ECONOMICS WORKING PAPER

The Prospects and Dangers of Algorithmic Credit Scoring in Vietnam: Regulating a Legal Blindspot

Nicolas Lainez
ISEAS - Yusof Ishak Institute
Email: nicolas_jose_lainez@iseas.edu.sg

January 2021

Abstract

Artificial intelligence (AI) and big data are transforming the credit market in Vietnam. Lenders increasingly use 'algorithmic credit scoring' to assess borrowers' creditworthiness or likelihood and willingness to repay loan. This technology gleans non-traditional data from smartphones and analyses them through machine learning algorithms. Algorithmic credit scoring promises greater efficiency, accuracy, cost-effectiveness, and speed in predicting risk compared to traditional credit scoring systems that are based on economic data and human discretion. These technological gains are expected to foster financial inclusion, enter untapped credit markets, and deliver credit to 'at-risk' and financially excluded borrowers. However, this technology also raises public concerns about opacity, unfair discrimination, and threats to individual privacy and autonomy. In Vietnam, the lending industry deploys this technology at scale but in legal limbo. Regulation is vital to delivering big data and AI promises in the financial services market while ensuring fairness and public interest.

The Prospects and Dangers of Algorithmic Credit Scoring in Vietnam: Regulating a Legal Blindspot

Nicolas Lainez[1]

1. Background

Big data and artificial intelligence (AI) are transforming the credit industry worldwide. This process is evident in Vietnam, an emerging lower-middle-income economy where credit markets are undergoing rapid liberalization and transformation. Lenders increasingly assess borrowers' creditworthiness – defined as the likelihood and willingness of individuals repaying their loans and meeting contractual obligations on time – using 'algorithmic credit scoring'. This expression uses a combination of big data and machine leaning algorithms to generate scorecards based on a statistical analysis of loan applicants' risk profile[2]. With tens of millions of unbanked and underbanked citizens short of credit history and access to banking services in Vietnam, algorithmic credit scoring emerges as a magic tool to mitigate high uncertainty and information asymmetry, unlocking credit growth potential.

Algorithmic credit scoring makes big promises. The prospects of efficiency, accuracy, cost-effectiveness, and speed in predicting credit risk augur the inclusion of millions of un(der)banked citizens in financial services markets, including the lending industry. However, algorithmic credit scoring also raises public concern about opacity, unfair discrimination, and the loss of individual privacy and autonomy (Aggarwal 2020). Regulation is critical to foster innovation while safeguarding public interests. Western countries mitigate anxieties about algorithmic credit scoring and meet the new challenges of the big data era by revising their credit laws, in particular the concepts of creditworthiness and discrimination (Danielle and Pasquale 2014; Hurley and Adebayo 2017). They also develop ethical guidelines for responsible AI that, in some cases, lead to AI governance frameworks (Jobin, Ienca, and

---

[2] Algorithmic decision-making is generally based on 'machine-learning' algorithms. In its most elementary form, an algorithm is a set of instructions designed to solve a problem. Machine learning algorithms learn from data and optimize their operations to predict outcomes in unseen data, thereby developing 'artificial intelligence' over time. Machine learning is a branch of 'Artificial Intelligence' (AI). AI refers to machines programmed to simulate human thinking, learn from experience, and perform human-like tasks. Machine learning algorithms are fed with 'big data' or large and complex sets of (un)structured data including social data, machine data, and transactional data.

Vayena 2019). With the exception of Singapore, most Southeast Asian countries lag in the regulation of algorithmic credit scoring. The lending industry deploys this technology at scale without a regulatory framework, which can be a source of concern.

In Vietnam, domestic and foreign (including Singaporean) 'credit scoring' startups in the financial technology (fintech) sector provide machine learning, big data-based credit scoring technology to lenders in a legal vacuum. The Vietnamese credit law is poorly prepared to regulate algorithmic credit scoring and the myriad of new digital lending services that depend upon it. Privacy law is piecemeal, under construction, and inadequately equipped to regulate big data mining and analytics. This study addresses these concerns from a legal and policy angle. After describing the promises and perils of algorithmic credit scoring in Vietnam, it suggests novel ways to regulate it. This legal proposal aims to provide oversight over algorithmic credit scoring, ensuring that the process from data collection to credit decision is transparent, accessible, accurate, and fair. This study's central argument is that adequate regulation is vital to delivering big data and machine learning's promises in the financial services market while ensuring fairness and public interest. The proposal calls for amending credit and data privacy legislation and develop ethical guidelines for responsible AI.

This study contributes to a small but growing body of works on fintech regulation in Asia. A recent edited volume titled 'Regulating Fintech in Asia: Global Context, Local Perspectives' shows how financial service providers, fintech startups and regulators strive for innovation but must deal with uncertainty brought by technological progress and regulatory unpreparedness. Responses like policy experimentation and regulatory sandboxes do not always reassure stakeholders (Fenwick, Van Uytsel, and Ying 2020). A chapter in this volume addresses fintech regulation in Vietnam. More than 150 fintech startups provide financial services including digital payment, crowdfunding, p2p lending, blockchain, personal finance management, and financial information comparators. To support their development, the Vietnamese government has set up a steering committee, designed a national financial inclusion strategy, and drafted a framework to operate regulatory sandboxes in five key domains (Nguyen 2020; see Do 2020 on regulatory sandboxes). In short, these discussions provide a broad context and policy recommendations to support fintech development, but leave algorithmic credit scoring aside[3]. Furthermore, there has not been any research done on the societal and political impact of algorithmic credit scoring in Southeast Asia. This study fills this gap by opening new lines of

---

[3] An exception is an article from Tinh Le Duc (2017) who recommends using a national credit database system to improve the efficiency of public credit registries in Vietnam (see the discussion on CIC and PCB below).

enquiry and policy debates that will become crucial as algorithmic credit scoring expands and reaches into people's daily lives.

This study is based on secondary sources, mainly materials collected from the websites of credit scoring startups and lenders operating in Vietnam and Singapore. Unfortunately, most startups simplify the information made available as part of their marketing strategy. Details on the nature of big data collected for risk prediction, machine learning protocols, customer base, product pricing and impact are trade secrets. An exception is CredoLab, a Singaporean credit scoring firm. This startup is highly open and transparent and provides abundant details about its mission, products, impact, and privacy policy in its website. The materials available in the Resources Section[4] proved valuable to understanding algorithmic credit scoring and writing this study. Other materials include news clips about credit, household debt, AI, algorithms, credit scoring, fintech, digital finance, regulation, and cybersecurity. Academic and policy literature on algorithms, big data, credit, and data privacy lays the foundation for the analysis of empirical data.

2. The Prospects of Algorithmic Credit Scoring

The following section describes the promises that algorithmic credit scoring makes to the lending industry and society. This technology is the tip of the spear of the global campaign for financial inclusion, which aims at delivering credit and other financial services to 'at-risk' un(der)banked citizens. By processing big data with machine leaning algorithms and improving efficiency, accuracy, speed, cost-reduction, it aims at boosting, if not revolutionizing, credit markets in the global South.

2.1 Big Data for Financial Inclusion

The financial industry is driven by information; in particular, lenders need accurate and up-to-date data on borrowers to determine loan pricing and terms, decrease exposure to bad debt, and maximize profit. However, lenders have imperfect information on their customers. To solve this problem, they collect data from traditional and non-traditional sources to assess applicants' creditworthiness and willingness to repay their loans. Traditional data includes 'hard data' such as credit history, income, labor, education level and tax records, and 'soft data' including opinions from loan officers and internal assessments (Abraham, Schmuker, and Tessada 2019,

---

[4] https://www.credolab.com, 'Resources' (accessed 20 November 2020),

1). These data can be hard to obtain in countries where financial exclusion prevails, i.e., citizens are not served or poorly served by financial institutions. Financial exclusion involves unbanked or 'no-file' borrowers with no bank account and no established credit history, and underbanked or 'thin-file' borrowers who may have a bank account but no credit history or an incomplete one. In both cases, these borrowers short of credit history pose a challenge to lenders who are unable to assess credit risk.

Financial exclusion is high in Southeast Asia but reliable figures are hard to come by. According to Bain & Company, Google and Temasek (2019, 12), 70 percent of the region's population is either unbanked or underbanked, totaling 458,5 million inhabitants out of 655 million. However, this figure is not definitive. It is difficult to characterize this group. Financial inclusion advocates describe un(der)banked workers as operating in a cash economy: they "save their money under the mattress and borrow from so-called 'loan sharks', paying exorbitant interest rates on a daily or weekly basis," argues Jose de Luna-Martinez (2016) in a post published in a World Bank blog[5]. Peter Barcak, CEO of CredoLab, a Singaporean credit scoring startup that operates in Asia, including Vietnam, argues that "underbanked consumers lack a voice. Most of them live off-the-grid, are employed in the informal economy, and get by on a hand-to-mouth existence where they do not even have enough to meet the minimum balance required by banks" (Barcak 2019). Financial exclusion prevails in Vietnam, a country that experienced an economic collapse in the mid-1980s. This financial crisis led to macro-economic reforms that put Vietnam on the path of a socialist-oriented market economy integrated in global exchanges. The economy and financial system have since developed immensely, helping Vietnam to emerge as a thriving lower-middle-income country with a GDP per capita of US$2,715 in 2019[6]. However, about 61 (Nguyen 2020, 124) to 70 (World Bank 2018a, 160) percent of 98 million Vietnamese still have no bank accounts and, therefore, credit history. This results in 59.8 to 68.6 million potential borrowers, who are mainly urban and rural low-income workers, locked out from formal financial services. Credit scoring startups and lenders assume that many of these individuals seek loans and have the ability and willingness to repay them. They thus target their marketing towards them.

---

[5] The use of the term 'loan shark' lending perpetuates the stereotype that poor and financially excluded borrowers only borrow money in the informal finance market. This market is epitomized by the 'evil' moneylender that preys on the poor. This stereotype ignores the diversity of the informal finance market in Southeast Asia and the fact that different credit products and debt regimes can produce an empowering (Lainez 2020) and disempowering impact over creditors and debtors (Lainez et al. 2020).

[6] Up from US$231 in 1985 (https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=VN) (accessed 20 November 2020)

High levels of financial exclusion in Southeast Asia and other regions of the world have laid the ground for a global campaign to boost financial inclusion[7]. This campaign is led by a coalition of governments, financial institutions, banks, development agencies, IT firms, donors, foundations, IGOs, NGOs and civil society organizations, which Daniela Gabor and Sally Brooks (2017, 424) refer to as a "fast-evolving fintech–philanthropy–development complex". The term 'financial inclusion' refer to "individuals and businesses hav[ing] access to useful and affordable mainstream financial services that meet their needs – transactions, payments, savings, credit and insurance – delivered in a responsible and sustainable way" (World Bank 2018b). Financial inclusion is a top priority in the 2030 Sustainable Development Goals adopted by all UN Member States in 2015, a road map to foster peace and prosperity, eradicate poverty, and safeguard the planet (United Nations 2015). The financial inclusion agenda places a strong emphasis on credit. Credit has many virtues according to its advocates: it improves resource allocation, promotes consumption and increases aggregate demand. In turn, a rise in demand boosts production, business activity and income levels. Consumer credit enables households to cope with their daily spending, compensate low wages that provide limited benefits and security, withstand economic shocks especially in times of pandemic and depression, invest in production and human capital, and foster upward mobility. Fintech is critical to enable the expansion of credit markets, especially digital lending. By assessing credit risk for millions of 'at-risk' financially invisible borrowers, credit scoring firms are able to tap on new and profitable markets and position themselves at the forefront of the global financial inclusion campaign.

2.2 'All Data is Credit Data': Leveraging the Digitalization of Life

To achieve financial inclusion, Southeast Asian credit scoring firms take an 'all data is credit data' approach[8]. They take advantage of the recent introduction of digital technologies and services, new electronic payment systems, and smart devices to collect alternative data, which capitalizes on high mobile phone and internet penetration rates. In 2016, Vietnam's internet penetration rate had reached 52 percent while smartphone ownership 72 and 53 percent in urban and rural areas, respectively (Nguyen 2020, 121). Furthermore, 132 million mobile

---

[7] The Vietnamese government ratified a National Financial Inclusion Strategy 2020-25. The goal is to ensure that 80 percent of adults have a bank account and access to basic financial services such as payment services, money transfer, saving, credit and insurance in the next five years (VNA 2020)

[8] The expression comes from Douglas Merrill, ZestFinance's CEO, who stated in an article published in the New York Time that "We feel like all data is credit data, we just don't know how to use it yet". He meant that predictive analytics can glean virtually any piece of data about a person and analyze it to assess creditworthiness through extrapolation (Hardy 2012; for discussions on "all data is credit data", see Aitken 2017; Rosamond 2016).

devices were in circulation in 2017 (CSIRO 2018, 3). These developments pave the way for credit scoring firms which mine alternative data from borrowers' mobile phones. These data are not directly relevant to their creditworthiness but generate detailed, up-to-date, multi-dimensional and intimate knowledge about their behavior, especially their likelihood to repay loans on time. Alternative data derives from telco and mobile data, social network data (Facebook, Twitter, WhatsApp, LinkedIn, Instagram, etc.), browser data, e-commerce, and financial transaction data. Credit scoring firms feed these data to machine learning algorithms to generate credit scorecards that rate and rank consumers above or below each other.

