

Information and Reality: Contributions from the Science and Religion Forum

with Finley I. Lawson, “The Science and Religion Forum Discuss Information and Reality: Questions for Religions and Science”; Niels Henrik Gregersen, “The God with Clay: The Idea of Deep Incarnation and the Informational Universe”; Michael Burdett and King-Ho Leung, “The Machine in the Ghost: Transhumanism and the Ontology of Information”; Marius Dorobantu and Fraser Watts, “Spiritual Intelligence: Processing Different Information or Processing Information Differently?”; Matthew Kuan Johnson and Rachel Siow Robertson, “A Co-Liberatory Framework for Big Data”; Peter M. Phillips, “Digital Theology and a Potential Theological Approach to a Metaphysics of Information”; and Andrew Jackson, “Peacocke Prize Essay—Towards an Eastern Orthodox Contemplation of Evolution: Maximus the Confessor’s Vision of the Phylogenetic Logoi.”

A CO-LIBERATORY FRAMEWORK FOR BIG DATA

by Matthew Kuan Johnson  and Rachel Siow Robertson 

Abstract. This article provides an account of the ethical issues that arise when digital technologies and online spaces are structured by Big Data algorithms. We show that although the uses of Big Data may be new, traditional theological and ethical categories are still applicable, including “the sins of the fathers” from hamartiology and the scholastic concept of haecceity. Using these resources, we map the overall ecosystem in which digital technologies are developed and used, identifying the relationships between the individuals and organizations involved. We show how these relationships are often characterized by various kinds of harm. We then argue for a reorientation of the tech ecosystem toward co-liberation—and ultimately, joy.

Keywords: AI ethics; big data; co-liberation; design; haecceity; hamartiology; joy; sins of the fathers; virtue theory

INTRODUCTION: THE PARADOXES OF BIG DATA

Despite early optimism about the digital age (e.g., Barlow 1996), recent discussions have grappled with the “paradoxes” (Richards and King 2013) that come with the large-scale collection and processing of personal data: for every positive, there is a shadow side. The following are some examples

Matthew Johnson is a researcher at the University of Oxford, Oxford, UK; e-mail: matthew.johnson@aya.yale.edu Rachel Robertson is an Assistant Professor at Hong Kong Baptist University, Hong Kong; e-mail: rsrobertson@hkbu.edu.hk

[*Zygon*, vol. 58, no. 3 (September 2023)]

www.wileyonlinelibrary.com/journal/zygon

© 2023 The Authors. *Zygon*® published by Wiley Periodicals LLC on behalf of Joint Publication Board of *Zygon*. ISSN 0591-2385
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

of these “paradoxes.” Big Data can be used to facilitate preference-satisfaction when we search for social connections, information, products, and entertainment (Ricci, Rokach and Shapira 2011), but can also steer us toward insular, addictive, or risky interactions and can leave us vulnerable to having our behavior and beliefs manipulated—whether deliberately by companies (Leonard 2013) or unintentionally as a result of algorithmic bias (Rainie and Anderson 2017). Big Data fuels efficient financial transactions and services (Carrière-Swallow and Haksar 2019), as well as advances in medical research and healthcare (Mittelstadt, Fairweather and Shaw 2014). However, it can also drive a ruthless pursuit of productivity at the cost of wellbeing (especially in the workplace), as well as dangers of privacy breaches, lack of consent, and discriminatory “penalization” when it comes to major decisions on jobs, loans, criminal sentencing, and insurance (ALLEA, FEAM and EASAC 2021).¹

We are two philosophers who want to live well as our daily routines become increasingly bound up in Big Data structures. We feel excited about the many opportunities supported by the use of Big Data, such as medical advancements and increased social connections. However, we struggle with being inundated by misinformation and malicious actors whenever we enter spaces shaped by Big Data. We feel especially prone to bad habits fueled by Big Data, such as compulsively checking social media feeds and “doomscrolling” through constantly updating bad news. Therefore, these paradoxes are personal for both authors as end-users.

Although the paradoxes that the use of Big Data involves us in are novel, we contend that virtue theoretical and theological concepts can still be applied, allowing us to understand and illuminate this new context according to a typical ethical and theological landscape of vice and virtue, and sin and sanctification. Using these resources, we will argue for a co-liberatory framework for Big Data: one which recognizes the ways in which our commitments, behavior, and characters have become bound up in others’ commitments, behavior, and characters in such spaces—and in this way, oppressive structures harm all involved (including those in ostensibly “privileged” positions of power). Under our co-liberatory framework, any solutions to the paradoxes of Big Data must be formulated with a view to helping all people to escape from oppressive cycles of vice—not only those who are most visibly oppressed by current uses of Big Data, but also the technologies, structures, institutions, and individuals responsible for the oppression themselves.

A key concept within our framework is joy—a concept which characterizes co-liberation. We hope that our framework will help not only members of faith communities but also those from nonreligious groups to work toward joy in their communities through the use and regulation of Big Data, whatever their place within the tech ecosystem.

As a brief roadmap of our article: in Section 2, we argue that relationships within spaces shaped by Big Data are often characterized by harm, and that viewing the current situation through the lens of hamartiology (models of sin) helps to show that we are all seriously implicated. In Section 3, we call for a more holistic “ecosystems” approach within tech ethics, with an emphasis on co-liberatory solutions. In Section 4, we suggest how co-liberation can happen by reorienting the tech ecosystem around thicker forms of engagement characterized by joy.

HARM AND HAMARTIOLOGY

How can Big Data go wrong, and who is responsible? To explain our focus, we are interested in Big Data algorithms, which can be understood as prediction machines, which are fed large amounts of data about different variables, in order to find patterns among these relationships. Using these patterns, they create models that can make predictions. As an example, say you want to decide which candidates to admit into university. You feed in data on past students and the algorithm finds out which variables correlate with academic success at university—for example, high exam scores in school. You could then input candidates’ exam scores and the algorithm will tell you whether to accept them. Of course, Big Data algorithms take in much more data about many more variables, so the model probably will not be representable in 2D or 3D space, but the basic task is similar. At the moment, these algorithms form the basis of emerging digital technologies and now-familiar terms such as machine learning and Artificial Intelligence. They underlie many design and User Experience (UX) approaches to digital technologies, which cluster around “user-engagement” frameworks in which massive sets of personal data are combined with algorithms that use previous patterns of user behavior to predict how to keep us clicking, looking, swiping, and buying.

