Journal of Business Research 131 (2021) 441–452

Contents lists available at ScienceDirect

Journal of Business Research

journal homepage: www.elsevier.com/locate/jbusres

Leveraging machine learning in the global fight against money laundering and terrorism financing: An affordances perspective

Ana Isabel Canhoto

Brunel Business School, Brunel University London, Kingston Lane, Uxbridge, Middlesex UB8 3PH, United Kingdom

ARTICLE INFO

Keywords:
Big data
Artificial intelligence
Machine learning
Algorithm
Customer profiling
Financial services
Anti-money laundering
United Nations
Sustainable development goals

ABSTRACT

Financial services organisations facilitate the movement of money worldwide, and keep records of their clients' identity and financial behaviour. As such, they have been enlisted by governments worldwide to assist with the detection and prevention of money laundering, which is a key tool in the fight to reduce crime and create sustainable economic development, corresponding to Goal 16 of the United Nations Sustainable Development Goals. In this paper, we investigate how the technical and contextual affordances of machine learning algorithms may enable these organisations to accomplish that task. We find that, due to the unavailability of high-quality, large training datasets regarding money laundering methods, there is limited scope for using supervised machine learning. Conversely, it is possible to use reinforced machine learning and, to an extent, unsupervised learning, although only to model unusual financial behaviour, not actual money laundering.

1. Introduction

In 2015, the United Nations General Assembly set out a global agenda for sustainable development consisting of 17 goals which are globally referred to as the Sustainable Development Goals (SDGs). Each individual goal is concerned with a particular social, economic or environmental issue, ranging from poverty elimination (Goal 1) to the strengthening of global partnerships (Goal 17) (U.N., 2019). Together, the goals constitute an ambitious development agenda (Economist, 2015), which will require the concerted efforts of governments and private institutions across the world (Madsbjerg, 2017), and across all goals, in the period leading up to the year 2030.

Given that economic development is negatively correlated with crime (Donfouet, Jeanty, & Malin, 2018), one of the SDGs is specifically concerned with fighting crime. Namely, as part of Goal 16, the U.N. has set out a range of targets aimed at reducing criminal activity around the world, such as significantly reducing all forms of violence, ending trafficking, promoting the rule of law and combating organised crime (U.N., 2019). Crime reduction is an essential step in paving the way for sustainable development, because doing so will support the creation of stable societies, enhance effective governance and promote peoples' well-being (UNODC, 2019).

Money is a key motivator for those engaging in illegal activities (Byrne, 2011). Human trafficking, for instance, generates an estimated U.S.$150.2 billion per year for criminal organisations, through activities such as forced labour, sexual exploitation and organ harvesting (FATF, 2018). Money is also needed to plan and execute criminal operations. In the case of human trafficking, money is needed to move the victims across locations; to run the places and operations where those human beings are exploited; and to bribe the various intermediaries that assist or, at least, condone this criminal activity. Given the strong link between money and crime, most governments pursue initiatives to curtail the movement of money to and from criminal organisations, in an attempt to reduce the criminals' incentive and their ability to engage in illicit behaviour (Ball et al., 2015). These programmes are generally referred to as anti-money laundering and terrorism financing initiatives, or AML programmes for short. The importance of AML programmes in the global fight against crime is such that several Heads of State have joined the Financial Action Task Force on Money Laundering (known as the FATF), with the purpose of sharing intelligence on money laundering and terrorism financing techniques, and setting out measures to combat this activity (FATF, 2019).

Since its inception, the FATF has advocated the use of technology to profile and detect money laundering and terrorism financing activity. Financial transactions (other than direct cash payments) leave electronic traces, and these can be processed and analysed in order to develop insight about the financial behaviours of those engaging in illicit activity, or even to prove criminal association (De Goede, 2012). Hence, it is no surprise that technological solutions such as big data analytics, natural language processing or distributed ledger technology have been

E-mail address: ana.canhoto@brunel.ac.uk.

https://doi.org/10.1016/j.jbusres.2020.10.012
Received 31 August 2019; Received in revised form 1 October 2020; Accepted 3 October 2020
Available online 17 October 2020

touted as an essential component of money laundering detection (e.g. Grint, O'Driscoll, & Paton, 2017, chap. 54). In particular, there is a growing interest in exploring the potential of artificial intelligence (AI) and, specifically, machine learning in supporting AML programmes (Kaminski & Schonert, 2018) and, thus, the global fight against crime. Advocates highlight machine learning's ability to handle large volumes of data, both structured and unstructured, and its potential to discover the patterns of financial behaviour adopted by those engaging in illicit activity (e.g. Banwo, 2018; Fernandez, 2019). Machine learning can also assist in analysing user-generated online content, such as Twitter conversations or YouTube videos, using sentiment analysis techniques, to identify supporters of terrorist groups, affiliation with extremist views or even plans to commit criminal activity (Ahmad, Asghar, Alotaibi, & Awan, 2019; Azizan & Aziz, 2017; Cunliffe & Curini, 2018; García-Retuerta, Bartolomé, Chamoso, & Corchado, 2019).

However, the industry remains cautious, and the use of these technologies is, so far, more experimental than systematic (Zimiles & Mueller, 2019). AI and machine learning are seen as costly technological solutions whose benefits remain unproven, as far as AML programmes are concerned (Grint et al., 2017, chap. 54). Moreover, there is a lack of expertise in understanding and operating AI and machine learning (Grint et al., 2017, chap. 54), which, associated with the lack of transparency of the algorithms underpinning them (Crosman, 2019), creates risks for the organisations relying on them for AML, as well as for the individuals whose financial behaviours are being probed.

In order to reconcile these two opposing views regarding the potential of machine learning technology for crime reduction, via its inclusion in AML programmes, this research investigates the following research question: *To what extent can machine learning algorithms be leveraged to assist with the detection and prevention of money laundering and terrorism financing?*

To pursue this goal, we adopt a socio-technical perspective which explicitly accounts for the technical features of information systems, as well as the social context within which such systems are developed and used (Loebbecke & Picot, 2015; Markus & Topi, 2015). By doing so, we can move beyond a discussion of the potential of machine learning for AML programmes, and start unpacking the variety of technological and social factors, such as the '*arguments, decisions, uncertainties and the processual nature of decision making*' (Bowker & Star, 1999, p. 135), which may support or hinder the performance of the machine learning solution. Specifically, we use the theory of affordances to identify the technical features of an approach to AML powered by machine learning, as well as the social behaviours impacting on the solution's use, and how the two condition each other.

The issues and concerns being expressed in relation to the use of machine learning for financial crime detection mirror those expressed in terms of using this technology more generally. For instance, many senior managers are concerned with their organisations' lack of expertise in handling big data (Merendino et al., 2018), while numerous companies are delaying adoption of AI because they are unsure about how it can help their firms (Bughin, Chui, & McCarthy, 2017). Likewise, there is growing awareness of the risks of AI for individuals, organisations and society (Cheatham, Javanmardian, & Samandari, 2019), including rising evidence about the negative impacts of algorithmic decision-making for organisations and individuals (see Newell & Marabelli, 2015). Hence, the findings from our study are relevant beyond the specific context of the U.N.'s SDGs; they talk to the issues at the heart of today's surveillance society (Zuboff, 2015). The ubiquity of algorithms, and the scale and scope of their impact in everyday life, have led Diakopoulos (2014, chap. 33) to describe them as '*the new power brokers in society*' (p. 2), and to urge researchers to investigate the sources and '*contours of that power*' (p. 29). This paper addresses Diakopoulos (2014, chap. 33), Constantiou and Kallinikos (2015) and others' calls for research, by investigating how the algorithms used in money laundering detection are developed and used to sort and classify financial transactions, and the scope for using machine learning algorithms for that end.

The paper is organised as follows. The next section provides a brief overview of the central role of transaction data and profiling technology in the fight against crime, and the challenges of modelling money laundering behaviour. This is followed by an exposition of the theory guiding this research – the theory of affordances – and its application to AI and machine learning. Subsequently, the details of the approach adopted in our empirical investigation are presented, and this is followed by the empirical findings. After discussing the findings, we reflect on the contributions of our paper to theory and practice, as well as areas for further research.

2. The role of financial profiling in the international fight against crime

Given the central role of money in enabling and even motivating criminal activity (Byrne, 2011), initiatives that hinder the movement of money to and from those individuals engaging in illicit activity are seen as one of the key tools in the international fight against crime (Ball et al., 2015). Financial services organisations are the main point of entry of cash in the financial system, as well as major facilitators of the movement of money globally. Moreover, the movement of cash through the financial system generates records, which can be analysed to understand, prove or even anticipate how money is used or how it was generated (De Goede, 2012). Therefore, governments worldwide have passed legislation ordering financial service providers to analyse how their customers are using the firms' financial products and services, in order to develop intelligence which can assist with crime reduction (Ball et al., 2015).

Developing intelligence about money laundering is a challenge, however, because of the nature of the phenomenon being modelled. Strictly speaking, money laundering does not correspond to one specific behaviour; rather, it can relate to any type of predicate crime, from small-scale tax evasion to the trafficking of weapons of mass destruction. It also includes the case where the money has a legitimate origin (e.g. a salary), but it is used to fund criminal activity (Kaufmann, 2002, chap. 10), as in the case of charitable donations to organisations that support terrorism.

The money launderer may also commit several crimes simultaneously. For instance, human traffickers also commit bribery and tax evasion (FATF, 2018). Moreover, money laundering may involve a varying number of actors, from sole traders to highly sophisticated organised crime groups with their own financial director (Bell, 2002). That is, unlike other decision-making scenarios where knowledge-based systems are modelling a specific behaviour with well-defined boundaries and participants, AML modelling systems need to account for a very broad phenomenon, with many possible behavioural manifestations and combinations of actors.

