HAL open science

Understanding the Impact of Transparency on Algorithmic Decision Making Legitimacy

David Goad, Uri Gal

To cite this version:

David Goad, Uri Gal. Understanding the Impact of Transparency on Algorithmic Decision Making Legitimacy. Working Conference on Information Systems and Organizations (IS&O), Dec 2018, San Francisco, CA, United States. pp.64-79, 10.1007/978-3-030-04091-8_6 . hal-02083587

HAL Id: hal-02083587

https://inria.hal.science/hal-02083587v1

Submitted on 29 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Understanding the Impact of Transparency on Algorithmic Decision Making Legitimacy

David Goad and Uri Gal

Department of Business Information Systems
University of Sydney Business School, Sydney Australia

Abstract. In recent years the volume, velocity and variety of the Big Data being produced has presented several opportunities to improve all our lives. It has also generated several challenges not the least of which is humanities ability to analyze, process and take decisions on that data. Algorithmic Decision Making (ADM) represents a solution to these challenges. Whilst ADM has been around for many years, it has come under increased scrutiny in recent years because of concerns related to the increasing breadth of application and the inherent lack of Transparency in these algorithms, how they operate and how they are created. This has impacted the perceived Legitimacy of this technology which has led to government legislation to limit and regulate its use. This paper begins the process of understanding the impact of Transparency on ADM Legitimacy by breaking down Transparency in Algorithmic Decision Making into the components of Validation, Visibility and Variability and by using legitimacy theory to theorize the impact of transparency on ADM Legitimacy. A useful first step in the development of a framework is achieved by developing a series of testable propositions to be used in further proposed research regarding the impact of Transparency on ADM Legitimacy.

Keywords: ADM, Algorithmic Decision Making, Transparency, Legitimacy

1 Introduction

Over the last decade, we have experienced an exponential increase in the volume of produced digital data. Current estimates are that 2.5 exabytes of data are being created each day and the number is doubling every 40 months [1]. It is projected that by 2025 we will be creating 163 zettabytes (i.e., one billion terabytes, or one trillion gigabytes) of data globally each year [2]. In many cases a whole new infrastructure has been built to handle the volumes of data being generated. For example, the new Square Kilometer Array in Western Australia, which will be the world's largest radio telescope, is building a vast storage, data and communications infrastructure to handle the data collection requirements [3].

The velocity or speed at which data is being created and updated is also increasing: geospatial data/ locational data derived from the IoT, which often needs to be analysed real time to have any value, is currently ranked as the third largest data type undergoing analysis by commercial organizations [4].

In addition to high volume and velocity, data is being collected from an increasing variety of sources, such as GPS devices, social media feeds, financial history, and wearable technologies. This broad scope of data allows data owners to construct comprehensive representations of relevant events, processes, and people.

This Big-Data, characterized by high volume, velocity and variety, is being used to facilitate algorithmic decision-making (ADM). We define ADM as the use of computer programming to automate the decision-making process. ADM utilizes complex statistical techniques and other tools such as neural networks and deep learning to support or replace human decision-making. It has been argued that human decision making is often suboptimal, as humans employ heuristics or mental shortcuts to make decisions [5] and will revert to reducing effort over achieving the most optimal decision [6].

Many believe that data-backed algorithms can be used to render the decision-making process less biased and more rational, increase the effectiveness of decisions made, and help decision makers infer future trends and human behavior with a high degree of accuracy. For instance, retailers have predicted the health conditions of their customers based on historical purchase patterns in an effort to determine what offers might be of most interest to them [7]. Similarly, some financial institutions rely on ADM to predict consumers' financial behavior and make loan decisions.

The use of ADM to make decisions has been increasing over the course of time. In larger part, this is due to advancements in Artificial Intelligence (AI) and Machine Learning (ML), which are key components of ADM but also due to an increase in the availability of the Big Data which feeds ADM. As the use of ADM increases, so too do the concerns about the accuracy of algorithmic processing, the inaccessibility of algorithms, and the ethical implications of their use (Beer, 2016). Some have directly called into question the legitimacy of ADM (O'Neill, 2017) whilst others have chosen to ascribe a nefarious nature to the use of ADM in business, describing them as tools that "undermine both economic growth and individual rights" [8]. Several authors have described ADM as inherently discriminatory [9] and as a set of tools that promotes security over privacy, increases societal control and also the dependence of humanity

on technology [10]. In many ways then the Legitimacy of ADM, which is defined as the degree to which one's actions comply with social norms, values and beliefs [11], is being called into question.

In this paper, we posit that many of the critiques of ADM have less to do with the nature of the technology as such; rather, they stem from issues of transparency around the application of ADM. By addressing these issues, ADM can be used in a way that would alleviate the practical and ethical concerns typically associated with it. We acknowledge that increased transparency can have both positive and negative effects on ADM legitimacy. We view the relationship between transparency and ADM legitimacy as a complex one where various types of transparency can have several effects on different types of legitimacy.