In this section, we will explore how harms can result from the use of Big Data, employing an external/internal distinction which we draw from Lisa Tessman (2005). External harms involve the imposition of adverse circumstances on people, such as introducing hardships or removing resources or opportunities; internal harms damage moral character, impeding intellectual or moral formation. After outlining these harms, we will use the resources of hamartiology to suggest that we are all implicated in such harms.

External Harms: Imposing Adverse Conditions

Let us start with external harms. There is plenty of excellent work on the ways in which Big Data has been used to impose adverse conditions, so we will only provide an overview in this section (we are particularly indebted

to the work done by Noble 2018; Benjamin 2019; and D'Ignazio and Klein 2020).

The first harm has to do with explainability. The issue here is that the more complicated the algorithm becomes as it is fed with variables and data, the less we will be able to understand or explain how it arrives at the answer—even if it is the right answer. To understand the potential harms involved, consider the example of job application screening. In 2016, it was reported that over 70% of applications are never seen by human eyes in the first round of the job selection process, but are instead fed into algorithms which predict the suitability of candidates (Mann and O'Neil 2016)—this percentage has likely increased in the years since this study. This means that a recruiter may not be able to provide an applicant with an explanation for why they were not progressed past the first stage in their application, since the algorithm that rejected their application may be too complicated to provide a human-understandable explanation for how it arrived at its decision. The lack of explainability is one way in which harms are done through the use of Big Data, as reasons may not be available for some of the most influential decisions in people's lives.

Explainability closely relates to the issue of algorithmic bias: if the training data is biased, the algorithm itself will be biased. In other words, the principle of “garbage in, garbage out” applies. We want to highlight two areas of concern when it comes to algorithmic bias: biased decision-making and biased information-curation.

The first area of concern is decision-making: because of the data that they are fed, Big Data algorithms can become biased toward certain variables as predictors of success. For example, they may be biased toward the variables of “white” and “male” for job candidates, because “successful” candidates who were hired in the past in many companies were, for historical and structural reasons, predominantly white and male. This bias is hard to screen out, as other markers such as educational background can represent race and gender by proxy. The algorithm might be trained to ignore “white” or “male,” but still weigh Ivy League educations highly in its calculations, which will function as a proxy for those markers (as those who attend Ivy Leagues have been disproportionately white and male—recall that Princeton and Yale only admitted women starting in 1969). Harms as a result of biased decision-making can be widespread, as algorithms can replicate and perpetuate biases in major decisions about job applications, housing availability, credit ratings, insurance pricing, child custody, bail, and criminal sentencing.

The second area of concern when it comes to algorithmic bias is information curation. Negative, salacious, or false content is correlated with greater user engagement (Kaiser and Rauchfleisch 2018; Ribeiro, Ottoni and West 2020; Spinelli and Crovella 2020). The previous choices of end-users to consume such content can therefore indirectly bias data sets,

thereby training algorithms aimed at promoting user engagement to provide more false or outrageous content. Moreover, there are plenty of malicious actors online (e.g., bots), so there are many direct attempts to bias data sets by engaging with and creating false content. Problems then arise when Big Data algorithms are relied upon to inform narratives, as with the promotion of content at the top of search results and newsfeeds. There are risks of perpetuating misrepresentations and contributing to a form of epistemic injustice known as “hermeneutical injustice” (Fricker 2007), which is injustice that happens when significant aspects of one’s experience are obscured from understanding due to prejudicial or oppressive norms, structures, patterns of behavior, or epistemic resources. One striking example in the literature was the discrepancy between Google Image search results of mugshots for the query: “three Black teenagers,” and wholesome portraits for the query: “three white teenagers” (Noble 2018). As this example suggests, algorithms reflect the data they have been trained on, which makes them largely a reflection of the society in which they are situated. This runs the risk of further engraining the biases of the society of which these algorithms are a part.

Another closely related issue is over-surveillance, where harm is done through excessive and inappropriate data collection. To give “informed consent” in accepting terms and conditions for data collection, people need to understand what they are agreeing to. However, as expectations of privacy are often misplaced, violations of privacy and consent are common. This was brought out in a highly publicized case where the store Target used individual purchase history data in combination with a pregnancy prediction model, to predict a teenager’s pregnancy—consequently sending her pregnancy-related targeted advertising before she had even told her own father (Noble 2018). Data collection is also often disproportionate for certain groups in society, especially the most vulnerable—as this example of targeting a pregnant teenaged girl suggests. This means that over-surveillance feeds back into the problem of bias. For example, certain neighborhoods, such as those with more people of color, end up being overly scrutinized when algorithms inform policing decisions (Benjamin 2019). This links to the issue of heightened security risks—data breaches will impact more on those vulnerable groups who have more data at stake.

A further harm done through the use of Big Data is in the imposition of adverse working conditions. Companies have leveraged Big Data algorithms toward a pursuit of efficiency at the cost of wellbeing, as workers are subjected to surveillance and judged according to the resulting data. For example, in the course of perfecting near-instantaneous delivery, Amazon fulfillment center and delivery workers pay the cost by being at increased risk of miscarriages due to their working conditions (Gurley 2021) or being “forced to urinate in plastic bottles because they cannot go to toilet on shift” (Drury 2019).

There are also environmental costs associated with the physical presence of data and the training of Big Data algorithms, which use a tremendous amount of electricity and contribute to the acceleration of the rate of climate change. Such climate change impact is disproportionately borne by those in the Developing World (Dhar 2020).

Lastly, all these harms interact with the “technological halo effect”: the tendency to trust algorithms because we think they are impartial, and a more reliable guide than humans (Benjamin 2019). This perceived trustworthiness can further reinforce the biased conditions under which algorithms are trained and deployed, contributing to hermeneutical injustices as narratives shaped by algorithms are believed over stories told by people about their own experience. The technological halo effect also allows people to escape responsibility. It is easy to blame the algorithm or “cold, hard math” when harm is done in the above ways.