Not only is it difficult to develop money laundering models, but it is also very difficult to test their performance. The predicted outputs produced by the model would need to be compared with confirmed cases of money laundering in order to fine-tune the model and to improve its accuracy (Zimiles & Mueller, 2019). However, it takes a long time (many months, possibly years) for a suspected case of money laundering flagged by a financial services provider to be formally investigated by law enforcement and, eventually, convicted.

Moreover, money launderers change their modes of operation frequently. For instance, the closure of national borders and the restriction of movement caused by the COVID-19 pandemic is leading to a decrease in the street sale of drugs, and a turn to online sales coupled with courier or mail delivery (Coyne, 2020). Criminals are also likely to take advantage of new financial products or trading strategies, such as using mobile payments (Whisker & Lokanan, 2019) or virtual currencies (Vandezande, 2017).

Therefore, any evidence which may be available to guide modelling gets outdated very quickly. That is, in the case of AML profiling, financial services providers are mostly engaging in speculative

modelling (Kunreuther, 2002).

The third challenge faced by financial services providers concerns the volume and type of data to be analysed. The typical financial organisation will produce, daily, a large volume of transaction records, in addition to structured and unstructured data produced by the organisation's many customer touch points – from login data, to biometric information or chatbot conversations (Fernandez, 2019). AML efforts, thus, require that financial services organisations invest in powerful technical systems to help them process and make sense of such data. In the UK alone, firms invest around £5 billion a year in customer profiling and transaction monitoring technology to assist in AML efforts, according to the latest estimates by the regulator (Arnold, 2018), although there are suggestions that the cost of investing in AML technology, plus the operational costs of AML compliance, outweigh any related benefits, such as improved processes or customer insight (Balani, 2019).

AML systems not only need to be powerful, but they also need to meet other criteria such as stringent data security, customer privacy and identity verification requirements (Grint et al., 2017, chap. 54). Moreover, by law, financial service providers must always be able to prove that the technologies that they use do not unfairly discriminate against certain customers (Crosman, 2019). These requirements mean that financial services organisations are wary of adopting technologies where they lack complete control over use of customer data, or whose workings they do not fully understand, as in the case of black-box type of algorithms. That is, while the AML area seems ripe for machine learning deployment, and some industry players are investing in this technology (Zimiles & Mueller, 2019), there are also various organisational and technical barriers to consider. To research these technical and organisational factors, we draw on the theory of affordances, as outlined next.

3. Theoretical background

The value of machine learning in AML comes not from what the technology is, but from what it enables users to do. Hence, in order to investigate the research question previously presented, we need a lens that accounts for both the technical and the social dimensions, such as the theory of affordances.

The theory of affordances originates from direct perception psychology (namely Gibson, 1979), and studies how the real and perceived characteristics of artefacts condition their use. One of its fields of application is the study of perceptions and use of information technology in organisations (e.g. Leonardi, 2013; Volkoff & Strong, 2013), and the effect of such usage in those organisations (e.g. Markus & Silver, 2008; Sebastian & Bui, 2012).

The term *affordance* refers to the patterns of user behaviour made possible by the properties of an artefact, used in a particular setting. For example, the realisation of the affordance 'surfing the web' results from the interaction between the properties of a web browser and the characteristics of the user (De Moor, 2002). The functional and relational aspects of the artefact are preconditions for activity (Greeno, 1994). That is, they create possibilities for action (Leonardi, 2011). For instance, a switch connected by a wire to a power source enables actors to turn the electricity on and off. The characteristics of the artefact also constrain action (Hutchby, 2001). Staying with the switch example, if the switch is positioned very high on a wall, it limits the ability to be switched off by small persons, such as young children.

In order for these possibilities for action – the real affordances – to be realised, the actor (e.g. an organisation's employee, team or unit) needs to recognise the affordance (Davern, Shaft, & Te'eni, 2012a, 2012b) and enact it. For instance, for the connectivity characteristic of a web browser to enable an internet user to access information on a remote server, the user needs to understand what the browser is for and how to use it.

The actor may recognise the affordance by virtue of the features of the artefact – for instance, the presence of "on" and "off" labels on the switch. In addition, the recognition of affordances is conditioned by the organisational systems in which the artefact is deployed. For example, Leonardi (2013) reported how employees from different departments in one organisation used a training simulation software in markedly different ways. Contextual features which may impact on the recognition of the affordance include the organisational and environmental structures and demands; attitudes and perceptions towards the artefact; the level of effort required from the actor; the actors' skill, ability and understanding; and the actors' ultimate goal (Bernhard, Recker, & Burton-Jones, 2013; Volkoff & Strong, 2013).

Affordances are relational – that is, the realisation of an affordance is both technology- and actor-specific (Strong et al., 2014). Therefore, their study requires the investigation of the technical features of the artefact, the social features related to the user and how the two impact on each other (Volkoff & Strong, 2013; Zammuto, Griffith, Majchrzak, Dougherty, & Faraj, 2007), In this way, the theory of affordances rejects the notion of either a technological or an organisational imperative (Zammuto et al., 2007), and focuses, instead, on the iteration between the two (Leonardi, 2013; Strong et al., 2014).

3.1. The affordances of AI and machine learning

Artificial intelligence is an assemblage of technological components which collect, process and act on data in ways that simulate human intelligence (Canhoto & Clear, 2020). AI can handle large volumes of data, including unstructured inputs such as images or speech, which makes it extremely relevant – or even essential – in the age of Big Data (Kietzmann, Paschen, & Treen, 2018).

The core component of an AI solution is the machine learning algorithm, which processes the data inputs (Skiena, 2012). What distinguishes machine learning from classical programming is that, in the former, the goal of the computational procedure is to find patterns in the data set, i.e. the rules that link the inputs to the outputs. In contrast, in classical programming, the rules are developed *a priori*, and the goal of the computational procedure is to apply those rules to input data, in order to produce an output.

There are various types of machine learning, each applicable to a different type of problem. Supervised machine learning is indicated for situations whereby there are known inputs and known outputs – such as patterns of cell variations vs. stages of cancer (Tucker, 2018). The analyst gives the computer training datasets, with data labelled as either input or output. The function of the algorithm is to learn the patterns that connect the inputs to the outputs, and to develop rules to be applied to future instances of the same problem. The opposite approach is unsupervised machine learning, which is indicated for data sets where it is not known which data points refer to inputs and which ones refer to outputs – for instance, a basket of items frequently bought together. The analyst gives the computer a training dataset with no labels. The algorithm's task is to find the best way of grouping the data points, and to develop rules for how they may be related. An intermediate approach, reinforced machine learning, should be applied to problems where certain courses of actions produce better results than others – for instance, playing a game (Mnih et al., 2013). The analyst gives the computer a dataset plus a goal, as well as rewards (or penalties) for the actions that it takes. The algorithm's task is to find the best combination of actions to attain that goal. To achieve that, the algorithm sorts through possible combinations of data, and analyses the rewards for different combinations, to find the patterns that maximise the overall goal.

The choice of type of algorithm to use should be based on fit with type of problem (Skiena, 2012). However, in practice, the choice is often determined by pragmatic reasons, such as the analyst's skills, compatibility between programming languages (Calvard, 2016) or processing power available (Agarwal, 2014).

Data are integral to the development of machine learning algorithms and, hence, to the system's performance. Without data, algorithms have been described as mathematical fiction (Constantiou & Kallinikos,

2015). Depending on the technical characteristics of the system, this may be only structured data (namely numeric data), or also include unstructured data such as images or voice (Paschen, Pitt, & Kietzmann, 2020).

Datasets may be collated from historical databases, such as shipping addresses or the type of IP connection used (O'Hear, 2016); real time data, collected via physical sensors or online tracking; or knowledge data, such as whether previous product recommendations were accepted or rejected. Moreover, data can be sourced internally or externally.

The choice of which type of data to use, or how much data, is often constrained by the need for compatibility between the different elements of the AI solution. Standardisation increases the ability to use multiple data sources, but also reduces the system's flexibility and limits its contextual richness (Alaimo & Kallinikos, 2017). Another important issue concerns the quality of the training data set, namely how the data were collected, their recency and whether they are representative of the population at large (Hudson, 2017). This problem is particularly relevant in the case of external data, when firms are unable to access and assess the underlying assumptions and data sources (Khan, Gadalla, Mitchell-Keller, & Goldberg, 2016).

Once data have been processed through the machine learning algorithm, the system produces an output, which may vary in terms of type and autonomy from human intervention (Canhoto & Clear, 2020). The examples of machine learning that tend to be featured in the media are those where the system has autonomy to act on the basis of the results of the computational process – such as steering a car without human intervention (Goodall, 2016). However, the system's output could be something as simple as a score, with no performative value until an analyst acts on it (e.g. Elkins, Dunbar, Adame, & Nunamaker, 2013). The output can also be re-entered into the training data set, to further the algorithm's development. For instance, AlphaGo Zero has mastered the board game Go by playing against itself over and over again (Silver et al., 2017). This means that machine learning algorithms have the capacity to learn over time, and to adapt to changes in the environment (Russell & Norvig, 2016). However, it also means that machine learning can create self-reinforcing feedback loops, quickly becoming so complex that analysis can no longer explain how they work. An example was Facebook's AI negotiation bots, which developed their own, incomprehensible-to-humans, language (Lewis, Yarats, Dauphin, Parikh, & Batra, 2017). Self-reinforcing loops can also spread biases and mistakes. For example, AI-powered bots that automatically aggregate news feeds' content can spread unverified information and rumours (Ferrara, Varol, Davis, Menczer, & Flammini, 2014), while automatic trading algorithms have been blamed for creating flash crashes in the U. S. stock market (Varol, Ferrara, Davis, Menczer, & Flammini, 2017). This problem is particularly relevant in the case of predictive analytics, where analysts are unable to assess the quality of the output prior to implementation and scaling (Mittelstadt, Allo, Taddeo, Wachter, & Floridi, 2016).