How are social network and telco data used to predict creditworthiness? Risk prediction is assessed through correlation, or the "tendency of individuals with similar characteristics to exhibit similar attitudes and opinions" (Tan and Phan 2018, 11). If connected borrowers act similarly, social network data should be able to predict their behavior their willingness to repay loans regardless of income and credit history. LenddoEFL, a digital credit scoring firm, gleans social network data to achieve its vision: "to provide 1 billion people access to powerful financial products at a lower cost, faster and more conveniently"[9]. Based in Singapore, the company operates in the Philippines, Mexico, Colombia and, more recently, Vietnam through partnership with Orient Commercial Joint Stock Bank (PRweb 2018). Its flagship solution, LenddoScore, is an onboarding app that generates a scorecard ranging from 0 to 1,000. This app collects social network data, mobile data, browser data, telco data, e-commerce transaction data, financial transaction data, psychometric data, and form filling analytics (Lenddo 2015). For social network data, LenddoScore gleans data from Facebook, LinkedIn, Yahoo, Google (including emails) and Twitter to map relationships with 120 social media profiles (L.P. 2016)[10]. The predictive power of social network data is based on the principle of 'who you know matters'. In other words, borrowers' personality and relationships are pivotal to assess risk.

Another popular source of alternative data is telco data provided by mobile operators[11]. TrustingSocial, a startup based in Vietnam and Singapore, collects telco data for risk profiling. This firm uses call/SMS metadata (when, where and duration), top-up data and value-added-service transaction data to determine borrowers' income, mobility patterns, financial skills,

---

[9] https://lenddoefl.com (accessed 20 November 2020)

[10] To ensure repayment, LenddoEFL also applies a social collateral system by which borrowers select a 'Trusted Network' of three friends that are accountable in case of default. If borrowers fail to repay a loan, their trusted friends are penalized with a lower score, which may limit their ability to borrow from LenddoEFL's and/or associated lenders (L.P. 2016).

[11] Many studies show that the effectiveness of mobile phone usage data (Ots, Liiv, and Tur 2020), credit-call records (Óskarsdóttir et al. 2019), and airtime recharge data (Shema 2019) for credit scoring purposes.

consumption profile, social capital, and life habits (Huynh 2017). Similarly, KT CRDP, a credit scoring firm based in South Korea with operations in Vietnam, also uses Telco data. Its K-Telco Score solution gleans phone subscription data (profile, duration, products, bundling), phone usage data (call patterns, porting history, missed calls, mobile device, micro payments, and paid TV), bill payment data (postpaid overdue, top-up pattern and credit, plan suspension and recovery), and proxies such as location, overseas roaming, and download of financial apps. These data offer glimpses of borrowers' characteristics in terms of: 1) financial capacities for repaying credit based on their income level, economic stability, and consumption pattern, 2) credit management abilities based on repayment behavior, and suspension management, 3) life pattern estimations based on residence stability, employment, and call networks, and 4) appetite for financial services (KT Global Business Group 2019). In practice, prepaid top-up card phone users who make regular phone calls of a certain duration may indicate that they earn enough income to recharge their plan frequently and maintain their phone routines. The frequency in which they top-up their phone card is another proxy for their income level. However, lenders will trust top-up card users less than those who take postpaid plans with monthly payments, which indicate a better economic situation. Another proxy, call location, provides insights about subscribers' income level, job stability, job rotation, and so on. If s/he calls from a residential area, their risk profile will be more credit favourable. In short, telco data provides 24-hour life insights about phone users that can be leveraged for risk prediction.

2.3 From 'Old' Credit Registries to 'New' Algorithmic Scoring Firms

Digital credit scoring firms make a strong argument for efficiency by comparing credit scoring based on alternative data and machine learning algorithms and scoring based on traditional data and less sophisticated statistical tools. Their marketing presents the latter as a system of "archaic credit data and scoring technologies of the 70s"[12]. This 'outdated' model leads not only to "lengthy paperwork and unpleasant customer experiences" (PR Newswire 2019), but also financial exclusion as "arbitrary lending rules and criteria will in most cases disqualify [a] potential customer" (CredoLab 2017). In general, traditional credit scoring registries use past credit performance, loan payments, current income, and the amount of outstanding debt to profile borrowers. These data are fed to linear statistics protocols (Aggarwal 2020; Burton et al. 2004, 5). In many cases, the final decision on a loan application and pricing depends on human discretion. In the US, the scoring model popularized by FICO (Fair, Isaac and

---

[12] https://www.trustingsocial.com/about-us/ (accessed 20 November 2020)

Corporation) until recently relied on three economic proxies that accounted for 80 percent of a single numeric score: consumers' debt level, length of credit history, and regular and on-time payments (Wei et al. 2016, 235). However, these data are problematic for several reasons. First, they do not render the economic profile of thin-file borrowers adequately. Second, they are not produced frequently enough to integrate the latest developments in borrowers' lives. Third, they are not easily and quickly accessible at an affordable cost (Guo et al. 2016, 2). These limitations lower risk prediction accuracy, convenience, and utility for lenders. Another problem is that assumptions underlying traditional scoring models may be outdated. These models approach borrowers as workers with documented regular income and a linear and stable employment path. In post-Fordist countries, labor flexibilization, neoliberal reforms and austerity policies have given birth to the 'precariat', a new class of 'denizens' who live in 'tertiary time', lack labor stability and economic security, and whose career paths are non-linear (Standing 2014; 2013)[13]. In emerging countries like Vietnam, the linear model bears little relevance to the lives of many precarious workers who live on a day-by-day basis (Lainez 2019). Only a minority of highly-educated, middle-class and urban workers build linear and future-oriented careers (Earl 2014). In brief, traditional scoring models show limitations in data sampling and credit risk modeling.

In Vietnam, the Credit Information Center (CIC) under the State Bank is the official public credit registry. The CIC collects, processes and stores credit data, analyzes, monitors and limits credit risk, scores and rates credit institutions and borrowers, and provides credit reports and other financial services. It gleans credit data from 670,000 companies and 30.8 million individuals with existing credit history and works with over 1,200 credit institutions including 43 commercial banks, 51 foreign banks' branches, over 1,100 people's credit funds, and 27 finance and leasing firms[14]. The CIC operates alongside the Vietnam Credit Information Joint Stock Company (PCB), a smaller private credit registry that 11 banks created in 2007. The State Bank of Vietnam granted a certificate of eligibility for credit information operation to PCB in 2013. This registry delivers financial services to its clients, including credit reports (Bartels 2020).

Both the CIC and PCB provide essential services to the financial community. However, they use traditional credit data and scoring methods that limit their reach and scope. Although the

---

[13] The precariat lives on 'tertiary time', a temporal regime that compels them to live permanently on standby, to labour excessively, to "work-for-labour" to secure casual jobs, and to give away their time to capricious employers.
[14] https://cic.org.vn (accessed 20 November 2020)

CIC is well suited to provide credit ratings about legal entities, it takes several days for them to deliver reports on individuals and lack a national credit database system with an individual profile for every citizen based on the new 12-digit personal ID number (Le 2017, 44). Moreover, the CIC uses linear data that does not reflect unstable and multidirectional career paths that is a norm in Vietnam. In a move to enhance transparency and efficiency and reduce costs and waiting time, the CIC had opened a portal to allow borrowers to track their credit data, scores and levels, prevent fraud, receive advice to improve their rating, and access credit packages from participating institutions (VNA 2019). Borrowers are able to access this portal online or through a mobile phone app called CiCB. On the whole, the CIC and PCB are poorly equipped to compete with credit scoring firms that pave the path to a lending ecosystem driven by big data and machine learning algorithms.

2.4 More Data, Better Accuracy

Credit scoring firms focus on efficiency – especially accuracy, speed, cost-effectiveness, and customization – to set themselves apart from credit registries and promote their advanced technology. Their predictions offer higher accuracy than traditional scoring by leveraging on a much larger number of data points relevant to risk behavior and analyzing them with AI. CredoLab describes how its technology increases efficiency by collecting anonymous metadata: SMS, phone, email and network communication activity (frequencies, ratios, intervals between actions, distribution), phone numbers and contacts (to correlate the address book with the communication activity), phone device features (model, display size, RAM size, storage size and utilization, age of device), browsing history (browsing patterns, preferences and intent to apply for a lending product), installed software (lending apps, office applications, e-wallets and 'suspicious' anonymizing Internet tools like Tor and VPN[15]), geolocation, personal data including IP address for KYC (Know Your Customer[16]) contactless verification and anti-fraud procedures contracted by banks [17]. In the non-anonymous mode that improves KYC's effectiveness, onboarding solutions like CredoApp and CredoSDK also collect the content of text messages, phone numbers, names, and details of contacts and other personal data.

---

[15] Tor and VPNs are secure networking technologies aimed at protecting online privacy. The former directs internet traffic through different servers or nodes located across the world to ensure anonymity. The latter hides a user's IP address, encrypts online connection and routes it through an intermediary server belonging to the VPN company to ensure privacy.

[16] Traditional KYC (Know Your Customer) involves a set of checks that lenders carry out with borrowers to verify their identity through paperwork and personification. E-KYC (Electronic Know Your Customer) refers to the digitalization of this process.

[17] https://www.credolab.com/faq, 'What kinds of data do you collect?' (accessed 20 November 2020). See also Privacy Policy (https://www.credolab.com/privacy-policies/english) (accessed 20 November 2020).

CredoLab's flagship solution is CredoApp, which once installed on a smartphone, gleans alternative data to predict risk. Loan applicants must grant access permission in their device – "the more permissions granted, the more accurate the score, the higher your chances of getting a better score"[18] – to generate a credit scorecard.

The harvesting of this vast amount of data increases predictive accuracy significantly, scoring startups claim. Peter Barcak, CredoLab's CEO, argues that "usually a scorecard consists of 20 to 25 features with the highest information value and features with the highest predictive power. But we have built a huge pool of 1.3 million features over the last four years to choose from" (Gaglia 2020). This difference impacts the Gini metric used by lenders to evaluate risk prediction accuracy. The Gini coefficient ranges between 0 and 1, indicating 50 percent and 100 percent accuracy rates respectively. Lenders in Asia are accustomed to traditional credit scoring systems with a 0.25-0.30 Gini coefficient or a 62-65 percent accuracy. However, CredoLab's claims that its digital scorecards have a Gini coefficient of 0.40-0.50 or 70-75 percent accuracy[19]. The 8-13 percent gain makes a difference for lenders, argues Peter Barcak (ibid.). For lenders, it translates into a 20 percent increase in new customer approval, a 15 percent drop in loans defaults, and a 22 percent decrease in fraud rate (ibid.). It also reduces the time from scoring request to credit decision to a few seconds.

2.5 'Fast and Easy Credit': Toward Real-Time Credit Decision

The emphasis on shorter times for credit loan approval reflects another area where credit scoring startups outperform traditional scorers: fast performance. They aim to provide 'real-time credit decision', as stated in CredoLab's privacy policy[20]. In a discussion on partnerships with banks in Southeast Asia, Peter Barcak stated that "we're helping them make fast credit decisions because our scores are available within one second. So our clients have the advantage because they can make their decisions fast" (Gaglia 2020). TrustingSocial also markets speed. Geoffrey See, head of identity, emphasizes that, "we are working to enable '5 minutes, anytime,

---

[18] https://www.credolab.com/faq, 'How many data points are collected from my smartphone device?' (accessed 20 November 2020)

[19] A few studies compare predictive performances for traditional and machine learning credit scoring models. A study with data from a Chinese fintech loan company highlights that 1) the machine learning scoring model based on alternative data outperforms traditional scoring methods, 2) alternative data improves risk prediction accuracy, 3) machine learning models are more accurate in predicting losses and defaults after a shock, 4) both models show efficiency as the length of the relationship between the bank and the customer increases, but traditional models do better in the long term (Gambacorta, Huang, and Wang 2019). Another study shows that the combination of email usage, psychometric and demographic data measures creditworthiness with greater accuracy than a model based on demographic data alone, which is conventionally used to assess risk (Djeundje et al. 2021).

[20] https://www.credolab.com/privacy-policies/english (accessed 20 November 2020).

anywhere' access to financial services, which describes going from the point a customer decides to sign up for with a financial service to the point of cash disbursement" (T. T. Le 2020). Geoffrey See refers to automated lending platforms that digitize the loan application process and shortens the time for scoring, decision-making, approval and disbursement to a few minutes. The race for speed does not only involves the collection of alternative data and its analysis by machine learning algorithms, but the entire process of contacting a potential borrower, processing an application, generating a scorecard, making a decision, determining pricing and terms, and disbursing the loan to a bank account.