Internal Harm: Characterological Damage

We have discussed harms which involve the imposition of adverse circumstances, inhibiting opportunities to access and own resources, explanations, and narratives. We turn now to a second kind of harm: the internal, characterological harms sustained by end-users and technologists through the use of Big Data.

End-Users. We have covered characterological harms done to end-users of digital technologies in detail in previous work (Robertson and Johnson 2023)—consequently, what follows is a summary of the issues we identified. Our claim is that end-users cannot navigate their way around many spaces shaped by Big Data algorithms, and the associated external harms, without harm also being done to their character. Specifically, we argue that Big Data structures can introduce constraints on three kinds of integrity: epistemic integrity (receptivity to the way things are), self-efficacy (capacities to align actions with desires and commitments), and self-unity (the internal integration of beliefs, desires, commitments, and identities).

Constraints on epistemic integrity hinder receptivity to the way the world is and the responsibilities placed upon us by the world. These constraints come about as end-users interact in spaces shaped by algorithmic bias. It is difficult to get beyond the vast quantities of false or otherwise misleading information promoted at the top of newsfeeds and search results, especially if we do not have the time or money needed to access higher quality information. Inundated with such content, users may be led to change their beliefs despite their original convictions and even their intellectual virtues. For example, users who search for the keyword “Asian” are more likely to be exposed to unwanted pornographic content that depicts Asians in hyper-sexualized terms, due to the previous behaviors of

other users who deliberately searched for or consumed this kind of content. Users may then come to unjustly misrepresent Asians and pass these schemas on to others—that is, they may contribute to hermeneutical injustice in society more broadly (Fricker 2007; on the causes and societal effects of viewing Asian women in hyper-sexualized terms, see Woan 2008; Kim 2021).

There is also a reinforcing effect with epistemic constraints, as each user's choices feed Big Data sets on which further algorithms are trained. This reinforcing dynamic contributes to “nudging,” where users are led into increasingly fringe communities online, and “filter bubbles,” where online groups form and ossify around certain beliefs and values. Filter bubbles can introduce further constraints on epistemic integrity, as members can be pressured to accept in-group beliefs and values, whilst setting unreasonable evidential standards for any claims made by out-group members—in effect, ignoring or downplaying out-group members' concerns.

A similar issue occurs for self-efficacy, where Big Data algorithms introduce constraints on capacities to align one's actions with one's desires and commitments. Instead of beliefs being at stake, here the issue is limitations on action. For example, Big Data algorithms will promote content and communities which are most likely to keep a user engaged. A user who is committed to avoiding certain vices in their past, such as a gambling addict, could repeatedly encounter content and communities encouraging those vices. In turn, data collected about this user's behavior will likely be used to promote harmful content and communities to others who are similarly vulnerable.

Big Data structures also introduce constraints on self-unity, disrupting the internal integration of beliefs, desires, commitments, and identities. As the above examples suggest, users can be pressured into displaying certain intellectual or moral vices that are in tension with their (offline) virtues, especially where “nudging” and “filter bubbles” occur. Another threat to self-unity comes when activities are subjected to data collection and judged accordingly. This may cause motivations to shift to achieving rewards which are limited to what can be collected and measured. In the case of the Amazon warehouse workers, the threat of losing their jobs leads to pursuing a data-driven productivity, even at the cost of wellbeing. Another example is someone who uses a wearable device to monitor their exercise because they want to get fit, but ends up prioritizing running at the cost of a more well-rounded exercise routine, as that is the only activity that their device is capable of recording and rewarding (Nguyen 2020).

Tech Leaders and Technologists. We have focused so far on harms done to end-users, but what about those who collect and use Big Data? Insofar as technologists are typically end-users as well, they face the harms

described above, but there is a further way they can experience characterological harm through their engagement with Big Data.

Data analysts, UX researchers, advertisers, and software engineers spend their time scraping data, purchasing data from third party data brokers, and running “A/B” tests to figure out how to keep us clicking, liking, looking, and buying. These activities encourage them to view end-users as clusters of data points to be scraped, bought and sold, and then manipulated for profit. This includes using data to make decisions and to provide tailored advertisements to persuade users to buy products and ascribe to certain beliefs. This highly reductive viewpoint extends beyond end-users—the example of the exploitation of Amazon employees shows that even others within tech companies come to be viewed as data points. An outcome of all this is a dehumanized view of human beings, who are reduced to clusters of data points to be fed into algorithms which impact on some of the most important parts of their lives.

We argue that this reductive, dehumanizing viewpoint is a failure of “inner virtue” (Bommarito 2018)—most notably of “loving attention,” which the philosopher and novelist Iris Murdoch describes as “a just and loving gaze directed upon an individual reality,” which is “the characteristic and proper mark of the active moral agent” (1970, 34). This failure of loving attention mirrors the characterological harms done to end-users because it is also a failure of integrity. In particular, technologists can—in this way—fail to recognize the moral responsibilities that they bear toward the intentional objects of their gaze. In this regard, “loving attention” can be seen as a form of “recognition respect” (Darwall 1977) or “consideration respect” (Frankena 1986; Cranor 1983). This constraint on moral formation may also be a reason why technologists often overlook the ways in which their technologies affect end-users, such as the spread of misinformation or even the environmental impact of Big Data algorithms. We will unpack more of the implications of “loving attention” shortly, but our point for now is that technologists risk doing characterological harm to themselves if their technologies inflict harm on end-users in ways that they overlook—and the technologists are perceptually formed to overlook these consequences for end-users.

HAMARTIOLOGY

So far, we have built a picture of the extensive internal and external harms that can occur in the course of using Big Data algorithms. We now want to consider the responsibilities of those who are involved in the use of Big Data. The tendency may be to blame the algorithm, in light of its perceived authority (as discussed with the technological halo effect). However, in the above sections, one repeated theme is that human action and decision-making are central to Big Data. We have mentioned

the deliberate activities of technologists as they design and implement Big Data algorithms. We have also mentioned the activities of end-users, who contribute to the “garbage in” side of the “garbage in, garbage out” dynamic, as their behaviors are collected into data sets to further train algorithms. In this section, we show how theological categories used to explore sin can help us to think through where responsibilities really lie.