We also emphasize that we take no view as to the efficacy of Transparency as a phenomena or its ability to actually achieve organizational accountability, increase organizational knowledge, or impact organizations in other ways. Rather we are strictly focusing our research on the impact of Transparency on the Legitimacy of ADM in organizations and note that Transparency can have an impact on Legitimacy without necessarily affecting organizations in other ways.

Accordingly, our approach to ADM legitimacy is a pragmatic one and focuses on articulating ideas that engender desirable outcomes (Goldkuhl, 2012). We maintain that increasing ADM legitimacy is important because, in a very real sense, the application of algorithms to big-data is an unavoidable part of contemporary society. This technology relies on a vast ecosystem of enabling infrastructures, from data centers, to communication protocols, to computer languages, to business and consumer applications, all of which are geared towards deepening and expanding the use of ADM. Consequently, ADM plays a broad role in allowing decision-makers to cope with the mounting quantities of data at their disposal. We aim to increase ADM legitimacy by scrutinizing its potentially harmful elements that result from the effects of transparency that often characterizes its application.

In what follows we define the concept of ADM and outline the main concerns around its use. After reviewing the literature on transparency and legitimacy we propose several ways in which ADM transparency impacts the pragmatic, moral and cognitive elements of legitimacy. These propositions break down ADM transparency into issues of validation, visibility and variability, and present the potential to introduce processes for improving ADM legitimacy. We illustrate the theoretical concepts developed through an applied example.

2 Background

2.1 ADM applications and underlying technology

ADM is not a new phenomenon. For years we have trusted algorithms to run our nuclear power plants and to fly our planes. It is estimated that 90% of commercial airline flight time is done on autopilot [12]. Automated production lines – from cars to integrated circuits – are controlled by algorithms. Similarly, supervisory control and data

acquisition systems, which are used extensively in manufacturing, rely on complex process control algorithms to make the everyday products we use safely.

Algorithms are characterized as a combination of logic and control and defined as "a set of defined steps to produce particular outputs" [13]. We consider algorithms in a software programming context and define ADM as the use of computer programming to automate the decision-making process.

Broadly speaking, ADM can be applied in two ways [14]. The first is to facilitate the processing of data to allow analysis by humans. In this scenario, algorithms make provisional recommendations, but the final decision is made by a human being. An example is the algorithm that presents an Uber driver with a ride which he / she can choose to accept or reject. A second application of ADM is where the final decision is made by an algorithm and may or may not be subject to human judgement or evaluation. For example, algorithms used by banks to place a hold on credit cards when certain suspicious activity is identified.

ADM is commonly applied through Artificial Intelligence (AI) by utilizing software code that detects internal and external factors through various sensors and then takes action or interacts with its environment through some form of automation to achieve pre-specified goals. In this way, AI is designed to mimic the natural intelligence displayed by humans.

One of the underlying techniques AI relies on is machine learning (ML). ML refers to mathematical models based on historical information that are constructed to predict the impact a specific action the machine takes will have on its environment. ML encapsulates a large set of mathematical modelling techniques which includes amongst other techniques the deep learning used in neural networks. The term neural networks denotes that the mathematical models are designed to represent in a coarse sense the function of a neuron or brain cell. The term deep learning implies that the models created using neural networks are highly complex using a number of hidden layers to represent various aspects of a system.

These distinctions should be noted as some of the criticisms of ADM (e.g., bias in the learning data set and opacity of the decisions taken) are specific to one or a few of its underlying technologies but not necessarily to ADM as a whole.

2.2 The concerns and challenges of deploying ADM

There are several reasons that can make the deployment of ADM concerning or challenging. The first involves the quality of data used by ADM applications. Often ADM is applied to understand human behavior and categorize it based on limited data. For instance, some banks use ADM to assess clients' loan applications by analyzing their social media profiles[1]. Similarly, several health insurers use data from wearable fitness devices to determine the coverage they offer to their customers[2]. In both

---

[1] https://www.smh.com.au/business/banking-and-finance/how-your-social-media-account-could-help-you-get-a-loan-20171219-p4yxw0.html

[2] https://www.usnews.com/news/national-news/articles/2017-02-17/could-fitbit-data-be-used-to-deny-health-insurance-coverage

instances, limited data is used as a proxy to subtle and complex human behaviors (people's credibility and general health). Such simplistic proxies are used because they are readily available, not because they present a comprehensive view of the behaviors they purport to model.

Another issue with ADM is the creation of self-fulfilling prophecies where decision-makers that act on predictions borne out of ADM, can create the conditions that realize those very predictions. For example, a company may use an algorithm to predict the performance of its recently-hired salespeople. Such an algorithm might draw on data from the salespeople's standardized tests, reviews from previous employers, and demographics. This analysis can then be used to rank new salespeople and justify the allocation of more training resources to those believed to have greater performance potential. However, this is likely to produce the very results that the initial analysis predicted. The higher-ranked recruits will perform better than those ranked lower on the list because they have been given superior training opportunities.