On the ground, banks and consumer lending companies also race to provide a 'fast and easy' credit experience. Kalidas Ghose, CEO of FE Credit, the leading consumer lender in Vietnam, the consumer credit branch of VPBank, addresses the issue of convenience in a PowerPoint presentation entitled 'Financial Inclusion: Leveraging the Mobile Device – The FE Credit Experience' (Ghose 2018). This document describes the power of big data to foster financial inclusion and credit growth. FE Credit takes advantage of high internet and smartphone penetration to target lower, middle and upper "mass" workers earning from US$75 to US$347 per month. Ghose describes the "infrastructural obstacles" that hamper the development of consumer lending, in particular the lack of information about unbanked segments and the heavy reliance on human labor and discretion by credit bureaus. He underscores the potential of telco data in assessing risk for the unbanked. These data are provided by Viettel (63.6 million users), Mobifone (20.5 million) and Vinaphone (34.6 million). Ghose describes the "enhanced process for instant application in 1 click" that takes less than two days: an SMS is sent to the customer, a "fair price" is calculated based on a "highly reliable score", the customer applies for a loan with "1 click", the sales team calls back in less than 30 minutes, an appointment is made is less than 24 hours to close the deal, and the loan is disbursed within 12 hours. Ghose cites the example of a construction worker who needed several hundred million VND to renovate his home. For this borrower, "applying for a bank loan is out of option with his current income. A few hundred million is a big amount to borrow from relatives. 'Loan shark' is a last resource that he would never dare to take." This borrower clicked on a FE Credit link received by SMS at 8am, went through the different stages of the application, and received a loan the following morning to renew his home. Overall, FE Credit shortened the loan application and delivery process to a couple of days, while other banks take several days.

FE Credit then sought to improve this performance to "align with its corporate mission of delivering fast and easy credit" by developing $NAP. This automated lending platform

digitizes the loan application process and shortens the time for approval and disbursement to 15 minutes (Saigon Times 2019)[21]. $NAP reduces customer recruitment cost and lowers the risk of losing them due to cumbersome and time-consuming procedures. The app achieves this performance by incorporating technology such as "facial recognition, AI-based optical character recognition (OCR) and Optical Ink Character Recognition (OCR+ICR) to verify the customer's identity, authenticate the documents submitted by customers and assess their creditworthiness using their phone cameras, voice- based virtual assistant, speech to text, device-based scoring, telco data scoring, eSignature, and many more" (Ghose 2019). A "strong risk management process", based on "extensive usage of big data analytics", has resulted in "an ability to enjoy higher exposure with better customers, deliver top-up loans earlier in customers' lifecycle, and expand the targeted base by using alternate behavioral data in partnership with telcos, e-commerce and utility companies" (ibid.). According to FE Credit, $NAP has attracted many news customers. It has also set the path for a 'fast and easy credit' experience in Vietnam (Lainez 2021). Machine leaning algorithms and big data are critical to achieve this goal. They deliver predictions in a snap and cut down bureaucracy and human discretion. They also facilitate the automation of lending processes.

Algorithmic credit scoring also allows scorecards to be customized by modifying the way credit risk is assessed for specific financial institutions and loan applicants. CredoLab develops scorecards for unsecured lending including credit cards, personal loans, Point of Sale (POS) loan, payday loans, two-wheeler loans and auto loans[22]. Customization is also the core business of Gobear, a popular online 'financial supermarket' that personalizes banking and insurance products down to customers' identified needs. The platform operates in Hong Kong and six ASEAN countries including Vietnam. It was designed to allow users to compare and select over 2,000 financial products, principally credit and insurance plans. In Vietnam, CredoLab has launched Easy Apply with GoBear. Easy Apply is a smartphone app that gathers digital footprints and consumer behavior data to help financial institutions, banks and consumer lenders associated with Gobear to assess risk, tailor financial products and pricing, and ensure fast services and high approval rates. Once loan applicants select customized financial products in the Easy Choices function on the GoBear platform, they can download Easy Apply and install it onto their device to generate a credit scorecard. Algorithmic credit scoring therefore

---

[21] FE Credit has also launched '$HIELD', a digital insurance app (Ghose 2019).
[22] https://www.credolab.com/faq, question 'What types of scorecards can we develop?' (accessed 20 November 2020)

gives lenders the ability to customize rating systems to best fit the unique risk characteristics of their company, partners, customers, and lending market.

To sum up, credit scoring firms promote their predictive technology as more accurate, faster, cheaper and more customizable than traditional scoring models. However, they do not reject credit registries and human discretion. On the contrary, they argue that big data and machine learning-based models can complement traditional scoring and human intervention. Their suggestion is to integrate both systems to improve prediction and convenience (see CredoLab 2019; Gillis 2020, 35 note 107). In the end, credit scoring firms pride themselves in being more efficient and better equipped than traditional scorers to include millions of un(der)banked customers and access an untapped and profitable lending market. However, they do not engage in a war against traditional scorers, quite the contrary.

3. The Dangers of Algorithmic Credit Scoring

Algorithmic credit scoring promises to foster financial inclusion and make customized credit products available to the un(der)banked. But can machine learning algorithms and big data live up to their promise of distributional fairness? Can they protect individual autonomy and privacy? There is no data showing that algorithmic credit scoring fosters financial inclusion and consumer lending in Vietnam yet. There is also hardly any evidence showing that Vietnamese borrowers are concerned or even aware of the existence of algorithmic credit scoring. However, public concerns about the dangers of opacity, unfair discrimination, and the loss of autonomy and privacy due to algorithmic governance have been growing increasingly in the West. The only common concern shared across the world is cybersecurity. This chapter will map the dangers of algorithmic credit scoring and examine their relevance for Vietnam and Asia more generally.

3.1 Biases and Discrimination: seeing through the thickness of 'black box' algorithms

There are conflicting public perceptions of big data and AI. A common assumption is that the technology is truthful, objective and neutral. This positive view gives AI authority over humans to take on heavy responsibilities and make vital decisions. On the other hand, the general public increasingly worry that algorithmic governance and decision-making may be biased and discriminate against historically vulnerable or protected groups. These groups include the poor, women and racial, ethnic, religious and sexual minorities (Noble 2018). The hidden nature of machine learning algorithms exacerbates these anxieties. Algorithms are popularly described

as 'black boxes' because they run autonomously, especially unsupervised learning. The reason is because they operate sometimes without disclosing, not even to their programmers, how they calculate, the datasets or combinations of datasets used, which combinations of datasets are significant to predict outcomes, and how they achieve specific outcomes. For creditworthiness assessment, borrowers do not understand how algorithms make predictions based on alternative data from call and SMS logs, network size of Facebook, Twitter or LinkedIn friends, and handphone device owned. A widespread fear is that 'opaque' algorithms will standardize past prejudices and biases into discriminatory rules that will negatively affect loan applicants. However, they ignore that the danger of discrimination lies not in algorithms but humans who classify and rank other humans, and design, program, and train algorithms to perform certain operations and reach determined outcomes.

Humans have a tendency to categorize themselves permanently. They measure behavior quantitatively with metrics, classify individuals in categories that carry rewards and sanctions, and segregate societies. The problem with measurement, classification and segregation is that humans may not realize, admit, or even be aware of their motives for discriminating themselves. As Jon Kleinberg, Jens Ludwig, Sendhil Mullainathan and Cass Sunstein (2018, 125) put it, "human cognition is the ultimate 'black box,' even to ourselves". Individuals may be aware of their actions' inputs and outputs but lack conscious access to the true motives and processes that drive their behavior (Card 2017). Credit scoring systems are tools for classifying, ranking, and discriminating borrowers. In the US, the law conceptualizes credit discrimination as 'disparate impact' and 'disparate treatment'. These two notions permeate anti-discrimination law, in particular the Constitution's Equal Protection Clause, civil rights laws, and the Equal Credit Opportunity Act (ECOA) that prohibits credit discrimination based on protected characteristics (Kleinberg et al. 2018, 121; see also Gillis 2020). The article §1691 of ECOA prohibits credit discrimination "(1) on the basis of race, color, religion, national origin, sex or marital status, or age (provided the applicant has the capacity to contract); (2) because all or part of the applicant's income derives from any public assistance program; or (3) because the applicant has in good faith exercised any right under this chapter"[23].

---

[23] https://www.govinfo.gov/content/pkg/USCODE-2011-title15/html/USCODE-2011-title15-chap41-subchapIV.htm (accessed 20 November 2020)

Disparate treatment refers to deliberate discrimination. An example would be a policy that purposely denies credit access to African-Americans or women, or that uses race and gender to discriminatory characteristics. Disparate impact refers to seemingly neutral and fair rules, practices and requirements that lead to unapproved discrimination against vulnerable and/or protected groups. A lender who only offers $100,000 mortgages to well-off families excludes families with lower income and home values. If these families belong to a protected group, the credit policy in question generates disparate impact. It will be considered unlawful unless the lender provides a legitimate business justification[24]. Disparate treatment and disparate impact can lead to redlining. A lender refusing to grant credit to certain communities or residents living in specific areas based on characteristics like race, origin, gender, and so forth is a case of redlining. Nowadays, overt redlining is unusual and well-hidden, especially in countries like the US, where it is explicitly repressed. However, disparate treatment and disparate impact can still be observed through pricing. In this case, lenders propose overpriced, unfavorable and personalized subprime loans to borrowers, especially no-file and thin-file applicants, to minimize risk and maximize profit from interest rates and fees. These borrowers do not lack access to credit but lack access to prime and favorable loans (Havard 2010, 244–45). In short, humans are the first source of bias. Their classificatory systems based on risk assessment and credit policies may inadvertently or deliberately lead to discrimination against vulnerable and/or protection borrowers.

How are human biases incorporated into machine leaning algorithms? This question requires some basic knowledge of algorithms. What we commonly refer to as an 'algorithm' is actually made up of two parts. The first is a 'screener' that "takes the characteristics of an individual (a job applicant, potential borrower, criminal defendant, etc.) and reports back a prediction of that person's outcome" that "feeds into a decision (hiring, loan, jail, etc.)". The second is a 'trainer' that "produces the screening algorithm" and decides "which past cases to assemble to use in predicting outcomes for some set of current cases, which outcome to predict, and what candidate predictors to consider" (Kleinberg et al. 2018, 132). In creditworthiness assessment, the trainer processes training data to establish which inputs to use and their weights to optimize the output and generate the screener. In turn, the screener uses inputs and combines them in ways determined by the trainer to produce the most accurate risk evaluation. Disparate

---

[24]https://www.occ.treas.gov/topics/consumers-and-communities/consumer-protection/fair-lending/index-fair-lending.html (accessed 20 November 2020)

treatment will arise if the screener includes variables like race and gender that are statistically informative to predict the outcome. Disparate impact will occur if the screener includes non-problematic proxies that will become problematic if they disadvantage vulnerable or protected groups excessively (Kleinberg et al. 2018, 139). It is important to underline that an algorithm that has been trained to exclude forbidden characteristics such as race, gender or ethnic origin may include them in the assessment by considering correlated proxies such as zip code, income, and school attendance. This detail has legal implications as we shall see in the next chapter.

Unfair outcomes may come from input, training, programming, and outcome bias. Kevin Petrasic and colleagues define the first three main sources of biases as follows: "Input bias could occur when the source data itself is biased because it lacks certain types of information, is not representative or reflects historical biases. Training bias could appear in either the categorization of the baseline data or the assessment of whether the output matches the desired result. Programming bias could occur in the original design or when a machine learning algorithm is allowed to learn and modify itself through successive contacts with human users, the assimilation of existing data, or the introduction of new data" (Petrasic et al. 2017, 2). An outcome bias "describes the phenomenon by which evaluators tend to take information about the outcome into account when evaluating the quality of a decision itself" (König-Kersting et al. 2020, 1). This approach is problematic because "First, the evaluator has available a different information set than the decision maker, who typically faces uncertainty at the time of her decision. Second, a good outcome might derive from a bad decision, and a bad outcome might derive from a good decision" (ibid.). Unfair discrimination may also arise when algorithms consider factors stemming from past discrimination, which becomes a liability if it can be connected to disparate treatment or disparate impact (Kleinberg et al. 2018, 139). Lastly, an algorithm might also produce an unexpected imbalance that people may find unacceptable, such as gender disparity. It will then require assessment, even if the disparity does not concern the law (ibid.).

In brief, the popular perception of algorithms as 'black boxes' originates from the public's inability to scrutinize calculations. It is true that reading an algorithm's code does not tell one what they do, nor does it indicate discrimination, especially disparate impact and disparate treatment. Contrary to beliefs, algorithms are more like 'glass boxes' that allow observers to look inside and inspect what each component does (Card 2017). Most importantly, they allow one to scrutinize inputs and test outputs (Kleinberg et al. 2018, 114). Algorithms are thus

predictable because humans design them and determine the outcome, the set of input variables and the training procedure that will allow the screener to "select the subset of input variables that optimize the resulting estimate of the outcome measure" (Kleinberg et al. 2018, 143). If the outcome is unfair discrimination, humans and their choices, actions and decisions are to blame, not algorithms. This argument applies to traditional scoring systems where humans, in particular loan officers, assess borrowers' creditworthiness using economic data and personal and moral evaluations. In the end, finance and societies must deal with biases ingrained in both human and automated credit scoring systems and decision-making processes.

3.2 Credit Scoring in Everyday Life: Nudging Behavior and Limiting Autonomy

Algorithmic credit scoring raises significant public concern about consumer autonomy and privacy. Autonomy refers to consumers' ability to make choices and decisions free from outside influence. The use of algorithmic credit scoring by consumer markets and states to monitor and influence human behavior is a double-edged sword. It can have a positive impact as shown in the previous chapter. However, it can also have detrimental effects on people's lives, choices, and agency. The struggle for behavior control is an uneven battle as scorers and lenders have diverse and powerful tools to impose it on borrowers, whereas borrowers have limited means to resist it.

