Communications of the Association for Information Systems

Volume 46 Article 24

5-2020

When Programs Collide: A Panel Report on the Competing Interests of Analytics and Security

Jacob A. Young
Bradley University, jayoung@bradley.edu

David P. Biros
Oklahoma State University

Ryan M. Schuetzler
University of Nebraska-Omaha

Tyler J. Smith
Bradley University

Paul R. Stephens
Bradley University

See next page for additional authors

Recommended Citation

Young, J. A., Biros, D. P., Schuetzler, R. M., Smith, T. J., Stephens, P. R., Syler, R. A., & Zheng, S. H. (2020). When Programs Collide: A Panel Report on the Competing Interests of Analytics and Security. Communications of the Association for Information Systems, 46, pp-pp. https://doi.org/10.17705/1CAIS.04624

When Programs Collide: A Panel Report on the Competing Interests of Analytics and Security

Authors

Jacob A. Young, David P. Biros, Ryan M. Schuetzler, Tyler J. Smith, Paul R. Stephens, Rhonda A. Syler, and Shawn H. Zheng

Communications of the Association for Information Systems

Panel Report DOI: 10.17705/1CAIS.04624 ISSN: 1529-3181

When Programs Collide: A Panel Report on the Competing Interests of Analytics and Security

Jacob A. Young
Bradley University
jayoung@bradley.edu

David P. Biros
Oklahoma State University

Ryan M. Schuetzler
University of Nebraska-Omaha

Tyler J. Smith
Bradley University

Paul R. Stephens
Bradley University

Rhonda A. Syler
University of Arkansas

Shawn H. Zheng
Bradley University

Abstract:

The increasing demand for business analytics and cybersecurity professionals provides an exciting job outlook for graduates of information systems programs. However, the rapid proliferation of devices and systems that spurred this trend has created a challenging ethical dilemma for the individuals responsible for educating future generations of information technology professionals. Many firms collect and store as much data as possible in the hope that technology might uncover useful insights in the future. This behavior results in an ever-increasing challenge for those charged with protecting organizational assets and exerts pressure on executives seeking an analytical edge to remain profitable in a hyper-competitive marketplace. With this dilemma in mind, a recent panel discussion at the 14th Annual Midwest Association for Information Systems Conference explored the delicate balance between unleashing the power of analytics and securing the sensitive data it consumes while respecting consumer privacy. This paper reports on that discussion and its insights.

Keywords: Analytics, Privacy, Security, IS Education.

This manuscript underwent editorial review. It was received 06/24/2019 and was with the authors for 2 months for 1 revision. Christoph Peters served as Associate Editor.

Jacob Young conceived the topic, invited panelists, moderated the discussion, and contributed to the panel report. The subsequent authors contributed equally to the panel report and are, therefore, listed alphabetically.

1 Introduction

The increasing demand for business analytics and cybersecurity professionals provides an exciting job outlook for graduates of information systems (IS) programs (Mandviwalla, Harold, & Purnama, 2019). Despite daily headlines about data breaches and governmental inquiries into questionable behavior by some of the world's most recognizable companies, the collective appetite for data has not slowed down. Many firms collect and store as much data as possible in the hope that technology might uncover meaningful insights in the future (LaValle, Lesser, Shockley, Hopkins, & Kruschwitz, 2011). This phenomenon results in an ever-increasing challenge for information security personnel charged with protecting organizational assets (Rawat, Doku, & Garuba, 2019) and consumer privacy (Barocas & Nissenbaum, 2014) while at the same time intensifying the mounting pressure for executives seeking an analytical edge to remain profitable in a hyper-competitive marketplace (LaValle, Hopkins, Lesser, Shockley, & Kruschwitz, 2010).

Several challenges propagate from these competing forces that have implications for industry and educators as we prepare business analysts and information security experts for the workforce. For example, many companies believe they protect consumers by anonymizing sensitive data. Despite these efforts, advanced analytical techniques can possibly reattribute anonymized data to specific individuals (Barocas & Nissenbaum, 2014; Sweeney, Abu, & Winn, 2013).

Decision makers might also rely too heavily on analytics and, thus, interpret data in biased, incorrect, or harmful ways. Many people naively view analytics as a magic bullet that will unlock the mysteries of the universe, which drives their appetite to consume even more data. However, if one does not recognize how bias can undermine results, one will find oneself wanting to collect more data than one should while simultaneously making decisions drawn from errant conclusions. Depending on the context, these decisions could have disastrous implications for privacy, security, and individual liberty.

The rapid proliferation of devices and systems that spurred these trends has created a challenging ethical dilemma for the individuals responsible for educating future generations of information technology (IT) professionals. With this dilemma in mind, we organized a panel that included faculty with expertise in cybersecurity, analytics, and law to discuss these issues. In this paper, we report on that discussion and bring forward the culminating thoughts to the broader community in hopes to continue the dialogue.

2 Motivation

The impetus for exploring these issues began when several panel members, all management information systems faculty at Bradley University, began to revise an undergraduate curriculum in 2016 in response to industry needs. As a result, the faculty added concentrations and minors in business analytics and cybersecurity to the curriculum from the 2018-2019 academic year. Throughout the curriculum-development process, the faculty had to address issues at points where the two programs converge. The questions lack straightforward answers, and, with the programs underway, the discussion continues.

The faculty identified several trends and issues while developing the program. First, organizations want more employees to have access to data to help them make operational, tactical, and strategic decisions. However, to use data analytics in a way that often best leverages its power—that is, to discover seemingly unrelated relationships—employees need access to various data and information sources that might fall outside the scope of their job titles. Yet, information security professionals need to protect and control access to data and information. For example, many data science and business intelligence techniques create insights by leveraging non-obvious relationships, but doing so requires one to bring together disparate data sources to be effective and provide meaningful insights. Potential conflict with information security access policies can hinder the value that an organization's analytics program provides. Thus, organizations need to understand how they can identify and deploy a balance between providing employees access to data and mitigating potential harm.

Further, due to the rapid rise in data and increased analytical power, fresh graduates from analytics programs commonly find that their organizations already consider them de facto experts, not apprentices. Although new hires can bring expertise in applying contemporary analytical techniques, they will likely lack the ability to foresee long-term consequences due to their limited experience. This reality led some panel members to struggle with touting analytics' power and capabilities given its potential threats to privacy and security. Although the MIS faculty at Bradley University recognized the clear need to develop highly skilled analytics professionals, they feared producing graduates who would become the next poster child for

unethical or negligent behavior. Thus, they felt it would irresponsible to develop any curriculum, analytics or otherwise, without exposing students to these issues throughout the program. Clarke (2016) explains this responsibility in his paper on the risks associated with big data:

> *Computer science and information systems academics and professionals have direct responsibility in relation to the technical aspects of data collection, storage and access. Our responsibility in relation to data analysis is shared with other disciplines and professions, but not to the extent that our responsibility is extinguished. The human aspects inherent in big data risks can, on the other hand, only be addressed through risk assessment and risk management, by ensuring that business process design incorporates safeguards, compliance audits and enforcement activities. Once again, computer science and information systems researchers and professionals bear some of the responsibility to ensure that problems are identified, publicized and addressed. We have moral responsibility, potentially translated by the courts into legal responsibility, to blow the whistle on hyperbole, and on undue reliance on technology.* (p. 88)

Therefore, we conducted the panel to ensure that we do not duck our responsibility as educators to prepare students to anticipate and address emerging ethical issues related to analytics and big data as we train future data analytics experts. Ultimately, we must answer the following question: how do we ensure that we fully equip our graduates to perform cutting-edge work while ensuring that they do not contribute to the next data breach or consumer privacy scandal?