A further issue with the use of ADM is the complexity of the mathematical and statistical techniques that they are based on. Because they are designed using computational language that is only understood to specialized data scientists and computer programmers, their logic is black-boxed from most of the population and, in most cases, from the businesses, people, and communities whose lives are impacted by ADM. In some cases, the development of ADM unfolds over time, through multiple design iterations and the use of patched code by multiple programmers. In these cases, the algorithmic logic is hard to decipher even for those who were involved in its development. As a result, often even data scientists cannot explain how the ADM application that they have built makes a prediction or comes to a decision.

Another concern with ADM is the breadth and variety of the data that is used in the models to analyze and predict behavior. In an era of constant connectivity through computers, and mobile and wearable devices, we leave behind us a digital trace which can be used to determine our location, personality, behavior, financial status, work performance, and health. This data can be used in ADM for such activities as credit risk assessment, the calculation of insurance rates, and identification of preferences for product advertising. It is the use of data that is potentially sensitive and discriminatory which has induced the European Union to ban the use of ADM in certain scenarios and to require disclosure as to how a decision is made in others [9].

The challenges described above have led to increasing concerns around the deployment of ADM. These concerns have been particularly acute given the increasing visibility and pervasiveness of ADM in recent years. The visibility of ADM has increased as it has become part of the general discourse. For example, discussions around the role of ADM in transportation[3], health[4], and military[5] have received wide media attention. ADM has also become increasingly pervasive. Whereas initially used

---

[3] https://www.independent.co.uk/life-style/gadgets-and-tech/ethical-challenges-self-driving-cars-driverless-vehicles-tempe-arizona-crash-a8287776.html

[4] https://www.engadget.com/2018/05/20/uk-promises-funding-for-ai-cancer-detection/?guccounter=1

[5] https://www.techrepublic.com/article/googles-ai-pact-with-pentagon-sparks-resignations-highlighting-leadership-disconnect/

predominantly in computer science and specialized business settings, today ADM is applied across various domains, which impact us in multiple ways: from whether we receive a loan, to what news we read, to what people we date, to how quickly we board our flight, to whether we get hired, promoted or fired.

3 The Importance of Transparency to ADM Legitimacy

As argued at the outset of this paper, we maintain that many of the concerns surrounding ADM have to do less with its underlying technology and more to do with the transparency around its use. We posit that transparency impacts the various facets of the legitimacy of ADM in both positive and negative ways, that there are factors which moderate this impact, and that there is a complex interrelationship between the various types of transparency and legitimacy. To elaborate on this further, we next explore the concepts of transparency and legitimacy in the literature and build a nomological scaffold upon which to make several testable propositions about the relationship between transparency and legitimacy.

Fig. 1. Proposed High-Level Nomological Structure between Transparency and Legitimacy

3.1 Transparency Theory as an organizational concept

There is a broad body of research on transparency, much of which examines transparency in the context of information provision and accuracy, and focuses on the social and communicative processes around it [15]. Transparency has variously been defined as "a matter of information disclosure… the conditions of transparency are directly linked to the quality and quantity of information, its transmission, and the resulting degree of observability" [15]. Transparency has also been described "as a social process constituted within a set of socio-material practices" [15].

The myth that is organizational transparency has been around for some time [16]. The idea that increased transparency improves insight and therefore accountability and thus organizational governance has pervaded common culture. And yet research has disputed this arguing that information does not necessarily equate to insight [17] and that transparency does not necessarily equate to information [16]. Further, it has been argued that increased information can lead to a distancing of individuals from their surroundings, making them less capable of comprehending the world in which they live [18]. Moreover, increased information can also lead to less trust in institutions that lose their revered status when there is increased transparency and understanding of how

they actually operate [18]. All of this literature points to the potentially negative impacts of transparency and increased information on the organisation. But none of it draws a direct relationship between transparency and legitimacy specifically.

3.2 Transparency Theory as it is applied to ADM

Various authors have suggested that transparency is not necessarily positive in ADM and have argued against its benefits, specifically with regards to the goal of accountability in organizational governance [19]. These authors maintain that transparency in ADM can be harmful, that it can intentionally occlude by presenting non-relevant information, that it can create false binaries, that it does not infer increased understanding, and that it has technical and temporal limitations. The literature also argues that more transparency does not necessarily ensure more accountability in ADM based processes [20].

Considering the issue of transparency in ADM, Burrell [21] defines three types of opacity (the opposite of transparency) based on the cause of the opacity (intentional concealment, specialist knowledge, model complexity) and discusses the issue of opacity specifically as it relates to classification and discrimination. Transparency has also been categorized based on the level of human involvement, the data used, the model involved, the inferences made by the ADM and the algorithmic presence [20].

De Fine Licht [22] demonstrated that in Sweden greater transparency in allocation of public health care resources did not necessarily guarantee procedural acceptance or decision acceptance. This study points to several moderating factors which have an impact on the effectiveness of transparency such as the framing of the information provided through transparency. This concept of transparency moderating factors is explored in more detail later in this paper.