Advantage over Justice

According to the Judeo-Christian tradition, the human tendency to try to blame-shift and escape responsibility runs up through the beginning of human history, starting in Genesis 3, where the man blames the woman for giving him the fruit from the tree, and the woman blames the serpent. The scholastic theologian Duns Scotus, in his theological anthropology, builds on Anselm’s work on the will and posits that we have two basic affections: the affection for advantage (*affectio commodi*) and the affection for justice (*affectio iustitiae*) (Hare 2001, 55–59). The affection for advantage is a tendency toward one’s own happiness, whereas the affection for justice is a love of the intrinsic goodness of things for their own sake. Where the affection for advantage does not compete with the affection for justice, it is good—but when the two are in competition, it is sinful to rank the affection for advantage over the affection for justice. Human sinfulness consists in our having a kind of “default setting” or innate tendency to consistently rank the affection for advantage over the affection for justice.

Our temptation with the technological halo effect is to offload decisions onto the algorithm, along with responsibility for any harmful outcomes of those decisions. As we have a disordered tendency to rank the affection for advantage over the affection for justice, algorithms pose a very real challenge to human agency and responsibility, because of the kind of escape route they provide. Indeed, an updated retelling of Genesis 3 might be “The algorithm which you gave to me told me to do it, and I ate.”

Failure to Recognize Haecceity

We can also add a theological understanding to the failure of epistemic integrity on the part of both end-users and technologists. We have suggested that in using people’s data in the exploitative ways that they often do, technologists can fail to pay “loving attention” to those around them. We have also suggested that end-users in “filter bubbles” can place higher evidential standards on out-group members, thereby ignoring their concerns. Both cases constitute a failure to perceive others in the right way.

Returning to Murdoch’s notion of “loving attention,” a helpful expansion comes from adapting the scholastic concept of *haecceity*, which is the property that individuates us and makes us unique. In its theological form, *haecceity* involves an appropriate recognition of how a

particular individual is loved by God in all of their uniqueness (Hare 2015, 145–46). Indeed, the philosophical theologian, John E. Hare (2001, 77) associates this unique-making or individuating property of haecceity with the promise in Revelation that God will give each person a white stone upon which is written “a new name that no one knows except the one who receives it” (Revelation 2:17, NRSV). Since “God’s call to us is to grow into this individual character” (Hare 2001, 77), “loving attention” is the capacity by which we recognize how another individual is loved in all of their uniqueness by God, and how they are coming to enter more fully into this uniqueness.

Returning to the use of Big Data with these hamartiological resources, one might say that ignoring the concerns of outgroup members (in the case of end-users) or treating people as mere data points (in the case of technologists) are not just instances of mere oversight. They represent a failure to perceive others as God perceives them and intends for us to perceive them, and so come under a kind of failure of “joint attention with God,” which is constitutive of sin (Stump 2012).

The Sins of the Fathers

Finally, the “sins of the fathers” concept (Milgrom 2001; Krašovec 1994) provides a model of how we are all bound together, end-users and technologists alike, such that our “sin” promotes online structures that encourage others to sin.

The “sins of the fathers” model suggests that the sinful decisions of previous persons can incline present individuals toward repeating the same vicious behavior. To clarify, a seeming tension in Hebraic hamartiology is that at times it appears that the “sins of the fathers” refers to how children are punished for the iniquity of the parents (e.g., Exodus 20:5; Ex.34:7), while at other times it is written that “A child shall not suffer for the iniquity of a parent ... the wickedness of the wicked shall be his own” (Ezek. 18:20). The apparent tension is often solved by pointing to the inescapable nature of the cycles of generational sin in question, such that the children suffer for their own freely committed iniquity, but the cycles of sin trap them such that they will ultimately repeat the sins of their parents (Milgrom 2001, 461; Krašovec 1994). Indeed, other Hebraic texts express how the children are only punished for the sins of the fathers “if they hold the deeds of their fathers in their hands” (*b. Sanh.* 27b), and the *Sipra* qualifies the verse “Our fathers sinned and are no more; and we must bear their guilt” (Lamentations 5:7, NRSV) with “whenever they adhere to their fathers’ deeds” (Behuqotay 8:2; see also Milgrom 2001, 461). Robert Alter comments on the underlying logic of the hamartiological concept of the sins of the fathers by explaining that “it is often the way of the world for sons to follow the path of their fathers” (Alter 2018, 568). In other

words, the sinful behavior of the ancestors sets up structures and schemas that constrain the behavior of subsequent generations, nudging them to replicate this sinful behavior.

With algorithmic bias, there is an inevitability to the cycles we are trapped in: if previous users clicked on biased or harmful sources of information, and these get ranked at the top of our feeds, we will ultimately (freely) click on those sources, further perpetuating the cycle for subsequent users. The biases and behavior of subsequent users are then captured in data sets and used to train further algorithms, showing the inescapability of these cycles.

“The sins of the fathers” concept reveals how responsibility and moral constraints come together. This model also suggests that individual moral agency is constrained in ways that cannot be escaped by current machine learning approaches—the “patterns of sin” simply become more engrained. Indeed, an updated rephrasing of the Hebraic proverb, “The parents have eaten sour grapes, and the children’s teeth are set on edge” (c.f. Jer.31:29; Ezek.18:2) may be, “The parents have clicked on fake news, and the children’s Google results are set on edge.”

In sum, although our tendency is to avoid blame for the harms done through the use of Big Data, our involvement at any level (whether as user or technologist) constitutes us freely entering into cycles of sin which can serve to entrap us, and others along with us.

A CO-LIBERATORY FRAMEWORK

Having now established our catalogue of internal and external harms and a hamartiological diagnosis, we want to set our approach alongside current ethical approaches to Big Data, challenging and expanding upon the usual frameworks.

Major approaches at the moment include computing design approaches, which seek technical solutions to ethical issues (e.g., solving the “black box” problem of machine learning algorithms by making them more “transparent”). Consequentialist or deontological normative ethical approaches articulate moral principles or rules, using concepts such as outcomes and impacts, rights, fairness, accountability, and transparency, with the aim of regulating technologies. Virtue theoretical approaches articulate the moral traits and capacities which will help us to use technology well—for an influential virtue theoretical approach and survey of other approaches, see Shannon Vallor’s *Technology and the Virtues* (2016).