3 Panel Organization

The panel discussion took place on May 21, 2019, at the 14th Annual Midwest Association for Information Systems Conference, which the University of Wisconsin Oshkosh hosted. To provide a panel that could address the multi-faceted perspective necessary to have a robust discussion, the lead author recruited a diverse group of panelists with overlapping relevant expertise in analytics, security, business law, and IS curriculum design. The panel focused on exploring the delicate balance between unleashing the power of analytics and protecting the sensitive data it consumes.

To tackle this organizationally complicated and potentially politically charged issue, we organized the panel discussion into four primary themes: privacy, security, legal considerations, and ethics. We provide the themes and related questions in Table 1. In the following subsections, we synthesize the overarching issues in each area before providing potential solutions for industry and implications for curriculum development in the following sections.

Table 1. Identified Themes and Related Questions

Theme	Questions
Privacy	1) What role does education play in helping consumers (the tracked) understand the value of their data to those collecting it? 2) How can organizations better communicate to consumers what data is being generated and how it is collected or stored? 3) What does privacy mean in the face of ubiquitous tracking and analytics?
Security	1) What are the potential information security issues created by the lack of standardization and the use of open source tools by data analytics professionals? 2) Does using analytics in security introduce any risks to employees (e.g., considering a pro-social whistleblower a threat to the organization)?
Legal	1) Where do current laws fall short on each side? 2) How do the two sides (analytics and cybersecurity) currently further their respective interests while abiding by current laws? 3) To what extent should analytics and cybersecurity professionals be responsible for ensuring the intended goals behind privacy laws are met?
Ethical	1) What responsibility does a company have for disclosing to consumers how they use customer data? 2) Is collecting everything ethical? Should we prohibit some data sets/types? 3) While analytics data generally falls outside an institutional review board's purview, should it?

4 Privacy

In security, we frequently teach that the greatest security vulnerability in any organization constitutes the people who work there. Social engineering, mistakes, and insider threats all work together to counteract technical and policy measures that organizations implement. We might also say that people constitute the greatest threat to their own privacy. For example, studies on the privacy paradox show that people consistently overstate how much they care about privacy when compared to their privacy-related behaviors (e.g., Barnes, 2006; Kang, Dabbish, Fruchter, & Kiesler, 2015). To mitigate this weakness, people need to learn about the dangers associated with large amounts of personal data residing on networks throughout the Internet. Additionally, they must understand how data aggregation impacts their personal privacy and security.

4.1 Consumer Education

While one customer data point may have little value on its own, millions have significant value to an organization. Most individuals recognize the former but not necessarily the latter and, therefore, find it difficult to understand why organizations view their personal information all that valuable. This perception leads consumers to hold largely agnostic opinions toward providing personal data. Since they do not know enough to feel concern, they will likely choose convenience over privacy when they want something immediately.

Consumer education can help individuals understand how widely organizations collect data, how they use it, and how they exchange it with other parties. Individuals may then better understand the true value of their personal information and want to have more control over it. A knowledgeable citizen can then make informed decisions about privacy and security.

In addition, by understanding how, when, and what kind of personal data they give away, when organizations use their personal data, and the consequences associated with such usage, consumers may develop more trust in organizations. Consumers want to know how and when organizations sell/exchange their personal data and what implications the transactions may have. Consumers will be able to more accurately assess the trust they place in organizations once they better understand the tradeoff between convenience and risk. Efforts to enhance consumer awareness will also empower consumers to make informed decisions, protest against unethical data practices, regain control of their data, and seek help when necessary.

4.2 Consumer Awareness through Communication

Unless governments force organizations to educate consumers, they will not likely do it. If (or when) most consumers demand such education, first-mover organizations might see an opportunity for competitive advantage. Until that time, organizations will see reduced transparency as an advantage (for evidence, one need only look at the differences between transparency practices in the US versus Europe). When organizations can self-regulate, they choose not to better communicate with consumers.

To protect users' privacy, one must first teach them how valuable companies find their data. The quote "if you are not paying for the product, you are the product" can often help users understand their data's value. While one may most commonly hear this refrain in technology circles relating to Facebook and Google's user data collection, it also applies to many services across the Web given that user tracking, profiling, and advertising form their primary revenue source. In any case, what form the education takes and who does it remain topics up for debate. Further, the education needs to consider how individuals learn and their needs (Wisniewski, Knijnenburg, & Lipford, 2017) or it will be ineffective.

Companies can—and should—take steps to educate users on how they will use their data. Presenting a privacy policy by default (Steinfeld, 2016) in language that average users can clearly understand (Milne, Culnan, & Greene, 2006) will help them make informed decisions about their data. Besides the language itself, privacy policies and terms of service should present the information users care about in a way they can understand. Organizations often write these policies in order to defend themselves legally, which means they seldom include what users care about. Services such as "Terms of Service; Didn't Read" (https://tosdr.org/) can help users make informed decisions about their data and privacy but only if users know about and care enough to use it.

Currently, companies provide little to no communication on how they specifically use personal data. Companies should start conversations regarding who owns, controls, or grants access to this data and

acknowledge the problems that exist with existing practices, which range from collecting personal data that goes far beyond traditional marketing needs to explaining the poor security and privacy management efforts that led to past data breaches. The issue also extends to underregulated buy/sell/trade practice guidelines from law and government. The "fine print" on the bottom of the page fails to fully inform consumers, and organizations should think about how they can better communicate key information effectively.

4.3 What Does Privacy Mean Today?

Most technologies default to open access to personal data. When asked, most consumers opt for convenience over privacy and security. Unless individuals learn to value their privacy more than convenience, privacy may become a quaint concept from history. In the US, we currently have an open season on privacy. The real question concerns what happens when the backlash occurs? Will consumers become so fed up with the unregulated landscape to the point the pendulum will swing in the other extreme direction? Can we anticipate a future where the government or other parties control data collection so much that it loses its value to organizations? Unregulated data collection and use leads to abuse. As an example, robocalls have become so intrusive that many have begun to scrutinize the technology itself. If given a choice, organizations would prefer regulated tracking and analytics over these activities becoming illegal someday.

Innovation always involves risks. With the development of location-based tracking technology and aggressive data collection, privacy requires rigorous security measures for primary usage. Any other activities that involve transferring or trading data would need to extensively anonymize individuals' identity. All these procedures should also have accordant governing mechanisms such as policies and regulations. The consumer should comprehensively understand how organizations will use their data when they opt-in to data collection and be allowed to maintain complete control over the data that the organizations collect thereafter.

Unfortunately, we already have wearable devices, the Internet of things (IoT), auto manufacturers that can track vehicles, and smart home devices. The data that these embedded systems generate presents a direct threat to user privacy and security. The news is littered with reports of hacked devices and compromised information. For example, smart watches capture so much data on user behavior that they will even ask individuals why they deviate from their normal patterns when they do so, such as not going to work. We have also seen ubiquitous tracking and analytics introduce tremendous threats to military operational security. In November, 2017, Strava, a fitness tracking company, revealed secret U.S. military bases and patrol routes to the world when three trillion individual GPS data points of Fitbit workout data that users uploaded were displayed on a heat map (Hern, 2018). In this case, anyone could readily access the open source data. Such situations highlight the privacy concerns that emerge when petabytes of tracking information get into the wrong hands.