In the development field, behavior change is the bedrock for digital financial inclusion. Credit scoring firms and lenders leverage on big data to discipline borrowers' behavior through rewards and sanctions, especially no-file and thin-file loan applicants. In 'Mind, Society, and Behavior', the World Bank (2015) develops a behavioral approach to poverty. It assumes that gaining a deep understanding of human behavior is key to fostering development. Poverty "is not only a deficit in material resources but also a context in which decisions are made. It can impose a cognitive burden on individuals that makes it especially difficult for them to think deliberatively" (World Bank 2015, 13). Poverty is therefore a "cognitive tax" that makes the poor react automatically – rather than purposefully – to economic stimuli, which leads to error and failure. In particular, "high consumer debt often results from a form of thinking automatically, in which individuals attach much more weight to current consumption through borrowing than to the loss of consumption that will occur when they have to pay back a loan in the future" (World Bank 2015, 14–15). This suboptimal behavior renders borrowers unpredictable and risky for lenders. Ultimately, it generates market instability and poverty. The World Bank's solution to this issue is not to change structural conditions like poverty and

inequality that impair human behavior but to correct the poor's behavior through surveillance, behavior nudging, and financial education. Overall, the poor are blamed for their failure and made responsible for improving their situation through "highly cost-effective behavioral interventions" (World Bank 2015, 13; for a critique of this paradigm, see Berndt and Boeckler 2016; Gabor and Brooks 2017; Mader 2016). Algorithmic credit scoring facilitates these interventions. It helps financial markets determine borrowers' inclination for financial services, monitor and shape their (repayment) behavior through pricing, redlining and personalization, and collect up-to-date data about risk that lenders can use to improve risk management.

Consumer lending markets use credit pricing and customization to discipline behavior as well. Loan prices and conditions result from the sorting and slotting of people in "market categories" based on their economic performance (Fourcade and Healy 2013, 561). In the US, these market categories are linked to credit scores that "quantify individual performance, determining which services can be obtained, in terms of type (home equity, credit card, or payday loans), volume (how much credit is extended), and price (the interest rate, required origination or balloon payments, and other fees)" (Fourcade and Healy 2013, 566). Credit scores are "moving targets" that continuously change based on borrowers' endeavors (Fourcade and Healy 2013, 568). They offer incentives for compliance (wider range of loans available, higher principals, lower prices, rates, and fees) and sanctions for failure (a small range of loans, lower principals, higher prices, rates, and fees). Consumers must therefore continuously manage their credit identity to be rewarded prime loans, and most crucially, to enhance their "life chances" (Fourcade and Healy 2013). The reason is that credit scores have become metrics to measure not only creditworthiness but also trustworthiness. In the US, a favorable credit score is a requirement to purchase a home, a car, a cell phone on contract with no security deposit, seek higher education, start a new business, and secure a job with a good employer. Conversely, bad scores hamper one's 'life chances'. In short, credit scores make and undo fates. However, scored individuals are often unable to determine what optimal behavior they should observe to improve their 'life chances'. In traditional credit scoring, the likelihood of repayment is based on assessing income and timely loan repayments from the past. This link between past and future behavior makes it easy for borrowers whose loan application is rejected to review their credit records and improve their behavior and future chances of approval. Furthermore, scorers and lenders can advise borrowers on how to improve their credit records and behavior. However, this transparency is missing in algorithmic credit scoring. Scorers and lenders are often unable to explain the motive of rejection and what data has proven significant in a

particular decision. As a result, an advice industry is growing in the US to help people adjust their behavior and improve their score and 'life chances' (see, i.e., Warren 2017; Weston 2007).

In Vietnam, algorithmic credit scoring is at its early stage of deployment. Banks and consumer lending companies increasingly use it to assess creditworthiness and make traditional and digital credit widely accessible to banked and un(der)banked borrowers (Lainez 2021). They also take the initiative to educate borrowers on credit scoring and advice on how to adjust their behavior to improve their scores and chances of obtaining loans (see FE Credit 2016; Tran 2019). The media pays limited attention to algorithmic credit scoring. A particular Vietnamese article is worth mentioning. In 'Credit Score: An important factor in the life of international students in the US' (my translation), the journalist orients the behavior of Vietnamese students who seek to study in the US and borrow money to fund their studies and consumption (Duong 2018). He shares "the reasons why you should accumulate credit points and how to learn how to build credit effectively and safely while living in the US". It also describes the factors affecting credit scoring and how to build it by "using your credit card responsibly" and "regularly moving money into your savings account [which] shows you take your finances seriously and are figuring out for the long-term future." Even if it is too early to observe behavior change among consumers due to credit scoring, these materials reveal the gradual emergence of algorithmic credit scoring and its normative power for guiding behavior.

3.3 Social Credit and Mass Surveillance

AI further threatens individual autonomy and privacy when governments co-opt big data to set panoptic mass surveillance systems designed to monitor and discipline human behavior. An example is the social credit system currently tested in China. This system uses traditional (public records) and non-traditional data to classify individuals, companies, social organizations, and institutions in red and black lists that carry rewards and sanctions respectively. The concept of 'social credit' (*xinyong*) blends creditworthiness and trustworthiness (Zhang 2020). The system's goal is to improve governance and market stability and foster a culture of trust as a bulwark against deviance, especially fraud, corruption, counterfeit, tax evasion and crime (Wong and Dobson 2019, 220). At this early stage of development, a few dozen local governments implemented the program with eight IT companies, including Tencent, Baidu, and Alibaba's Ant Financial. Local governments collect traditional data from the Public Security Ministry, the Taxation Office and legal financial institutes, whereas IT firms collect big data from their users, often through loyalty programs.

Their shared goal is to develop algorithms and procedures that can be centralized and applied nationally in a near future (Wong and Dobson 2019, 222).

Alibaba's Ant Financial runs a program that has generated considerable attention and controversy in the West: Sesame Credit (Zhima Credit in China). It leverages the vast database of Alipay, a payment app with over a billion users. It collects data through its built-in features to rank voluntary users based on "credit history; behavior and preference; fulfillment capacity; identity characteristics; and social relationships" (Chong 2019, 5). It rewards 'trustworthy' users who act 'responsibly' with high scores and privileges such as a fast-track security lane at Beijing Airport, fee waivers for car rental and hotel bookings, and higher limits for Huabei, Alipay's virtual credit card. Low-score users may encounter difficulties in lending, renting, finding jobs, and so forth (Chong 2019, 6). While social credit scoring systems such as Sesame Credit induce deep anxiety about data mass surveillance in the West, Chinese citizens perceive it as an innocuous, trustworthy, and secure payment app that helps produce personal and social good (Chong 2019, 14). As for firms, those red-listed with good regulatory compliance history enjoy rewards such as "lower taxes, fast-tracked bureaucratic procedures, preferential consideration during government procurement bidding, and other perks" and those black-listed endure sanctions that include "punishments and restrictions by all state agencies") (Trivium China 2019, 3). It should be noted that a few of the local governments that are testing the social credit system are skeptical about its capacity to develop a prosperous, trustworthy and harmonious society, as claimed by the central government. Some of these governments even contest its legality and show resistance in its application (Zhang 2020, 18).

Could Vietnam follow the Chinese social credit system? Administrative, cultural and political structures, policies and practices could lay the foundation. In China, the social credit system aligns with the state's long tradition of monitoring people and guiding their behavior. It is also inspired by an old 'public record' system (*dang'an*) that contains "diplomas and degrees, academic transcripts, professional qualifications, work appraisals, political affiliations and major political activities, any awards and disciplinary sanctions received, and his or her history of employment, promotions, transfers, dismissals and retirement" (Chen and Cheung 2017, 375). Vietnam has its own *dang'an* system called 'personal resume' (*sơ yếu lý lịch lý lịch*).

Local authorities collect these data where citizens have their household registration (*hồ khẩu*)[25]. The pending digitalisation of the household registration system, of which the 'personal resume' system depends on, will make it much easier for the government to collect personal data. It may also give rise to the possibility of the government exercising some form of social credit, should they decide to replicate the Chinese social credit system. Like China, Vietnam has an authoritarian regime that restricts basic freedoms such as speech and protest, imprisons dissidents, and exerts strict control over the media. It also subsidizes and collaborates with state corporations such as FPT (Corporation for Financing and Promoting Technology), including FTP-AI, to develop AI and embrace the fourth industrial revolution (see next chapter). The Vietnamese government has already been leveraging on AI to develop monitoring systems such as facial recognition to track citizens and human activity. In the near future, it could also co-opt credit scoring technology from FPT-AI and private startups to run an experimental social credit system. At this stage, however, the government has not shown any interest in taking this path. It gains legitimacy by ensuring economic growth and political stability and supporting public companies. Replicating a look-alike Chinese social credit system to further control and suppress political freedom could tarnish Vietnam's international image. This risk, in turn, could affect investment, growth, and political stability. Regardless, we should keep a watchful eye as the government could reverse its decision in. a near future.

To sum up, concerns about the loss of autonomy, behavior control, and mass surveillance are well-founded. However, these concerns emerge from the Western world, where AI is more advanced and embedded in people's lives. In Southeast Asia, credit scoring startups refrain from addressing these thorny matters. In an article published in Bankingtech, Peter Barcak describes two key principles CredoLab applies to protect privacy. The first is to refrain from asking blanket permissions to access different parts of a user's phone (calendar, contacts, SMS, storage, etc.), a common procedure that grants data collectors open access to data owners' digital footprint and lives. The second is data anonymization from extraction to evaluation to ensure data integrity. (Barcak 2019). This response on consumer autonomy is unsatisfactory. It ignores that users can give consent agreement without being fully cooperative or informed about the privacy agreement and the implications of big data collection and machine learning

---

[25] Personal data includes – address, ethnicity, religion, family composition after the land reform, current position of the family, educational level, language proficiency, data and place of admission to the Vietnam communist party, occupation and qualification, salary level, family background, social activity participation, education and school attended, and rewards received (see Lan 2020).

analytics (Aggarwal 2020)[26]. It also overlooks pressing debates about the power of big data and AI with regards to monitoring the population and the ability it provides to make and undo fates. By supporting pervasive data collection that gives lenders the power to know more about borrowers than borrowers might know about themselves, borrowers inflict subjective harm to themselves and give up much autonomy and privacy to fintech firm, (credit) markets and states.

3.4 Cybersecurity: A Growing Threat

Another concern about big data is shared worldwide: cybersecurity. This term refers to the fraudulent use of data for profit, in particular, identity theft or new account fraud, synthetic identity fraud, and account takeover to make fraudulent purchases and claims. Fraudsters leverage on networks, big data and the dark web, and replicate good customer behavior to game the system. The number of cyberattacks have been growing steadily with the expansion of digital banking, transactions, networks, and devices in circulation, and the decrease in face-to-face contact between customers and financial providers. A survey conducted by KPMG's Global Banking in 2018 across 43 retail banks worldwide found that over half of respondents experienced an increase in fraud value and volume. Unfortunately, they recovered less than a quarter of their losses (KPMG 2019, 5).

Vietnam is a good case in point as new wealth breeds cybercrime. With rapid economic and digital growth and a growing and impressive number of internet (66 percent out of 98 million inhabitants) and social media users (60 percent), Vietnam has been described as an "El Dorado for cyber-offenders" (Tech Collective 2019). Cyberattacks and data breaches are widespread. In 2019, the Government Information Security Commission reported 332,029 access attacks (using improper means to access a user's account or network) and 21,141 authentication attacks (using fake credentials to gain access to resources from a user) (Vietnam Security Summit 2020). Due to the absence of survey data on cybercrime, the only figures available are from law enforcement agencies which comprise of only reported prosecutions and publicized cases. In a country where cybercrime remains largely underreported, these cases form just the tip of the iceberg (Luong et al. 2019, 294). The Global Security Index compiled by the U.N. International Telecommunication Union ranked Vietnam 101st out of 195 countries in 2017.

---

[26] For Nikita Aggarwal (2020, 16), "this is in part due to the ambiguity of standard-form online contracts (privacy policies and terms of service) that notoriously include vague, catch-all clauses enabling service providers to collect, store and re-use customer data for related purposes". Jathan Sadowski (2019) describes common forms of data collection and accumulation as 'data extraction' based on little regard for consent and compensation.

This is far behind its ASEAN peers; Thailand, Malaysia and Singapore were ranked in the top 20 (VnExpress 2017). In 2018, Vietnam made progress in the fight against cybercrime and rose in ranks to the 50th position. However, Vietnam still falls short behind Singapore, Malaysia, Thailand, and Indonesia (VnExpress 2019).