Our discussion thus far suggests that the use of Big Data involves all of us in a complicated ecosystem, such that our commitments, behavior, and character are bound up in one others’ commitments, behavior, and character. We call this an “ecosystems” approach: one that maps all of the relationships at work in the development, deployment, and use of digital

technologies. This approach reveals how interconnected we are, and how currently these relationships are often primarily characterized by external and internal harm.

Returning to the existing approaches, we suggest that computing design needs to go further upstream to optimize for user integrity. Technical fixes will not solve the crucial issues if the whole ecosystem is structured in ways which promote harm. Our approach also accommodates a virtue theoretical frame alongside the deontological and consequentialist frames. Here is a brief sketch of how we see concepts of rights, fairness, accountability, and transparency within the virtue theoretical frame.

End-users have *rights* to true information and to be formed in virtuous ways (Watson 2021). It is therefore a violation of *fairness* if tech companies supply algorithms that hinder the moral and epistemic formation of end-users. It is also a violation of fairness for end-users to behave viciously online, as this constrains the behavior of other end-users, either directly (through sending vicious content directly) or indirectly (through biasing the data sets that train algorithms that constrain subsequent users). End-users and technologists must therefore be held *accountable* for their behavior in order to promote virtuous end-users. For end-users, this accountability can be to the government (e.g., laws regulating online hate speech), to the tech company that created the platform or product (e.g., community guidelines and content regulation), to other end-users (e.g., social censure by other end-users), or to algorithms (e.g., algorithms that automatically detect false content and suspend accounts). One way to hold technologists accountable is for their algorithms to be *transparent*—that is, they must honestly (and accessibly) disclose how they operate, at various levels of explanation depending on the expertise of the entity to which they are accountable.

Our approach also broadens the virtue theoretical frame in two ways. First, there are serious structural constraints on the development and exercise of virtue in end-users, which shape which virtues are appropriate to recommend for navigating the use of digital technologies. For example, we cannot simply recommend end-users to be more diligent on social media if newsfeeds currently do not reward diligent scrolling with true content. Second, despite these constraints on virtue, we want to emphasize that individuals are responsible for developing their own characters and working to resist the conditions of the ecosystem which constrain themselves and others. The extent of responsibility is not equal—this will depend on one's specific positionality within the ecosystem. Nevertheless, each member of the ecosystem still has responsibility for their part in it and their impact on other people.

To summarize our approach, we call for more than an account of the right virtues to develop or of the right normative ethical principles or rules to adopt. Full characterological formation and flourishing will not happen

for any individual or group within the tech ecosystem until it happens for all those involved. We therefore endorse a movement away from models involving dyads of perpetrators and victims (or of privileged and oppressed), and toward models of co-liberation (D'Ignazio and Klein 2020).

BUILDING FOR CO-LIBERATION, BUILDING FOR JOY

A co-liberatory framework approaches the collection and use of Big Data as the collaborative project of freeing ourselves and each other from oppressive structures in the tech ecosystem. But what are we freeing ourselves for, and what kind of use of Big Data are we aiming for instead? What does a co-liberatory approach look like in practice? In this last section, we propose that the aim should be to build for joy when using (or refusing to use) Big Data.

Joy is a concept which represents the antithesis of the current orientation of the most prominent Big Data structures. Instead of aiming for thin engagements characterized by external harm and burdens on the integrity of end-users and technologists, we should be building for joy: an intense feeling of fulfilment resulting from the recognition of integrity—a deep alignment between some good in the world and ourselves (Johnson 2020a; 2020b; Johnson and Robertson *in prep*). We could think of the joy of a “Eureka” moment of a scientific discovery, the creation of a beautiful work of art, or the celebration of a wedding. Joy as a theological category is connected with visions of human flourishing and typically involves normative elements, such as an appropriate recognition of being in the right kind of moral, spiritual, or metaphysical relationship with some significant intentional object (Johnson 2020a; 2020b). As a psychological category, joy is connected with building cognitive and social skills (Fredrickson 2004; 2009). Joy is therefore conceptually thicker than happiness or pleasure. In what remains, we will explore two theologically-informed recommendations for promoting joy in the course of pursuing co-liberation.

Pursue Radical Dispossession

Our first recommendation is to pursue radical dispossession, which involves being set free more generally from the control of harmful Big Data algorithms. Dispossession, in its theological form, works by relinquishing one's own desires that conflict with the desires God intends for them to have, and that conflict with the desires of those one is oppressing—instead placing one's attention upon God and others. Through this redistribution of attention through radical dispossession, one's desires for God and for the flourishing of others undergo a kind of intensification (Coakley 2013; 2015). In other words, before “loving attention” and a respect for the haecceity of others is possible, one must first lay the groundwork through which one can more fully turn one's attention to God and others.

Dispossession means laying down privileges and benefits and comes in many forms depending on one's place in the ecosystem.

Consider how this would work for end-users: some end-users will not use algorithms that make their lives easier (e.g., e-commerce recommender algorithms, which promote acquisitiveness), those with unhealthy relationships with technology (e.g., pornography or gambling addicts) will redouble their efforts to divest, and so on. Radical dispossession will not be easy, and may well involve developing “burdened virtues” that in some ways run counter to individual flourishing. “Burdened virtues,” drawing on work from Lisa Tessman (2005), are traits and skills which are *virtuous* in the sense that they enable resistance to constraints, but are *burdened* in the sense that they are not fully morally right (in an unqualified sense) and come at a cost. To provide a couple of examples, you could develop *burdened virtues of dishonesty and sensitivity* (for a deeper discussion of these examples, see Robertson and Johnson 2023). One example of such dishonesty is putting in false details whenever you are using online platforms. This helps to preserve your self-efficacy—you are holding back information according to your commitment to privacy, and you know that any true information entered in could be used to influence you (and others like you) in targeted advertising. However, this dishonesty challenges epistemic integrity, and it also challenges self-unity (as motivations and behaviors for self-efficacy and epistemic integrity are brought into conflict). Another example would be committing to being sensitive to the way things are outside of your own “filter bubble,” and to bearing this truth within your own group. This commitment supports epistemic integrity but could again challenge self-unity. Sensitivity to others' suffering is painful and may come with difficult realizations about yourself, and your standing in your own group may suffer as a result of referring to views outside of the group. In other words, resisting constraints in the course of pursuing dispossession may be hard work, and may involve doing characterological damage to oneself in other areas.

