ETHICS IN DATA SCIENCE PROJECTS: CURRENT PRACTICES AND PERCEPTIONS

jeff saltz
Syracuse University, jsaltz@syr.edu

Follow this and additional works at: https://aisel.aisnet.org/ecis2019_rip

Recommended Citation
https://aisel.aisnet.org/ecis2019_rip/68

This material is brought to you by the ECIS 2019 Proceedings at AIS Electronic Library (AISeL). It has been accepted for inclusion in Research-in-Progress Papers by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
Ethics in Data Science Projects: Current Practices and Perceptions

Research in Progress

Saltz, Jeffrey, Syracuse University, Syracuse New York, USA, jsaltz@syr.edu

Abstract

Data science projects continue to grow in number and importance. At the same time, the number of potential ethical conundrums that teams might encounter is also increasing. However, little has been done to investigate the perceived challenges and benefits of incorporating ethics oversight and analysis within a project. To address this gap, this paper reports on current practices, as well as the perceived benefits and challenges, with respect to how data scientists consider ethics analysis within a project. The results of the study show that most teams do not have an explicit process, and that potential ethical issues are identified in an ad hoc manner. The main challenges of incorporating ethics within a data science project were focused on the complexity and compatibility of being able to implement an ethics review process, while the benefits focused on the relative advantage of an improved analysis and the fact that others will know that the project has considered its ethical implications. While more work is required to validate and refine these findings, this analysis can help teams as they consider integrating ethics analysis within their data science project.

Keywords: Data Science, Big Data, Ethics.

1 Introduction

Data science is the analysis of data to solve problems and develop insights. It is an emerging discipline that combines expertise across a range of domains, including software development, data management and statistics. While there are many views on what to include in the field of data science, this research adopts Saltz & Stanton’s (2017) definition, which includes the collection, preparation, analysis, visualization, management, and preservation of large collections of data. This definition embraces the notion that data science is more than just analytics. Moreover, the field of big data, where the data sets are too large and/or complex for traditional data analysis techniques, can be viewed as a subset of data science.

With the increasing ability to collect, store and analyze an ever-growing diversity of data that is being generated with increasing frequency, the field of data science is growing rapidly. As the field of data science grows, data scientists, just as professionals in other fields, will face pressure to deliver results. In trying to deliver results, the question of what is appropriate or ethical should arise. Specifically, based on a literature review of ethical conundrums data scientists might encounter, Saltz & Dewar (2019) identified two general paths to cause harm. First, with respect to data related challenges, the preparation, storage and dissemination of data could impinge on the privacy or anonymity of the subject, or cause bias in the resulting analytics (ex. just because data is available, it does not mean it is ethical to use that data). Second, with respect to model related challenges, a data science model might operate incorrectly or operate correctly, but the objective that the model is inherently unfair to some subjects (ex. some
subjects could be misclassified, resulting in harm). As one can see, these challenges are different than ethics situations that a software development team might encounter. In other words, some ethics issues are specific to data science. More broadly, while data science can bring objectivity to decision making, there is subjectivity within data science modeling in that decisions must be made about which algorithm to use, which data sources to use, whether one data point should be used as a proxy for a missing fact, and how to interpret results (Sandvig et al, 2014).

As an example of an ethical situation that data scientists might have to contemplate, one might ask if it is acceptable for an organization to develop a model that predicts the health care cost of a prospective employee, such as by exploring an employee’s eating habits and exercise routine (Gumbus and Grodzinsky, 2016). In order to address this type of project, data scientists, and the management of that organization, need to understand a range of underlying ethical issues such as fairness (what training data should be used to ensure there is no gender bias in an algorithm used to rank job applications) and privacy (is it okay to data mine “public” social media data to train models to infer personal attributes and identity). With respect to these types of questions, data scientists and managers need to work together to approach these dilemmas thoughtfully.

However, the perceived challenges and benefits of incorporating ethics within a data science project have not been studied. Furthermore, the challenges and benefits might be different than other information system efforts. This potential difference from other domains is due to the fact that while data science projects have parallels to other domains, there are also differences compared to these other types of projects. For instance, compared to software development, data projects have a broader range of ethics questions such as the potential for bias (Guan & Zhou, 2017). This helps to explain why Shin (2015) notes that traditional technology acceptance models require modifications in the context of new and emerging trends and technologies.

Hence, this article explores the current use of ethics analysis within a data science team as well as the perceived challenges and benefits of incorporating ethics within a data science project. Specifically, this research focuses on the following research questions:

**RQ1**: Do data science projects integrate the analysis of possible ethics conundrums within their project? And if so, how?