5 Security

In addition to the apparent privacy concerns related to consumer data, using analytics might also introduce new security issues. In this section, we discuss how poor standardization and reliance on certain indicators might result in unanticipated negative outcomes.

5.1 Technology Standardization Issues

Data analytics professionals use various tools in their work. However, the technologies and software that support data analytics lack standardization both in and across industries, which creates potential information security issues. For example, consider database management systems. Organizations may use many different databases and database products. Database administrators play the dual role of security and analytics professionals. These professionals must follow database-hardening best practices that cover a plethora of activities designed to secure databases.

Database administrators need to define user roles, set permissions, and maintain password hashes. They also look after management and reporting policies. At the same time, database administrators must provide analytic professionals with timely access to data to help them make decisions. Once analysts have the data in their hands, load it into different software applications, and manipulate it, data administrators no longer directly control it. And, if the data source models combine external and internal data sources, analysts do not necessarily know how systems might share internal data with external

sources. Today's cloud environment for database platforms adds even more information security implications.

The inherent challenges with performing analytical work are not limited to the lack of standardization for analytical tools. Generally, data mining/analytics requires large data sets to derive patterns to predict new data points. The process itself does not concern any single data point in the training data set. However, large data sets introduce more risks, even with anonymization. Sometimes analytics can identify data points without identifiers. At the same time, one can often access so much information about that data point (profiling, educated guess, etc.). The challenge lies in the amount of data available. Different data sets potentially contain many common identifiers, which makes anonymization extremely difficult if not impossible.

5.2 Security Analytics: Risks to Employees

Many employees already understand that their workplaces monitor their activities. On the other hand, typical employees likely do not accurately understand what their organization monitors and how it may use the information it gathers against them in unanticipated situations (Krouse, 2019).

For example, some employees might trust their organization's in-house "anonymous hotline" only to find out the individuals they talk about can retroactively identify them. The increased rate at which organizations collect and store logs to feed into security information and event management (SIEM) systems could result in organizations identifying and retaliating against more whistleblowers, especially since organizations need such data to identify true insider threats (Agrafiotis et al., 2015). Obviously, all forms of whistleblower retaliation are unacceptable, but especially since organizations also run the risk of retaliating against the wrong employee if they rely too much on analytics. In any case, whistleblowers must weigh and consider the fact that someone may discover their identity. However, in the US, the courts and government agencies do offer some protection for whistleblowers. Under the False Claims Act, the Department of Justice can pay rewards to whistleblowers. Further, some law firms such as The Employment Law Group specialize in protecting whistleblower rights i.

Despite offering these legal protections, many whistleblowers suffer mental and emotional anguish while navigating through costly court battles before they have a chance to receive a reward or vindication (Alford, 2001). Therefore, they must successfully raise their concerns anonymously to truly avoid retaliation. Unfortunately, naive technology users may find it extremely difficult to maintain their anonymity in today's information age (Marcum & Young, 2019). Ultimately, the rapid advancement in data analytics has the potential to further silence individuals who can put a stop to unethical behavior before it gets out of control.

6 Legal

The privacy and security issues we discuss above should have legal consequences, but we also rarely see significant judgments. Many companies who have violated consumer privacy or failed to protect sensitive data continue to conduct business as usual often without a noticeable change in their business practices. In this section, we discuss the shortcomings of existing laws and explore responsibilities for analytics and cybersecurity professionals.

6.1 Legal Shortcomings

Data analytics and cybersecurity have competing interests. The laws supposed to protect the information each side values for different reasons are largely ineffective even when present. Currently, the US does not have an overarching federal cybersecurity law, nor does it have an overarching information privacy law. Rather, the sectoral approach regulates information and data privacy, which means that different laws that overlap in some areas and leave gaps in others affect each economic sector. As a result, one often does not know how and when certain laws apply. For example, the Health Insurance Portability Accountability Act (HIPAA) and the Family Educational Rights and Privacy Act (FERPA) have generated confusion when it comes to school healthcare providers disclosing protected health information. In some situations that involve the same providers, HIPAA applies, and, in other situations, FERPA applies. Further, each law defines basically the same information in different ways and includes multiple exceptions, which muddies the waters further.

Adding to the confusion, over 21 different federal statues mandate that organizations disclose or provide notice in some form to consumers about how they use their data. Data privacy regulations that prescribe "mandated disclosure" or "notice and choice" have many pitfalls. When purchasing something such as a house or software or applying for credit, the law requires organizations to make a dizzying array of disclosures. These disclosures use language in such a way that people who need to know the information cannot learn it simply by being told. For example, speaking on credit card disclosures, U.S. Senator Elizabeth Warren said: "I teach contract law at Harvard Law School and I can't understand my credit card contract. I just can't. It's not designed to be read." (Brancaccio, 2009). Lawmakers can easily mandate disclosure and then claim consumers have notice. Yet, it does not actually solve the problem.

6.2 Legal Responsibilities for Analytics and Cybersecurity Professionals

Regulation best comes from within. Even when external regulations do not mandate that organizations protect consumers' privacy, organizations should still do for several reasons. First, it represents the ethical choice. Many consumers simply do not have the time, patience, or knowledge to take steps to protect their privacy, and companies have the best position to protect sensitive data. Second, in anticipation that lawmakers will implement more comprehensive privacy laws, companies can beat the curve by upping their standards. The General Data Protection Regulation (GDPR) that the European Union implemented in 2018 is an overarching data privacy regulation that marks an important evolution in data privacy regulations and rules. Under the GDPR, all companies that handle E.U. citizen data must comply with a single set of regulations no matter where they reside. By default, the GDPR mandates that consumers must opt-in and give consent to companies to use their data in contrast to the preexisting law that mandated companies send customers notices that gave them the option to opt-out. In anticipation of the GDPR, many companies implemented changes to get a head start on compliance.

7 Ethical

Our discussion thus far leads to several ethical questions, especially as it pertains to collecting and using consumer data. How should organizations disclose what they do with consumer data? What restrictions should they follow when collecting data? How might institutional review boards (IRBs) assist in strengthening protections?

7.1 Company "Use of Data" Disclosure Policy

Given what we know about customer behavior, it seems clear that people will allow organizations to use their data for all kinds of activities if they feel they will benefit from it. As such, an organization can easily exploit users' trust (or lack of awareness) and gain access to data users may not even understand. The analysts and the companies using this data must have some ethical obligations towards the individuals whose data they analyze. While an organization may be legally in the clear with an appropriately worded privacy policy, ethics should extend beyond the legal minimums. Organizational policies such as the now-deprecated Google slogan "don't be evil" can help guide decisions even if a legal reason not to do something does not exist. Organizations could share, in clear language designed for consumer reading, the ways in which and the business model behind how they will use data. In doing so, organizations could clearly and positively define their data goals.