To sum, transparency has been examined from various perspectives and in different contexts for organizations in general and as it relates specifically to ADM. The desirability of transparency as an end state has been questioned as has its impact on organizational accountability. However, only one study has considered the relationship between transparency and legitimacy as a factor separate from accountability and this study only offered a crude conceptualization of legitimacy.

It is important for the subsequent discussion to emphasize that we take no view as to the efficacy of Transparency as a phenomena or its ability to increase accountability. Rather we will strictly focus on the effect of transparency on Legitimacy perception.

3.3 ADM Transparency as a Process

We define ADM Transparency as a process not as an end state in that we believe that complete transparency can never be achieved and to main a level of transparency requires ongoing organizational effort. In our analysis we add to the existing literature on transparency by defining Algorithmic Transparency as a process that involves Validation, Visibility and Variability. That process is moderated by the qualities of

Experience, Volition and Value which impact the degree of transparency and its impact on the ADM process. We define each of these terms in this section, set out a basis of Legitimacy Theory in the next section and then in the subsequent section discuss the impacts of each form of ADM Transparency on ADM Legitimacy, using Legitimacy Theory [11, 23, 24] and Technological Legitimation as the basis.

To begin we define *Validation* as the degree to which the actual decisions made by algorithms are reviewed and approved by humans. We acknowledge that the nuclear power plant is run by a computer or that a plane is flown by an auto-pilot but we also ensure that there is a physical human being constantly reviewing the decisions made by these algorithms.

The Validation of these algorithms can be further subdivided based on a temporal view into forward validation and backward validation. Forward Validation involves a decision made by an algorithm which is validated by a human before it takes effect. Frequently, these decisions are presented as recommendations to people, whom then must determine whether or not to follow the algorithmic recommendation. Backward Validation is where actions taken by algorithms are reviewed by human after they have been carried out. Following the review, the algorithm can be adjusted to improve the performance of the decision. To be effective those conducting either forward or backward validation should be independent third parties with sufficient skill sets to perform the validation but motivations independent of those who have developed the algorithms.

Drawing on Johnson, Dowd [23] and Binz, Harris-Lovett [24], ADM Validation can also be characterized based on the scope of application. Validation regarding a specific application of ADM we call Local Validation whereas validation across a whole class of ADM application is known as General Validation.

We define *Visibility* as the degree to which there is an effective presentation layer which distils down the large volumes of data involved in algorithmic calculations to an understandable representation, which then allows us to consider how an algorithm has come to a decision. Visibility impacts the degree to which we can critique the decisions made by the algorithms we use to run our businesses and our lives. Often referred to as Business Intelligence this presentation layer is often the first step in the transition of either a business and process management systems to an ADM supported process.

Variability is the degree to which decisions about people, processes or systems can vary based on the diversity of data sets provided. Does an Algorithm based on big data with a large variety of information actively reject a person's loan application because they live in a certain geography or does it simply flag the issue for the follow up and human supervision. Variability reflects the degree to which people are treated differently and the impact on those people of how an Algorithm is designed. Increased Variability present in the Algorithmic models increases the potential for accusations of bias and drives much of the concern presented in recent regulation used to limit the use of Algorithms and to increase the transparency of their use.

Even with the best tools to render and present the information under consideration, people reviewing that information must have the Experience to understand what is being presented. A lack of experience can contribute to a lack of transparency and represents a moderating factor in the relationship of transparency to legitimacy [14].

For example, a person who is not a trained pilot would likely find the outputs from algorithmic decisions made by the plane's auto pilot system as they are presented on the cockpit display screen to be opaque and ambiguous. The operators which nuclear facilities often have years of training to before being allowed to supervise the algorithms running nuclear power plants.

We further define Volition as the degree to which an organisation willingly provides Visibility into their ADM practices to allow for Local and General Validation as another moderating factor. For example, some organizations such as Google and Uber restrict access to their algorithms as they regard them as a source of competitive advantage. This lack of volition therefore moderates the impact of Visibility and Validation on ADM Legitimacy.

Value is the level of importance assigned to the underlying system impacted by ADM and represents another moderating factor. For example, the ADM used in a chat bot to answer common call center queries regarding billing may be perceived to have less value than the ADM in a driverless car that is making regular decisions about a drivers physical safety. This factor provides a further moderator of the effect of transparency in ADM.

To examine the impact of transparency and legitimacy in further detail, we next elaborate on the concept of legitimacy as a multi-faceted phenomenon.

3.4 Organizational Legitimacy as a Concept

Legitimacy is commonly understood to be "a generalized perception or assumption that the actions of an entity are desirable, proper, or appropriate within some socially constructed system of norms, values, beliefs and definitions" Suchman [11]. Legitimacy Theory has been used in the context of corporate reporting around social and environmental issues [25]. Legitimacy Theory gives consideration to the expectations of various parts of society and the implicit social contract between an organisation, its actions within that social contract, and the society in which that organisation operates. As social norms change, so do societal expectations and organizations must adapt accordingly. As this adaptation process takes time, there will be a gap between organizational actions and practices and societal expectations. Organizations must seek to actively minimize this gap in order to be allowed to operate freely within society. Organizations can do this by changing their practices, influencing societal norms, values, and beliefs about those practices, or by the better communication and positioning of their practices within those societal norms, values and beliefs.

