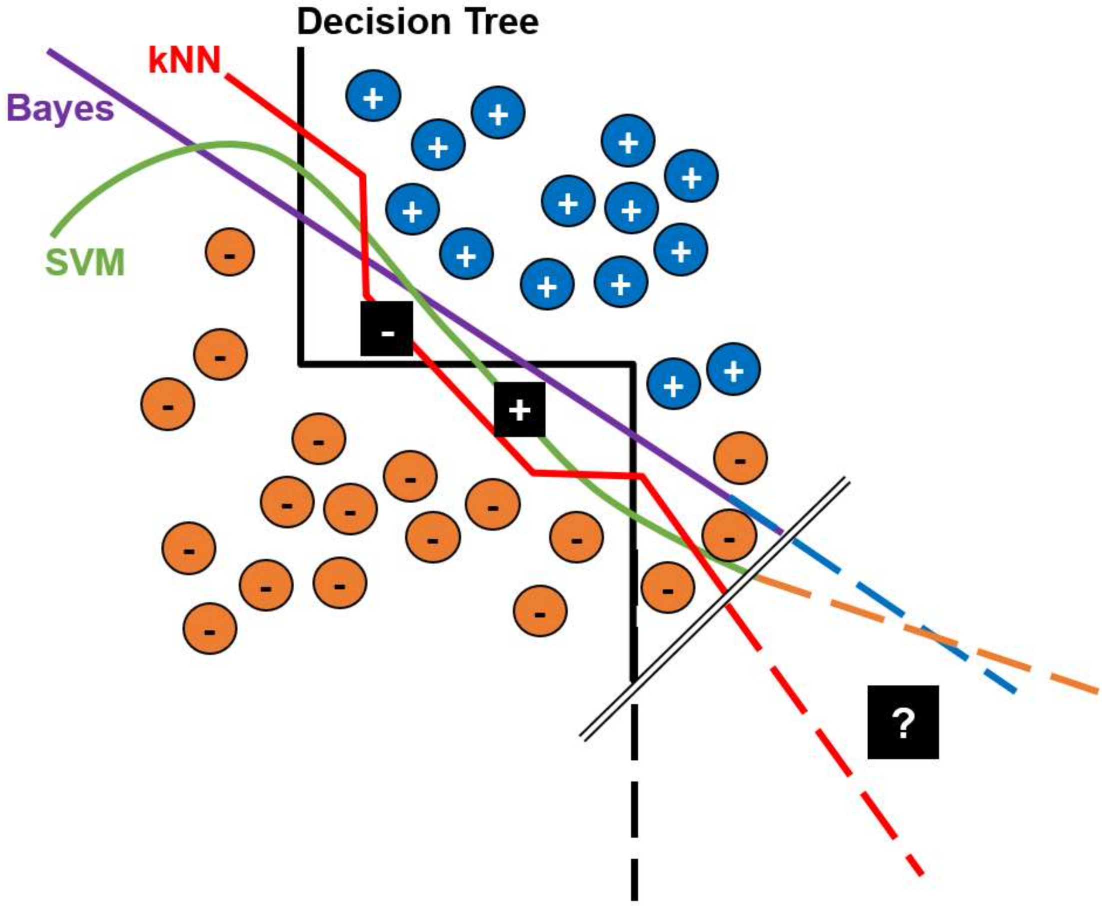

Figure 5: Illustration of machine leaning as 'statistical classifier' with the examples of Support Vector Machines (SVM), k-Nearest Neighbour (kNN), Naive Bayes, and Decision Trees. The training data are 'blue' circles with '+' and 'orange' circles with '-' and new events as squares. Events near the frontiers (or hyperplane) can cause statistical estimations depending on the selected method. Events outside the scope of the training data set (market as '?') exceed the scope of the classifiers.

The figure exemplifies how the input dataset (recorded data for training of the AI as a statistical classifier) determines the scope of the classifier for a new event.

The graph is adopted from [Domingos, 2012].

These principles of statistics apply to every statistical classification whether done by traditional approaches, conventional artificial intelligence (as in Figure 5), or advanced ‘deep learning’ - and likewise any statistical classifier requires a diligent approach for - *inter alia* - representativeness or, vice versa, no naivety in data analysis[11]. Nevertheless, a proper understanding of the decision-making process (see Figure 1) and all relevant parameters is essential to apply statics right. Vice versa, it remains unclear why the AIA proposal did not specify statistical quality criteria, as these criteria are standard scientific concepts. Consequently, the AIA proposal is neither transparent and consistent regulation, nor good statistical science. Any regulation should be agnostic towards technologies and methods, but either establish general principles cross-industry or specific approach to defined industrial products or services with a clear justification by measurable evidence that this regulation is necessary.