7.2 Restrictions on Data Collection

Since collecting data and using data refer to two different things, informed consent must account for both. If organizations collect data with users' consent (even via the fine print nobody reads), it should be ethical provided that they use the information in ways in which users intended. From a research perspective, if organizations implement rigorous protocols, such as being approved by an institutional review board (IRB), there should not be any limitation on dataset/types. However, the issue is consumers often do not understand how organizations could use their data. That said, while some organizations want to collect everything about everyone because they can do so relatively easily and cheaply or because they see an opportunity for monetization, collecting everything is neither desirable nor ethical.

Dave Eggers' dystopian fiction novel *The Circle* shows where such behavior might lead. The book features a powerful technology company called The Circle whose founders advocate for complete transparency and universal lack of privacy in order to discourage individuals from acting unethically. However, when reading the book, one slowly learns that the founders seek to monopolize the market,

collect all data, and monetize it. Compare that to the public messages from companies in the real world. One can possibly see *The Circle* as a response to Google's original mission statement ("Our mission is to organize the world's information and make it universally accessible and useful") and mantra "don't be evil". Despite Google's stated intentions, its critics likely see its behavior as matching The Circle's founders' behavior.

Unfortunately, as a society, we have not given these issues enough attention. When a company asks people to voluntarily submit their DNA and the marketing professionals at these firms present it to them as a wonderful thing, too few loudly object. Organizations should have some limits on what data they can collect, and our best minds should provide some guidance in that regard. These concerns will have to trickle down to the public conscience and onto the plates of local, state, and federal legislators, but change would require a long and drawn out process, and practice proceeds policy by years.

7.3 Analytics Data and the Institutional Review Board (IRB)

Data analytics opens new possibilities for discovery. For example, Cerner conducted analytics on healthcare data that resulted in reducing sepsis cases in hospitals across the country. Analytics has also proven useful at predicting the survival rate of cancer patients with comorbidities (Zolbanin, Delen, & Hassan Zadeh, 2015). While both studies produced positive outcomes and their authors redacted patients' personally identifiable information (PII), researchers need to stay vigilant to prevent unintended adverse outcomes. IRBs can help in this area by helping researchers understand the negative consequences associated with analyzing HIPAA-controlled data. Of course, we need to make sure those who serve on our IRBs understand analytics' true power and capabilities or they will not be able to recognize the risks to human subjects. At some point, someone will make a mistake and PII or leak patient healthcare information, which could harm not only patients but also researchers and academic institutions.

8 Proposed Solutions

Given our rapid technological progress, much technical content we teach today may become obsolete tomorrow. Consequently, we need to prepare our graduates for lifelong learning. Therefore, we took a holistic approach in identifying possible solutions. Ultimately, we organized them into the following categories: curricular, technical, and legal/ethical. We address each category by offering proposed solutions for both academia and practice (see Table 2).

Table 2. Proposed Solutions for Academia and Practice

Category	Proposed solutions for academia	Proposed solutions for practice
Curricular	• Increase attention on data ethics • Improve multi-disciplinary integration • Provide leadership from curriculum task forces and accrediting bodies • Account for culture differences	• Renew focus on continuing education • Create and deploy graduate certificate programs • Administer Implicit Association Test • Develop and provide mindfulness training
Technical	• Develop data taxonomy • Increase research on anonymization and privacy enhancing technologies, such as zero-knowledge systems, end-to-end encryption, and self-sovereign identity • Conduct behavioral research on consumer privacy • Formally assess privacy practices	• Follow best practices for security and data anonymization • Perform privacy and security audits and embrace bug bounty programs • Adopt a privacy-by-design approach • Become less financially dependent upon consumer data • Leverage analytics to identify and protect against threats to information assurance
Legal/ethical	• Assist lawmakers in better understanding current and emerging technology • Future-proof legislation • Mandate privacy and security audits and bug bounty programs • Improve on and expand the use of IRBs	• Strengthen and enforce codes of ethics • Empower professional organizations to hold practitioners responsible for unethical behavior • Develop global industry standards for the analytics profession

8.1 Curricular

In this section, we suggest possible curricular solutions to enhance traditional academic programming. First, we suggest integrating data ethics throughout the curriculum and pursuing multi-disciplinary opportunities. We also call on special interest groups and accrediting bodies to provide leadership on ethics education. Second, we encourage educators to better inform practice by offering additional continuing education opportunities.

8.1.1 Academia

Analytics professionals have learned to understand and work with data. However, we have rarely seen analytics programs in the US mention ethics. While some textbooks include chapters on ethics in big data and analytics, many instructors skip such chapters to make more time for technical content. Ethical concerns must either rise in importance so that they rank higher than, say, a programming course, or instructors must integrate them in some way into every class and each topic.

We can enhance the extent to which data sciences programs cover ethics and, thus, the extent to which students understand ethical principles and issues by identifying how analytics professionals should act when handling sensitive data and integrating that knowledge into these programs' key learning objectives.. These should be key learning objectives in programs about data science. Instructors could teach material in a standalone ethics course, incorporate it into discussions in other applied courses, or, ideally, both. Further, students in data analytics courses could benefit from briefly discussing the ethical implications related to the practices and techniques that they cover. Researchers have already begun to publish papers on how to better incorporate these discussions into curricula (Lester & Dalat-Ward, 2019).

While not desirable, the current situation does present a perfect opportunity to work across disciplines to develop curricula on the ethics of big data and analytics. We need to systematically ensure the overall curriculum for both analytics and cybersecurity helps students recognize these dilemmas and learn how to balance it (legally and ethically). Since many graduates will likely face ethical dilemmas early in their career, we cannot trust that industry will protect them from making poor decisions. For example, the Facebook-Cambridge Analytica scandal had its roots in academic research on quantifying personality traits that the University of Cambridge's Psychometrics Centre developed (Kosinski, Stillwell, & Graepel, 2013; Schwartz et al., 2013). Shortly thereafter, Christopher Wylie, at age 24, began using this research to profile individuals for both elections and psychological operations for the United Kingdom's Ministry of Defense and the U.S. Department of Defense (Cadwalladr, 2018). Through its parent company SCL Elections, Cambridge Analytica actively participated in over 100 election campaigns in over 30 countries (Ghoshal, 2018). However, Cambridge Analytica did not become a household name until 87 million unassuming Americans had their data used in micro-targeted advertising campaigns to influence likely voters leading up to the 2016 election (Isaak & Hanna, 2018).

While various parties used targeted marketing in political campaigning before the Cambridge Analytica scandal, it demonstrates how technological capabilities outpace the general public's ability to recognize evolving threats (Privacy International, 2017). The documentary *The Great Hack* and the testimony that witnesses provided to the House of Commons in the UK about Cambridge Analytica highlight the global impact that pervasive collection and unethical use of data can have on society. Our graduates will likely face similar ethical dilemmas early in their careers. Therefore, we must find a way to ensure that every syllabus for data science and analytics courses includes ethics as a key component. We should look at cross-functional curricular-development teams that include analytics and security professionals along with business ethicists. When one couches data ethics as information security, analytics professionals will likely pay more attention simply because they care about data access, quality, and reliability. They should understand that data security represents an important facet of their work. Additionally, the literature has suggested ways to educate analytics students about privacy (Schwieger & Ladwig, 2016). Therefore, we encourage faculty to give these issues more attention in their programs.