The media regularly report data breaches. In 2019, the Maritime Commercial Joint Stock Bank (MSB) suffered a personal data leak. A list comprising of two million account holders' names, ID, phone numbers, addresses, birthdates, gender, emails, and occupations was posted on a website that traded stolen data. It is unclear if the data was leaked by an employee from the bank or a hacker who attacked the bank's database (Bao 2019). In 2018, the police dismantled a syndicate of 12 cybercriminals that broke into the server of an unspecified bank and accessed the accounts of customers who had not subscribed to online banking services. After fraudulently subscribing to these online banking services, they impersonated bank officers and called customers to ask for a one-time password (OTP) to disburse a 'loan'. They used OTPs to log into the accounts of 560 customers and stole VND43 billion (US$1.8 million) (VNS 2018). In 2016, a massive cybersecurity breach made the news headlines. An alleged group of Chinese hackers stole personal data from 410,000 VIP member accounts of Vietnam Airlines' Lotusmiles fidelity program. They published their names, birthdays and addresses online. Concerned that the account holders' credit card details were also stolen and used to make online purchases, two Vietnamese banks, Techcombank and VietinBank, disabled their online payment feature as a cautionary measure to reassure clients. In addition to stealing personal data, the hackers compromised the flight information display and loudspeaker system in Ho Chi Minh City and Hanoi's airports to diffuse offensive messages about Vietnam in relation to the South China Sea dispute (Tuoi Tre News 2016).

Credit scoring firms take cybercrime seriously and heavily market their anti-fraud tools. CredoLab and TrustingSocial proposed face retrieval and customer identification solutions aimed at deterring fraudsters that impersonate customers by using masks and digital images. They digitalize KYC procedures to increase security and reduce costs, time, resources and bureaucracy when verifying customer's identity. According to Geoffrey See, Head of Identity at TrustingSocial, AI models trained with specific groups perform badly with different groups. To overcome this hurdle, TrustingSocial trains AI models with Vietnamese faces to deploy them in Vietnam and with Indian faces to deploy them in India (Huynh 2017). Furthermore, credit scoring firms encrypt data from end to end when they share it with lenders and do not

store them within their servers. CredoApply, CredoLab's one-stop-shop onboarding solution gleans their metadata from their mobile phone that are instantly sent to the lender through a safe API channel. CredoLab does not store these data in their servers or share them with third parties[27]. By applying these safety measures, credit scoring firms seek to reassure corporate customers and society of their secure services. However, they are still generating cyber risks by collecting, commodifying, and sharing big data with lenders and third parties. Their actions may therefore have unintended and far-reaching consequences for society, as shown by the persistent growth of cybercrime figures in Vietnam and elsewhere.

4. Regulating a Legal Blind Spot

Big data and AI make big promises but pose significant challenges as well. To rectify these challenges, oversight, safeguards and regulations are required. However, in its current state, Vietnamese law is ill-equipped to regulate big data and AI, including algorithmic credit scoring. This chapter addresses this challenge by proposing regulatory measures that could amend existing credit and personal information protection regulations. These measures encompass three areas: creditworthiness assessment, data collection and use, and privacy. In addition, developing a set of ethical guidelines for responsible, transparent, and safe AI use is also highly recommended. These ethical principles could lead to an AI governance framework in line with that of Singapore.

4.1 Amending Creditworthiness Requirements in Credit Law

Algorithmic credit scoring predicts borrowers' creditworthiness using big data and AI. As shown in the previous chapter, this technology raises new challenges that regulators should address, starting with the concept of creditworthiness in credit law. In Vietnam, credit law is scattered across several legal instruments. The 2010 Law on Credit Institutions is the general framework, but it only addresses creditworthiness tangentially. Other rules can be found in circular 39/2016/TT on lending transactions of credit institutions and/or foreign banks, circular 43/2016/TT on stipulating consumer lending by financial companies, and circular 02/2013/TT on conditions of debt restructuring.

Creditworthiness assessment is regulated in the following manner. Circular 39/2016/TT requires credit institutions to request customers for documents that prove their financial

---

[27] https://www.credolab.com/faq, 'What type of data does CredoApply collect?' (accessed 20 November 2020)

capability to repay their debt before granting or extending credit and establishing the limit, interest rate and terms of a loan (art 7, 9, 12, 13, 17, 31). 'Financial capability' is understood as the "capacity with respect to capital, asset or financial resources" (art 2). If a credit institution rejects a loan application, the customer must be informed of the reasons why (art 17). Credit institutions must also request customers to report the loans' intended use and prove that loans are used legally and adequately (art 7). For this purpose, they may inspect and supervise consumers' use of the loan (art 94 of the 2010 Law on Credit Institutions, art 10 of the circular 43/2016/TT). The Circular 02/2013/TT regulates the collection of personal data for generating internal risk rating systems used to rate customers and classify loans based on risk. It specifies that customers' risk assessment rests on financial qualitative and quantitative data, business and administration situation, prestige, and data provided by the CIC. Credit institutions must update internal risk ratings and submit them regularly to the State Bank of Vietnam in-charge of CIC to expand and keep the CIC database up-to-date. In brief, the law requires creditors and regulators to continually assess and categorize debtors and loans in risk categories.

These regulations on creditworthiness assessment were designed before the current growth of credit markets and the deployment of big data and AI in Vietnam. As a result, they contain gaps and inadequacies. First, lenders are required to assess borrowers' creditworthiness using traditional, financial and (non)credit data ("qualitative and quantitative, business and administration situation"), which they collect from borrowers and the CIC. This framework is inadequate to regulate the collection and processing of alternative data by credit scoring firms. As shown, alternative data raise new challenges because of its nature. Examples of issues include its collection, use, transfer and storage. Data property rights of the scorers, lenders and third parties are unclear while the relevance of the data and accuracy and fairness in human and automated credit decisions are also called into question. By leaving these issues unregulated, the legislator puts the lending industry and borrowers at risk.

Second, credit law does not ban discrimination and especially disparate treatment and disparate impact against vulnerable borrowers through sensitive characteristics like race, religion, national origin, sex, marital status, and so forth. Vietnam does not have any specific law about discrimination in general and credit discrimination in particular. The term appears in the article 16 of the Constitution, which stipulates that "All citizens are equal before the law. No one shall be discriminated against based on his or her political, civic, economic, cultural or social life". Discrimination is also evoked in the Law on Gender Equality that is mostly used in labor issues. The article 12 on "Gender equality in the field of economy" stipulates that "man and woman

are equal in setting up a business, carrying out business and production activities, managing business and are equal in accessing information, capital, markets and labour sources". The lack of regulation on credit discrimination results from the fact that consumer finance has only recently taken off in Vietnam. Credit discrimination might have not existed or been prevalent widely enough yet to instigate new regulation. However, big data and AI are a game changer: if credit scoring firms collect big data comprising sensitive characteristics and utilize them to predict risk, they may (in)advertently (re)produce unfair discrimination.

Third, credit regulation requires credit institutions to provide borrowers with a reason for rejecting loan applications. However, the law does not grant borrowers the right to correct credit data, appeal against human or automated decisions, request for explanations on how decisions are made, and receive suggestions on how to improve their credit record and scores to avoid future rejections. Overall, legal provisions on creditworthiness assessment leave lenders unaccountable for their decisions and adds an extra layer of opacity to credit scoring and decision-making, thereby putting borrowers at disadvantage.

How should credit regulation be amended to enhance customers' rights and protection with the onset of the digital credit revolution in Vietnam? Discussions and insights about credit law in the US can be useful for reference (Danielle and Pasquale 2014; Gillis 2020; Hurley and Adebayo 2017). In the US, two laws regulate consumers' risk assessment: the 1970 Fair Credit Reporting Act (FCRA) and 1974 Equal Credit Opportunity Act (ECOA). The FCRA is a federal law that governs the collection and report of credit data about consumers. It aims at preserving fairness in credit scoring and reporting, making sure that credit decisions are accurate and relevant. It also seeks to protect consumers' privacy by limiting how data is collected, disclosed, kept, and shared. Furthermore, the FCRA grants consumers the right to access, review and edit data on their credit scores, and understand how lenders use their data for making financial decisions. However, the FCRA has several limitations when it comes to big data. First, it does not restrict the types of data that scorers and lenders can collect and use for scoring purposes. Second, the FCRA does not require credit bureaus to disclose their processing techniques for inspection to protect trade secrets. Third, the FCRA grants consumers the right to access credit data to check accuracy and dispute decisions. However, this right may be impossible to exercise effectively with big data, as machine learning algorithms process thousands of variables that cannot be disaggregated and, in some cases, identified. All-in-all, consumers are powerless against a law that throws them the burden of locating unfairness and having to challenge lenders' credit decisions in court. The ECOA is the

second law that regulates creditworthiness assessment. It prohibits discrimination against loan applicants based on proxies such as race, color, religion, national origin, sex, marital status, age, and dependence on public aid. However, borrowers must again prove disparate treatment, a daunting task due to trade secrecy clauses and the opacity of machine learning algorithms[28]. In addition, lenders are free to inflict discrimination for business necessity, in which case the burden of proof shifts back to the discriminated. In short, the FCRA and ECOA give limited protection and recourse to borrowers, while it protects scorers and lenders' business interest.

How do legal scholars overcome FCRA and ECOA's limitations? Mikella Hurley and Julius Adebayo (2017, 196–200) drafted a model bill – the Fairness and Transparency in Credit Scoring Act (FaTCSA) – that aims to enhance transparency, accountability and accuracy, and limit biases and discrimination[29]. It proposes that the regulator grant borrowers the right to inspect, correct and dispute sources of data collected for scoring purposes, the datapoints and nature of the data collected, and credit scores that inform loan decisions. Consumers would acquire the rights to oversee the entire process, be provided with clear explanations on how the scoring process operates and motives for rejections, appeal rejections, and revise their credit data to improve their record. The goal is to make all credit data, machine learning technology, calculation and decision-making processes open, accessible and inspectable by the public, regulators including the Federal Trade Commission, and third parties through audits and grant licensing. Mikella Hurley and Julius Adebayo (2017, 199) also suggested shifting the burden of proof to scorers to make sure that they do not use predictive models and datasets that discriminate vulnerable groups. To prevent discrimination based on sensitive variables or characteristics, they could adhere to 'industry best practices'. Understandably, this proposal may come under fire from industry players as they may object that transparency requirements could provide borrowers with critical knowledge to 'game' credit scoring models. Moreover, the disclosure of trade secrets could hamper innovation and competition in the industry. According to Mikella Hurley and Julius Adebayo (2017) and Danielle Citron and Franck Pasquale (2014), these risks are overridden by the urgent need for transparency, accountability and limitation of bias and discrimination in a world where AI is increasingly used to classify,

---

[28] The burden of proof is high for plaintiffs. According to Hurley and Adebayo (2017, 194) "Historically, in order to make a prima facie case of disparate impact, plaintiffs were required to show three things: 1) a specifically identifiable practice or policy; 2) a statistically significant disparity in treatment between a protected group and other groups; and, 3) a causal link between the disparity and the practice or policy. It has never been sufficient for a plaintiff to simply show an imbalance between a protected group and a non-protected group, no matter how stark."

[29] The principles resonate with the framework proposed by Danielle Citron and Franck Pasquale (2014).

rank, and govern societies. As argued by Nikita Aggarwal (2020), the goal for industry players from the fintech and finance sectors, the regulator, and society should be to negotiate for normative trade-offs between efficiency and fairness, innovation and public interest.

This proposal is relevant to Vietnam, where credit law overlooks transparency and discrimination. The regulator could issue a decree to enshrine the principles described above, updating the 2010 Law on Credit Institutions and circulars 39/2016/TT, 43/2016/TT and 02/2013/TT. The decree could limit the type of data collected, how it will be used, circumstances in which it can be transferred, stored and sold, and whether it should be shared with the CIC. The decree could also ensure that consumers are given the possibility to oversee the entire scoring process and have the right to explanation in the case of rejection, appeal against rejection, and correct incorrect credit data to improve their chances of approval in the future. It could also detail safeguards and procedures on how to open the scoring system for inspection by a public regulatory body – for instance the State Bank of Vietnam – and audits by private and external actors. The decree could also address the fundamental issue of bias and discrimination, in particular the risk of disparate impact and disparate treatment against vulnerable groups. Such groups include the urban and rural poor and ethnic and religious minorities that have been excluded from credit markets since the launch of macro-economic reforms in the early 1990s. The decree could ban the use of sensitive characteristics such as race, color, religion, national origin, sex, marital status, age, and dependence on public aid, and other proxies absent from ECOA but could be relevant to Vietnam[30]. The regulator could design this decree and negotiate optimal normative trade-offs between efficiency and fairness with industry players and civil society organizations.

---

[30] Adding to the fundamental notion of discrimination in credit law is relevant as traditional and credit markets are rapidly expanding in Vietnam. Today, new credit markets target working and middle-class borrowers in urban centres, especially Ho Chi Minh City and Hanoi. Consumer lending remains rather inaccessible in rural areas, where some 60 million inhabitants live, comprising of the majority of the Vietnamese population. However, adopting an input-centred approach that would prohibit algorithmic credit scoring models from using sensitive characteristics like class, gender and ethnicity, as suggested by several legal scholars, is technically challenging if not impossible (Gillis 2020; Kleinberg et al. 2018). The reason is that algorithms may use correlated proxies such as zip code, income, education, phone use patterns, and so forth if they enhance predictive accuracy to assess creditworthiness. According to Talia Gillis (2020, 2), the approaches that "exclude protected characteristics and their proxies, limit algorithms to pre-approved inputs, or use statistical methods to neutralize the effect of protected characteristics […] fail on their own terms, are likely to be unfeasible, and overlook the benefits of accurate prediction". Instead, she proposes to "shift to outcome-focused analysis" as "when it is no longer possible to scrutinize inputs, outcome analysis provides a way to evaluate whether a pricing method leads to impermissible disparities" (ibid.). However, this shift toward empirical testing of algorithmic outcomes raises legal and technical challenges as well (see also Kleinberg et al. 2018) .