10. Conclusion

The simplified assessment in this paper shows that neither sector-specific regulation and supervision (in case of the USA with systemic discrimination as social reality) nor omnibus regulation (in Europe and especially with the AIA proposal following the *Zeitgeist* but leading to legal uncertainty) do fit the challenges of digitalization in the 21st century. Strangely, the autocratic approach in China (with ‘a laissez faire’ approach until any challenge for the regime emerges) seems to be rather beneficial for consumers and ‘effective’ concerning inclusion in the digital economy.

While no test of advanced technologies to support credit scoring and to provide a ‘second chance’ for consumers with insufficient traditional credit score values seems to be feasible in Europe - especially with the proposal of AIA qualifying any algorithmic credit scoring as ‘high-risk’ - the recently reported pilot of large U.S. banks (Rudegeair and Andriotis, 2021) to test the use of account transaction data could be a benchmark to help financial inclusion and provide benefit for non-/underbanked people with state-of-the-art technology.

[11] Illustrating examples were provided in recent work of Heinrich-Hertz-Institut (HHI, 2019) for pattern recognition with artificial neural networks. Although images could be classified correctly, a tool can lack reliability when context determines the outcome, as for example ‘ships’ were classified due to surrounding water, ‘trains’ due to railways, or ‘horses’ due to copyright watermarks on the images (as training pictures with horses came from a source with such watermarks). No ‘artificial intelligence’ can - for the time being - replace human intelligence especially concerning gathering, selection and preparation of data.

On another side there is an increasing scepticism concerning technology, fear of ‘autonomous’ artificial intelligence, and concerns about algorithms, which could ‘petrify’ inequality in the society persistently. This debate touches many aspects of our social and economic life and reveals much about the complicated socio-technical nexus of the digital age. A summary of the examples given in this paper is gathered in the appendix (below) as a synopsis of different situation, problem and potential solutions.

Future quantitative research with real applications of advanced credit scoring technology instead of constructed possibilities of hypothetical harm is needed, but such tests would require an openness to technology to be tested before attributed as ‘high-risk’ without evidence.

Remark

This article reflects the author’s opinions only.

Appendix: Comparison of the different examples

Situation	Problem	Potential Solution
Un- or underbanked minorities in the USA (BIS, 2020) with non or 'thin' credit files	Historical legacy of systemic discrimination in the USA	Issue of the society / systemic discrimination Second chance with 'alternative data in credit underwriting' (CFPB, 2019) i.e. payment transaction history especially for 'false negative' results
'Rooney Rule' of the U.S. NFL	Low representation of African-Americans in head coaching positions	Protocol - instead of a fixed quota - that at least one of the finalists for an open position should be from an underrepresented minority
Findings on the Apple Card	Misunderstanding, while 'underwriters are not required to treat authorized users the same as account holders, and may consider many other factors.' (DFS, 2021)	In terms of gender … that Apple Card applications from women and men with similar credit characteristics generally had similar outcomes. (DFS, 2021)
Berkeley admission paradox (see Pearl, 2018)	In total, fewer women were accepted at university level, as women preferred to apply to departments with a lower probability for acceptance at all (independent of gender), whereas men preferred engineering et cetera.	Issue of the society that more women apply to 'soft' sciences but avoid 'hard' sciences ('MINT'). The university decided on objective criteria for comparable situations (i.e. application for same faculty)
Credit Scoring* with Gender Pay Gap (model in this paper)	Gender Pay Gap in Germany as example for the social problem of different payments (even for an adjusted case with comparable job situations)	Issue of the society that there is - still - a Gender Pay Gap, but the credit decision will be similar for similar economic situations ('statistical twin').
'Matching' of current application to an existing / historical structure** (e.g. in recruitment of new employees)	Hiring of new employees 'like the existing employees' is dependent on the history and sensitive to (i) discrimination and (ii) statistical noise leading to clustering.	Awareness for the problem (risk of clusters) and existing EU regulation (Framework Employment Directive, 2000). Independent decisions can be made based on objective criteria or assessments

*) The credit decision (if-the-else) is based on an estimation of re-payment (or vice versa default of a consumer), which is derived from a data-set of financial conditions versus recorded re-payments/defaults, but not from former credit decisions. The score value is an 'independent and identically distributed' variable (i.i.d.), which makes every decision independent from other decisions and generates similar decisions for similar values (economic situation).