Unfortunately, programs have limited capacity, and many educators see data ethics issues, no matter how important, as tangential to developing content focused on the knowledge, skills, and abilities needed for the profession. Thus, it will likely require a concerted effort by the Association for Information Systems Special Interest Group for Education (AIS SIGEd) or the Association for Computing Machinery's (ACM) curriculum task force to synthesize and integrate such suggestions into a cohesive model curriculum. Top-down leadership from accrediting bodies, such as the Association to Advance Collegiate Schools of Business (AACSB) and the Accreditation Board for Engineering and Technology (ABET), would further

strengthen these efforts. Applying ethics across a data science (or any IT) curriculum can help students realize ethics constitutes more than a one-course topic and that they must consider ethical issues throughout their careers.

Lastly, educators must account for varying privacy perspectives across cultures and governments. Individuals in Western-style democracies view privacy differently to individuals in other parts of the world (Bellman, Johnson, Kobrin, & Lohse, 2004). For example, some privacy advocates see China's plan to establish a social credit system to evaluate its citizens' everyday activity and render rewards or punishments based on observed behavior as troubling (Botsman, 2017; Kobie, 2019; Liang, Das, Kostyuk, & Hussain, 2018). While privacy and legislation concerns have risen since China launched the social credit system back in 2014, the country has continued to use the system with a wide scope (Chen & Cheung, 2017). However, from a cultural perspective, Chinese citizens are more likely to accept these developments since they have a lower desire for privacy and stand to gain a safer overall public environment (Chen & Cheung, 2017). On the other hand, research suggests that Japanese citizens exhibit higher privacy concerns than Americans due to societal differences in relational mobility (Thomson, Yuki, & Ito, 2015). Therefore, educators need to address these cultural differences to ensure they prepare students for the global marketplace.

8.1.2 Practice

We initially decided to conduct our panel due to a need to revise an undergraduate IS curriculum. However, we recognize that academia's role in society must extend beyond delivering traditional higher education programs. In particular, the IS discipline needs to contribute to society in as many ways as possible given technology's rapid development and proliferation. Much content in undergraduate programs can become obsolete in a decade. Therefore, we believe that IS educators and researchers also have a responsibility to inform practice more efficiently. We contend that IS faculty should make external outreach a higher priority to provide continuing education to local constituents. Offering additional continuing education opportunities and graduate certificate programs, which have lower time and financial commitments, would allow working professionals to remain aware of current trends and issues.

For example, academics can assist today's data analytics professionals with respect to bias. The results that analysts derive from data analytics depend on their own interpretations more than people realize. On one single dataset, different analysts could have completely different models, algorithms, and results based on personal preferences and parameter settings. Clearly, research has demonstrated that data analysts, who explain analytical results to decision makers, can let bias interfere with their interpretations (e.g., Kahan, Peters, Dawson, & Slovic, 2017). Unconscious bias affects every decision we make, but the opportunity for bias has increased now that organizations mandate data-driven decision making. Data-driven decision making possibly has some unwarranted credibility simply because it relies on data. However, the data, the analysis, and the credibility that go with data-driven decision making can result in harm to others if not handled properly.

Before analysts can take steps to mitigate unconscious bias, they must first recognize its existence. For example, one can use the implicit association test to demonstrate the presence of bias (Greenwald, McGhee, & Schwartz, 1998). Recognizing when bias may be present or what situations could trigger bias can help individuals consciously mitigate it rather than unconsciously applying it. Mindfulness training may be another effective way to help medical practitioners overcome their biases (Teal, Gill, Green, & Crandall, 2012). Academics can provide such training to today's analytics professionals through continuing education programs to help mitigate bias.

8.2 Technical

The appetite for big data has resulted in organizations storing massive databases, which only increases their attractiveness to malicious actors. We have seen the damage that significant data breaches can cause, such as with Equifax and the U.S. Office of Personnel Management. Many data issues we face today result from poor practices established well before the Internet age that technology has exponentially expounded.

For example, various agencies in the US have long used social security numbers (SSN) to both identify and authenticate individuals, but such use does not match its original intended purpose. Had agencies not adopted SSNs as the de facto universal identifier for Americans, a SSN breach would not be as devastating as it is today. It will require drastic action to resolve these deeply embedded business

practices once and for all. For example, the U.S. Social Security Administration could publish all SSNs by a certain date and, thus, render them useless as a form of identification or authentication (Chapple, 2019). As such, organizations would have to replace the use of SSNs with contemporary identification and authentication methods, which would ultimately increase privacy and security for everyone.

In this section, we discuss how researchers can push for these much-needed changes through continued research on anonymity and privacy enhancing technologies. For practice, we encourage organizations to consider new approaches to managing data. In our recommendations for practice, we first encourage IT professionals and business leaders to rethink their market strategy and revenue models to become less dependent on exploiting consumer data by adopting a zero-knowledge approach. Second, we discuss controls necessary to minimize the impact of data breaches before touching on how analytics can help organizations protect assets and reduce the chance that breaches will occur.

8.2.1 Academia

In collecting primary data, researchers can find it difficult to maintain subjects' confidentiality. If researchers use data in ways beyond its original purpose and a common identifier exists across different data sets, the risks associated with confidentiality and integrity grow exponentially. We contend that researchers should develop a taxonomy of data types. They should strictly predict certain types, such as sensitive personal information. Other types of data have less sensitivity. Once they have established such a taxonomy, they should develop different standards for each data type. Analysts can then decouple the common identifier or de-identify the data for any usage other than the original purpose, such as buying, selling, or trading. However, they need to anonymize the data properly.

First, to protect against de-anonymization, analysts need to recognize that the issue exists. When Netflix first released its Netflix prize dataset, it anonymized the data from its perspective by removing identifying information. However, it failed to consider other public datasets with similar information (such as IMDb), which meant one could identify some of the data (Narayanan & Shmatikov, 2008). In particular, analysts must consider de-anonymization threats and process data accordingly in situations where they share raw data or have small sample sizes.

Since data analysts look at full datasets (the forest, not the individual trees), they typically do not focus on individual privacy and anonymity in their day-to-day work. Yet, analytic professionals do often clean and prepare data to ensure data quality and reliability. We believe that the data preparation and cleansing process should also include a privacy and security audit. The audit would ensure the data that analysts make available to others meets organizational security requirements and complies with privacy policies. Further, supplementing formal audits with continuous bug bounty programs would allow analytics professionals to swiftly address unforeseen issues. Therefore, we encourage researchers to regularly assess data-anonymization techniques to ensure they still achieve their stated goals and to help guard against analytics professionals becoming complacent.

In addition, we encourage researchers to evaluate privacy practices, such as following the methodology that "Terms of Service; Didn't Read" employ, which would provide consumers with simplified assessments. Organizations that receive poor marks could be publicly shamed to encourage reform. Behavioral researchers should continue to assess consumer and organizational attitudes towards privacy and assess approaches to privacy awareness training and education (Crossler & Bélanger, 2019; Kokolakis, 2017; Lowry, Dinev, & Willison, 2017; Smith, Dinev, & Xu, 2011). As the general population becomes more aware either through negative experiences or privacy education, the demand for privacy-enhancing technologies (PET) will grow. We recommend that researchers continue pushing forward concepts such as zero-knowledge proof systems (Goldwasser, Micali, & Rackoff, 1989), privacy by design (Cavoukian, 2009, 2012), self-sovereign identity (SSI) (Tobin, Reed, Windley, & Foundation, 2017), and the Dark Internet Mail Environment (DIME) (Levison, 2018). When implemented, these efforts empower users to reclaim privacy control.