4.2 Enforcing Transparency and Delimiting Data Collection in Privacy Law

As we have seen, big data and AI raise public concerns over privacy, transparency, and accountability[31]. These concerns are central to the regulation on personal data protection in Vietnam. However, the dispersed laws and decrees complicates the enforcement of privacy regulations. The reference legislation on these matters is the Law 86/2015/QH13 on Cyber-Information Security (CIS). Other relevant regulations can be found in the Law 67/2006/QH11 on Information Technology (IT), the Law 51/2005/QH11 on E-transactions, the Decree 52/2013/ND-CP on E-commerce, and the Law 59/2010/QH12 on protection of consumers' rights (Nguyen 2019). For the sake of simplicity, I will mostly be referring to CIS and IT in my discussion.

The CIS defines "personal data" as "information associated with the identification of a specific person" (art 3.15) and the "processing of personal data" as "the performance of one or some operations of collecting, editing, utilizing, storing, providing, sharing or spreading personal information in cyberspace for commercial purpose" (art 3.17). The CIS and IT require data collecting companies to obtain consent from data owners prior to personal data collection. For the CIS, consent requirement includes the "scope and purpose of collection and use of such information" (art 17a). According to the IT, companies must "inform those people of the form, scope, place and purpose of collecting, processing and using their personal information" (art 21.2a). However, some exceptions apply. The Decree 52/2013/ND-CP on e-commerce stipulates that e-commerce businesses are not required to obtain data owners' consent when data is collected to "sign or perform contract of sale and purchase of goods and services" (art 70.4b) and "to calculate the price and charge of use of information, products and services on the network environment" (art 70.4c). If lenders operate under e-commerce licenses, this article could allow for a collection of personal data without data owners' consent for assessing creditworthiness and determining loan pricing.

Personal data protection regulation grants rights to data owners, including the right to "update, alter or cancel their personal information collected or stored" by companies and to stop these companies from "providing such personal information to a third party" (CIS, art 18.1). It also requires data owners to "delete the stored personal information when they have accomplished

---

[31] Since privacy is a relative concept, citizens perceptions of it are unsurprisingly conditioned by what they are used to. It raises the question of whether and why Western values should be used as a yardstick to define and regulate privacy in other countries and regions.

their use purposes or the storage time has expired" (CIS, art 18.3). Art 22 of the CIS stresses that data owners have a right to "inspect, correct or cancel such information" (see also art 22 of the IT). In addition, the CIS (art 18.1) and IT (art 22.1) prohibit companies from "providing, sharing or spreading to a third-party personal information they have collected, accessed or controlled, unless they obtain the consent of the owners of such personal information or at the request of competent state agencies". The CIS also requires companies to delete the stored personal data once it has fulfilled its purpose or the storage time has expired and notify the user about this operation (art 18.3). Overall, the law provides data owners a comprehensive set of rights and protections against companies that collect personal data. It also promotes transparency and accountability, the two guiding principles in the proposed decree for amending credit law (see previous section).

However, as comprehensive as the law is, it does not delimit what data companies can collect to protect privacy. This gap is problematic with regards to algorithmic credit scoring given the severe implications big data and AI have in people's lives. In her brilliant analysis of algorithmic credit scoring norms in the UK, Nikita Aggarwal (2020, 2) calls for "substantive restrictions on the processing of (personal) data by credit providers, through legal as well as technical measures". She argues that "consumer credit markets need just the right amount of consumer privacy: not too little such that consumer autonomy is threatened, yet not too much such that the efficiency and fairness gains due to personal data processing are forsaken" (Aggarwal 2020, 17). Her proposal is to ban the collection and "processing of certain types of personal data – such as relationship, health and social media data – that are considered intrinsic to a consumer's identity and autonomy" (ibid.). This suggestion raises a challenging question considering that privacy is not a universal concept, but one that is historically, socially, and politically situated. The question is how to determine what data is intrinsic – and therefore non-commodifiable – for preserving data owners' privacy and 'intrinsic identities'. In a similar vein, the World Bank and the Consultative Group to Assist the Poor (CGAR) (2018, 13–14) have called for "data minimization" to protect the "vital interests of data subjects" and to limit the collection of data to what is strictly necessary for creditworthiness assessment. The task of determining what data are relevant for this purpose would fall on the regulator. This proposal leads to Aggarwal's conundrum: how to determine what data can be collected and commodified to reach normative trade-offs? Aggarwal and the World Bank leave this thorny question open as it requires extensive political, societal and legal negotiation and compromise.

This debate is relevant for Vietnam. The government could begin with launching a consultation with industry players, international bodies, data protection organizations, and civil society organizations to determine what personal data should or should not be collected and commodified for credit scoring purposes. The goal would be to find a balance between efficiency in risk assessment and credit decision-making while preserving data owners' privacy, intrinsic identities and vital interests. The results of this consultation could be made enforceable by issuing a decree that would amend the CIS, IT and other privacy regulations. Meanwhile, the authorities should ensure that credit scoring firms and local and international lenders operating in Vietnam follow the law on personal data protection to respect privacy. In particular, companies should always obtain consent from data owners before collection, grant them the right to update, alter or cancel credit data, delete the data once it has been used for risk prediction, and inform them when data are transferred to third parties. These principles overlap with those described in the proposed decree to amend credit law.

Credit scoring startups remain discreet on how they approach privacy. CredoLab's Privacy Policy provides some guidance. The section 'What rights do I have over my data?' describes the rights that CredoLab grants to data owners:

- "to ask us about the data we process about you, the purpose and nature of the processing, and to provide information on who we share it with."
- "to request that we update or delete (assuming that this does not impact the services we are providing to you) the data we have collected about you at any time […] Unless you request us to delete your data, please note that we may keep your data after you stop being a user (but we typically keep your data no longer than is reasonably necessary given the purposes for which the data was collected)."
- "to expect us to protect your data and keep it safe."

To summarize, data owners must consent to data collection. They also have the right to enquire about the data that CredoLab collects about them and request to update or delete them under specific circumstances. It appears that this privacy policy falls short if the goal is to give data owners the right to update, alter, cancel, and delete their personal data and be informed of transfers to third parties. The regulator, industry players and society could bring together their goals and policies and negotiate a normative trade-off.

4.3 Setting up an AI Ethical Principles Guideline

Apart from amending credit and personal data protection legislation, the Vietnamese government should support the development of ethical principles for responsible AI. AI has become a priority for rich and transitional countries that position themselves as the world and regional leaders in applying machine learning algorithms to strategic sectors such as health, education, finance, transport, and services. To assuage the fears caused by the rapid growth of big data and AI, numerous governments and international organizations have set up ad hoc committees to draft ethical guidelines. Notable examples include the Council of Europe's Ad Hoc Committee on AI, the OECD Expert Group on AI in Society, and the Singapore Advisory Council on the Ethical Use of AI and Data. These committees gather key actors from the public and private sectors. In Singapore, the Advisory Council on the Ethical Use of AI and Data brings together representatives from the government, leading technology firms such as Google, Microsoft and Alibaba, local firms using AI, and social and consumer interests advocates (Monetary Authority of Singapore 2018). The private sector develops ethical guidelines as well, especially leading technology firms (Google, Microsoft, IBM, etc.), professional associations (Institute of Electric and Electronical Engineers, etc.), universities and research centers (Montreal University, Future of Humanity Institute, etc.) and NGOs (The Public Voice, Amnesty International, etc.). A recent review of the global landscape of AI ethical guidelines published in Nature Machine Intelligence identified 84 documents published worldwide in the past five years (Jobin, Ienca, and Vayena 2019). Although these ethical principles are not legally binding regulations, they carry advantages as they steer the private sector into behaving in the desired way and become regulation in the future.

It is in the best interests of Vietnam to develop an AI ethical guideline, or at least, a set of best practices' standards for responsible AI. The country has recently embraced the 'fourth industrial revolution', an expression that refers to the fusing of the physical, digital and biological worlds by incorporating new technologies such as AI, the Internet of things, and robotics. In 2019, the Vietnamese government signed Resolution 52-NQ/TW to kickstart the Industry 4.0 race. This document issues guidelines to invest in science and technology and develop the digital economy, promote rapid and sustainable development, foster innovation, ensure national defense and security, and protect the environment (Vietnam Law & Legal Forum 2020). The resolution also lays the ground for developing regulation in areas ranging from the digital economy to the financial and monetary systems, consumer privacy, and intellectual property. However, there is no timeline set for the development of this ambitious

legal framework. Setting up an ad hoc committee, launching discussions and negotiations with private stakeholders and society, and publishing an AI ethical guideline for responsible AI could be a step forward for the development of an AI governance model in a near future, an essential tool for competing in the Industry 4.0 race.

What would this ethical guideline look like? Anna Jobin Macello Lenca and Effy Vayena (2019, 391–95) identified five primary ethical principles that converge in the corpus of ethical guidelines published globally: transparency, justice and fairness, non-maleficence, responsibility, and privacy. These terms show high semantic and conceptual discrepancies in how these concepts are interpreted, used and associated to particular recommendations. Despite the limitations and overlaps, they could lay a foundation for preparing an AI ethical guideline in Vietnam. They could be defined as follows:

Transparency

This concept is ubiquitous in the literature reviewed. It includes actions to increase explainability, interpretability, communication, and disclosure of AI. The goal is to minimize harm, improve AI performance and foster public trust. What should be disclosed varies significantly from one guideline to another: "use of AI, source code, data use, evidence base for AI use, limitations, laws, responsibility for AI, investments in AI and possible impact" (Jobin, Ienca, and Vayena 2019, 391).

Justice and fairness

Justice refers to fairness and "prevention, monitoring or mitigation of unwanted bias and discrimination" (Jobin, Ienca, and Vayena 2019, 394). The advice is to acquire and process accurate, complete and diverse data, especially training data, to mitigate bias. The Singapore Model AI Governance Framework (PDPC 2019, 35–50) provides recommendations to limit unintended discrimination with models that use biased or inaccurate data or are trained with biased data. These propositions involve understanding the lineage of data, ensuring data quality, minimizing inherent bias, using different datasets for training, testing and validation, and conducting periodic reviewing and dataset updates.

Non-maleficence

This notion calls for safety and security and the avoidance of predictable or accidental risks and harm. Harm mainly refers to discrimination, violation of privacy and bodily harm. Alternatively, it can also refer to the loss of trust or skills, the overshadow of regulation by

technological progress, and long-term societal and political impact. Harm-reduction measures involve policy strategies related to AI research, design, development and deployment, and technical solutions like in-built data quality evaluations, security and privacy measures, and industry standards. If damages are inevitable, risks should be assessed and minimized, and liability should be clearly stipulated and attributed (Jobin, Ienca, and Vayena 2019, 394).

Responsibility and Accountability

These terms are frequently used in the corpus of ethical guidelines on AI, but are seldom defined. They lead to a diverse set of suggestions that propose acting with integrity, specifying responsibility and liability in contracts, whistleblowing and focusing on the processes that generate harm, and bringing ethics into science and technology. In this case, harm is attributed to AI developers, designers, industries, and institutions (Jobin, Ienca, and Vayena 2019, 394–95).

Privacy

Privacy is another key concept that often remains undefined but is extensively used in discussions on data protection. It refers to "a value to uphold and as a right to be protected" and, on some occasions, freedom and trust (Jobin, Ienca, and Vayena 2019, 395). Privacy can be preserved through technical measures (differential privacy, privacy by design, data minimization, access control), research and sensibilization campaigns, and regulation (legal compliance, certification, production, or revision of laws adapted to AI characteristics).

These five ethical principles are relevant to algorithmic credit scoring. Transparency is often present in debates on algorithmic credit scoring regulation, harm reduction, and individual privacy and autonomy protection. It is closely related to responsibility and accountability. Legal scholars propose that consumers have the right to inspect, correct, and dispute sources of traditional and non-traditional data collected for scoring purposes, the datapoints and nature of the data, and credit scores used to make loan decisions (allocation or rejection, loan pricing). Borrowers should be allowed to exercise oversight over the entire process going from big data collection to decision-making. They should also be able to understand why their loan applications are accepted or rejected, review their credit data, appeal rejections, and be given clues on how to improve their credit record and score. The goal is to make all credit data, AI-

based predictive technology, calculation, and decision-making open, explainable, inspectable, auditable, and interpretable by the public, regulators and third parties.

External controllers and AI may be used to for monitoring and to ensure fair decisions and non-harm. For the US, Danielle Citron and Franck Pasquale (2014, 20) argue that the U.S. Federal Trade Commission should be given access to credit scoring systems, especially datasets, source codes and notes on variables, correlations and inferences, for testing bias, arbitrariness, and unfair mischaracterization. In Vietnam, the State Bank could be granted similar privileges to endorse a similar supervisory role and ensure transparency, fairness and accountability. Last but not least, algorithmic credit scoring models could also be subject to licensing and audit imperatives. The regulator could also be authorized to use machine learning technology to pinpoint and solve biases (Aggarwal 2019, 4; Kleinberg et al. 2018).