**) As the hiring decision depends on a 'match' to a given structure (or portfolio) based on former decisions, this is no independent process, and the central limit theorem (for independent events) does not apply.

References

Jones, David (2006) 'The Bankers of Puteoli: Finance, Trade and Industry in the Roman World: Financing Trade and Industry in the Roman World', Tempus Publishing Limited, Feb. 1, 2006.

Longerich, Peter (2021) 'Antisemitismus: Eine deutsche Geschichte' (in German), Siedler Verlag, München, 19. April 2021.

European Commission (2021) ‘Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (ARTIFICIAL INTELLIGENCE ACT) and Amending Certain Union Legislative Acts’, Brussels, April 21, 2021 (available at: https://op.europa.eu/en/publication-detail/-/publication/e0649735-a372-11eb-9585-01aa75ed71a1, accessed 6.5.2021).

Mau, Steffen (2020) "Numbers matter! The society of indicators, scores and ratings", International Studies in Sociology of Education, Vol.29/1-2, pp. 19-37.

Eubanks, Virginia (2017) “Automating Inequality. How high-tch tools profile, police and punish the poor” St. Martins Press, New York, USA.

Datenethikkommission (2019) ‘Gutachten der Datenethikkommission’, Datenethikkommission der Bundesregierung - Bundesministerium des Innern, für Bau und Heimat, 23.10.2019, p. 168 (available at: https://www.bmi.bund.de/SharedDocs/downloads/DE/publikationen/themen/it-digitalpolitik/gutachten-datenethikkommission.pdf, accessed 14.4.2021).

Pearl, Judea with Dana Mackenzie (2018) ‘The Book of Why’, Basic Books/Hachette Book Group, N.Y., 2018 (15.5.2018).

Mitchell, Shira, Eric Potash, Solon Barocas, Alexander D’Amour and Kristian Lum (2021) 'Algorithmic Fairness: Choices, Assumptions, and Definitions', Annual Review of Statistics and Its Application, Vol. 8, pp. 141-163 (available at: https://www.annualreviews.org/doi/pdf/10.1146/annurev-statistics-042720-125902, accessed 20.3.2021).

Kleinberg, Jon, Sendhil Mullainathan and Manish Raghavan (2017) 'Inherent Trade-Offs in the Fair Determination of Risk Scores', 8th Innovations in Theoretical Computer Science Conference, ITCS 2017 (available at: https://drops.dagstuhl.de/opus/volltexte/2017/8156/pdf/LIPIcs-ITCS-2017-43.pdf, accessed 20.3.2021).

Karimi, Amir-Hossein, Gilles Barthe, Bernhard Schölkopf, Isabel Valera (2021) 'A survey of algorithmic recourse: definitions, formulations, solutions, and prospects', arXiv, 1. March 2021 (version 2 available at https://arxiv.org/abs/2010.04050, accessed 17.10.2021).

Kügelgen, Julius von, Amir-Hossein Karimi, Umang Bhatt, Isabel Valera, Adrian Weller and Bernhard Schölkopf (2021) 'On the Fairness of Causal Algorithmic Recourse', ICML 2021 Workshop on Algorithmic Recourse, July 2021 (version 4 available at https://arxiv.org/abs/2010.06529, accessed 17.10.2021).

Obermeyer, Ziad, Brian Powers, Christine Vogeli, and Sendhil Mullainathan (2019) "Dissecting racial bias in an algorithm used to manage the health of populations", Science, Vol. 366, pp. 447–453, 25.10.2019.

Stigler, George (1971) 'The Theory of Economic Regulation', Bell Journal of Economics and Management Science, Vol. 3, pp. 3–18.

Aggarwal, Nikita (2021) 'New Morality of Debt', Finance & Development, IMF, March 2021 (available at: https://www.imf.org/external/pubs/ft/fandd/2021/03/pdf/new-morality-of-debt-aggarwal.pdf, accessed 16.3.2021).

FTC (2021) 'A Special Note To Women', Consumer Information - Your Equal Credit Opportunity Rights (available at: https://www.consumer.ftc.gov/articles/0347-your-equal-credit-opportunity-rights#women, undated, accessed 30.3.2021).