**RQ2**: What are the perceived challenges and benefits of incorporating ethics evaluation within a data science project?

The rest of the paper is structured as follows. In Section 2, some additional background is presented, which is followed, in section 3, with a description of the methodology used in this study. Section 4 describes the findings from this research and section 5 presents a discussion of the results. Finally, in section 6, a synthesis of the research results is presented as well as noting some limitations and planned next steps.

## 2 Background

### 2.1 Reviewing Ethics and Data Science

Like many significant scientific advancements, the use of data science has raised a number of significant ethical challenges and the need for ethics in data science has been frequently noted (Floridi & Taddeo, 2016; Fong, 2016). In fact, emerging ethical dilemmas are already touching the practices of computing professionals who use data science to solve problems, forcing them to make difficult decisions (Guan & Zhou, 2017). Hence, it is not surprising that Tiell and Metcalf (2016) have argued that data science introduces new classes of risk to organizations and that others have noted that none of the existing codes of conducts sufficiently cover the full range of potential ethical challenges a data science team might encounter (Tractenberg et al, 2015; Saltz, Dewar & Heckman, 2018).
For example, data science techniques, such as machine learning algorithms, have shown the capability of inheriting racial (Angwin et al, 2016; Chouldechova, 2017) and gender (Bolukbasi et al, 2016; Datta et al, 2015) biases. Further, these types of algorithms have been used to predict, and thus disclose, private attributes of users (Kosinski et al, 2013) or even target their beliefs and psychological traits (Harris, 2016; Metcalf and Fiesler, 2018). Without exploring these questions, the unethical use of data science could impact the reputational and economic well being of an organization, such as the public’s well publicized reaction to Target’s alleged prediction of a teenager’s pregnancy (Someh et al, 2016). In order to address these types of questions, the data science team, and the management of that organization, need to be aware of the possible ethical situations a project might encounter, so as to at least be able to consciously explore the potential ethical dilemmas.

From a broader perspective, ethics has been found to be a key component that can help determine the acceptance of new technologies (Stahl, Timmermans & Mittelstadt, 2016). Thus, so as to not stunt the adoption of data science, it is important that data scientists consider the harm that might arise from their work while still allowing for the novel adoption of data science algorithms.

To help guide teams to incorporate ethics analysis within their data science project, some ethics frameworks have recently been proposed for data science efforts (Saltz, 2018; Guan & Zhou, 2017). However, there has been no research that has been identified that explores the challenges, or the benefits, of incorporating ethics within the process a data science team uses to execute their project. This work helps to address this gap.

2.2 Theoretical Foundations

From an IT perspective, innovation refers to a new practice or operational idea (Vahtera, 2008; Lind and Zmud, 1991). Hence, from a theoretical perspective, the use of data science ethics explicitly within a data science project can be thought of as a process innovation. Oliveira and Martins (2011) noted that most studies on IT adoption, including process adoption, leverage one of two frameworks, either the Technology-Organization-Environment (TOE) framework (Tornatzky and Fleischer, 1990) or the Diffusion of Innovation (DOI) framework (Rogers, 1995).

The DOI framework was selected since it has been extensively used to study information systems process innovation. For example, at one end of the spectrum, DOI was used to examine IS process adoption over a period spanning four decades (Mustonen-Ollila & Lyytinen, 2003). More recently, DOI was leveraged to understand the acceptance factors for using a new agile software development technique (Schlauderer, Overhage & Fehrenbach, 2015). Thus, the Diffusion of Innovation (DOI) Theory (Rogers, 1995) is an appropriate lens to examine these results.

The DOI framework defines innovation as an idea, practice, or object that is perceived as new by an individual. Specifically, DOI describes the factors that determine the assimilation (or adoption) of an innovation. To summarize DOI, the theory states that the way potential adopters perceive the attributes of the innovation impacts the willingness of those individuals to assimilate (or adopt) that innovation. The five perceived attributes of the innovation are (Rogers, 1995; Moore & Benbasat, 1991):

- **Relative advantage** - the degree to which the innovation is perceived to be better than the idea it supersedes.
- **Compatibility** - the degree to which the innovation is perceived to be consistent with the existing values, past experiences and needs of potential adopters.
- **Complexity** - the degree to which the innovation is perceived to be difficult to understand and use.
- **Observability** - the degree to which the results of the innovation are visible to others.
- **Trialability** - the degree to which an innovation may be experimented with on a limited basis. This may include trying parts of a process or being able to watch others using the new process.
Previous research has noted some common impacts of these attributes (Schlauderer and Overhage, 2015). Specifically, the higher the perceived relative advantage, the more likely it is that the innovation will be adopted. In addition, if the process is amenable to being used on a limited bases, the trialability is also positively related to the process being adopted. However, if the innovation is perceived as an extreme change, then it will not be compatible with past experiences and is less likely to be adopted. In addition, if the process innovation is perceived as complex, it is also less likely to be adopted. Finally, if the observed effects are perceived to be small or non-existent, then this low observability reduces the likelihood of adoption.