Organizations and individuals may choose not to work with other individuals or organizations due to a perceived lack of legitimacy associated with the organization as a whole, or with its practices, technologies, and products. Governments may undertake legislative action when the perceived gap between an organisation and societal expectations becomes too big and when organizational practices or technologies are perceived to be illegitimate. This was recently demonstrated when the European Union enacted legislation to reduce the use by organizations of ADM and give consumers a

"right to explanation" to understand how the ADM in question has made a decision about them [9].

Legitimacy can take three broad forms: *Pragmatic Legitimacy*, which provides value to the organization's interested constituents (e.g., customers, shareholders and employees) and can be gained through organizational policy (Dowling and Pfeffer [26]; *Moral Legitimacy* which is developed from a positive normative evaluation of organizational practices by external stakeholders, and; *Cognitive Legitimacy* which is based on external stakeholders' understanding of organizational practices and which is neither self-interested nor evaluative in nature [27]. Cognitive Legitimacy occurs when an organisation pursues activities and goals which have become "taken for granted" by society.

To illustrate the three forms of legitimacy in the context of ADM, consider a bank which uses ADM for its home loan origination and approval process. *Pragmatic Legitimacy* might be achieved with a bank's shareholders when they see that the use of ADM in loan origination reduces administrative effort and operational costs, consequently improving profit and raising the stock price. *Moral Legitimacy* might be gained for the same loan origination process when it is clear to external stakeholders that the use of ADM does not unfairly disadvantage individuals from a specific background; for instance, that people's racial or religious backgrounds are not used to assess their eligibility for a loan. Finally, *Cognitive Legitimacy* may be achieved when stakeholders accept that the use of ADM in loan origination and approvals is common place with financial industry as a whole.

3.5 The Process of Technological Legitimation

Technological legitimation is the process of narrowing the legitimacy gap between an organization and its external stakeholders, specifically as it concerns technologies employed by the organization. Technological legitimation is described as a cumulative, non-linear process that moves through the stages of innovation, local validation, diffusion and general validation [23, 24].

In this process there is a technological innovation. Organisational actors attempt to link that new technology to existing organizational activities, hoping to passively validate that new technology and/or hoping that it does not get challenged [23]. Challenging of that technology may be on pragmatic, moral or cognitive grounds. If local validation is achieved then the new technology diffuses to other applications and contexts either within that organisation or to other organizations. Over the course of time as this diffusion of that technology continues it "increasingly interferes with more broadly shared normative, regulative and cognitive rules. The relevant audience is no longer restricted to an isolated project or community, but rather comprises the general public that assesses the legitimacy of both the technology and the 'industry' that emerges in the new field" [24]. In so doing the technology begins to go through a process of general societal legitimization. In order to continue to diffuse through organizations and eventually society the technology needs to have Moral Legitimacy before it is applied and it needs to achieve Pragmatic Legitimacy as part of each local

validation process. As the technology diffuses it either is adapted to or adapts societal values through an ongoing process of normative re-evaluation. Once the application of the technology becomes broad based, society begins to take for granted the new innovation and Cognitive Legitimacy is achieved.

An useful illustration of the process of Technological Legitimation is found in the work of Binz, Harris-Lovett [24]. As a case study they analyze the practice of introducing purified wastewater into surface or underground drinking commonly known as Indirect Potable Reuse (IPR). Up until the year 2000 the practice of IPR was restricted due to public perceptions. Yet after 2010 many reports demonstrate a significant uptake of the technology [28]. What occurred in between was a series of activities to achieve IPR Legitimacy through increased diffusion, demonstrating that even highly undesirable technological practices of organizations can achieve Legitimacy.

3.6 The Complex Relationship between Transparency and ADM Legitimacy

Summarizing the previous sections, we posit that ADM transparency impacts ADM legitimacy in a complex process that involves multiple components of both transparency and legitimacy, and that this process is moderated by several factors (Figure 2):