AI and machine learning can be deployed to perform mechanical, analytical, intuitive or even empathetic tasks, although, given the current state of development of the technology, they are better suited for the former than the latter ones (Huang & Rust, 2018).

In summary (Fig. 1), the potential of machine learning to be used in different scenarios is shaped by technical features such as its ability to learn patterns in data, process various types of data and act autonomously. Moreover, it is shaped by contextual features such as the type of problem to which it is applied, the analyst's skills, system compatibility, processing power, variability of data, quality of the data set, acceptability of the output, comprehensibility, risk of unchecked biases and mistakes and the nature of the task.

Defendants of machine learning use in AML highlight the potential of this technology to discover novel patterns in financial transaction data, and to do so in a cost-effective manner (e.g. Fernandez, 2019). However, whether that potential is realised or not depends entirely on the interplay between the technical and contextual features of AML programmes. We investigated this problem empirically, as described next.

4. Research design

Given the relational and dynamic nature of affordances, they are best studied via qualitative methods (Bernhard et al., 2013). The explanatory case study methodology is particularly well suited for affordances' research (Leonardi, 2013), as it enables researchers to identify the genesis of change, in context (Dubé and Paré, 2003). Case study methodology is also indicated to study the development of algorithms, to eliminate the possible effects of spurious correlations which can mask which variables are used, how and why (O'Neil & Schutt, 2013).

Negotiating access for this type of study is extremely difficult. First, the development and use of algorithms is usually shrouded in secrecy (Beer & Burrows, 2013), with most analysis of algorithmic decision making relying on reverse engineering of algorithms (O'Neil & Schutt, 2013). It is particularly difficult to get access to financial services organisations, due to the secretive nature of this sector and particularly since the banking crisis of 2009 (Canhoto et al., 2017). Moreover, the subject matter of this research (money laundering detection) is deemed

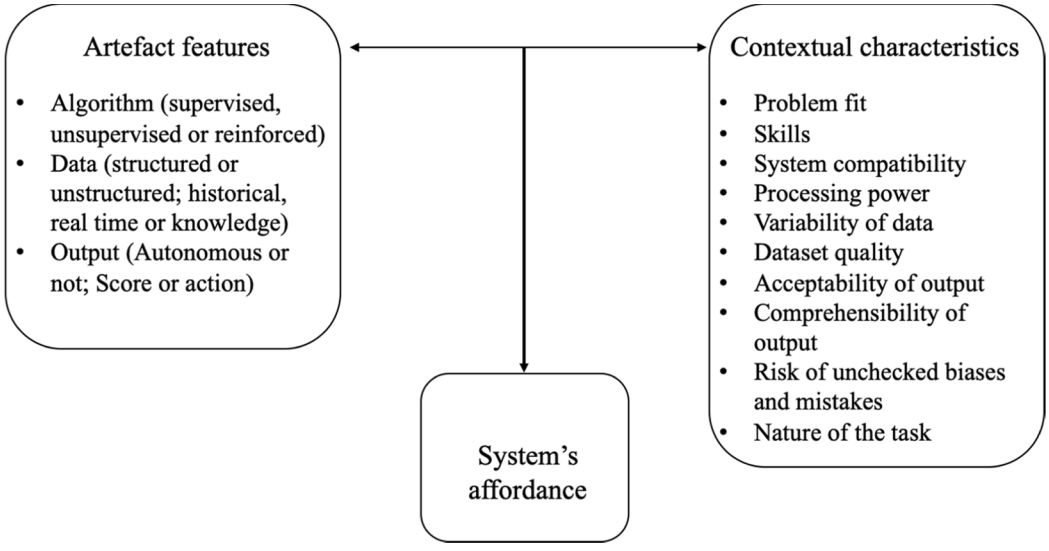

Fig 1. The link between machine learning's features, context and affordances.

by many financial services organisations to be highly sensitive: financial institutions are very reluctant to discuss their approach to money laundering and terrorism financing detection for legal, strategic and operational reasons (Ball et al., 2015). For all these reasons, we used a single, embedded case study.

The focus on a single organisation, while limiting in terms of variability of observations and generalisability of the findings, offers a rich and holistic perspective (Creswell, 2003) that is very much needed in this under-researched problem area. Moreover, the use of multiple sources of data and of types of evidence in the case study offers a depth of insight into thinking and doing processes not available when using other, mono-data collection instrument approaches (Woodside, 2010).

The unit of analysis was a UK-based financial services organisation, to be referred to as BANK. BANK is part of one of UK's largest financial services groups. Its largest business unit is retail banking, contributing to over three-quarters of the group's profit. The retail business includes the provision of current accounts, savings, personal loans and mortgage services, long term investment products and credit cards, among others. In line with the framework articulated in the previous section, data collection had two foci: the technical features (the algorithms and data used, and the type of outputs developed) and the contextual features (the type of problem, skills, etc…) of money laundering profile development at BANK. The data collected is summarised in Table 1. All interviews were recorded and transcribed, while contemporaneous notes were taken during the observations.

The data collected were first analysed according to the data collection instrument (e.g. electronic documents), and then across methods (e.g. interview with vs. observation of system administrator). This approach helped to recognise converging findings, and increased the robustness of the analysis (Jick, 1979). The coding process followed the approach outlined in Miles and Huberman (1994), whereby an initial list of codes was developed based on the theoretical categories depicted in Fig. 1, and applied deductively to the data. This list was, subsequently, augmented with codes emerging inductively during the hermeneutic process of data analysis. Finally, a detailed case history was developed and presented to the research participants, to confirm the accuracy of the findings.

5. Findings

Like other financial institutions in the UK, BANK needs to analyse the records of financial transactions of its customers, in order to identify those that might be linked to underlying criminal activity. The regulator does not provide financial service providers with models for querying the database, so it is up to BANK to develop its own.

An important factor to consider is that the financial transactions are not, usually, illegal. Indeed, unless the client is defrauding BANK, the transactions themselves are legitimate, and part of the normal business of a financial services organisation. That is, what BANK is actually trying to achieve with AML profiling is to identify the patterns of behaviours followed uniquely by customers that are attempting to disguise the illegal source or illegal intended use of their money:

> '*We are trying to find out, first of all, a very basic profile… For instance, a finding may be that a customer aged between 25 and 40 years old is twice as likely to be [involved in criminal activity than] the entire customer base.*' (Interview, Systems manager)

The favoured approach to develop algorithms for AML detection is by drawing on factual information provided by law enforcement agencies regarding confirmed cases of criminal activity. For instance, when prosecutors secure a conviction, some information is made public about the convicted person(s)'s characteristics and financial behaviour. Hence, BANK has access to historical data about confirmed pairings between a specific crime and pattern of transactions. This information is valued by BANK because it provides confirmed inputs (the person's characteristics and behaviours) and confirmed outputs (what crimes the person was convicted of). It is, therefore, amenable to analysis through supervised learning (Fig. 2a). For instance, confirmed reports that a number of terrorist financers had lived in a particular geographical area and been involved in a specific type of business activity led BANK to investigate the transaction patterns of that type of business account:

> *'There is an area (…) with two particular postcodes in which there are lots of [particular type of business mentioned in conviction reports]. One piece of intelligence that we had was that the only two people who were ever convicted for being members of al-Qaida, in the UK, were actually from that area. We know that area and a lot of these [businesses] (…) We can look at the customers who live in that area.*' (Interview, Head of FI team)

However, this type of data is limited in both number and value. In number, because, due to legal and operational constraints, not all details of the convictions are released. Moreover, there is usually a gap of many months – and, often, several years – between the criminal activity, its conviction and the subsequent release of information. Therefore, the training data set is very small. In value, because the information that is made available, by its nature, deals with specific events and, often, unique behaviours. In the case mentioned, the information related to the source of funding of a particular international terrorist organisation associated with al-Qaida. Other terrorist organisations are known to use different sources of funding, such as trafficking or gambling. Moreover, following the conviction mentioned by this interviewee, there was a change in the law that curbed the activities of the type of business mentioned in the reports; and, therefore, limiting terrorist organisations' ability to be funded this way. Finally, while BANK had customers with the characteristics mentioned, this is not always the case. Hence, the training data are not always relevant.

The second favoured approach to develop money laundering detection algorithms at BANK is based on Court Production Orders (CPOs). CPOs are mechanisms used by the court to gain access to specific information about someone who is being investigated for suspected criminal activity. The court contacts financial institutions where the suspect has accounts, asking for their transaction history:

> *'We are asked to provide a lot of witness statements. We often get production orders (…). Our action there is reactive.*' (Interview, Head of FI team)

Following the receipt of a CPO, BANK investigates the pattern of financial transactions for the account(s) flagged (Fig. 2b):

> *'I told them about the scam and the referral we had. When we investigated, it was a garage, and the only thing that was happening was money coming into the account from [country x], and then going out. But there was nothing else: no salary payments, no bills… (…) We later learned that it was all stolen vehicles that went to [country x] and other countries that [drive on] the left.*' (Interview, Trainer)

Table 1
Empirical material collected.