3 Methodology

This effort consisted of a two-phased approach. First, was to understand if teams were currently integrating ethics within their data science project, data scientists were surveyed to understand current practices with respect to ethics analysis within a data science project. Then, in the second phase, the focus shifted to what were the perceived challenges and benefits of integrating ethics analysis within a data science project. The focus of the second phase was on the perceived challenges and benefits that can drive a team to start (or not start) integrating ethics within their data science project (as opposed to the challenges or benefits when teams actually integrate ethics within their project). Hence, for the second phase, the research did not observe data science teams trying to incorporate ethics analysis, but rather, leveraged semi-structured interviews with data scientists to understand the perceived challenges and benefits that might drive (or hinder) the adoption of an effort to integrate ethics analysis within a data science project. For the second phase, the use of semi-structured interviews was leveraged since semi-structured interviews provides both breadth and depth of the discussions, while also enabling the comparison of results between the discussions by following a common outline of questions (Yin, 2003).

3.1 Current Approach to Ethics in Data Science Projects

In phase 1, to understand how current data science practitioners incorporate ethics into their projects, 78 professionals were surveyed across multiple organizations – from industry, academia and not-for-profit organizations. The participants were identified via direct outreach at the IEEE Big Data 2017 conference. In the survey, the data scientists were asked how they thought about ethical situations in their project. For this question, participants where provided four choices (Work through questions/issues as they arise, Ethics doesn't really come up in our projects, Use a published framework to think about ethical issues or Other). In addition, open-ended questions were presented to provide the opportunity for additional context and insight into the answers provided.

With respect to the 78 respondents, 54% of the participants identified themselves as a data scientist, 22% as a data science team lead and 24% as a data engineer. In terms of team size, 13% worked mainly by themselves, 40% worked in a team of 2-3 people, 37% worked in a team of 4–6 people and 10% worked in a team with more than 6 people.

3.2 Incorporating Ethics within a Data Science Project

In phase 2, to better understand the potential barriers and perceived benefits of incorporating ethics into a project (for teams that do not yet have a defined process of exploring potential ethics situations), 106 semi-structured interviews were conducted. Specifically, respondents provided free form answers to questions such as “what do you perceive as the key benefits in incorporating ethics within a data science project?” and “what do you perceive as the key challenges in incorporating ethics within a data science project?”.

Data science graduate students were selected to be the participants in this phase of the study. Others have also used students to understand information system related process adoption. For example, Brock
& Khan (2017) used graduate students to understand the acceptance factors of organizations thinking of using big data analytics. Further justifying the use of graduate students, as part of their analysis, Brock and Khan noted that there was “no reason to believe these students do not represent the average profile of IT professionals”. Specifically, all 106 of the respondents had taken at least one graduate level data science course and eight-five percent of the students were in an information management or applied data science graduate program, with the remaining students in programs ranging from business administration to public policy. Forty percent of the students were female and on average, students had two years of previous IT/Data Science work experience.

Since most of the graduate students had previous experience working within an IT organization, consistent with Brock and Khon, they were well suited research participants. For example, when responding to the survey questions, students were able to leverage both their previous work experience, as well as being able to leverage the knowledge gained while they were graduate students – which included foundational information on potential ethical situations that might be encountered within a data science project. This helped to ensure that the participants were exposed to some of the potential challenges in using an ethics framework within a data science project (from their time in industry), as well as having some foundational knowledge on the importance of thinking through potential ethical conundrums that data science teams might encounter (from their time in their data science course).

The information collected from the interviews was used to identify perceived benefits and challenges with respect to incorporating ethics within a data science project. Specifically, the data collected from the surveys was analyzed in a manner consistent with Miles and Huberman (1999). First, open coding was used to search the statements to help identify recurring topics and to get an overview of the key points contained in the surveys. Second, the identified concepts were grouped by similar topics that were repeatedly articulated in the surveys. By using this process, it was possible to identify concepts that were consistently mentioned in the surveys.