Technologists will have to let go of exploitative approaches to data. Computing design needs to go further upstream to optimize for user integrity, aiming at promoting and measuring *quality* rather than *quantity* of engagement and user satisfaction. Although initially this may seem to come at a financial cost to technologists, a potential upside is that the technologists who prioritize quality may be seen as more trustworthy by end-users, which could in turn contribute to financial benefit.

There may also be the possibility to adapt theological principles to move toward more radical forms of dispossession. One possibility is applying the logic of the Jubilee year. The Jubilee year is commanded in Leviticus, and happened every 49 years (or possibly every 50, according to certain interpretations)—occurring at the end of seven cycles of Sabbatical years (which themselves occurred every seven years). The Jubilee prescriptions primarily concern property and labor rights. In this year, slaves and

captives would be released, debts forgiven, and property returned, as described in the book of Leviticus: “You shall . . . proclaim liberty throughout the land to all its inhabitants. It shall be a Jubilee for you: you shall return, every one of you, to your property and every one of you to your family.” (Lev. 25:10) In other words, if someone “falls into difficulty” (Lev. 25:25) and sells a piece of property, or their home, or even themselves and are unable to redeem it or themselves—all of these things are returned in the year of Jubilee.

The application of Jubilee principles is especially salient in the tech ecosystem because technologists are predicting that with a move to the metaverse, some people’s full-time employment would involve spending time online and providing their personal data to companies. Consequently, personal data is very much viewed in economic terms. Applying Jubilee principles would involve personal data being returned, by being deleted. This would also allow us to break free of the digital structures that hold us captive, as it would hit the reset button on all of the ads that previously drew us into the kinds of behaviors we were trying to avoid. The gambling addict, rather than being bombarded with advertisements promoting online gambling, would be faced with different advertisements and have the chance to chart a way out. Of course, if they fall back into old patterns of online gambling, the advertisements they are faced with will then center upon gambling again, but this just shows the necessity of Jubilee years occurring regularly—we need second, third, fourth, *n*th chances.

We acknowledge that dispossession may be difficult. The burdened virtues involve characterological harm, and regular deletion of personal data could impact on carefully curated personal profiles and the efficiency of recommender systems. Nevertheless, in dispossession, joy is found, as Jonathan Tran argues (2021, 238): “Joy without dispossession is escapist. Dispossession without joy is sadist. The two together order the Christian life. At least the one offered by Jesus ‘who for the joy that was set before him endured the cross, despising its shame, and has sat down at the right hand of the throne of God’ (Heb. 12:2, WEB), and the one racial capitalism provokes.”

Of course, this form of radical dispossession is not easily reconcilable to the current global capitalist order in which the tech ecosystem is situated (to which Tran alludes in the quote above). Considerations of the bottom line fundamentally drive how technologies are designed and deployed. Still, even within this order, there is room for tech leaders and tech companies to practice radical dispossession. In fact, Tran (2021) provides one such example to support his account, through his ethnography of the California Bay Area tech company, Dayspring Partners. Dayspring charts an alternative path through the current global capitalist order, by focusing on relationships and the well-being of all involved stakeholders, rather than on maximizing financial margins for the company. Through significant investment in and service to their community, their financial

profitability arises from their prioritizing relationships and well-being over profits, as their approach increases consumer trust. Additionally, they enjoy other, nonfinancial goods, such as high employee well-being, strong relationships with their community, and joy through their practice of radical dispossession.

Finally, recapitulatory atonement models provide resources toward another way to escape harmful cycles and enter into the practice of radical dispossession. On these models, human nature fell through the iniquity of the first man and woman, which constrained human beings into committing subsequent iniquities. The fallenness of human nature is perpetuated and engrained through the sins of all subsequent human beings. Each person is free in any given moment to not sin; however, because they have a fallen human nature, there is no possible world in which they will be able to live their entire life having successfully chosen to not sin in every decision—they suffer from “transworld depravity” (Plantinga 1974). Recapitulatory atonement models suggest that Christ, by joining Himself to human nature and living a sinless life, restores the potential for human nature to live without sin. Through some degree of participation in Christ in the present, believers begin to take on this redeemed human nature and have a greater possibility of avoiding sin—although complete freedom from iniquity will not be achieved until there is full participation in redeemed human nature, in the eschaton.

A computer scientific analogue of restored or redeemed human nature could involve synthetic data: artificially generated data that represents fully virtuous online behavior. Synthetic data can be fed into data-sets in order to train out the bias from algorithms. With algorithms that contain less bias, users will be surrounded by higher quality information, which will in turn enable them to cultivate and exhibit a higher level of virtue online. Their online behavior will feed into subsequent training of the algorithms, which will enable virtuous behavior in further end-users, and so on. Instead of relying on gathering personal data and consigning users to cycles of algorithmic bias, synthetic data provides one pathway toward co-liberation. Indeed, if algorithms trained through this kind of synthetic data could “nudge” human behavior away from cycles of algorithmic bias and vice, and toward radical dispossession and virtue, they could contribute to healthier online lives, deeper human connection, and joy. Technologists and regulators will have to think carefully about what restored human nature looks like, however. This is an area in which dialogue with philosophers and scholars of religion would be particularly useful.

Pursue Embodied Relationships

Alongside pursuing radical dispossession, our second recommendation is to pursue embodied relationships. There are twin dangers with radical

dispossession: one danger would be to think of the call to dispossession as involving a kind of denial of the body and detachment from the world, including the tech ecosystem; another would be to embrace the move to the digital world by accepting increasing levels of abstraction and disembodiment, such that consequences for others and ourselves seem small. We are arguing against both of these views, for a form of radical asceticism that involves a move more deeply into our embodiment, and into an integration between online and offline life.