CFPB (2019) 'Interagency statement on the use of alternative data in credit underwriting', Consumer Financial Protection Bureau, Dec. 3, 2019 (available at: https://files.consumerfinance.gov/f/documents/cfpb_interagency-statement_alternative-data.pdf, accessed 22.3.2021).

Rudegeair, Peter and AnnaMaria Andriotis (2021) 'JPMorgan, Others Plan to Issue Credit Cards to People With No Credit Scores', Wall Stree Journal, May 13, 2021 (available at: https://www.wsj.com/articles/jpmorgan-others-plan-to-issue-credit-cards-to-people-with-no-credit-scores-11620898206, accessed 17.5.2021).

Gambacorta, Leonardo, Yiping Huang, Zhenhua Li, Han Qiu and Shu Chen (2020) 'Data vs collateral', BIS Working Papers, No 881, Sep. 1, 2020 (available at: https://www.bis.org/publ/work881.htm, accessed 23.3.2021).

Dang, Elaine, Samuel Huang, Adrian Kwok, Harrison Lung, Michael Park and Emily Yueh (2020) 'COVID-19 and advancing Asian American recovery', McKinsey &

Company, Aug. 6, 2020 (available at: https://www.mckinsey.com/industries/public-and-social-sector/our-insights/covid-19-and-advancing-asian-american-recovery, accessed 3.4.2021).

BIS (2020) 'III. Central banks and payments in the digital era', Annual Economic Report 2020 by chapter, pre-release 24.6.2020 (available at https://www.bis.org/publ/arpdf/ar2020e3.pdf, accessed 25.6.2020).

Quillian, L., Lee, J.J. and Honoré, B. (2020) 'Racial Discrimination in the U.S. Housing and Mortgage Lending Markets: A Quantitative Review of Trends, 1976–2016', Race and Social Problems, Vol. 12, pp. 13–28.

Department of Justice (2012) 'Justice Department Reaches Settlement with Wells Fargo Resulting in More Than $175 Million in Relief for Homeowners to Resolve Fair Lending Claims', U.S. Department of Justice, Office of Public Affairs, Press Release, July 12, 2012 (available at: https://www.justice.gov/opa/pr/justice-department-reaches-settlement-wells-fargo-resulting-more-175-million-relief, accessed 24.3.2021).

NAS (2021) 'Motivating Local Climate Adaptation and Strengthening Resilience: Making Local Data Trusted, Useful, and Used', National Academies of Sciences, Engineering, and Medicine, The National Academies Press, Washington, USA (available at https://doi.org/10.17226/26261, accessed 14.9.2021).

Blattner, Laura and Scott Nelson (2021) "How Costly is Noise? Data and Disparities in Consumer Credit", 17.5.2021 (available at: https://arxiv.org/abs/2105.07554, accessed 19.6.2021).

Cressey, Donald R. (1953) 'Other people's Money: A study of the social psychology of embezzlement', Glencoe/Free Press.

DFS (2021) 'DFS Issues Findings on the Apple Card and its Underwriter Goldman Sachs Bank', New York State Department of Financial Services, March 23, 2021 (available at: https://www.dfs.ny.gov/reports_and_publications/press_releases/pr202103231, accessed 22.3.2021).

Milkau, Udo (2021) 'The Risks of Artificial Intelligence used for Decision-making in Financial Services', Capco Journal of Financial Transformation, Vol. 52, 6.1.2021, pp. 128-141.

KPMG (2021) 'Algorithmic bias and financial services - A report prepared for Finastra International', March 2021 (available at: https://www.finastra.com/sites/default/files/documents/2021/03/market-insight_algorithmic-bias-financial-services.pdf, accessed 10.3.2021).

Fuster, Andreas, Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther (2020) 'Predictably Unequal? The Effects of Machine Learning on Credit Markets', October 1, 2020 (available at: https://ssrn.com/abstract=3072038, accessed 22.3.2021), earlier version as CEPR Discussion Papers 12448 (2017).

Wang, J. Christina and Charles B. Perkins (2019) 'How magic a bullet is machine learning for credit analysis? An exploration with FinTech lending data, Working Papers, No. 19-16, Federal Reserve Bank of Boston, Boston, MA (available at: http://hdl.handle.net/10419/227931, accessed 14.3.2021).