## 4 Findings

### 4.1 Current Approach to Ethics

When asked how they thought about ethical situations within their projects, as noted in Table 1, 46% responded that they “work through questions/issues as they arise”, 32% responded that they thought that ethics “Doesn’t really come up in our projects” and 21% left the answer blank, suggesting that they had not really thought about ethics in their data science project context. This suggests that there is, at best, typically an ad-hoc approach for thinking through potential ethical situations that might arise. The open-ended survey questions yielded minimal insight, in that most of the responses were blank. Thus, a second phase of the research was needed to specifically identify key challenges and benefits of integrating ethics analysis within a data science project, the results of which are described in the rest of this section.

<table>
<thead>
<tr>
<th>Answer</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Work through questions/issues as they arise</td>
<td>46%</td>
</tr>
<tr>
<td>Doesn’t really come up in our projects</td>
<td>32%</td>
</tr>
<tr>
<td>No Answer (answer left blank)</td>
<td>21%</td>
</tr>
<tr>
<td>Use a published ethics framework</td>
<td>0%</td>
</tr>
</tbody>
</table>

Table 1: Responses to “How do you think about ethical situations”?
4.2 Data Science Ethics Analysis Acceptance Model

Using the DOI framework, an analysis of the 106 survey responses from the semi-structured interviews, that focused on the challenges and benefits of incorporating ethics analysis within a data science project, identified the following perceived adoption benefits and challenges.

4.2.1 Key Benefits of Incorporating Ethics Analysis

Three key benefits where identified (mitigate risk, achieve better results, and improve public perception of using data science), which are each described below Mitigate Risk - One of the most common themes that emerged was the perception that, as one person noted, “the key benefit is the avoidance of legal liability”. Sometimes the specific risk was noted, such as a comment from different participant, who stated that consciously incorporating ethics would “help mitigate risks involved while working on sensitive data sets”, while still others noted “ensuring no bias in the analysis”. Yet a different person thought of the risk at a personal level, by noting that it would help “keeping my job”. While that response was not typical, it does highlight the perceived risk to the data scientist and their organization when ethics is not fully considered.

Achieve Better Results - Another key theme that emerged was the perception that incorporating ethics into the project would help provide better results. For example, one response noted that it would help “ensure that the results are not biased”. In other words, results with, for example bias, are not as useful to the client. This bias might be due to input data sets having bias, which could cause bias towards a class of the society. Better results were also perceived to be due to improved oversight of the data scientist, as noted by a different participant “there is a fair amount of human judgment and assessment that goes into creating an analysis. Hence, ethics supervision should be engrained in the organization (e.g., to help ensure that human judgment does not have bias)”.

Improve Public Perception of Using Data Science – A very different benefit was the belief that if the public knew that thoughtful ethical considerations were specifically part of the data science process, then the public would be more accepting of the use of data science by creating, as noted by one response, a “positive public reputation”. In other words, the public’s knowledge of an ethics review could help offset some of the recent negative publicity with respect to the use of data science. For example, one person noted that it could “allow the data scientist to be an enabler of progress as opposed to a point of controversy”, or stated a different way within a difference response, “if done ethically, data science should be revealing real and actionable decisions as opposed to creating controversial or politically charged situations”. Others focused on the client requesting the analysis, and noted that explicitly exploring ethic situations “instills a sense of confidence in the client on the integrity of the analysis, since there would be fairness, accountability and transparency in the analysis”.

4.2.2 Key Challenges of Incorporating Ethics Analysis

Based on the analysis of the 106 survey responses from the semi-structured interviews that focused on the challenges and benefits of incorporating ethics analysis within a data science project, two key challenges where identified (implementing an ethics process, challenge to know what’s unethical). These are each described below:

Implementing an ethics process – One key challenge that was often noted was the fact that it would be difficult to incorporate ethics checkpoints or evaluation within the current process the team uses to execute their data science project. For example, there was a perceived increase in costs that might be associated with this effort, and some “organizations would not want to incur the added cost”. In addition, others noted that it would be “difficult to determine who is responsible for what output, and how such responsibilities relate to each other”. More generally, “accountability with regard to data ownership and sharing as well as data analysis would be difficult and/or expensive”.