Instrument	Data collected
Documents *(including electronic files)*	• Data file descriptions • Queries used • Suspicious transactions dashboard • Intranet • Marketing and training materials from the system's provider • Internal guidance documents • Meeting notes
Interviews	• Managers: 5 • Systems administrator: 4 • Analysts and other staff: 11
Observations	• Systems administrator: 2 × 1-hour sessions • Analysts: 6 × 1-hour sessions

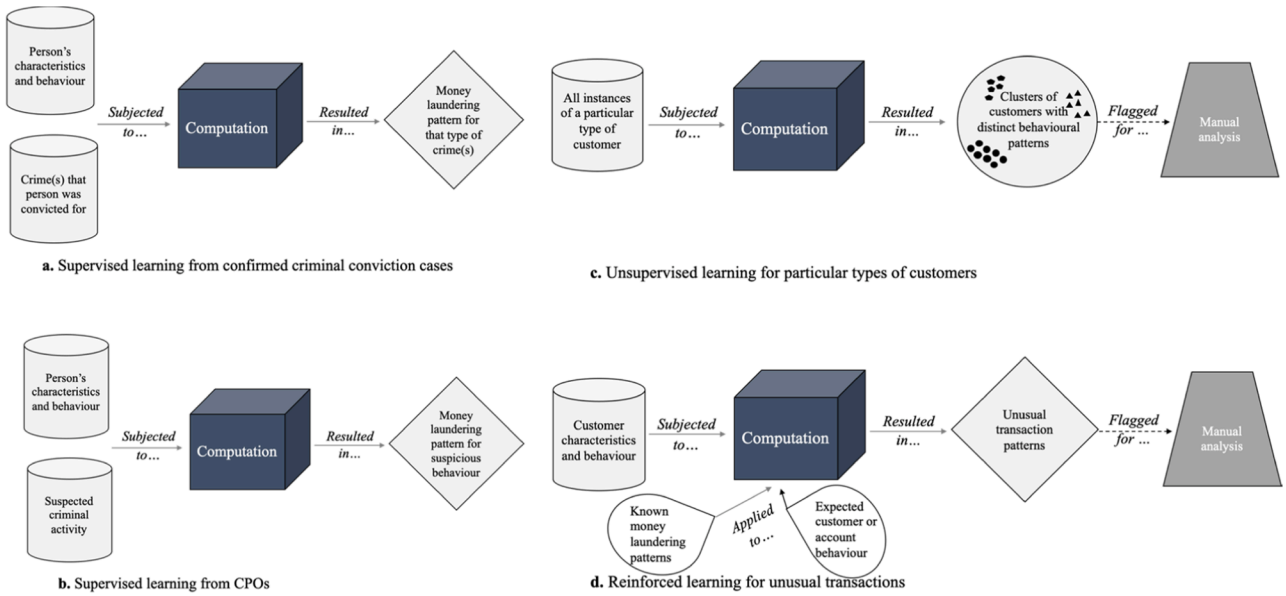

Fig 2. The application of machine learning to AML profiling.

The advantages of CPOs are that they are more frequent and timelier than convictions, which increases their value as training datasets. However, BANK does not know what crime the customers are being investigated for, or even whether they will end up being convicted. Therefore, there is a low level of certainty that any particular pattern identified corresponds to actual criminal activity, or which crime, which limits their value as training datasets.

The third approach used by BANK consists of looking for variations in the financial behaviour of particular types of customers, which is suitable for unsupervised learning. For instance, BANK used this technique to analyse the financial behaviour of business accounts associated with a particular type of trade suspected of engaging in tax evasion. The analysis identified two clusters of accounts with distinct behavioural patterns in terms of cash deposits – one cluster, with a large number of accounts, where traders usually deposited round amounts (e.g. £1,230); and another cluster, with a small number of accounts, where traders tended to deposit exact amounts (e.g. £1,237.50). These clusters were, subsequently, subjected to further probing by the analysts, to identify those customers that might be deliberately attempting to avoid paying tax (Fig. 2c):

> *"Normal behaviour dictates that [these customers] only deposit exact amounts. Is it possible that, therefore, the money launderer might leave it at the exact pounds and pence? Should we be targeting the unusual end instances?"* (Interview, Systems manager)

This kind of analysis is fairly frequent at BANK. Because it focuses on a large number of accounts and current behaviours, this approach avoids the problems of small and dated training data sets that characterise the two approaches previously discussed (Fig. 2a and 2b). Moreover, by alternating the focus of analysis – such as particular types of business activities (e.g. certain types of trade), particular types of customers (e.g. who might be vulnerable to identity fraud), particular types of accounts (e.g. dormant accounts) or particular types of transactions (e.g. international transfers) – it allows the organisations to develop insight about the usual patterns of behaviour of those types of account holders or products.

However, this approach assumes customers engaging in legitimate behaviour have markedly different financial behaviours from those handling the proceeds of crime. It also assumes that the majority of BANK's customers in any given category of analysis are not criminals. Moreover, the analysis of the outliers relies on speculation about the reasons underpinning the observed behaviours.

Another problem faced by BANK when using this approach is that, due to limited processing power, BANK can only run a specified number of queries at any one time. Hence, when the team wants to investigate a new type of behaviour (e.g. deposits in accounts held by traders), the systems manager needs to switch off one of the other queries in use (e.g. international transfers).

The final type of approach uses a combination of pattern analysis of a dataset, and criteria matching, to identify accounts with suspicious patterns of financial transactions. It is the approach used for routine analysis of financial transactions, and it is suitable for application of reinforced learning techniques. Every day, BANK uses this algorithm to analyse the transactions occurred in a given period, giving more weight to those that match known money laundering methods (e.g. depositing large quantities of cash, or quickly defunding an account), and/or that violate rules about the normal use of a given product and/or the expected behaviour for a type of customer. The output is a set of transactions that are deemed to follow an unusual pattern, and that are flagged for further investigation by the analysts (Fig. 2d):

> *'This rule targets [accounts] that have a total of cash deposits between [£X] and [£Y] over a [specific] period. And it is round figures, specifically.* (Interview, Systems manager)

As with the previous approach (Fig. 2c), this one applies to a large number of accounts, rather than relying on small and dated training data sets (as in Fig. 2a and b). Another advantage of this approach is that can be adjusted to reflect BANK's evolving knowledge about its customers and money laundering methods, which is reviewed every week by the financial intelligence team. For instance, the discovery of a money laundering scheme linked to caravan parks led BANK to create a filter for type of accommodation; while another filter that gave more weight to personal accounts without a known residential phone number was dropped when it became clear that more and more customers did not have fixed phones in their houses.

However, this approach is not focused on known criminal behaviour (unlike the approaches depicted in Fig. 2a and b), or on accounts with high likelihood of being linked to criminal activity (unlike the approach depicted in Fig. 2c). Rather, it flags a number of accounts that may have unusual – rather than suspicious – transaction patterns, and which need to be further investigated, manually, by the analysts team. This investigation, like the approach depicted in Fig. 2c, relies on speculation by

the analysts about the legitimacy of the behaviours observed. In turn, the need for the results to be filtered by analysts creates another challenge: queries that produce large amounts of flags undermine the unit's goal of staying within specific performance targets. So, the parameters may be fine-tuned because of the need to limit the number of output cases, rather than because of specific intelligence:

> '*We wrote a rule that says "from our personal account customers, tell us which accounts are receiving in excess of [£X] in a [N] day period". Initially, that prompted many cases, and over time, we brought that figure down to cases between [£Y] and [£Z] over N days.*' (Interview, Analysts' manager)

As with the third approach (Fig. 2c), due to limited processing power, BANK needs to switch off an existing filter whenever it wants to introduce a new one.

In addition to the specific technical and organisational challenges associated with the specific types of algorithms discussed above, there are some generic issues that condition BANK's ability to use machine learning in AML profiling.

In terms of input data, BANK can only use its own transaction databases. Hence, it will always have a limited view of a customer's financial behaviour. For instance, BANK may be aware that a customer has accounts in another financial services organisation, but is unable to access information about the transactions in those accounts. Moreover, due to the system's constraints, BANK only holds data for analysis for a certain number of months. That is, any automated machine learning exercise is done on records from the last X months of activity only (although analysts can query older databases, manually). Furthermore, due to compatibility issues, not all legacy systems can feed into the automated analysis system. One example is the mortgage database. Again, analysts can query the mortgage database individually, but not as part of an automated machine learning exercise. In addition, the system currently in place at BANK can only process numerical data and some types of non-numerical data (e.g. postcodes). It is unable to process free-form text fields, voice and other types of unstructured data:

> '*We don't carry details such as scratch pads, history names, notes that customer advisors might use, telephone conversations… Although some of the conversations could be useful, and we can't run queries on them. So, there is no use in having them into the system.*' (Interview, Systems manager)

Additionally, collecting and keeping updated data on all customers is a costly activity for BANK, and intrusive for the customer. Hence, BANK does not collect and does not routinely update all types of data that the analysts deem useful for AML profiling:

> *'If you went into our branch to open a simple savings account and deposit £500, and we cross-interrogated you about how much money you expect to pass through this account, from where the funds are coming from, what do you do for a living… They are going to say "Hey, do you want my money as a savings account, or not?"'* (Interview, Head of FI unit)

In summary, the type of evidence available to BANK as a training data set, the type of data available for querying, the type of systems in place and other resource constraints, mean that there are significant practical limitations to using machine learning for automated discovery of specific money laundering behaviour. Its main application potential seems to be in the case of speculative analysis of unusual behaviour requiring subsequent manual investigation by analysts. Though, due to system constraints, BANK needs to engage in focused discovery, keeping in mind that the data set may not be as broad, as complete or as up to date as desirable; and that the volume of output (i.e. flagged cases) needs to be manageable within the target deadline.