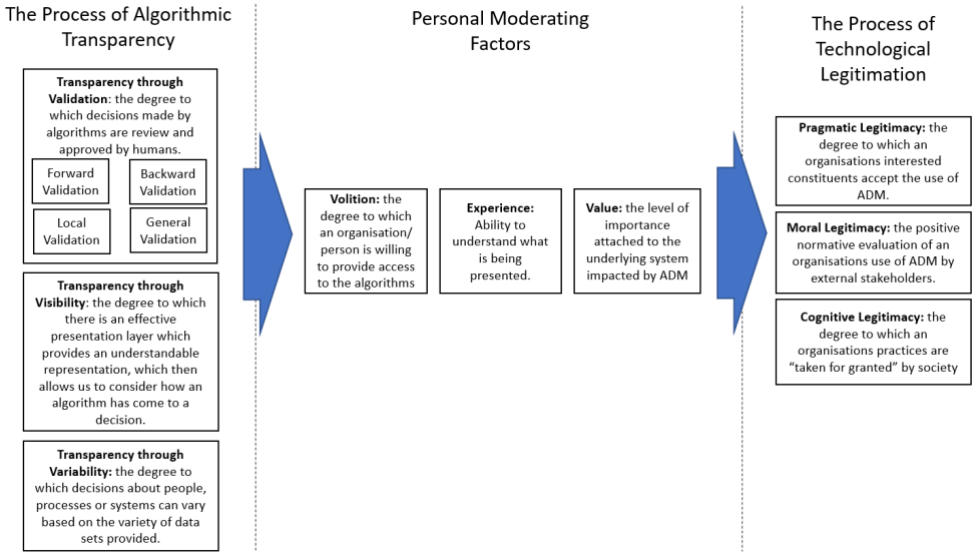

Fig. 2. Expanded Nomological Structure between Transparency and Legitimacy

Transparency is not the only factor which impacts legitimacy and transparency may or may not have impacts on other facets of an organisation (for example accountability). But for the purposes of this article, and due to space constraints, we focus on the interactions between transparency and legitimacy. We discuss these interactions in the subsequent section and make several propositions as to the nature of this relationship between ADM Transparency and ADM Legitimacy within this nomological structure.

4 Towards an Understanding of the Impact of ADM Transparency on ADM Legitimacy

After the previous discussions of ADM Transparency, Legitimacy, and Technological Legitimation we next make several propositions as to how ADM transparency affects the different forms of ADM legitimacy throughout the legitimation process. The purpose of these propositions is to lay the groundwork for further research into the impacts of ADM Transparency on ADM Legitimacy..

To begin, we look at the degree of ADM Validation. Reiterating that Validation has a temporal component (Forward Validation vs. Backward Validation) and scope of application component (Local Validation vs. General Validation) we consider the effects Validation may have on Pragmatic, Moral and Cognitive Legitimacy. As discussed in Binz, Harris-Lovett [24], organizational practice will gain local validation in a specific context. It will then diffuse to other similar contexts to eventually achieve

general validation. In our view of ADM legitimization there is a temporal component to validation as well as the contextual component of validation already presented in the literature. We also see that the temporal and application aspects of validation can be combined together in various combinations. For example we further propose that…

> ***Proposition 1:*** *An increase in the amount of Forward and Local Validation of an Algorithm will increase the Pragmatic Legitimacy of the algorithm.*

Referring back to the definition of Pragmatic Legitimacy as a phenomena involving an organizations immediate constituents, we perceive this form of legitimacy as a local phenomenon where validation will improve an ADMs perception of fitness for use but not necessarily its acceptance morally or cognitively. Further we propose that:

> ***Proposition 2:*** *General Backward Validation of an application of ADM is required for Cognitive Legitimacy.*

… as the "taken for granted-ness" required for Cognitive Legitimacy requires actual application of the ADM in the specific context for Cognitive Legitimacy to attach. In line with the Technological Legitimization process described by Binz, Harris-Lovett [24] we view that…

> ***Proposition 3:*** *Increased Local Validation of ADM in a greater variety of contexts will lead to increased General Validation of ADM across all forms of Legitimacy (Pragmatic, Moral and Cognative).*

This leads to our next proposition regarding transparency in ADM and the impact of Visibility on Legitimacy:

> **Proposition 4:** *An increase in Visibility will increase the Cognitive Legitimacy of the algorithm but not necessarily Moral Legitimacy or Pragmatic Legitimacy.*

We derive this view from the observation that greater visibility does not necessarily create moral acceptance and that the users of an expert system may lose trust in an expert system the more they understand it [18]. Further we present…

> **Proposition 5:** *That the effect of Visibility on the Cognitive Legitimacy of ADM will be moderated by the degree of human experience with the system being modelled and managed algorithmically.*

… in that information does not bring understanding without a degree of expert knowledge [29].

In terms of the moderative effects of Volition and Value we view that…

> **Proposition 6:** The Volition of an organisation to provide Transparency around ADM will be inversely proportional to the perceived Value assigned to the system the ADM is managing. Decreased Volition will decrease Visibility of the ADM which has an impact on all forms of legitimacy.

Clearly, even if an organisation is compelled by legislation to provide more transparency the way that transparency requirement is affected by the organisation moderate it's impact. Finally we present our last proposition which is that…

> **Proposition 7:** *The degree of Variation present in an Algorithm will have an inverse effect on the Moral Legitimacy of the Algorithm.*

… in that the more significant the difference in how people are treated through the ADM process the higher the likelihood than Moral Legitimacy will be impacted.

As previously discussed, these seven propositions based on Transparency and Legitimacy Theory and the interplay between the two in an ADM context are intended to be used and tested in future research on ADM Legitimacy.