To sum up, these ethical concerns and legal and procedural propositions are relevant to Vietnam, a country where algorithmic credit scoring is being rapidly deployed. However, the regulators and policymakers do not yet fully fathom the importance of regulating it. Furthermore, a look at how credit scoring startups and lenders deploy this technology suggest their unpreparedness or unwillingness to embrace basic ethical guidelines for responsible AI. For Vietnam's interests, it is our hope that the government sets up an ad hoc committee gathering public and private actors that could eventually produce an ethical guideline on responsible AI. Hopefully, the guideline will inspire a model for AI governance and enforceable regulations, and that credit scoring firms and lenders operating in Vietnam will follow them in spirit, if not to the letter.

5. Conclusion

This study seeks to cast light on algorithmic credit scoring in Vietnam. Assessing its impact is relevant in light of this technology's ambition – to include tens of millions of un(der)banked citizens in financial services markets – and the mixed feelings it has stirred worldwide. On the one hand, the fintech–philanthropy–development complex praises its efficiency and potential for fostering distributional fairness. On the other hand, algorithmic credit scoring fuels anxieties about unfair discrimination and the loss of individual autonomy and privacy. The study concludes that it is too early for a comprehensive assessment of the benefits and risks of algorithmic credit scoring in Vietnam. The outcome will most likely be ambiguous and result in a long process of technological and regulatory negotiation, experimentation, adaptation, and

normalization. A key finding is that this technology raises more questions than answers at this early stage of deployment, especially in emerging countries with limited financial infrastructure and awareness about privacy and data protection.

The study has revealed two critical areas that require monitoring and research in the years to come. The first is the thorny issue of distributional fairness and discrimination. Machine learning algorithms make biased decisions that discriminate against vulnerable groups when they are flawed and/or reproduce human biases. Whether algorithmic credit scoring will reproduce biases and exclusion entrenched in credit markets is relevant in countries where the consumer lending industry has a history of discrimination. However, in emerging countries like Vietnam, the consumer lending industry is still nascent. In the 1990s and 2000s, rural households could only borrow money from Agribank and the Bank for Social Policies for rural development and poverty alleviation. Consumer lending was inexistent back then. To fund consumption and contingencies, families relied on informal finance. They mainly borrowed funds from relatives, friends, moneylenders, employers, rotative credit associations, and pawnshops. The concern that algorithmic credit scoring may reproduce biases entrenched in old credit markets that did not exist before is therefore irrelevant. However, caution needs to be exercised in this technology to draw new consumer credit landscapes from scratch, as they may produce (rather than *re*produce) biases and discrimination against vulnerable groups, in particular rural workers, women and the poor. The rural-urban divide is already becoming apparent as consumer credit targets mainly city dwellers.

Another critical area to monitor is privacy and autonomy. Big data has already raised much anxiety in the West, contrary to Asia where it is only starting to seep into people's lives. A good example of this technology's ambivalent reception is the Chinese social credit system. In the West, programs like Sesame Credit stir anxieties about repressive data surveillance. In China, these concerns are "often superseded by celebratory discourses about convenience, creativity/innovation, cost-effectiveness, efficiency, and, above all, security" (Chong 2019, 2). Ironically, Chinese Alipay's users go as far as to consider that this app's "philanthropic efforts have helped cultivate the awareness of environmental protection and social and personal well-being" (Chong 2019, 14). Essential questions emerge from these remarks; for instance, the extent to which China will succeed in implementing the social credit system nationwide and the possibility that Vietnam may be tempted to replicate it. Another fundamental question is how will Vietnamese citizens and borrowers embrace or resist a technology that rates, ranks and classifies them and may shape their 'life chances' for better or worse. Furthermore, the

critical issue of how users build trust on technology has received limited attention. A recent study about Ant Credit Pay, Alipay's virtual consumer credit service based on Sesame Credit, shows that users appreciate its 'depersonalized' mode of lending as opposed to borrowing from banks, friends, or relatives where human discretion prevails. However, they re-personalize the app by attributing human-like logic to the scoring system, "ascertaining how the platform 'thinks' by embedding it within personal spending and ascribing its mode of operations to Jack Ma's charisma" (McDonald and Dan 2020, 10). This example shows that the success or failure of new fintech depends on human trust and perception to a great extent. The issue of trust begs for empirical research. It reminds us that the success of algorithmic credit scoring, fintech apps, and new financial products depends on efficiency and progress as much as on human appropriation and domestication (see Pellandini-Simányi, Hammer, and Vargha 2015; M. Nguyen 2020).

Finally, this study also highlights the importance of regulation to support fintech innovation and positive socioeconomic change. It suggests a cautious, pragmatic and gradual approach to algorithmic credit scoring. Given that this technology thrives in legal limbo in Vietnam, the regulator should address new challenges posed by big data and AI and reflect on what constitutes fairness and individual's 'intrinsic identities' and 'vital interests', and achieve normative trade-offs as soon as possible. It should also explore how to implement legal safeguards and human oversight to support efficiency and accuracy while preserving public interest and privacy. Equally vital is the need to develop ethical guidelines for responsible AI and eventually an AI governance model. The authorities will also have to make sure that domestic and foreign credit scoring firms and lenders enforce regulation and follow ethical principles. With the Vietnamese Central Bank in the midst of launching a pilot regulatory sandbox program for five key fintech sectors, we have reasons to be optimistic about the Vietnamese government's interest and capacity to regulate the fintech sector. This study calls for the government to appreciate the urgent nature of regulating this technology.

References

Abraham, Facundo, Sergio Schmuker, and Jose Tessada. 2019. "Using Big Data to Expand Financial Services: Benefits and Risks." 26. Research & Policy Briefs. World Bank Group.

Aggarwal, Nikita. 2019. "Machine Learning, Big Data and the Regulation of Consumer Credit Markets: The Case of Algorithmic Credit Scoring." In *Autonomous Systems and the Law*, edited by Nikita Aggarwal, Horst Eidenmüller, Luca Enirques, Jennifer Payne, and Kristin van Zwieten. Munich: C.H. Beck.

———. 2020. "The Norms of Algorithmic Credit Scoring." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3569083.

Aitken, Rob. 2017. "'All Data Is Credit Data': Constituting the Unbanked." *Competition & Change* 21 (4): 274–300.

Bain & Company, Google, and Temasek. 2019. "Fulfilling Its Promise: The Future of Southeast Asia's Digital Financial Services." Singapore: Bain & Company, Google, Temasek.

Bao, Lam. 2019. "Two Million Account Details from Major Vietnamese Bank Leaked Online," November 22, 2019. https://e.vnexpress.net/news/news/two-million-account-details-from-major-vietnamese-bank-leaked-online-4016176.html.

Barcak, Peter. 2019. "Two Keys Protecting the Privacy of the Underbanked in Asia." *Bankingtech.Com*, January 16, 2019.

Bartels, Joachim. 2020. "Meet Our Full Member Vietnam Credit Information J.S.C. (PBC)." *Business Information Industry Association*, January 2, 2020. https://www.biia.com/meet-our-full-member-vietnam-credit-information-j-s-c-pbc.

Berndt, Christian, and Marc Boeckler. 2016. "Behave, Global South! Economics, Experiments, Evidence." *Geoforum* 70 (March): 22–24.

Burton, Dawn, David Knights, Andrew Leyshon, Catrina Alferoff, and Paula Signoretta. 2004. "Making a Market: The UK Retail Financial Services Industry and the Rise of the Complex Sub-Prime Credit Market." *Competition & Change* 8 (1): 3–25.

Card, Dalias. 2017. "The 'Black Box' Metaphor in Machine Learning." Towards Data Science. July 6, 2017. https://towardsdatascience.com/the-black-box-metaphor-in-machine-learning-4e57a3a1d2b0.

Chen, Yongxi, and Anne Sy Cheung. 2017. "The Transparent Self under Big Data Profiling: Privacy and Chinese Legislation on the Social Credit System." *The Journal of Comparative Law* 12 (1): 357–77.

Chong, Gladys Pak Lei. 2019. "Cashless China: Securitization of Everyday Life through Alipay's Social Credit System—Sesame Credit." *Chinese Journal of Communication* 12 (3): 290–307.

CredoLab. 2017. "Alternative Credit Scoring: Banking for the Unbanked." *CredoLab*, November 13, 2017. https://www.linkedin.com/pulse/alternative-credit-scoring-banking-unbanked-peter-barcak.

———. 2019. "Credit Scoring for Improved Lending Decisions." *CredoLab Blog* (blog). June 18, 2019. https://www.credolab.com/blog/digital-scorecards-redefining-credit-scoring-for-improved-lending-decisions.

CSIRO. 2018. "Vietnam Today: First Report of the Vietnam's Future Digital Economy Project." Hanoi: Commonwealth Scientific and Industrial Research Organisation.

Danielle, Citron, and Franck Pasquale. 2014. "The Scored Society: Due Process for Automated Predictions." *Washington Law Review* 89 (1): 1–34.

Djeundje, Viani B., Jonathan Crook, Raffaella Calabrese, and Mona Hamid. 2021. "Enhancing Credit Scoring with Alternative Data." *Expert Systems with Applications* 163 (January): 1–12.

Do, Nhung. 2020. "Regulatory Sandbox: The Key to Fintech Success." *Vietnam Investment Review*, January 20, 2020. https://www.vir.com.vn/regulatory-sandbox-the-key-to-fintech-success-72864.html.

Duong, Anh. 2018. "Credit Score (Điểm Tín Dụng) – Yếu Tố Quan Trọng Trong Cuộc Sống Của Du Học Sinh Tại Mỹ [Credit Score: An Important Factor in the Life of International Students in the US]." *San Sang Du Hoc*, June 28, 2018. https://sansangduhoc.vn/tin-tuc/credit-score-diem-tin-dung-yeu-to-quan-trong-trong-cuoc-song-cua-du-hoc-sinh-tai-my.html.

Earl, Catherine. 2014. *Vietnam's New Middle Classes: Gender, Career, City*. Copenhagen: NIAS Press.

FE Credit. 2016. "Điểm Tín Dụng Xấu Là Gì? Làm Gì Khi Điểm Tín Dụng Xấu? (Phần I) [What Is a Bad Credit Score? What to Do When a Bad Credit Score? (Part I)]." May 23, 2016. https://fecredit.com.vn/taichinh/diem-tin-dung-xau-la-gi-lam-gi-khi-diem-tin-dung-xau-phan/.

Fenwick, Mark, Steven Van Uytsel, and Bi Ying, eds. 2020. *Regulating FinTech in Asia: Global Context, Local Perspectives*. Singapore: Springer.

Fourcade, Marion, and Kieran Healy. 2013. "Classification Situations: Life-Chances in the Neoliberal Era." *Accounting, Organizations and Society* 38 (8): 559–72.

Gabor, Daniela, and Sally Brooks. 2017. "The Digital Revolution in Financial Inclusion: International Development in the Fintech Era." *New Political Economy* 22 (4): 423–36.

Gaglia, Frances. 2020. "How CredoLab Reduces Banks' Credit Risks through Smartphone Metadata." *Asian Banking & Finance*, May 16, 2020. https://asianbankingandfinance.net/financial-technology/in-focus/how-credolab-reduces-banks-credit-risks-through-smartphone-metadata.

Gambacorta, Leonardo, Yiping Huang, and Jingyi Wang. 2019. "How Do Machine Learning and Non-Traditional Data Affect Credit Scoring? New Evidence from a Chinese Fintech Firm." BIS Working Paper n. 834. Basel: Bank for International Settlements.

Ghose, Kalidas. 2018. "Financial Inclusion- Leveraging the Mobile Device: The FE Credit Experience."

———. 2019. "FE Credit: We are the Transformational Market Leader in Vietnam." *EFMA*, November 26, 2019. https://www.efma.com/article/detail/32044.

Gillis, Talia B. 2020. "False Dreams of Algorithmic Fairness: The Case of Credit Pricing." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3571266.

Guo, Guangming, Feida Zhu, Enhong Chen, Qi Liu, Le Wu, and Chu Guan. 2016. "From Footprint to Evidence: An Exploratory Study of Mining Social Data for Credit Scoring." *ACM Transactions on the Web* 10 (4): 1–38.

Hardy, Quentin. 2012. "Just the Facts. Yes, All of Them." *The New York Times*, March 25, 2012.https://archive.nytimes.com/query.nytimes.com/gst/fullpage-9A0CE7DD153CF936A15750C0A9649D8B63.html.

Havard, Cassandra. 2010. "'On the Take': The Black Box of Credit Scoring and Mortgage Discrimination." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1710063.

Hurley, Mikella, and Julius Adebayo. 2017. "Credit Scoring in the Era of Big Data." *Yale Journal of Law and Technology* 18 (1): 147–216.

Huynh, Tuyen. 2017. "TrustingSocial- Financial Inclusion for One Billion Unbanked People." http://www.jaist.ac.jp/~bao/DS2017/Discussion-HNTuyen.pdf.

Jobin, Anna, Marcello Ienca, and Effy Vayena. 2019. "The Global Landscape of AI Ethics Guidelines." *Nature Machine Intelligence* 1 (9): 389–99.

Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, and Cass R Sunstein. 2018. "Discrimination in the Age of Algorithms." *Journal of Legal Analysis* 10: 113–74.

König-Kersting, Christian, Monique Pollmann, Jan Potters, and Stefan T. Trautmann. 2020. "Good Decision vs. Good Results: Outcome Bias in the Evaluation of Financial Agents." *Theory and Decision*, September. https://doi.org/10.1007/s11238-020-09773-1.