Jagtiani, Julapa and Catharine Lemieux (2019) "The Roles of Alternative Data and Machine Learning in Fintech Lending: Evidence from the LendingClub Consumer Platform, Federal Reserve Bank of Philadelphia, Working Paper 18-15, April 2018 - Revised January 2019 (available at: https://www.philadelphiafed.org/consumer-finance/the-roles-of-alternative-data-and-machine-learning-in-fintech-lending, accessed 14.7.2021).

Eccles, Peter, Paul Grout, Paolo Siciliani and Anna (Ania) Zalewska (2021) 'The impact of machine learning and big data on credit markets', Bank of England, Staff Working Paper No. 930, 9.7.2021 (available at: https://www.bankofengland.co.uk/working-paper/2021/the-impact-of-machine-learning-and-big-data-on-credit-markets, accessed 14.7.2021).

FED (2017) 'Federal Fair Lending Regulations and Statutes - Overview', Dec. 31, 2017 (available at: https://www.federalreserve.gov/boarddocs/supmanual/cch/fair_lend_over.pdf, accessed 29.3.2021).

Taylor, Winnie F. (2018) 'The ECOA and Disparate Impact Theory: A Historical Perspective', Journal of Law and Polic, Vol. 26/2, pp. 575-635 (available at: https://brooklynworks.brooklaw.edu/jlp/vol26/iss2/3; accessed 30.3.2021).

Michaelson, Andrew, Katherine Kirkpatrick, Brian M. Thavarajah and Margaret D. McPherson (2020) 'The Implications of a Revived Disparate Impact Doctrine Under a Biden CFPB', King & Spalding LLP, Dec. 17, 2020 (available at:

https://www.kslaw.com/news-and-insights/the-implications-of-a-revived-disparate-impact-doctrine-under-a-biden-cfpb, accessed 30.3.2021).

Marshall, Mickey (2020) 'Re: Request for Information on the Equal Credit Opportunity Act and Regulation B', Independent Community Bankers of America, Letter to CFPB, Dec. 1, 2020 (available at: https://www.icba.org/docs/default-source/icba/advocacy-documents/letters-to-regulators/comments-on-equal-credit-opportunity-act-and-reg-b.pdf, accessed 29.3.2021)

Kleinberg, J. and Raghavan, M. (2018) 'Selection Problems in the Presence of Implicit Bias', ITCS 2018, 12.1.2018, Leibniz International Proceedings in Informatics, available at: https://drops.dagstuhl.de/opus/volltexte/2018/8323/pdf/LIPIcs-ITCS-2018-33.pdf (accesses 21.1.2021).

Langenbucher, Katja and Patrick Corcoran (2021) 'Responsible AI credit scoring – a lesson from upstart', forthcoming in: European Company and Financial Law Review, 2021.

Viola, Roberto (2021) "A Human-Centric Approach to AI", Director-General of DG CONNECT, European Commission, interview at "A FRAMEWORK FOR TRUST-WORTHY AI" conference, 14.9.2021.

Wendehorst, Christiane and Yannic Duller (2021) 'Biometric Recognition and Behavioural Detection', Study requested by the JURI and PETI committees, 4.8.2021 (available at: https://www.europarl.europa.eu/RegData/etudes/STUD/2021/696968/IPOL_STU(2021)696968_EN.pdf, accessed 14.9.2021).

UNHCR (2021) "The right to privacy in the digital age", Report of the United Nations High Commissioner for Human Rights, 13.9.2021 (available at: https://www.ohchr.org/EN/HRBodies/HRC/RegularSessions/Session48/Documents/A_HRC_48_31_AdvanceEditedVersion.docx, accessed 16.9.2021).

destatis (2021) 'Gender Pay Gap 2020: Frauen verdienten 18 % weniger als Männer', Statistisches Bundesamt, 9.3.2021 (available at: https://www.destatis.de/DE/Presse/Pressemitteilungen/2021/03/PD21_106_621.html, accessed 1.4.2021).

Kortendiek, Beate (2021) 'Mehr Fairness bei Leistungsbezügen', Forschung & Lehre, Vol. 3/2021, 26.2.2021 (available at: https://www.forschung-und-lehre.de/karriere/professur/mehr-fairness-bei-leistungsbezuegen-3529/, accessed 1.4.2021).

Bauer, K. (2020) 'Algorithmic discrimination in the financial sector', SAFE Finance Blog, Leibniz Institute for Financial Research SAFE, 6.10.2020, available at: https://safe-frankfurt.de/news-latest/safe-finance-blog/details/algorithmic-discrimination-in-the-financial-sector.html (accesses 21.1.2021).