6. Discussion

Machine learning's ability to discover patterns in data, process various types of data and act autonomously promises to enable financial intermediaries to detect money laundering activity in a cost-effective manner (Fernandez, 2019). Through the use of multiple data collection tools, we researched AML algorithm development at a UK-based financial services organisation. We found that, as far as this type of organisation is considered, the real affordance of machine learning for AML detection falls short of the perceived one (e.g. Banwo, 2018. Fernandez, 2019). We also identified the technical and contextual features that constrained this organisation's ability to tap into machine learning's potential, as summarised in Table 2. Some of these constraints are specific to this organisation, while others are common across the sector, as discussed next.

One of the key criteria in choosing between alternative approaches to machine learning is the fit between the type of approach and the nature of the phenomenon being modelled (Skiena, 2012). Our analysis shows that, in AML profiling, there are, actually, two very distinct phenomena being modelled, each requiring a different approach. One phenomenon consists of developing knowledge about money laundering schemes, via descriptive profiling; the other of detecting attempts to launder money, via predictive profiling. While developing knowledge is an essential step in understanding the nature of the phenomenon, ultimately, to meet the goal of assisting with crime reduction, financial intermediaries need to be able to detect and prevent attempts to launder the proceeds of crime through their organisations (Ball et al., 2015).

In the empirical setting considered, the first type of profiling relies on historical datasets, and produces descriptive outputs. Supervised machine learning algorithms seem best suited for this type of phenomenon. In turn, the second phenomenon relies on real-time data and knowledge, some of which may be derived from the first type of profiling. It produces performative outputs – predictions of high-risk transactions, which need to be investigated by analysts. Unsupervised and reinforced machine learning algorithms fit the second type of phenomenon best. That is, not only are the two types of profiling problem best suited to different types of machine learning algorithm (as per Skiena, 2012), but they also use very different inputs, and produce different types of outputs.

Table 2
Impact of technical and contextual constraints on affordance realisation.

Profiling type:	Knowledge of ML patterns	Detection of ML attempts
Input	• Historic data	• Real time data • Knowledge
Machine learning algorithm	• Supervised	• Unsupervised • Reinforced
Output	• Description • Adds to knowledge database	• Prediction • Flagged for manual analysis
Constraints that impact all financial services organisations in the UK	• Limited view of the customer's transactions • Limited insight regarding reasons for the observed behaviour • Cognitive restrictions • Legal requirements	
	• Small training dataset • Relevant training dataset	• Unchecked assumptions • Unable to test predictions • Bias and stereotyping
Constraints that may be specific to BANK	• Unable to process unstructured data • Compatibility with legacy systems • Access to old(er) data records	
		• Limited capacity to analyse algorithmic outputs manually • Number of queries that can run simultaneously

Regardless of the type of profiling, financial service organisations face the constraint that their perspective – and, hence, ability to model – is limited to the financial transactions processed by the organisation. While some customers may use only one financial services provider for all of their banking needs, many – if not most – will use more than one provider. Hence, any individual intermediary will only process a subset of a customer's financial transactions. Moreover, given that financial organisations do not share information about their customers with each other, for both legal and strategic reasons, each organisation will always have an incomplete dataset of the customer's financial transactions. As a result, they may fail to recognise the importance of a particular transaction, or, conversely, give undue importance to another.

Financial service organisations also have limited insight regarding the reasons underpinning the observed behaviours, because they cannot always probe the customer about the reasons for the observed behaviour. This is particularly the case for online transactions, which have become the norm for around three-quarters of the UK population (Cherowbrier, 2019). The absence of such information, or doubts about the veracity of the information provided, result in inferences, which may be shaped by various cognitive restrictions (Desouza & Hensgen, 2005), and which become crystallised in subsequent decision-making (Bowker & Star, 1999).

We also need to consider the legal requirement for explicability of decision-making, and to prove that no customer has been unfairly discriminated through the use of technology (Crosman, 2019). This is particularly – though not exclusively – likely to occur in cases where the AI system has autonomy to act, and when there are self-reinforcing feedback loops (Canhoto & Clear, 2020), as well as when the algorithm is used for prediction rather than description (Mittelstadt et al., 2016). Based on the description of AML monitoring at BANK, this means that supervised learning might be the least likely to breach these criteria, because it is used for description not prediction, and there are no feedback loops. Generally, unsupervised learning is likely to have low explicability because it has the most potential to produce outputs that are not comprehensible to humans (Lewis et al., 2017). This characteristic puts the financial services organisation at risk of non-compliance with the sector's regulations. In turn, reinforced learning is the most likely to lead to feedback loops, particularly if the rules have been derived from previous unsupervised learning exercises.

Even where information about the reasons for the observed behaviour exists, the organisation may be unable to use it in algorithm development. In BANK, this was the case for information stored in the form of notes, recordings of conversations or other forms of unstructured data. The type of data that BANK's systems could process in practice was much less varied than vast array of data typically mentioned in the AI literature (e.g. Kietzmann et al., 2018). BANK was also unable to draw on all databases due to compatibility issues, or to use data older than a certain period due to system constraints. In theory of affordances' terminology, the realised affordance of AI is much narrower than its functional affordance, which limits its value in AML. Though, this observation reflects the extant literature (e.g. Grint et al., 2017, chap. 54; Zimiles & Mueller, 2019), this may not be the case in other financial services organisations. Other providers may have access to systems that can seamlessly integrate with more databases, and/or which can process all types of structured and unstructured data.

In addition to these generic constraints, there are others that relate to the particular type of profiling, or the approach, pursued. For descriptive profiling, these challenges are mostly related with the availability of high-quality, relevant training data in a timely manner. This is the case not just in AML, but also for other contexts using inductive approaches to model development (Staat, 1993). For instance, a similar problem was observed in the use of machine learning to diagnose those infected with the SARS-CoV-2 virus: even though this technology could potentially read lung scans in a fraction of the time required by a radiologist, initially it lacked sufficient quality images of lungs confirmed to be infected with this virus (vs. lung cancer, for instance) to form a useful training dataset (Ray, 2020).

A related challenge concerns the relevance of the available data. Criminals are constantly innovating their mechanisms of laundering money, such as using mobile payments (Whisker & Lokanan, 2019) or virtual currencies (Vandezande, 2017). Therefore, the limited training datasets available may quickly lose relevance and applicability (Sloman & Lagnado, 2005).

In the case of predictive profiling, the challenges are mostly related to the quality of the underlying assumptions and the inability to test prior to prior to scaling (Mittelstadt et al., 2016). The first key assumption is that the majority of BANK's customers are not engaged in money laundering. While this may be true of the general population (Zhang & Trubey, 2018), it may not be the case for individual organisations, or for all types of predicate crime – for instance, tax evasion or support for terrorist organisations may be very prevalent in certain geographical locations. The second dominant assumption is that the pattern of transactions of customers engaging in money laundering is very different from that of the other customers. Given the dynamic and broad nature of money laundering, this is not necessarily the case for all types of predicate crime and/or customers. Both assumptions are difficult to test, meaning that there are few, if any, opportunities to assess the quality of the models developed (Zimiles & Mueller, 2019).

Moreover, the analysis of outputs produced in the case of the predictive algorithms relies on deductive reasoning, whereby the analysts try to reason about how someone who is trying to use the financial system to launder money without being detected might use the various products and channels at their disposable. This approach is liable to be affected by biases (Pazzani, 2000), stereotyping (Bouissac, 2003) and other cognitive restrictions (Desouza & Hensgen, 2005).

These challenges are heightened by the fact that AML predictive modelling actually focuses on unusual transaction patterns among a specific client base, rather than actual criminal behaviour. If financial organisations were to treat the accounts flagged by the predictive algorithms as certain to be involved in money laundering, and reported them all to law enforcement, it would cause extensive disruption to customers, and could lead to customer complaints and possible financial losses (Ball et al., 2015). Therefore, manual analysis is required, which adds costs and delays to the process. In the case of BANK, the limited availability of manual analysts to scrutinise the algorithmic outputs was one of the key constraints shaping the use of AML predictive algorithms. Other organisations may not experience this constraint, although evidence from other institutions and even other sectors suggests that this is a generalised problem. For instance, it is one of the reasons why online providers such as YouTube or Facebook cannot completely prevent the publication of pornographic or extremist materials on their platforms.

Another constraint that may be specific to BANK refers to the number of queries that can run at any one time. This had to do with a number of technological and organisational reasons, common to many other organisations, such as the investment cycle in new technologies, or the limited appetite to invest in what is seen, across the industry, as a risky and costly solution (Grint et al., 2017, chap. 54). While others have observed similar limitations (e.g. Zimiles & Mueller, 2019), it is possible that other financial services organisations have access to systems that can run more queries simultaneously than BANK, for instance by using cloud services.

The issue of the cost of AI and machine learning technologies should not be underestimated. Small intermediaries may not have enough AML budget to buy sophisticated AI solutions, while large organisations operating across multiple jurisdictions need standardised solutions that few providers are able to offer (Grint et al., 2017, chap. 54). Other trade-offs to consider are whether to focus on processing speed, degree of confidence in the result or learning curve (Cormen, Leiserson, Rivest, & Stein, 2001). Adopting new AI technology may also require additional investment elsewhere in the organisation, such as updating legacy database systems to make them compatible with the new solution. Moreover, there are indirect costs to consider such as recruiting staff

with the necessary subject matter and technical expertise (Merendino et al., 2018), or trying to retain dissatisfied customers (Masciandaro & Filotto, 2001). That is, the calculation of the cost of AI solutions, and therefore the calculation of this technology's cost-effectiveness, is not straightforward. There are numerous trade-offs, and direct and indirect costs, to consider.