KPMG. 2019. "Global Banking Fraud Survey: The Multi-Faceted Threat of Fraud: Are Banks up to the Challenge?" Global Banking Fraud Survey. Zurich: KPMG.

KT Global Business Group. 2019. "Credit Scoring Solution Based on Telecom Data: CRDP & K-Telco ScoreTM." https://corp.kt.com/eng/attach/area/A000000561/C000003656v5.pdf.

Lainez, Nicolas. 2019. "Treading Water: Street Sex Workers Negotiating Frantic Presents and Speculative Futures in the Mekong Delta, Vietnam." *Time & Society* 28 (2): 804–27.

———. 2020. "Debt, Trafficking and Safe Migration: The Brokered Mobility of Vietnamese Sex Workers to Singapore." *Geoforum*. https://doi.org/10.1016/j.geoforum.2020.01.021.

———. 2021. "Navigating Vietnam's Lending App Maze." *Fulcrum*, January 6, 2021. https://fulcrum.sg/navigating-vietnams-lending-app-maze/.

Lainez, Nicolas, Vu Thuy Quynh Nguyen, Bui Thao Le, and Georges Blanchard. 2020. "'Easy to Borrow, Hard to Repay': Credit and Debt in Ho Chi Minh City's Sex Industry." Research Report no. 5. Ho Chi Minh City: Alliance Anti-Trafic.

Lan, Anh. 2020. "Sơ Yếu Lý Lịch Là Gì? Xin Xác Nhận Sơ Yếu Lý Lịch ở Đâu? [What Is a Resume? Where Can I Confirm My Resume?]." *Canavi.Com*, December 10, 2020. https://canavi.com/so-yeu-ly-lich-la-gi-xin-xac-nhan-so-yeu-ly-lich-o-dau.

Le, Duc Thinh. 2017. "An Overview of Credit Report/Credit Score Models and a Proposal for Vietnam." *VNU Journal of Science: Policy and Management Studies* 33 (2): 36–45.

Le, Thu Thuong. 2020. "TrustingSocial Aims to Revolutionize Credit Scoring in Emerging Markets." *KrASIA*, January 14, 2020. https://kr-asia.com/trustingsocial-aims-to-revolutionize-credit-scoring-in-emerging-markets.

Lenddo. 2015. "LenddoScore™: Disruptive Technology Solutions in Credit." Lenddo. https://lenddo.co.kr/pdfs/Lenddo-Scoring-Factsheet-201608.pdf.

L.P. 2016. "It's All about Who You Know: Lenddo Makes Credit Decisions Based on Your Social Network." *Technology and Operations Management- MBA Student Perspectives* (blog). November 18, 2016. https://digital.hbs.edu/platform-rctom/submission/its-all-about-who-you-know-lenddo-makes-credit-decisions-based-on-your-social-network/.

Luna-Martinez, Jose. 2016. "How to Scale up Financial Inclusion in ASEAN Countries." *East Asia & Pacific on the Rise* (blog). February 3, 2016. https://blogs.worldbank.org/eastasiapacific/how-to-scale-up-financial-inclusion-in-asean-countries.

Luong, Hai Thanh, Huy Duc Phan, Dung Van Chu, Viet Quoc Nguyen, Kien Trung Le, and Luc Trong Hoang. 2019. "Understanding Cybercrimes in Vietnam: From Leading-Point Provisions to Legislative System and Law Enforcement." *International Journal of Cyber Criminology* 13 (2): 290–308.

Mader, Philip. 2016. "Card Crusaders, Cash Infidels and the Holy Grails of Digital Financial Inclusion." *BEHEMOTH A Journal on Civilisation* 9 (2): 50–81.

McDonald, Tom, and Li Dan. 2020. "Alipay's 'Ant Credit Pay' Meets China's Factory Workers: The Depersonalisation and Re-Personalisation of Online Lending." *Journal of Cultural Economy*, May, 1–14. https://doi.org/10.1080/17530350.2020.1763424.

Monetary Authority of Singapore. 2018. "Principles to Promote Fairness, Ethics, Accountability and Transparency (FEAT) in the Use of Artificial Intelligence and Data Analytics in Singapore's Financial Sector." Singapore: MAS.

Nguyen, Ha Quyen Hoang. 2019. "Protecting Personal Data on Cyberspace: Enterprise's Obligations under the Laws of Vietnam." December 2, 2019. https://www.inhousecommunity.com/article/protecting-personal-data-cyberspace-enterprises-obligations-laws-vietnam/.

Nguyen, Hai Yen. 2020. "Fintech in Vietnam and Its Regulatory Approach." In *Regulating FinTech in Asia*, edited by Mark Fenwick, Steven Van Uytsel, and Bi Ying, 115–38. Perspectives in Law, Business and Innovation. Singapore: Springer.

Nguyen, Minh. 2020. "Portfolios of Social Protection, Labour Mobility and the Rise of Life Insurance in Rural Central Vietnam." *Development and Change*, July, 1–24. https://doi.org/10.1111/dech.12610.

Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.

Óskarsdóttir, María, Cristián Bravo, Carlos Sarraute, Jan Vanthienen, and Bart Baesens. 2019. "The Value of Big Data for Credit Scoring: Enhancing Financial Inclusion Using Mobile Phone Data and Social Network Analytics." *Applied Soft Computing* 74 (January): 26–39.

Ots, Henri, Innar Liiv, and Diana Tur. 2020. "Mobile Phone Usage Data for Credit Scoring." In *Databases and Information Systems*, edited by Tarmo Robal, Hele-Mai Haav, Jaan Penjam, and Raimundas Matulevičius, 82–95. Communications in Computer and Information Science. Cham: Springer International Publishing.

PDPC. 2019. "Model Artificial Intelligence Governance Framework: Second Edition." Singapore: Personal Data Protection Commission. https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf.

Pellandini-Simányi, Léna, Ferenc Hammer, and Zsuzsanna Vargha. 2015. "The Financialization of Everyday Life or the Domestication of Finance?: How Mortgages Engage with Borrowers' Temporal Horizons, Relationships and Rationality in Hungary." *Cultural Studies* 29 (5–6): 733–59.

Petrasic, Kevin, Benjamin Saul, James Greig, Matthew Bornfreund, and Katherine Lambert. 2017. "Algorithms and Bias: What Lenders Need to Know." Washington, D.C: White & Case LLP.

PR Newswire. 2019. "Trusting Social Sees the Future of Retail Banking in AI, Presented in a Co-Hosted Seminar with the Vietnam Banks Association." *Markets Insider*, May 30, 2019. https://markets.businessinsider.com/news/stocks/trusting-social-sees-the-future-of-retail-banking-in-ai-presented-in-a-co-hosted-seminar-with-the-vietnam-banks-association-1028242315#.

PRweb. 2018. "Vietnam's Orient Commercial Bank Adopts LenddoEFL Psychometric Credit Assessment to Serve the Underbanked Segment," January 22, 2018. https://www.prweb.com/releases/2018/01/prweb15094938.htm.

Rosamond, Emily. 2016. "'All Data Is Credit Data': Reputation, Regulation and Character in the Entrepreneurial Imaginary." *Paragrana* 25 (2): 112–24.

Sadowski, Jathan. 2019. "When Data Is Capital: Datafication, Accumulation, and Extraction." *Big Data & Society* 6 (1): 1–12. https://doi.org/10.1177/2053951718820549.

Saigon Times. 2019. "FE Credit Reshapes Vietnam's Consumer Finance Industry with Its Disruptive '$nap' Digital Lending Platform." *Saigon Times*, January 11, 2019. https://english.thesaigontimes.vn/65498/fe-credit-reshapes-vietnam's-consumer-finance-industry-with-its-disruptive-$nap-digital-lending-platform.html.

Shema, Alain. 2019. "Effective Credit Scoring Using Limited Mobile Phone Data." In *Proceedings of the Tenth International Conference on Information and Communication Technologies and Development - ICTDX '19*, 1–11. Ahmedabad, India: ACM Press.

Standing, Guy. 2013. "Tertiary Time: The Precariat's Dilemma." *Public Culture* 25 (1): 5–23.

———. 2014. *The Precariat: The New Dangerous Class*. London, UK ; New York, NY: Bloomsbury.

Tan, Tianhui, and Tuan Q. Phan. 2018. "Social Media-Driven Credit Scoring: The Predictive Value of Social Structures." *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3217885.

Tech Collective. 2019. "Vietnam Suffers the Most Southeast Asia Offline Cyber Attacks Q2 2019." *KrAsia*, August 9, 2019. https://kr-asia.com/vietnam-suffers-the-most-southeast-asia-offline-cyber-attacks-q2-2019.

Tinh, Le Duc. 2017. "An Overview of Credit Report/Credit Score Models and a Proposal for Vietnam." *VNU Journal of Science: Policy and Management Studies* 33 (2): 36–45.

Tran, Henry. 2019. "Easiest Ways To Improve Your Credit Score." *VietnamCredit*, December 11, 2019. https://vietnamcredit.com.vn/news/easiest-ways-to-improve-your-credit-score_13593.

Trivium China. 2019. "Understanding China's Social Credit System: A Big-Picture Look at Social Credit as It Applies to Citizens, Businesses and Government." Beijing: Trivium China / Social credit watch. http://socialcredit.triviumchina.com/wp-content/uploads/2019/09/Understanding-Chinas-Social-Credit-System-Trivium-China-20190923.pdf.

Tuoi Tre News. 2016. "Banks on the Defense Following Vietnam Airlines Data Breach." *Tuoi Tre News*, August 1, 2016. https://tuoitrenews.vn/business/36271/banks-on-the-defense-following-vietnam-airlines-data-breach.

United Nations. 2015. "Transforming Our World: The 2030 Agenda for Sustainable Development." Washington, D.C: United Nations. https://sustainabledevelopment.un.org/post2015/transformingourworld/publication.

Vietnam Law & Legal Forum. 2020. "Party Determines to Embrace Fourth Industrial Revolution." *Vietnam Law & Legal Forum*, October 6, 2020. https://vietnamlawmagazine.vn/party-determines-to-embrace-fourth-industrial-revolution-16916.html.

Vietnam Security Summit. 2020. "Cyber Security in the AI and Big Data Era: Event Brochure." Vietnam Security Summit. https://securitysummit.vn/docs/2020/Brochure_Vietnam_Security_Summit.pdf.

VNA. 2019. "CIC Launches Portal for Borrowers." *Vietnam Investment Review*, June 12, 2019. https://www.vir.com.vn/cic-launches-portal-for-borrowers-68467.html.

———. 2020. "PM Ratifies National Financial Inclusion Strategy until 2025." *Vietnam Investment Review*, February 3, 2020. https://www.vir.com.vn/pm-ratifies-national-financial-inclusion-strategy-until-2025-73554.html.

VnExpress. 2017. "Vietnam Ranks Bottom in Southeast Asia for Cybersecurity: UN Report." *VnExpress.Net*, July 6, 2017. https://e.vnexpress.net/news/news/vietnam-ranks-bottom-in-southeast-asia-for-cybersecurity-un-report-3609653.html.

———. 2019. "Vietnam Jumps 51 Places in Global Cybersecurity Rankings," August 12, 2019. https://e.vnexpress.net/news/news/vietnam-jumps-51-places-in-global-cybersecurity-rankings-3966278.html.

VNS. 2018. "Sophisticated Cyber Crime on the Rise in Việt Nam." *Viet Nam News*, August 14, 2018. https://vietnamnews.vn/society/463726/sophisticated-cyber-crime-on-the-rise-in-viet-nam.html.

Warren, Carolyn. 2017. *Repair your Credit like the Pros: How Credit Attorneys and Certified Consultants Legally Delete Bad Credit and Restore your Good Name*. Chicago: Bookmark Publishing.

Wei, Yanhao, Pinar Yildirim, Christophe Van den Bulte, and Chrysanthos Dellarocas. 2016. "Credit Scoring with Social Network Data." *Marketing Science* 35 (2): 234–58.

Weston, Liz Pulliam. 2007. *Your Credit Score: How to Fix, Improve, and Protect the 3-Digit Number That Shapes your Financial Future*. 2nd ed. Upper Saddle River, NJ: FT Press.

Wong, Karen Li Xan, and Amy Shields Dobson. 2019. "We're Just Data: Exploring China's Social Credit System in Relation to Digital Platform Ratings Cultures in Westernised Democracies." *Global Media and China* 4 (2): 220–32.

World Bank. 2015. *Mind, Society, and Behavior*. World Development Report 2015. Washington, DC: World Bank.

———. 2018a. *The Little Data Book on Financial Inclusion 2018*. Washington DC: World Bank. https://openknowledge.worldbank.org/handle/10986/29654.

———. 2018b. "Financial Inclusion." *World Bank (Understanding Poverty)* (blog). October 2, 2018. https://www.worldbank.org/en/topic/financialinclusion/overview.

World Bank, CGAP. 2018. "Data Protection and Privacy for Alternative Data." Washington,

D.C: World Bank, CGAP.

Zhang, Chenchen. 2020. "Governing (through) Trustworthiness: Technologies of Power and Subjectification in China's Social Credit System." *Critical Asian Studies*, September, 1–24. https://doi.org/10.1080/14672715.2020.1822194.