5 Affecting ADM Legitimacy through Transparency: An Illustrative Example

As discussed in previous sections, we maintain that the issues with Algorithmic Decision Making relate not necessarily to the automation of the decision but rather to the degree to which the decision making process is *Transparent*. This *Transparency* has many characteristics including the degree to which the Algorithmic Decisions are *Validated* by Humans, the *Visibility* humans have of these decisions as they are made by these ADMs and the impact of the potential *Variation* in the decisions implemented by these algorithms which in turn impacts human perceptions of the Algorithms itself. By designing Algorithmic Decision Making Processes to address the issue of *Transparency* it is proposed that it may be possible to affect their perceived Legitimacy. We illustrate the use of the previously proposed framework with an example.

Recent concerns about how Uber's App algorithmically manages its drivers has received significant news coverage [30]. Concerns that have been expressed include how Uber sets the rates, performance targets and schedule. We apply the concepts of *Transparency* and *Legitimacy* in this case to illustrate the nomological structure we have constructed in this article.

First, the practices of rate setting and scheduling have already achieve *Cognitive Legitimacy* in that they are already take for granted in the ride-sharing industry. Other ride-sharing applications such as LYFT, GET, and JUNO use ADM in similar ways

[31]. *Pragmatic Legitimacy* is demonstrated every time someone uses Uber accepts the quote for the ride sharing service. The ride sharing market was over $10 Billion USD in revenues in 2016 [32]. What is at question here is the *Moral Legitimacy* of the process. Is there a positive normative evaluation of Ubers practices by its drivers and by the general public that use Uber?

With regards to how Uber sets performance targets, those targets would likely be better received by its drivers if it was known that there was a level of *Validation* of the performance grading assigned to the drivers by human managers. Preferably this would be by *Foreword Validation* before the statistics were published, although *Backward Validation* would probably also have an impact. With regards to how Uber sets rates, it is reasonable to assume that perceived legitimacy would increase if the *Variation* in the rates and performance targets was reduced. With regards to the scheduling *Visibility* of both the surge pricing zones to the passengers as well as the drivers there would be less concerns about the process.

These Transparency processes are moderated by the *Experience* the Uber drivers and Uber ride users have with the Uber pricing and scheduling and would be impacted if the average *Value* for each ride remained small. In all cases the use of Algorithmic Decision Making is not precluded but the perceived Moral Legitimacy of the system on the part of its users is impacted. Overall Legitimacy and social acceptance of the ADM processes increases through transparency. Thus designing ADM processes with a view on transparency can have an impact on Legitimacy. As discussed previously, legitimacy theory tells us that organizations must always seek to minimize gap between an organizational actions and practices and societal expectations if for no other reason than to avoid government and market reaction. In this case it is clear the Uber needs to be more transparent with regards to its ADM practices in order to avoid further actions on the part of governments that are already actively seeking to address the disruption that Uber and other Ride Sharing applications have caused to the taxi industry.

6 Conclusions

Algorithmic Decision Making represents one path to increased benefits realization from Big Data and the Internet of Things. Algorithmic Decision Making can resolve many of the cognitive issues present in human decision making and takes advantage of the volume, velocity and variety of the Big Data presented by the IoT that human decision making cannot. But it has also been shown that a recent increase in the pervasiveness of Algorithmic Decision Making has led to several concerns which may limit its application. In some cases governments have already legislated against its use.

In review the theory on Transparency and Legitimacy we build a series of propositions on how we believe Transparency impacts the different forms of Legitimacy in an ADM context. We posit that in some cases this effect will be to increase legitimacy but in other cases transparency may decrease legitimacy and that there are moderating factors which may affect this impact. We use an example to illustrate the interplay between transparency and algorithmic decision making and how business practices can be redesigned to improve ADM Legitimacy. We identify that

ADM Transparency has many elements including the levels of Validation, Visibility and Variation. We reframe Transparency as a process as opposed to an end state. By unpacking the issues of transparency we create the foundations of a framework for understanding Algorithmic Technology Legitimization which can be used to analyze, redesign and thereby affect the acceptance and appropriation of ADM.

It is important to reiterate that the analysis presented is not intended to be a holistic or complete view of Algorithmic Decision Making Legitimacy. Rather it is intended to provide a key precursor for future research into this subject. The intention being to further refine and add to the theory through field research and the analysis of case studies where the level of transparency within an organisation has varied over time with corresponding changes in ADM appropriation due to improved levels of perceived ADM Legitimacy. In this way the described nomological structure for the impact of transparency on ADM Legitimacy could be further refined. This could then be further validated through empirical studies which track ADM adoption in specific industries, with specific use cases over time. Finally the interrelationship between legitimacy and technological acceptance could further be explored in subsequent works.