7. Concluding comments

This paper set out to investigate if and how machine learning can assist in money laundering detection and contribute to achieving goal number 16 of the U.N.'s SDGs. This question is of interest to both the scholar and the managerial communities (Kaminski & Schonert, 2018). However, there is a lack of empirical investigation regarding the actual use of machine learning in AML, as well as regarding the process of development of algorithms generally (O'Neil & Schutt, 2013). The theoretical framing and subsequent empirical investigation focused on AML monitoring by financial services organisations, given their role as enablers of the movement of cash globally, the legal requirement that they face to detect and prevent money laundering and the abundance of transaction data that they traditionally hold.

Through our consideration of the characteristics of machine learning, and of the phenomenon of AML profiling, we conclude that there are some opportunities for using machine learning to assist with identifying unusual transaction patterns, or even with suspicious behaviour more generally. However, this potential is severely curtailed by the current legal structures, the mechanisms for data sharing between law enforcement and financial services organisations and the relative high cost, complexity and perceived risk of these solutions. Moreover, we did not find any evidence of use of sentiment analysis of user generated online content. Hence, we concur with Arslanian and Fischer (2019) view that the potential for machine learning in AML is far behind that of other applications and other industries (e.g. Castelli, Manzoni, & Popovič, 2016; Fosso Wamba et al., 2017).

7.1. Practical contributions

In terms of this study's contribution regarding the contribution of machine learning to support Goal 16 of the United Nation's SDGs, we showed that its value is very limited at the level of individual financial services organisations. On the one hand, this is because of the nature of the phenomenon being modelled, namely a multi-dimensional phenomenon, characterised by secretive and deceptive behaviours, and which is constantly evolving (Whisker & Lokanan, 2019). On the other hand, this is because of the specific position of financial services organisations in the money laundering supply chain, the limited perspective that they have on their customers' transactions and the nature of the AML task that they are asked to perform (i.e. prevent money laundering).

While financial services organisations may be essential enablers of money laundering and, indirectly, criminal activity, their perspective is limited to the transaction data for their own customers and their own institution. Money laundering often involves multiple individuals and institutions, possibly in multiple jurisdictions, and may take place over an extended period of time. In particular, the type of transnational, organised crime mentioned in the U.N.'s SDG 16 may be difficult to detect via routine AML monitoring by any individual financial services organisation. In some jurisdictions, such as Italy, AML monitoring is done at the national level, rather than at the organisational level as is the case of the UK. It is possible that machine learning would be effective for AML at the national level, for either descriptive or predictive profiling, and further research should consider this specific empirical scenario.

Moreover, financial services providers hold a large volume of data about their customers' identity and behaviour (Fernandez, 2019). However, they lack timely, relevant and sufficient data about money laundering behaviours with which to train machine learning algorithms. As is the case with the novel SARS-CoV-2 virus, the key to unlocking the processing power of machine learning is the training dataset (Ray, 2020). Without such datasets, the actual value of this technology falls very much short of its potential, yet this aspect is largely absent from the discussion about the application of machine learning in AML (e.g. Kaminski & Schonert, 2018), or, indeed, other areas.

In summary, as far as individual financial services organisations are concerned, the short-to-medium-term potential of machine learning for AML has been somehow inflated in the commercial and technical literature, and will require the agreement of standardised approaches to transaction monitoring (Grint et al., 2017, chap. 54).

7.2. Theoretical contributions

While this study focused on the specific case of AML, the findings are relevant for other scenarios, and the study makes contributions to the broader literature on algorithmic decision making and customer surveillance, as discussed next.

The data-driven view of behavioural analysis and decision-making tends to present the use of AI as means of reducing the influence of the analyst on the process and, hence, bias by '*letting the data speak for itself*' (Williams, 2006). However, as this study showed, there is human influence at every step of the process: starting with the data that the organisation decides to collect, after considering the trade-offs between insight potential on the one hand, and collection costs or customer irritation on the other; to the evidence that is considered relevant when developing the models; to the interpretation of the links between data and the assumed underlying behaviour, or the fine-tuning of the models to cope with staff availability. Drawing on the speaking analogy, it is not the case that data speaks for itself. Instead, when algorithms use data, they do so with a vocabulary, a set of grammar rules and a range of assumed pragmatic meanings, that are not only socially construed and subjective, but also contextual (Constantiou & Kallinikos, 2015). In fact, it may be undesirable to eliminate the human element from the process. Subject matter expertise, intuition and social context are all useful in improving the quality of the decision-making. Similar effects were observed in relation to credit decisions during the sub-prime crisis of 2007–08, where there was a higher default rate among loans that were screened automatically than among those where the decision was made, at least in part, manually (Canhoto & Dibb, 2016). Outside of the financial services sector, it has been shown that manual input is essential to improve the quality of the training datasets to detect COVID-19 in lung images (Ray, 2020). Given the significant scope for human influence on data analysis, across different types of applications, further research could explore how different foci and forms of manual interventions impact on machine learning algorithm's deployment, use and performance.

If the models developed via reinforced machine learning wrongly deem a certain transaction as suspicious, there is a 'false positive' error, which generates unnecessary work for analysts and inconveniences the customers. If, on the contrary, the transaction is wrongly deemed legitimate, there is a 'false negative' error, and the financial institution faces the possibility of criminal prosecution, fines and reputation damage. That is, different classification errors impact on different stakeholders. Given the scale of continued expansion of customer surveillance by commercial organisations for their own strategic purposes (Zuboff, 2015), or on behalf of governments (Ball et al., 2015), the possibility of errors and associated consequences is hugely magnified. Yet, by and large, the cost of those errors are not part of the discussions or calculations of the cost and benefits of using AI. Further research could conceptualise the value of using AI in ways which not only consider the characteristics of these technologies, but also the context where they are used, and the consequences of their deployment for a broad range of stakeholders impacted by their use (Newell & Marabelli, 2015), including the possibility of discrimination and victimisation of certain groups, or the erosion of privacy.

Our study considered a broad range of contextual features described in the affordances' literature, and how they impact on the realisation of the affordances of AI. In this way, we contributed to the body of empirical work on the realisation of affordances. In particular, we investigated the constraining aspect of affordances, which is an area that tends to be neglected in empirical research (Volkoff & Strong, 2013), yet is absolutely critical to understand why technology sometimes fail to meet expectations. We also contributed to the body of work on affordance realisation processes at organisational level. Affordances research often adopts a first-person perspective, focusing on the perceptions and actions of individual actors. Yet, as shown by Capra and Luisi (2014), organisations manifest properties different from those of the sum of its groups or individuals. However, we acknowledge that our understanding of the phenomenon would have been more thorough if we had considered both the organisational and the individual perspectives. For instance, we only considered resources at the level of the organisation, yet the characteristics of individual actors within the organisation can also have an impact on the realisation of affordances, namely that the willingness and ability to perceive or realise the affordance may be influenced by the individual's attitudes, skills and previous experiences (Volkoff & Strong, 2013). For instance, it has been shown that key decision-makers' attitudes towards big data influence how, or indeed whether, the organisation takes advantage of this technological development (Merendino et al., 2018).

Finally, we responded to calls for investigating not only how AI is used, but also how algorithms are developed (e.g. Constantiou & Kallinikos, 2015). We followed the approach recommended by O'Neil and Schutt (2013), collecting data about the technology itself, as well as the process and the team in charge of its deployment and use. To be clear, due to the nature of the application (crime detection and prevention), and the conditions of access to the organisation, it was not possible to report, in this paper, on certain technical characteristics of the algorithms deployed by the financial service organisation, such as the proxy variables or the clustering techniques used. Doing so would have undermined the organisation's efforts to detect criminal activity, and would go against the conditions of access granted for this research. The inability to report on these aspects limited the technical contribution from this research. Nonetheless, this paper filled an important gap previously noticed in the literature on algorithmic decision-making (e.g. Beer & Burrows, 2013; Newell & Marabelli, 2015). It was very difficult to secure access to the empirical setting, particularly given the sensitive nature of the application. It required lengthy negotiations, a very detailed plan for the safe collection and storage of the data collected and numerous checks and reassurances. Hopefully, other researchers will be encouraged to use similar research strategies, and other organisations will facilitate access to their algorithms, because they appreciate the urgency of understanding the social, as well as the technical, dimensions of this phenomenon.

References

Agarwal, R. (2014). Editorial - Big data, data science, and analytics: The opportunity and challenge for IS research. *Information Systems Research, 25*, 443–448. https://doi.org/10.1287/isre.2014.0546.

Ahmad, S., Asghar, M. Z., Alotaibi, F. M., & Awan, I. (2019). Detection and classification of social media-based extremist affiliations using sentiment analysis techniques. *Human-centric Computing and Information Sciences, 9*(1), 1–24. https://doi.org/10.1186/s13673-019-0185-6.

Alaimo, C., & Kallinikos, J. (2017). Computing the everyday: Social media as data platforms. *Information Society, 33*(4), 175–191. https://doi.org/10.1080/01972243.2017.1318327.

Arnold, M. (2018). HSBC brings in AI to help spot money laundering. Financial Times. London. https://www.ft.com/content/b9d7daa6-3983-11e8-8b98-2f31af407cc8. Accessed 28 April 2019.

Arslanian, H., & Fischer, F., (2019). The Future of Finance. Cham: Palgrave Macmillan. 10.1007/978-3-030-14533-0.