**Challenge to know what’s unethical** – Perhaps the most difficult and frequently noted challenge was that there was no “rule book” on what is OK (what is ethical what is not ethical). This was perhaps best observed by one participant who stated “there is no preset formula for applying ethical consideration that works in all cases. This means that you can not just study a list of rules but instead must constantly be evaluating a number of different ethical issues (bias, data integrity, use of the results) to make sure that you're covering all that apply in each given situation”. In other words, it is difficult for data scientists to incorporate ethics, and “not having a ‘one size fits all’ solution to approaching ethics means that every situation will have to be more comprehensive”. A different response conveyed a similar idea, but focused on the potential unintended consequences by noting that “the biggest challenge appears to be effectively considering all of the ways that a project or method of utilizing data could be unethical, for example the concept of de-anonymizing members of a dataset by merging it with another dataset. This idea, which might not be predicted or considered when the original dataset was released, can still cause an inherently unethical result”. This is especially challenging since most data scientists have minimal training in how ethics applies to data science, and as noted by one participant “data scientist need to have ethical training to better understand ethical challenges in a data science project”.

One example to help demonstrate the difficulty in evaluating ethics within a data science project is the application of machine learning to criminal justice, specifically a recent controversy that involved a county in Florida in the United States, using the COMPAS recidivism prediction score to determine sentencing. This algorithm was found to have false-positive and false-negatives, which created a disparate impact for African Americans (Angwin et al, 2016). While some disagreed with this view (Floridi and Taddeo, 2016), recent studies deepened the understanding of trade-offs in how fairness is defined (Chouldechova, 2017) and demonstrated the tension between improving public safety and satisfying the prevailing notions of algorithmic fairness (Corbett-Davies, 2017). This example, combined with the previously noted example on predicting health care costs, helps to demonstrate the importance of considering the rights of different stakeholder groups whose lives may be disparately impacted by the outcomes of a data science project, as well as the challenge in being able to think through these issues thoughtfully.

4.2.3 Data Science Ethics Analysis Acceptance Model

Based on these findings, Figure 1 shows the derived data science ethics analysis acceptance model. As one can see, complexity and compatibility were seen as key challenges that are likely inhibiting the adoption of ethics analysis and observability and relative advantage were the positively perceived factors.
5 Conclusion

5.1 Limitations and Future Research

One key limitation was that a key aspect of the study was done with graduate students. While the students had previous industry experience, and the use of students enabled a large number of individuals that had some exposure to the potential ethics questions that might need to be addressed within a data science project to be included in this research, the results might be different if experienced data scientists were surveyed. To address this limitation, a follow-up study, leveraging these findings, is a planned next step. Specifically, a more comprehensive survey of industry professionals is planned. Next, a more in-depth case study is planned, that will help evaluate the actual challenges encountered within a team trying to incorporate ethics (as opposed to the perceived challenges that might cause a team not to try and incorporate ethics into their project framework).

In addition, for the roughly one third of the respondents in our first survey who thought that ethics did not come up within their project, it is not clear if that is actually the situation, or if the respondents had just not thought through the many potential ethical situations that might be applicable for their project. Exploring this, via surveys and case studies, could shed some light on if the challenges and benefits might vary by the type of organization, size of organization or by the type of big data project (Saltz, Shamshurin & Connors, 2017).

Finally, another avenue of future research is to create a framework that facilitates the analysis of the ethics within a specific data science project. This could help address one of the key identified challenges in this study, namely, that since there is no one set of rules that one can currently follow to know what are the possible ethical conundrums to consider and how to handle those ethical conundrums when they do arise.

5.2 Summary

There were two goals of this study. First, this research aimed to understand how current data science projects integrate ethics. Based on the findings of our initial survey, most teams think that either ethics is not applicable to their project or that the potential situations can be explored in an ad-hoc fashion as issues arise. Thus, the first research question was addressed (do data science projects integrate the analysis of possible ethics conundrums within their project). This finding suggests that, based on the need for ethics within data science efforts, in the future, teams will need to incorporate ethics evaluation within a data science project.

The second goal was to understand the perceived challenges and benefits of incorporating ethics evaluation within a data science project, which would help provide insight into why data science projects do not yet typically integrate ethics into their process methodology. Leveraging a Diffusion of Innovation framework, two key benefits were identified – focusing on relative advantage (mitigating risk, achieving better results) and observability (improved public / client acceptance of using data science). In addition, two key challenges were identified that focused on compatibility (Implementing an ethics process within the group) and complexity (Difficult to know what’s unethical). Hence, the second research question was also addressed (what are the perceived challenges and benefits of incorporating ethics evaluation within a data science project).

The results of this research can help managers and data science leaders as they start to consider incorporating a structured way to analyze potential ethical conundrums relating to their data science project.
References