8.2.2 Practice

First, professionals must use the same security controls to protect analytic data as they would in securing any other data. For confidentiality, data must be encrypted. Organizations must adhere to system-configuration requirements and have a policy in place to enforce their commitment to security. For integrity, organizations must use backups and strong hash functions to ensure that no one manipulates

the data. Unfortunately, since most organizations will not adhere to those controls, breaches will occur, and they will lose data (Posey, Raja, Crossler, & Burns, 2017).

Since network intrusions and data breaches will inevitably occur, despite our best efforts to prevent them, the best strategy involves reducing their impact. Employing a zero-knowledge, end-to-end encryption, and privacy-by-design approach all but eliminates the risk of exposing customer information. For example, services such as ProtonMail (protonmail.com) and MySudo (mysudo.com) do not access customer information. Both companies rely on paid accounts to fund their operations rather than exploit valuable user data. As such, users need to complete the necessary steps to back up their account and/or set up recovery options since neither company will be able to restore access. While this approach eliminates the opportunity for the company to perform analytics on consumer data, it is far more private and secure for the user, which builds trust and eliminates the temptation for the company to engage in unethical behavior. Of course, if a company fails to uphold their promise, they would damage their reputation. As data breaches continue to pile up, the market demand for zero-knowledge alternatives to popular services will continue to grow, which provides an excellent opportunity for organizations to establish themselves early and leverage privacy by design as a market differentiator.

Second, with the vast amounts of information traversing today's organizational information infrastructures, organizations will need analytics to process it all. Analytics can assist new security controls by quickly processing network data and identifying trends in real time. Understanding where the data came from, where it goes, who had access to it, and what routes it traversed while in transit all represent complex issues that will require data analytics.

Some organizations already use such data to prevent failure. For example, organizations can use data that they collect from production processes to design better products and prevent manufacturing errors. Error prevention can be a labor-intensive and expensive process. Engineers have traditionally used a process called failure modes and effects analysis (FMEA) to prevent product failures in the field. It has been a difficult, cumbersome, and labor-intensive but necessary process. Organizations now collect data in the field and machine learning and artificial intelligence (AI) to replace human-driven FMEA. The results save time and money and more effectively prevent design errors over time.

Information security requires real-time analysis, threat assessment, and prevention. Involving people in this process makes the error prevention in information security difficult. Organizations can apply machine learning and AI to information security just as they apply it to product design and manufacturing. We can easily imagine a scenario in which organizations use data to prevent failures in information security. With that said, AI itself could cause problems for organizations given its complexity.

8.3 Legal / Ethical

When analyzing the latest data breach or consumer privacy scandal in hindsight, we find it easy to criticize organizations while they undergo their public shaming on television and social media. Obviously, organizations would ideally regulate themselves especially since the patchwork of existing laws complicates the situation. However, if organizations could easily regulate themselves and learn from others' mistakes, we should see poor security and privacy practices decrease, not increase. Clearly, we need to do things differently. In this section, we discuss the role academics can play in these reforms and call on professional organizations to better educate their membership and hold practitioners responsible for unethical behavior.

8.3.1 Academia

Unfortunately, due to industry's failure to self-regulate, we have seen legislative chambers give increased attention to these areas. The U.S. Congress has tried since the 1970s to pass a more comprehensive federal privacy law that would supersede state privacy laws. Lobbying succeeded in ensuring these laws only applied to the government and not the private sector. As of 2019, congress has begun its fourth attempt, which the Cambridge Analytica scandal has spurred on. In the 2019 congress session, lawmakers have introduced at least six data privacy protection bills.

At the state level, California passed the California Consumer Privacy Act in 2018, which took effect in 2019, to make it clear that consumers have a right to know what personal information organizations collect about them, to know whether they sell that information, to say no to organizations selling their information, and to access their information. However, this law has received criticism for building on mandated disclosure.

In recent years, numerous countries around the world have passed laws and rules addressing personal data protection and data security. China, for example, much like the US, does not have a single comprehensive omnibus data protection law in place. Rather, various laws and regulations create a complex framework. In June, 2017, China implemented the first national-level law addressing cybersecurity and data privacy protection, the People's Republic of China (PRC) Cybersecurity Law (DLA Piper, 2019). In Europe, countries continue to work on implementing the General Data Protection Regulation (GDPR) that came into effect in 2018.

With these legislative efforts ongoing, academics have an opportunity to help lawmakers better understand the relevant technology and emerging issues. Without such collaboration, legislative processes will likely not effectively resolve or prevent undesirable outcomes. First, we encourage legislators to require organizations to undergo regular security and privacy audits and incentivize bug bounty programs. In doing so, they should avoid limiting legislation to issues we face today. Instead, legislators should future-proof laws by mandating that organizations conduct audits in accordance with current best practices, which would allow the law to evolve with technology as auditors can tailor their security and privacy assessments to contemporary threats.

Second, we believe that more organizations besides research institutions should use and improve on IRBs. An independent body should have to review how all organizations collect consumer data collect, use, and store that data. Knowledgeable professionals that can recognize weaknesses in proposed activities should conduct these reviews. Organizations should also ensure that their IRBs perform thorough reviews and do not simply become a rubber stamp process.

8.3.2 Practice

We believe professional organizations can do a lot more to better hold their members accountable. First, they can begin by requiring their members to adhere to a code of ethics. A code of ethics adds value to any profession. For example, the well-known Hippocratic Oath forms the basis for the code of ethics that medical practitioners must adhere to. In accounting, the Institute of Management Accountants and the Institute of Internal Auditors both follow a code of ethics based on four tenets: confidentiality, integrity, objectivity, and competency. These codes can help professionals ensure they act ethically. At a minimum, adopting a code of ethics will help organizations and professionals slow down and at least think about what they do. Just because they can exploit analytics for competitive gain does not necessarily mean they should.

Certified ethical hackers (so-called "white hat" hackers) also accept a code of ethics that includes such principles as being honest about their limitations and ensuring activities follow the law (https://www.eccouncil.org/code-of-ethics/). It would be a great idea for the data science community to identify similar principles that one should follow when analyzing, interpreting, and reporting data. Just as white hat hackers look for vulnerabilities in systems and networks, white hat data analysts can identify potential unintended consequences for organizations that have large data sets. They could then advise the organizations on ways to properly handle the data.

Second, professional organizations can provide training on their respective codes and educate their members using ethical scenarios drawn from the real world. Requiring a code of ethics review for continuing professional education (CPE) credits would encourage professional members to undergo regular training. Finally, they can make it clear in their by-laws that members who do not adhere to the code will lose their certification and be expelled from the association. However, the IT industry currently lacks an authoritative accrediting body, which limits the effectiveness of a code of ethics because formal sanctions would have little to no effect. This reality makes ethics a particularly important issue at the organizational level for IT (Ponelis & Britz, 2012).

We can see the challenges associated with fostering ethical behavior particularly in data science where formal training continues to emerge since many practitioners might have learned the art without going through a program that includes ethical components. While professional associations can adopt a code of ethics and encourage compliance, they currently lack the power to provide any meaningful enforcement across the industry. To gain that power, organizations would need to recognize and reward an accrediting body that includes and enforces a code of ethics.