Embodied relationships with others make claims on you as a whole person—not just upon some of the offline parts. When you have a concrete, embodied individual in front of you, they make claims upon you as your friend, your child, your lover, your student, and so on. They do not just want a part of you, and they certainly do not just want some data points. In order to be fully known, there must be integration between all parts of you: the online and offline parts. Consequently, while the majority of the algorithms that structure our online lives view us in reductive and instrumentalized ways, our offline, embodied lives involve friends who aim to view us in all of our haecceity and uniqueness—or to put it theologically, as God views us. We would do well to remember who we are—not just in our offline lives, but in all parts of our lives, including online.

This principle of seeking embodied relationships can also inform the design of algorithms. Technologists need to recognize that algorithms, in one sense, perform similar activities to humans when they learn. Human beings gather data, develop models of the data that are used to predict events, and then adjust these models where those predictions were shown to be inaccurate. Algorithms are just able to accommodate, process, and predict much more data and much more quickly than we can. Theological approaches can inform the boundaries and directions for human learning: there are forms of human learning that are beneficial, such as medicine and art; there are also forms of human learning that are dangerous and forbidden, such as torture. When we approach algorithms, we should apply these same categories. Furthermore, there are better and worse ways of learning. For instance, if we were to treat our teachers in instrumentalizing and reductive terms, we might still learn the information we want, but there will be some kind of deep moral and pedagogical remainder, because it would not have been obtained in the right way. Imagine coming to class and taking no interest in the personal life of your teacher, and just “downloading” the information and leaving. There would be deep human goods, such as friendship, opportunities to exercise service, and even joy, that would be denied to both you and your teacher. There could also be character deficits that arise, as you could become a more self-centered, small-hearted person.

If technologists develop their algorithms through treating end-users in dehumanizing and objectifying ways, treating people as clusters of data points to be manipulated for their own ends, there are goods and character

traits that technologists deny themselves, and there are goods denied to the end-users in being treated this way. Algorithms must be designed with something like *haecceity* in mind: they need to be optimized to promote things that will help each individual flourish and grow into their individual character. In all this, we need humility about the limitations of algorithms, appreciating that there are some things that Big Data does not currently capture. We need to keep looking to lived experience and to those people whose narratives are usually left out or misrepresented.

In summary, the work of tech ethics will require being in the right kind of listening, caring relationships to concrete, affected individuals—and continually adjusting to the concerns of these individuals. It may not be practical to be in these kind of relationships to concrete, affected individuals—carefully listening to and adjusting to affected stakeholders breeds inefficiency, which is especially challenging for an industry driven by the dictum “Move Fast and Break Things.” Nevertheless, this reorientation is needed for the tech ecosystem to support joy and deeper human flourishing, for it is in a deep alignment within and between our whole selves, others, and the world that joy can be found.

CONCLUSION

As we draw to a close, we want to take the opportunity to reflect on the role of religious communities in particular. We have suggested that theological sources of wisdom can illuminate the issues raised by Big Data. We hope that religious communities can provide resources and spaces to encourage technologists to consider how their work impacts on human flourishing, and to be creative in reimagining uses of Big Data.

Another thing to emphasize is continuity between online and offline lives for pastoral care. Real harms and hurts and addictions happen online, and real care and repentance and reconciliation need to happen offline in order for healing to occur. The same goes for online community: making meetings or services available online to those who could not otherwise access them is a positive development, but this does not relieve us of further responsibilities in how we care for them. The pursuit of deep relationships needs to continue in offline life. In general, we hope that religious communities can see engagements with online contexts structured by Big Data algorithms as relevant to religious life, and as one of the major ways in which communities can live out their faith. There may be a place for developing liturgies and healthy habits to shape these engagements, such as digital Sabbaths where we take a break from technologies—making room for habituation into virtue rather than vice.

There are complicated structural problems posed by the use of Big Data, with widespread harms which fall under serious, hamartiological categories. We have argued that the solution cannot merely be an issue of better

regulation or the development of virtues, but requires fundamentally restructuring how technologies are designed and developed. We therefore recommend a co-liberatory framework for approaching the use of Big Data—an approach which is sensitive to each individual's place in the tech ecosystem and their responsibilities within it. As a starting point, taking this approach will look like radical dispossession—the dispossession of those who benefit from the current arrangement of the tech ecosystem, and the dispossession of various conveniences and apparent benefits by those oppressed by it. Moreover, we have argued that dispossession and building deep, embodied relationships are two sides of the same coin. In this way, dispossession leads to joy: as we more deeply find ourselves, each other, and technologies that more fully contribute to our flourishing.

ACKNOWLEDGMENTS

The authors would like to thank The Faraday Institute for Science and Religion and the NViTA project at the University of St Andrews for their support of this research project.

FUNDING

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The authors' work on this publication was made possible through the support of grants from the John Templeton Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

NOTE

1. This is based on a report resulting from a joint initiative between All European Academies (ALLEA), the European Academies' Science Advisory Council (EASAC), and the Federation of European Academies of Medicine (FEAM).

REFERENCES

- ALLEA, FEAM, and EASAC. 2021. "International Sharing of Personal Health Data for Research." ALLEA, EASAC and FEAM joint initiative on resolving the barriers of transferring public sector data outside the EU/EEA. <https://doi.org/10.26356/IHDT>.
- Alter, Robert. 2018. *The Hebrew Bible: A Translation with Commentary*. New York: W. W. Norton.
- Barlow, John Perry. 1996. "A Declaration of the Independence of Cyberspace." Electronic Frontier Foundation. 1996. <https://www.eff.org/cyberspace-independence>
- Benjamin, Ruha. 2019. *Race After Technology: Abolitionist Tools for the New Jim Code*. Cambridge: Polity Press.
- Bommarito, Nicolas. 2018. *Inner Virtue*. Oxford: Oxford University Press.
- Carrière-Swallow, Yan, and Vikram Haksar. 2019. "The Economics and Implications of Data: An Integrated Perspective." Departmental Paper No. 19/16. International Monetary Fund. Strategy, Policy, and Review Department.