7 References

1. McAfee, A., E. Brynjolfsson, and T.H. Davenport, *Big data: the management revolution.* Harvard business review, 2012. 90(10): p. 60-68.
2. Reinsel, D., J. Gantz, and J. Ryding, *Data Age 2025: The Evolution of Data to Life-Critical.* IDC, 2017.
3. Luchetti, D. *The Square Kilometer Arrary*. Square Kilometer Array 2017 [cited 2017 21 Dec 2017]; Available from: http://www.ska.gov.au/Pages/default.aspx.
4. Heudecker, N.H., Jim, *Survey Analysis: Big Data Investments Begin Tapering in 2016.* Gartner Research, 2016.
5. Tversky, A. and D. Kahneman, *Judgment under uncertainty: Heuristics and biases.* science, 1974. 185(4157): p. 1124-1131.
6. Todd, P. and I. Benbasat, *The use of information in decision making: an experimental investigation of the impact of computer-based decision aids.* Mis Quarterly, 1992: p. 373-393.
7. Hill, K. *How Target figured out a Teen Girl was pregnant before her father did*. Forbes Magazine 2012 [cited 2017 12 Dec 2017]; Available from: https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/#1999ad516668.
8. Pasquale, F., *The black box society: The secret algorithms that control money and information*. 2015: Harvard University Press.
9. Goodman, B. and S. Flaxman, *European Union regulations on algorithmic decision-making and a" right to explanation".* arXiv preprint arXiv:1606.08813, 2016.
10. Newell, S. and M. Marabelli, *Strategic opportunities (and challenges) of algorithmic decision-making: A call for action on the long-term societal effects of 'datification'.* The Journal of Strategic Information Systems, 2015. 24(1): p. 3-14.
11. Suchman, M.C., *Managing legitimacy: Strategic and institutional approaches.* Academy of management review, 1995. 20(3): p. 571-610.
12. Cox, J. *Ask the Captain: How often is autopilot engaged?* USA Today 2014 [cited 2018 28 Jan 2018]; Available from:

https://www.usatoday.com/story/travel/columnist/cox/2014/08/11/autopilot-control-takeoff-cruising-landing/13921511/.
13. Kitchin, R., *Thinking critically about and researching algorithms.* Information, Communication & Society, 2017. **20**(1): p. 14-29.
14. Barocas, S., S. Hood, and M. Ziewitz, *Governing algorithms: A provocation piece.* 2013.
15. Albu, O.B. and M. Flyverbom, *Organizational transparency: conceptualizations, conditions, and consequences.* Business & Society, 2016: p. 0007650316659851.
16. Christensen, L.T. and J. Cornelissen, *Organizational transparency as myth and metaphor.* European Journal of Social Theory, 2015. 18(2): p. 132-149.
17. Baudrillard, J., *Simulacra and simulations. Poster, M.(Ed.). Jean Baudrillard: selected writings pp. 166-184.* 1988, Stanford (CA): Stanford University Press.
18. Tsoukas, H., *The tyranny of light: The temptations and the paradoxes of the information society.* Futures, 1997. 29(9): p. 827-843.
19. Ananny, M. and K. Crawford, *Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability.* new media & society, 2016: p. 1461444816676645.
20. Diakopoulos, N., *Accountability in algorithmic decision making.* Communications of the ACM, 2016. 59(2): p. 56-62.
21. Burrell, J., *How the machine 'thinks': Understanding opacity in machine learning algorithms.* Big Data & Society, 2016. **3**(1): p. 2053951715622512.
22. De Fine Licht, J., *Do we really want to know? The potentially negative effect of transparency in decision making on perceived legitimacy.* Scandinavian Political Studies, 2011. 34(3): p. 183-201.
23. Johnson, C., T.J. Dowd, and C.L. Ridgeway, *Legitimacy as a social process.* Annu. Rev. Sociol., 2006. 32: p. 53-78.
24. Binz, C., et al., *The Thorny Road to Technology Legitimation—Institutional Work for Potable Water Reuse in California.* Technological Forecasting & Social Change, 2016. 103: p. 249.
25. Deegan, C., *Methodological Issues in Accounting Research: Theories, Methods and Issues*. 2006.
26. Dowling, J. and J. Pfeffer, *Organizational legitimacy: Social values and organizational behavior.* Pacific sociological review, 1975. 18(1): p. 122-136.
27. Aldrich, H.E. and C.M. Fiol, *Fools rush in? The institutional context of industry creation.* Academy of management review, 1994. 19(4): p. 645-670.
28. Schroeder, E., et al., *Direct Potable Reuse: Benefits for public water supplies, agriculture, the environment and energy conservation.* National Water Research Institute White Paper, 2012.
29. Christensen, L.T. and G. Cheney, *Peering into transparency: Challenging ideals, proxies, and organizational practices.* Communication Theory, 2014. **25**(1): p. 70-90.
30. Rosenblat, A. *The Truth about how Uber's Application Manages Drivers*. Harvard Business Review 2016 [cited 2017 21 Dec 2018]; Available from: https://hbr.org/2016/04/the-truth-about-how-ubers-app-manages-drivers.
31. Coomes, K. *The Best Ridesharing Apps*. Digital Trends 2018 [cited 2018 22 Aug 2018]; Available from: https://www.digitaltrends.com/mobile/best-ride-sharing-apps/.
32. Hensley, R., A. Padhi, and J. Salazar. *Cracks in the Ridesharing Market and how to fill them* McKinsey Quarterly 2018 22 August 2018]; Available from: https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/cracks-in-the-ridesharing-market-and-how-to-fill-them.