Lastly, to truly protect against undesirable outcomes, such as biased analyses, we recommend that each industry adopt specific guidelines because generalizability may suffer from personal biases, contexts, and cultures. However, we need to answer several questions in developing such guidelines. Should

organizations require that multiple analysts interpret results? Must analysts consult external experts for critical decisions? We need to consider all techniques to reduce bias in analysis and decision making, but data's ubiquity poses a challenge.

9 Conclusion

We recognize that the IS discipline and the analytics profession face many challenging issues. Although we might be late in addressing the dilemmas that the big data explosion has spawned, we know we can do more to ensure future generations of analytics professionals can make better decisions. Above all, we believe that we need to pay more attention to ethically using information technology. Rather than relegating ethics to a chapter in a textbook, we must find ways to embed these lessons throughout our academic programming. While we do not claim to have a perfect solution to these problems, we hope that our panel discussion will help other faculty develop curricula at their institutions and, as a result, responsible IT professionals.

Acknowledgments

We thank Don Heath and the University of Wisconsin Oshkosh for organizing the conference and allowing us to discuss these important issues. We also thank the conference attendees who shared their insights following the panel discussion.

References

Agrafiotis, I., Nurse, J. R., Buckley, O., Legg, P., Creese, S., & Goldsmith, M. (2015). Identifying attack patterns for insider threat detection. *Computer Fraud and Security*, *2015*(7), 9-17.

Alford, C. F. (2001). *Whistleblowers: Broken lives and organizational power*. Ithaca, NY: Cornell University Press.

Barnes, S. B. (2006). A privacy paradox: Social networking in the United States. *First Monday*, *11*(9).

Barocas, S., & Nissenbaum, H. (2014). Big data's end run around procedural privacy protections. *Communications of the ACM*, *57*(11), 31-33.

Bellman, S., Johnson, E. J., Kobrin, S. J., & Lohse, G. L. (2004). International differences in information privacy concerns: A global survey of consumers. *Information Society*, *20*(5), 313-324.

Botsman, R. (2017). Big data meets big brother as China moves to rate its citizens. *Wired*. Retrieved from https://www.wired.co.uk/article/chinese-government-social-credit-score-privacy-invasion

Brancaccio, D. (2009). Elizabeth Warren on credit card "tricks and traps". *PBS*. Retrieved from http://www.pbs.org/now/shows/501/credit-traps.html

Cadwalladr, C. (2018). "I made Steve Bannon's psychological warfare tool": Meet the data war whistleblower. *The Guardian*. Retrieved https://www.theguardian.com/news/2018/mar/17/data-war-whistleblower-christopher-wylie-faceook-nix-bannon-trump

Cavoukian, A. (2009). Privacy by design: The 7 foundational principles: Implementation and mapping of fair information practices. *Information and Privacy Commissioner of Ontario.* Retrived from https://collections.ola.org/mon/24005/301946.pdf

Cavoukian, A. (2012). Privacy by design. *IEEE Technology and Society Magazine*, *31*(4), 18-19.

Chapple, M. (2019). Taking Social Security numbers public could fix our data breach crisis. *CNN*. Retrieved from https://www.cnn.com/2019/06/05/perspectives/labcorp-quest-diagnostics-data-breach-social-security-numbers/index.html

Chen, Y., & Cheung, A. S. (2017). *The transparent self under big data profiling: Privacy and Chinese legislation on the social credit system*. *The Journal of Comparative Law*, *12*(2), 356-378.

Clarke, R. (2016). Big data, big risks. *Information Systems Journal*, *26*(1), 77-90.

Crossler, R. E., & Bélanger, F. (2019). Why would I use location-protective settings on my smartphone? Motivating protective behaviors and the existence of the privacy knowledge-belief gap. *Information Systems Research*, *30*(3), 995-1006.

DLA Piper. (2019). *Law in China.* Retrieved from https://www.dlapiperdataprotection.com/index.html?t=law&c=CN

Ghoshal, D. (2018). Mapped: The breathtaking global reach of Cambridge Analytica's parent company. *Quartz*. Retrieved from https://qz.com/1239762/cambridge-analytica-scandal-all-the-countries-where-scl-elections-claims-to-have-worked/

Goldwasser, S., Micali, S., & Rackoff, C. (1989). The knowledge complexity of interactive proof systems. *SIAM Journal on Computing*, *18*(1), 186-208.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit assocation test. *Journal of Personality and Social Psychology*, *74*(6), 1464-1480.

Hern, A. (2018). Fitness tracking app Strava gives away location of secret US army bases. *The Guardian*. Retrieved from https://www.theguardian.com/world/2018/jan/28/fitness-tracking-app-gives-away-location-of-secret-us-army-bases

Isaak, J., & Hanna, M. J. (2018). User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer*, *51*(8), 56-59.

Kahan, D. M., Peters, E., Dawson, E. C., & Slovic, P. (2017). Motivated numeracy and enlightened self-government. *Behavioural Public Policy*, *1*(1), 54-86.

Kang, R., Dabbish, L., Fruchter, N., & Kiesler, S. (2015). “My data just goes everywhere”: User mental models of the Internet and implications for privacy and security. In *Proceedings of the Symposium on Usable Privacy and Security*.

Kobie, N. (2019). The complicated truth about China’s social credit system. *Wired*. Retrieved from https://www.wired.co.uk/article/china-social-credit-system-explained

Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & Security*, *64*(2013), 122-134.

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences*, *110*(15), 5802-5805.

Krouse, S. (2019). The new ways your boss is spying on you; It’s not just email. Employers are mining the data their workers generate to figure out what they’re up to, and with whom. There’s almost nothing you can do about it. *Wall Street Journal*. Retrieved from https://www.wsj.com/articles/the-new-ways-your-boss-is-spying-on-you-11563528604

LaValle, S., Hopkins, M. S., Lesser, E., Shockley, R., & Kruschwitz, N. (2010). Analytics: The new path to value. *MIT Sloan Management Review*, *52*(1), 1-25.

LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MT Sloan* Management Review. Retrieved from https://sloanreview.mit.edu/article/big-data-analytics-and-the-path-from-insights-to-value/

Lester, L.-J., & Dalat-Ward, Y. (2019). Teaching professionalism and ethics in IT by deliberative dialogue. *Informstion Systems Education Journal*, *17*(1), 4-17.

Levison, L. (2018). *Dark Internet mail environment*. Retrieved from https://sovrin.org/wp-content/uploads/2017/06/The-Inevitable-Rise-of-Self-Sovereign-Identity.pdf

Liang, F., Das, V., Kostyuk, N., & Hussain, M. M. (2018). Constructing a data-driven society: China’s social credit system as a state surveillance infrastructure. *Policy & Internet*, *10*(4), 415-453.

Lowry, P. B., Dinev, T., & Willison, R. (2017). Why security and privacy research lies at the centre of the information systems (IS) artefact: Proposing a bold research agenda. *European Journal of Information Systems*, *26*(6), 546-563.

Mandviwalla, M., Harold, C., & Purnama, M. (2019). *The Association for Information Systems and Temple University Information Systems job index.* Retrieved from https://isjobindex.com/downloads/1755/

Marcum, T., & Young, J. A. (2019). Blowing the whistle in the digital age: Are you really anonymous? The Perils and pitfalls of anonymity in whistleblowing law. *DePaul Business & Commercial Law Journal*, *17*(1), 1-38.