- <https://www.imf.org/en/Publications/Departmental-Papers-Policy-Papers/Issues/2019/09/20/The-Economics-and-Implications-of-Data-An-Integrated-Perspective-48596>.
- Coakley, Sarah. 2013. *God, Sexuality, and the Self: An Essay "On the Trinity"*. Cambridge: Cambridge University Press.
- . 2015. *The New Asceticism*. London: Bloomsbury.
- Cranor, Carl F. 1983. "On Respecting Human Beings as Persons." *Journal of Value Inquiry* 17 (2): 103–17.
- Darwall, Stephen. 1977. "Two Kinds of Respect." *Ethics* 88 (1): 36–49.
- Dhar, Payal. 2020. "The Carbon Impact of Artificial Intelligence." *Nature Machine Intelligence* 2 (8): 423–25.
- D'Ignazio, Catherine, and Lauren Klein. 2020. *Data Feminism*. Cambridge, MA: MIT Press.
- Drury, Colin. 2019. "Amazon Workers 'Forced to Urinate in Plastic Bottles Because They Cannot Go to Toilet on Shift.'" *The Independent*, July 19, 2019, News. <https://www.independent.co.uk/news/uk/home-news/amazon-protests-workers-urinate-plastic-bottles-no-toilet-breaks-milton-keynes-jeff-bezos-a9012351.html>
- Frankena, William K. 1986. "The Ethics of Respect for Persons." *Philosophical Topics* 14 (2): 149–67.
- Fredrickson, Barbara L. 2004. "The Broaden-and-Build Theory of Positive Emotions." Edited by F. A. Huppert, N. Baylis, and B. Keverne. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 359(1449):. 1367–77.
- . 2009. "Joy." In *The Oxford Companion to Emotion and the Affective Sciences*, edited by D. Sander and K. Scherer, 230. Oxford: Oxford University Press.
- Fricker, Miranda. 2007. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press.
- Gurley, Lauren Kaori. 2021. "Amazon Denied a Worker Pregnancy Accommodations. Then She Miscarried." *Vice*, July 20, 2021. <https://www.vice.com/en/article/g5g8eq/amazon-denied-a-worker-pregnancy-accommodations-then-she-miscarried>
- Hare, John E. 2001. *God's Call: Moral Realism, God's Commands, & Human Autonomy*. Grand Rapids, MI: Wm. B. Eerdmans Publishing Co.
- . 2015. *God's Command*. Oxford: Oxford University Press.
- Johnson, Matthew Kuan. 2020a. "Joy: A Reply to the Replies." *The Journal of Positive Psychology* 15 (1): 84–88.
- . 2020b. "Joy: A Review of the Literature and Suggestions for Future Directions." *The Journal of Positive Psychology* 15 (1): 5–24.
- Johnson, Matthew Kuan, and Rachel Siow Robertson. in prep. "Cultivating Joy."
- Kaiser, Jonas, and Adrian Rauchfleisch. 2018. "Unite the Right? How YouTube's Recommendation Algorithm Connects the US Far-Right." *D&S Media Manipulation*.
- Kim, Grace Ji-Sun. 2021. *Invisible: Theology and the Experience of Asian American Women*. Minneapolis: Fortress Press.
- Krašovec, Jože. 1994. "Is There A Doctrine of 'Collective Retribution' in The Hebrew Bible?" *Hebrew Union College Annual* 65, 35–89.
- Leonard, Andrew. 2013. "How Netflix Is Turning Viewers into Puppets." Salon. February 1, 2013. https://www.salon.com/2013/02/01/how_netflix_is_turning_viewers_into_puppets/
- Mann, Gideon, and Cathy O'Neil. 2016. "Hiring Algorithms Are Not Neutral." *Harvard Business Review*, December 9, 2016. <https://hbr.org/2016/12/hiring-algorithms-are-not-neutral>
- Milgrom, Jacob. 2001. *Leviticus 23–27*. New Haven, CT: Yale University Press.
- Mittelstadt, Brent, Ben Fairweather, Mark Shaw, et al. 2014. "The Ethical Implications of Personal Health Monitoring." *International Journal of Technoethics (IJT)* 5 (2): 37–60.
- Murdoch, Iris. 1970. *The Sovereignty of Good*. London: Routledge & Kegan Paul.
- Nguyen, C. Thi. 2020. *Games: Agency as Art*. Oxford: Oxford University Press.
- Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press.
- Plantinga, Alvin. 1974. *The Nature of Necessity*. Oxford: Oxford University Press.
- Rainic, Lee, and Janna Anderson. 2017. "Code-Dependent: Pros and Cons of the Algorithm Age." Pew Research Centre. <https://www.pewresearch.org/internet/2017/02/08/code-dependent-pros-and-cons-of-the-algorithm-age/>

- Ribeiro, Manoel Horta, Raphael Ottoni, Robert West, et al. 2020. "Auditing Radicalization Pathways on YouTube." In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 131–41.
- Ricci, Francesco, Lior Rokach, and Bracha Shapira. 2011. "Introduction to Recommender Systems Handbook." In *Recommender Systems Handbook*, edited by Francesco Ricci, Lior Rokach, Bracha Shapira, and Paul B. Kantor, 1–35. Boston: Springer US.
- Richards, Neil M., and Jonathan H. King. 2013. "Three Paradoxes of Big Data." *Stan. L. Rev. Online* 66: 41.
- Robertson, Rachel Siow, and Matthew Kuan Johnson. 2023. "Moral Education in and for Virtual Spaces." In *Moral Education in the 21st Century*, edited by Douglas W. Yacek, Mark E. Jonas, and Kevin H. Gary. Cambridge: Cambridge University Press.
- Spinelli, Larissa, and Mark Crovella. 2020. "How YouTube Leads Privacy-Seeking Users Away from Reliable Information." In *Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization*, 244–51.
- Stump, Eleonore. 2012. *Wandering in Darkness: Narrative and the Problem of Suffering*. Oxford: Oxford University Press.
- Tessman, Lisa. 2005. *Burdened Virtues: Virtue Ethics for Liberatory Struggles*. Oxford: Oxford University Press.
- Tran, Jonathan. 2021. *Asian Americans and the Spirit of Racial Capitalism*. Oxford: Oxford University Press.
- Vallor, Shannon. 2016. *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting*. Oxford: Oxford University Press.
- Watson, Lani. 2021. *The Right to Know: Epistemic Rights and Why We Need Them*. London: Routledge.
- Woan, Sunny. 2008. "White Sexual Imperialism: A Theory of Asian Feminist Jurisprudence." *Washington and Lee Journal of Civil Rights and Social Justice* 14 (2): 275.