D'Espallier, B., Guérin, I. and Mersland, R. (2011) 'Women and Repayment in Microfinance: A Global Analysis', Rural Microfinance and Employment project (RUME), World Development, Vol. 39/5, May 2011, pp. 758-772.

Zuiderveen Borgesius, F. (2018) 'Discrimination, artificial intelligence, and algorithmic decision-making', Council of Europe, Directorate General of Democracy, Strasbourg (available at: https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73, accessed 26.5.2020).

Bognanni, Massimo, Lea Busch und Peter Hornung (2020) 'Schufa will Kontoauszüge durchleuchten' (in German), Tagesschau, 26.11.2020 (available at: https://www.tagesschau.de/investigativ/ndr-wdr/schufa-111.html, accessed 30.3.2021).

Busch, Lea, Peter Hornung und Massimo Bognanni (2021) 'Schufa stellt "CheckNow" ein' (in German), Tagesschau, 24.3.2021 (available at: https://www.tagesschau.de/investigativ/ndr-wdr/schufa-checknow-einstellung-101.html, accessed 30.3.2021).

Börsen-Zeitung (2021) ‚Schufa-Fintech bringt den digitalen Kontocheck', 8.6.2021 (available at: https://www.boersen-zeitung.de/banken-finanzen/schufa-fintech-bringt-den-digitalen-kontocheck-25669234-c85c-11eb-acaa-66fcf86d8c5c (restricted); accessed 9.6.2021).

Berg, Tobias, Valentin Burg, Ana Gombović and Manju Puri (2019) 'On the Rise of FinTechs: Credit Scoring Using Digital Footprints', The Review of Financial Studies, Vol. 33 (2020), pp. 2845–2897.

EDPB (2020) 'Guidelines 06/2020 on the interplay of the Second Payment Services Directive and the GDPR, Version 2.0, adopted on 15 December 2020, European Data Protection Board (available at:

https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_202006_psd2_afterpublicconsultation_en.pdf, accessed 5.4.2021).

Gambacorta, Leonardo, Yiping Huang, Han Qiu and Jingyi Wang (2019) 'How do machine learning and non-traditional data affect credit scoring? New evidence from a Chinese fintech firm', BIS Working Papers, No. 834, Dec. 19, 2019 (available at: https://www.bis.org/publ/work834.htm, accessed 5.4.2021).

Frost, Jon, Leonardo Gambacorta, Yi Huang, Hyun Song Shin and Pablo Zbinden (2019) 'BigTech and the changing structure of financial intermediation', BIS Working Papers, No. 779, April 8, 2019 (available at: https://www.bis.org/publ/work779.htm, accessed 5.4.2021).

Brunnermeier, Markus K. (2021) 'The Resilient Society', 23.8.2021, Endeavor Literary Press, Colorado Springs, CO, USA.

Wuermeling, Joachim (2021) 'Exploring DORA – the Digital Operational Resilience Act and its impact on banks and their supervisors', Speech at the European Savings and Retail Banking Group, 23.9.2021 (available at: https://www.bundesbank.de/en/press/speeches/exploring-dora-the-digital-operational-resilience-act-and-its-impact-on-banks-and-their-supervisors-876006, accessed 26.9.2021).

Clinton, Bill (1995) Introduction to "The National Homeownership Strategy: Partners in the American Dream", 2.5.1995, U.S. Department of Housing and Urban Development (available at: https://globalurban.org/National_Homeownership_Strategy.pdf, accessed 11.9.2021).

Ant Financial (2015) ‘Ant Financial Unveils China’s First Credit-Scoring System Using Online Data’, 28.1.2015 (available at: https://www.alibabagroup.com/en/news/article?news=p150128; accessed 15.5.2020).

Gill, Chris (2020) ‘WeChat launches personal credit rating for 600m users’, Asia Times Financial, May 12, 2020 (available at: https://www.asiatimesfinancial.com/wechat-launches-personal-credit-rating-for-600m-users, accessed 6.4.2021).

Kharpal, Arjun (2020) 'China regulator puts country’s fintech giants on notice hinting at more rules', CNBC, Dec. 8, 2020 (available at: https://www.cnbc.com/2020/12/08/china-regulator-hints-at-further-regulation-for-fintech-giants.html, accessed 6.4.2021).