Milne, G. R., Culnan, M. J., & Greene, H. (2006). A longitudinal assessment of online privacy notice readability. *Journal of Public Policy & Marketing*, *25*(2), 238-249.

Narayanan, A., & Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. In *Proceedings of the IEEE Symposium on Security and Privacy* (pp. 111-125).

Ponelis, S. R., & Britz, J. J. (2012). The elephant in the server room: Confronting the need for an ethics officer in the IT function. *Journal of Information Ethics*, *21*(1), 27-39.

Posey, C., Raja, U., Crossler, R. E., & Burns, A. J. (2017). Taking stock of organisations’ protection of privacy: Categorising and assessing threats to personally identifiable information in the USA. *European Journal of Information Systems*, *26*(6), 585-604.

Privacy International. (2017). *Cambridge Analytica explained: Data and elections.* Retrieved from https://privacyinternational.org/feature/975/cambridge-analytica-explained-data-and-elections

Rawat, D. B., Doku, R., & Garuba, M. (2019). Cybersecurity in big data era: From securing big data to data-driven security. *IEEE Transactions on Services Computing*, *1374*(c), 1-1.

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M., Shah, A., Kosinski, M., Stillwell, D., Seligman, M. E. P., & Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open-vocabulary approach. *PloS One*, *8*(9), e73791.

Schwieger, D., & Ladwig, C. (2016). Protecting privacy in big data: A layered approach for curriculum integration. *Information Systems Education Journal*, *14*(3), 45-54.

Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: An interdisciplinary review. *MIS Quarterly*, *35*(4), 989-1015.

Steinfeld, N. (2016). "I agree to the terms and conditions": (How) do users read privacy policies online? An eye-tracking experiment. *Computers in Human Behavior*, *55*, 992-1000.

Sweeney, L., Abu, A., & Winn, J. (2013). *Identifying participants in the personal genome project by name (a re-identification experiment)*. Retrieved from https://ssrn.com/abstract=2257732

Teal, C. R., Gill, A. C., Green, A. R., & Crandall, S. (2012). Helping medical learners recognise and manage unconscious bias toward certain patient groups. *Medical Education*, *46*(1), 80-88.

Thomson, R., Yuki, M., & Ito, N. (2015). A socio-ecological approach to national differences in online privacy concern: The role of relational mobility and trust. *Computers in Human Behavior*, *51*, 285-292.

Tobin, A., Reed, D., Windley, P. J., & Foundation, S. (2017). The inevitable rise of self-sovereign identity (white paper). *Sovrin Foundation*. Retrieved from https://sovrin.org/wp-content/uploads/2018/03/The-Inevitable-Rise-of-Self-Sovereign-Identity.pdf

Wisniewski, P. J., Knijnenburg, B. P., & Lipford, H. R. (2017). Making privacy personal: Profiling social network users to inform privacy education and nudging. *International Journal of Human-Computer Studies*, *98*, 95-108.

Zolbanin, H. M., Delen, D., & Hassan Zadeh, A. (2015). Predicting overall survivability in comorbidity of cancers: A data mining approach. *Decision Support Systems*, *74*, 150-161.

About the Authors

Jacob Young is an Assistant Professor of Management Information Systems in the Foster College of Business and the Director of the Center for Cybersecurity at Bradley University. He earned his D.B.A. in Computer Information Systems from Louisiana Tech University. Dr. Young conducts research on privacy, security, and anonymity issues related to information systems with a primary focus on anonymous whistleblowing systems. He serves as the Senior Advisor on Cybersecurity at the National Whistleblower Center in Washington, D.C. His work has been published in *AIS Transactions on Human-Computer Interaction*, the *Journal of the Midwest Association for Information Systems*, the *DePaul Business & Commercial Law Journal*, and other journals and conference proceedings.

David Biros is an Associate Professor of Management Science and Information Systems and Fleming Chair of Information Technology Management at Oklahoma State University. A retired Lieutenant Colonel of the United States Air Force, Dr. Biros' last assignment was as Chief, Information Assurance Officer for the AF-CIO. His research interests included deception detection, insider threat, and information system trust. His work has been published *MIS Quarterly, the Journal of Management Information Systems, Decision Support Systems, Group Decision and Negotiation, MISQ Executive, the Journal of Information Systems Education, the Journal of the Midwest Association for Information Systems, the Journal of Digital Forensics Security and Law* and other journals and conference proceedings.

Ryan Schuetzler is an Assistant Professor in the Department of Information Systems and Quantitative Analysis at the University of Nebraska at Omaha. He completed his Ph.D. at the University of Arizona, where he completed his dissertation on the effects of automated conversational agents on human behavior and attitudes in interviews. His primary research focuses on human-computer interaction with conversational agents. Other research interests include interpersonal deception, collaboration, and behavioral analysis. Ryan serves as president of the Midwest chapter of the Association for Information Systems. His research has been published in the *Journal of Management Information Systems*, *Decision Support Systems*, *Computers in Human Behavior*, *Group Decision & Negotiation*, and others, as well as numerous conferences.

Tyler Smith is an Assistant Professor of Entrepreneurship, Technology, and Law in the Foster College of Business at Bradley University. He earned his J.D. from Indiana University Robert H. McKinney School of Law and his LL.M. from Notre Dame Law School. He conducts legal research on information privacy, constitutional law, and labor-management relations. His work has been published in several journals, such as the *Fordham International Law Journal*, the *New York International Law Review*, and the *Indiana International & Comparative Law Review*.

Paul Stephens is an Associate Professor in the Department of Entrepreneurship, Technology, & Law in the Foster College of Business at Bradley University. His career in academia began after 10+ years of experience in various engineering and manufacturing management positions in industry. He has taught in a variety of areas including technology entrepreneurship, business intelligence/data analytics and operations management. He has many research interests that include technology self-efficacy, information systems ethics, quality management and improving pedagogy with technology. His work has been published in *Decision Sciences*, the *Journal of Business Ethics*, the *International Journal of Information Systems in the Service Sector* and other journals and conference proceedings. He has presented his research at regional, national and international conferences. He earned his Ph.D. in Operations Management from the University of Cincinnati.

Rhonda Syler is an Associate Director of Enterprise Systems and faculty in Information Systems in the Sam M. Walton College of Business at the University of Arkansas and received her PhD from Auburn University. Dr. Syler's experience includes co-developing an Internet of Things Lab at Louisiana Tech University designed to provide a learning and innovation environment for big data, security, ERP, mobile development, and cloud computing. Her analytic experiences include working on data- and analytic-driven curriculum projects with Fortune 500 companies and conducting workshops across the globe. She teaches in the business analytics and ERP spaces. Her research focuses on the organizational impact and business value of disruptive technologies such as the Internet of Things, blockchain, and 5G particularly in terms of the value of data analysis and implications of security challenges.

Haoran (Shawn) Zheng is an Assistant Professor of Management Information System in the Foster College of Business at Bradley University. He earned his Ph.D. in Information Systems and Business Analytics from Chapman Graduate School at Florida International University. Dr. Zheng's conducts research in adoption, integration, and assimilation of e-health systems, organizational changes and modern privacy issues related to big data and analytics.