Azizan, S. A., & Aziz. (2017). Terrorism detection based on sentiment analysis using machine learning. *Journal of Engineering and Applied Sciences, 12*(3), 691–698. https://doi.org/10.36478/jeasci.2017.691.698.

Balani, H. (2019). Assessing the introduction of Anti-Money Laundering regulations on bank stock valuation: An empirical analysis. *Journal of Money Laundering Control, 22*(1), 76–88. https://doi.org/10.1108/JMLC-03-2018-0021.

Ball, K., Canhoto, A. I., Daniel, E., Dibb, S., Meadows, M. & Spiller, K. (2015). The Private Security State? Surveillance, Consumer Data and the War on Terror. Frederiksberg: Copenhagen Business School Press.

Banwo, A. (2018). Artificial intelligence and financial services: Regulatory tracking and change management. *Journal of Securities Operations & Custody, 10*(4), 354–365. https://doi.org/10.1002/9781119208365.ch10.

Beer, D., & Burrows, R. (2013). Popular culture, digital archives and the new social life of data. *Theory, Culture & Society, 30*(4), 47–71. https://doi.org/10.1002/9781119208365.ch10.

Bell, R. E. (2002). An introductory who's who for money laundering investigators. *Journal of Money Laundering Control, 5*(4), 287–295. https://doi.org/10.1108/eb027309.

Bernhard, E., Recker, J., & Burton-Jones, A. (2013). Understanding the Actualization of Affordances: A Study in the Process Modeling Context. In: Proceedings of the 34th international conference on information systems (ICIS 2013) (pp., 1–11). Association for Information Systems (AIS).

Bouissac, P. (2003). Bounded Semiotics: from Utopian to Evolutionary Models of Communication. In H. W. M. Gazendam, R. J. Jorna & R. S. Cijsouw (Eds.), Dynamics and change in organisations: Studies in organisational semiotics. Boston MA, Kluwer: 17-39. 10.1007/978-94-010-0161-8_2.

Bowker, G. & Star, S. L., 1999. Sorting things out: Classification and its Consequences. Cambridge, Mass.: MIT Press. 10.7551/mitpress/6352.001.0001.

Bughin, J., Chui, M., & McCarthy, B. (2017). What Every CEO Needs to Know to Succeed with AI. McKinsey, https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/digital-blog/what-every-ceo-needs-to-know-to-succeed-with-ai. Accessed September 12 2017.

Byrne, E. (2011). Business ethics should study illicit businesses: To advance respect for human rights. *Journal of Business Ethics, 103*(4), 497–509. https://doi.org/10.1007/s10551-011-0885-y.

Calvard, T. S. (2016). Big data, organisational learning, and sensemaking: Theorizing interpretive challenges under conditions of dynamic complexity. *Management Learning, 47*(1), 65–82. https://doi.org/10.1177/1350507615592113.

Canhoto, A. I., & Clear, F. (2020). Artificial intelligence and machine learning as business tools: Factors influencing value creation and value destruction. *Business Horizons, 63*(2), 183–193. https://doi.org/10.1016/j.bushor.2019.11.003.

Canhoto, A. I., & Dibb, S. (2016). Unpacking the interplay between organisational factors and the economic environment in the creation of consumer vulnerability. *Journal of Marketing Management, 32*(3–4), 335–356. https://doi.org/10.1080/0267257x.2015.1123759.

Canhoto, A. I., Meadows, M., Ball, K., Daniel, E., Dibb, S., & Spiller, K. (2017). The role of customer management capabilities in public-private partnerships. *Journal of Strategic Marketing, 25*(5–6), 384–404. https://doi.org/10.1080/0965254x.2016.1148769.

Capra, F., & Luisi, P. L. (2014). The Systems View of Life: a Unifying Vision. Cambridge: Cambridge University Press. 10.1017/cbo9780511895555.

Castelli, M., Manzoni, L., & Popovič, A. (2016). An artificial intelligence system to predict quality of service in banking organisations. *Computational Intelligence and Neuroscience, 2016*, 1–7. https://doi.org/10.1155/2016/9139380.

Cheatham, B., Javanmardian, K. & Samandari, H. (2019). Confronting the Risks of Artificial Intelligence. New York: McKinsey Analytics. https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/Confronting-the-risks-of-artificial-intelligence?cid=other-eml-alt-mcq-mck&hlkid=4c0ac948c366411d8ca259da848fe6e6&hctky=1980318&hdpid=4c9baef0-8a99-4bb9-bdff-0327ade1db12. Accessed 27 April 2019.

Cherowbrier, J. (2019). Share of People using Internet Banking in Great Britain 2007-2019. Statista. September. https://www.statista.com/statistics/286273/internet-banking-penetration-in-great-britain/. Accessed 20 April 2020.

Constantiou, I. D., & Kallinikos, J. (2015). New Games, new rules: Big data and the changing context of strategy. *Journal of Information Technology, 30*(1), 44–57. https://doi.org/10.1057/jit.2014.17.

Cormen, T. H., Leiserson, C. E., Rivest, R. L., & Stein, C. (2001). *Introduction to algorithms* (2nd ed., p. 984). London: McGraw-Hill.

Coyne, J. (2020). Pandemic will force organised crime groups to find new business models. The Strategist. https://www.aspistrategist.org.au/pandemic-will-force-organised-crime-groups-to-find-new-business-models/. Accessed 7 May 2020.

Creswell, J. W. (2003). Research Design: Qualitative, Quantitative, and Mixed Method Approaches. London: SAGE Publications.

Crosman, P. (2019). Can AI's 'black box' problem be solved? American Banker. 184 (January). https://www.americanbanker.com/news/can-ais-black-box-problem-be-solved. Accessed 29 September 2020.

Cunliffe, E., & Curini, L. (2018). ISIS and heritage destruction: A sentiment analysis. *Antiquity, 92*(364), 1094–1111. https://doi.org/10.15184/aqy.2018.134.

Davern, M., Shaft, T., & Te'eni, D. (2012a). Cognition matters: Enduring questions in cognitive IS research. *Journal of the Association for Information Systems, 13*(4), 273–314. https://doi.org/10.17705/1jais.00290.

Davern, M., Shaft, T., & Te'eni, D. (2012b). More enduring questions in cognitive IS research: A reply. *Journal of the Association for Information Systems, 13*(12), 1012–1016. https://doi.org/10.17705/1jais.00317.

De Goede, M. (2012). Speculative Security. The Politics of Pursuing Terrorist Monies. Minneapolis, MN: University of Minnesota Press.

De Moor, A. (2002). Language/action meets organisational semiotics: Situating conversations with norms. *Information Systems Frontiers, 4*(3), 257–272. https://doi.org/10.1023/A:1019946318686.