Yu, Sun and Ryan McMorrow (2021) 'Beijing to break up Ant’s Alipay and force creation of separate loans app', Financial Times, 13.9.2021 (verfügbar unter https://www.ft.com/content/01b7c7ca-71ad-4baa-bddf-a4d5e65c5d79, abgerufen am 14.9.2021).

Zhou, Xiaochuan (2020) 'China's choices for a digital currency system', seminar at Peking University on Nov. 27, 2020 (available at: https://asia.nikkei.com/Spotlight/Caixin/Zhou-Xiaochuan-China-s-choices-for-a-digital-currency-system, 22.2.2021; accessed 6.3.2021).

Gao, Dingxin (2020) 'DC/EP and e-RMB in China for Payments Modernization in the Digital Era', presentation at The Sino-German Center of Finance and Economics (SGC) at Frankfurt School of Finance & Management, 2.12.2020 (available at: https://sgc.frankfurt-school.de/wp-content/uploads/2020/12/Dingxin-Gao.pdf, accessed 6.3.2021).

Gigerenzer, Gerd (2021) Interview in Frankfurter Allgemeine Sonntagszeitung (in German), Oct. 10, 2021, page 40.

Winkler, Vincent (2021) ‘Rechte an Daten im Zivilrecht‘ (in German), Mohr Siebeck, Tübingen, 2021.

Ding, Yijie, Nils Klages, Peter Leibküchler, and Knut Benjamin Pissler (2020) 中华人民共和国民法典 (German Translation of the 'Civil Code of the People's Republic of China', promulgated on May 28, 2020), German Journal of Chinese Law (ZChinR), Vol. 27, No. 3 & 4, pp. 207-417, December 2020, Max Planck Private Law Research Paper No. 20/22 (available at SSRN: https://ssrn.com/abstract=3747931, accessed 7.4.2021).

Mu, Changchun (2021) 'Navigating the uncharted waters of retail CBDC', panel discussion, BIS Innovation Summit 2021(https://www.bis.org/events/bis_innovation_summit_2021/agenda.htm, accessed 5.4.2021).

Chen, Yongxi (2021) 'Die Entwicklung des Social Credit Systems: zu drei seiner Funktionen aus Sicht des Öffentlichen Rechts' (in German), in: Harald Gapski and Stephan Packard (edts.) 'Super-Scoring? Datengetriebene Sozialtechnologien als neue Bildungsherausforderung', Schriftenreihe zur digitalen Gesellschaft NRW, Band 6, kopaed verlagsgmbh, München, Germany.

Domingos, Pedro (2012) ‘A few useful things to know about machine learning’, Communications of the ACM, Vol. 55/10, pp 78-87.

Milkau, Udo (2021) ‘The Risks of Artificial Intelligence used for Decision-making in Financial Services’, Capco Journal of Financial Transformation, Vol. 52, pp 128-141.

HHI (2019) ‘Der Blick in Neuronale Netze’, Forschung Kompakt (in German), Fraunhofer-Institut für Nachrichtentechnik, Heinrich-Hertz-Institut (HHI), 1.7.2019 (available at: https://www.fraunhofer.de/de/presse/presseinformationen/2019/juli/der-blick-in-neuronale-netze.html; accessed 3.5.2020).

About the author

Udo Milkau, Digital Counsellor, Frankfurt, Germany, udo.milkau@web.de

Udo Milkau is a 'digital dinosaur' with first experiences in digital technology in 1974, many innovation projects including the first European securities online-brokerage in 1995 and working as a Digital Counselor now. For three decades he holt management positions with automotive industry, professional services firms and transaction banking, served customers in Asia and Europe, the European banking industry, and was Chief Digital Officer, Transaction Banking until 2020. After his academic education in physics, he worked as a research scientist in large collaborations at different European research centers incl. CERN, CEA de Saclay, and GSI. He was chairman of the European Association of Co-operative Banks (EACB) Digital and Data Working Group, member of the EACB Payment Services Working Group and member of the European Central Bank's Operation Managers Group (ECB OMG). Udo Milkau published nearly 100 papers including payments strategy, digitalization of banking, risk management / risk culture, digital economies, blockchain and 'law & digitalization'. He lectured at Goethe University Frankfurt am Main, Frankfurt School of Finance and Management, WHU - Otto Beisheim School of Management (Vallendar) and Baden-Wuerttemberg Cooperative State University (DHBW in Mosbach).