Desouza, K. C. & Hensgen, T. (2005). Managing Information in Complex Organisations: Semiotics and Signals. London, M.E. Sharpe.
Diakopoulos, N. (2014). Algorithmic Accountability Reporting: On the investigation of black boxes. New York: Columbia Journalism School Tow Centre for Digital Journalism.
Donfouet, H. P. P., Jeanty, P. W., & Malin, E. (2018). Analysing spatial spillovers in corruption: A dynamic spatial panel data approach. *Papers in Regional Science, 97* (S1), 63–78. https://doi.org/10.1111/pirs.12231.
Dubé, L., & Paré, G. (2003). Rigor in IS positivist case research: Current practices, trends, and recommendations. *MIS Quarterly, 27*(4), 597–635. https://doi.org/10.2307/30036550.
Economist. (2015). The 169 Commandments. The Economist. 26 March. https://www.economist.com/leaders/2015/03/26/the-169-commandments. Accessed 29 September 2020.
Elkins, A. C., Dunbar, N. E., Adame, B., & Nunamaker, J. F., Jr (2013). Are users threatened by credibility assessment systems? *Journal of Management Information Systems, 29*(4), 249–262. https://doi.org/10.2753/mis0742-1222290409.
FATF. (2018). Financial Flows from Human Trafficking. Financial Action Task Force: 71, Paris. https://www.fatf-gafi.org/media/fatf/content/images/Human-Trafficking-2018.pdf. Accessed 28 April, 2019.
FATF (2019). History of the FATF. Financial Action Task Force: 72, Paris. https://www.fatf-gafi.org/about/historyofthefatf/#d.en.3157. Accessed 26 April, 2019.
Fernandez, A. (2019). Artificial intelligence in financial services. *Economic Bulletin: Banco de Espana, 2019*(2), 1–7. https://doi.org/10.23977/ceed.2019.049.
Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2014). The rise of social bots. *Communications of ACM, 59*(7), 96–104. https://doi.org/10.1145/2818717.
Fosso Wamba, S., Gunasekaran, A., Akter, S., Ren, S. J-f, Dubey, R., & Childe, S. J. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research, 70*, 356–365. https://doi.org/10.1016/j.jbusres.2016.08.009.
García-Retuerta, D., Bartolomé, Á., Chamoso, P., & Corchado, J. M. (2019). Counter-terrorism video analysis using hash-based algorithms. *Algorithms, 12(5), 110*, 1–9. https://doi.org/10.3390/a12050110.
Goodall, N. J. (2016). Away from trolley problems and toward risk management. *Applied Artificial Intelligence, 30*(8), 810–821. https://doi.org/10.1080/08839514.2016.1229922.
Greeno, J. G. (1994). Gibson's affordances. *Psychological Review, 101*(2), 336–342. https://doi.org/10.1037/0033-295x.101.2.336.
Grint, R., O'Driscoll, C., & Paton, S. (2017). New technologies and anti-money laundering compliance. *London: Financial Conduct Authority*. https://doi.org/10.1007/978-1-4302-6161-2_9.
Huang, M.-H., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research, 21*(2), 155–172. https://doi.org/10.1177/1094670517752459.
Hudson, L. (2017). Technology Is Biased Too. How Do We Fix It? FiveThirtyEight. https://fivethirtyeight.com/features/technology-is-biased-too-how-do-we-fix-it/amp/. Accessed 29 September 2020.
Hutchby, I. (2001). Technologies, texts and affordances. *Sociology, 35*(2), 441–456. https://doi.org/10.1177/s0038038501000219.
Jick, T. D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly, 24*(4), 602–611. https://doi.org/10.2307/2392366.
Kaminski, P. & Schonert, J. (2018). Monitoring Money-Laundering Risk with Machine Learning. McKinsey Quarterly 2. New York: McKinsey & Company, Inc.
Kaufmann, D. (2002). Governance in the Financial Sector: The Broader Context of. Money Laundering and Terrorist Financing. Washington DC: World Bank.
Khan, I., Gadalla, C., Mitchell-Keller, L. & Goldberg, M. S. (2016). Algorithms: The New Means of Production. D!gitalist. https://www.digitalistmag.com/executive-research/algorithms-the-new-means-of-production. Accessed 29 September 2020.
Kietzmann, J. H., Paschen, J., & Treen, E. (2018). Artificial intelligence in advertising: How marketers can leverage artificial intelligence along the consumer journey. *Journal of Advertising Research, 58*(3), 263–267. https://doi.org/10.2501/jar-2018-035.
Kunreuther, H. (2002). Risk analysis and risk management in an uncertain world. *Risk Analysis, 22*(4), 655–664. https://doi.org/10.1111/0272-4332.00057.
Leonardi, P. M. (2011). When flexible routines meet flexible technologies: Affordance, constraint, and the imbrication of human and material agencies. *MIS Quarterly, 35* (1), 147–168. https://doi.org/10.2307/23043493.
Leonardi, P. M. (2013). When does technology use enable network change in organisations? A comparative study of feature use and shared affordances. *MIS Quarterly, 37*(3). https://doi.org/10.25300/misq/2013/37.3.04.
Lewis, M., Yarats, D., Dauphin, Y. N., Parikh, D. & Batra, D. (2017). Deal or No Deal? Training AI bots to Negotiate. Facebook Code. https://code.fb.com/ml-applications/deal-or-no-deal-training-ai-bots-to-negotiate/. Accessed 29 September 2020.
Loebbecke, C., & Picot, A. (2015). Reflections on societal and business model transformation arising from digitization and big data analytics: A research agenda. *Journal of Strategic Information Systems, 24*(3), 149–157. https://doi.org/10.1016/j.jsis.2015.08.002.
Madsbjerg, S. (2017). A New Role for Foundations in Financing the Global Goals. New York: the Rockefeller Foundation. https://www.rockefellerfoundation.org/blog/new-role-foundations-financing-global-goals/. Accessed 26 April 2019.
Markus, M. L., & Silver, M. S. (2008). A foundation for the study of IT effects: A new look at DeSanctis and Poole's concepts of structural features and spirit. *Journal of the Association for Information Systems, 9*(10), 609–632. https://doi.org/10.17705/1jais.00176.
Markus, M. L., & Topi, H. (2015). *Big Data, Big Decisions for Government, Business, and Society*. Report on a Research Agenda Setting Workshop Funded by the U.S. National Science Foundation.
Masciandaro, D., & Filotto, U. (2001). Money laundering regulation and bank compliance costs: What do your customers know? Economics and the Italian Experience. *Journal of Money Laundering Control, 5*(2), 133–145. https://doi.org/10.1108/eb027299.
Merendino, A., Dibb, S., Meadows, M., Quinn, L., Wilson, D., Simkin, L., & Canhoto, A. I. (2018). Big data, big decisions: The impact of big data on board level decision-making. *Journal of Business Research, 93*(December), 67–78. https://doi.org/10.1016/j.jbusres.2018.08.029.
Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis*. Thousand Oaks, CA: Sage.
Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2). https://doi.org/10.1177/2053951716679679.
Mnih, V., Kavukcuoglu, K., Silver, D., Graves, A., Antonoglou, I., Wierstra, D. & Riedmiller, M. (2013). Playing Atari with Deep Reinforcement Learning. arXIV: 1312.5602 (cs).
Newell, S., & Marabelli, M. (2015). Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of 'datification'. *Journal of Strategic Information Systems, 24*(1), 3–14. https://doi.org/10.1016/j.jsis.2015.02.001.
O'Hear, S. (2016). Fraugster, a startup that uses AI to detect payment fraud, raises $5M, TechCrunch. https://techcrunch.com/2017/01/16/fraugster/.
O'Neil, C., & Schutt, R. (2013). *Doing data science: Straight talk from the frontline*. Sebastopol, CA: O'Reilly Media.
Paschen, J., Pitt, C., & Kietzmann, J. (2020). Artificial intelligence: Building blocks and an innovation typology. *Business Horizons, 63*(1), 147–155. https://doi.org/10.1016/j.bushor.2019.10.004.
Pazzani, M. J. (2000). Knowledge discovery from data? *IEEE Intelligent Systems & Their Applications, 15*(2), 10–12. https://doi.org/10.1109/5254.850821.
Ray, T. (2020). AI runs Smack up Against a Big Data Problem in COVID-19 Diagnosis. ZDNet. 4 April. https://www.zdnet.com/article/ai-runs-smack-up-against-a-big-data-problem-in-covid-19-diagnosis/. Accessed 20 April 2020.
Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: A modern approach* (3rd ed.). London: Pearson.
Sebastian, I., & Bui, T. (2012). The influence of IS affordances on work practices in health care: A relational coordination approach. *International Conference on Information Systems, ICIS, 2012*(5), 4270–4280.
Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., … Hassabis, D. (2017). Mastering the game of Go without human knowledge. *Nature, 550*(October), 354–359. https://doi.org/10.1038/nature24270.
Skiena, S. S. (2012). *The algorithm design manual*. London: Springer.
Sloman, S. A., & Lagnado, D. A. (2005). The problem of induction. In K. Holyoak, & R. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning*. Cambridge: Cambridge University Press.
Staat, W. (1993). On abduction, deduction, induction and the categories. *Transactions of the Charles S Peirce Society, 27*, 197–219.
Strong, D. M., Johnson, S. A., Tulu, B., Trudel, J., Volkoff, O., Pelletier, L. R., … Garber, L. (2014). A theory of organisation-EHR affordance actualization. *Journal of the Association for Information Systems, 15*(2), 53–85. https://doi.org/10.17705/1jais.00353.
Tucker, I. (2018). AI Cancer Detectors. The Guardian. https://www.theguardian.com/technology/2018/jun/10/artificial-intelligence-cancer-detectors-the-five. Accessed 29 September 2020.
U.N. (2019). About the Sustainable Development Goals. United Nations. https://www.un.org/sustainabledevelopment/sustainable-development-goals/. Accessed 26 April 2019.
UNODC. (2019). Crime Prevention, Criminal Justice, the Rule of Law and the Sustainable Development Goals. United Nations Office on Drugs and Crime. https://www.unodc.org/e4j/en/mun/crime-prevention-and-sdgs.html. Accessed 26 April 2019.
Vandezande, N. (2017). Virtual currencies under EU anti-money laundering law. *Computer Law & Security Review, 33*(3), 341–353. https://doi.org/10.1016/j.clsr.2017.03.011.
Varol, O., Ferrara, E., Davis, C. A., Menczer, F. & Flammini A. (2017). Online Human-Bot Interactions: Detection, Estimation, and Characterization. arXix:1703.03107.
Volkoff, O., & Strong, D. M. (2013). Critical realism and affordances: Theorizing it-associated organisational change processes. *MIS Quarterly, 37*(3), 819–834. https://doi.org/10.25300/misq/2013/37.3.07.
Whisker, J., & Lokanan, M. E. (2019). Anti-money laundering and counter-terrorist financing threats posed by mobile money. *Journal of Money Laundering Control, 22* (1), 158–172. https://doi.org/10.1108/jmlc-10-2017-0061.
Williams, G. (2006). Data Mining: Desktop Survival Guide. Sydney: Togaware.
Woodside, A. G. (2010). Case study research: Theory, methods, practice. Bingley, WA: Emerald.
Zammuto, R. F., Griffith, T. L., Majchrzak, A., Dougherty, D. J., & Faraj, S. (2007). Information technology and the changing fabric of organisation. *Organisation Science, 18*(5), 749–762. https://doi.org/10.1287/orsc.1070.0307.
Zhang, Y., & Trubey, P. (2018). Machine learning and sampling scheme: An empirical study of money laundering detection. *Computational Economics*, 1–21. https://doi.org/10.2139/ssrn.3161436.
Zimiles, E. & Mueller, T. (2019). How AI is transforming the fight against money laundering. Geneva, Switzerland: The World Economic Forum. https://www.weforum.org/agenda/2019/01/how-ai-can-knock-the-starch-out-of-money-laundering/. Accessed 26 April 2019.
Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology, 30*, 75–89. https://doi.org/10.1057/jit.2015.5.

Ana Isabel Canhoto is a Reader in Marketing. Her research focuses on the use of digital technology in interactions between firms and their customers. One stream of work looks at the use of digital technology on customer insight, such as digital footprints or social media profiling. The other focuses on the impact of technology on targeted interactions, such as the popularisation of algorithmic decision making in customer interactions, or the potential of wearables and beacons for personalisation.