Data Ethics in Catholic Health Systems

Rachelle Barina, Becket Gremmels, Michael Miller, Nicholas Kockler, Mark Repenshek, Christopher Ostertag, Kirsten Dempsey

Abstract. The Catholic moral tradition has a rich foundation that applies broadly to encompass all areas of human experience. Yet, there is comparatively little in Catholic thought on the ethics of the collection and use of data, especially in healthcare. We provide here a brief overview of terminology, concepts, and applications of data in the context of healthcare, summarize relevant theological principles and themes (including the Vatican's *Rome Call for AI Ethics*), and offer key questions for ethicists and data managers to consider as they analyze ethical implications pertinent to data governance and data management. *National Catholic Bioethics Quarterly* 22.2 (Summer 2022): 289–317.

Rachelle Barina, PhD, HEC-C, is the chief mission officer for the Hospital Sisters Health System in Springfield, IL. Becket Gremmels, PhD, is the system vice president of theology and ethics at CommonSpirit Health in Dallas-Fort Worth, TX. Michael Miller, HEC-C, is the system vice president for mission and ethics for SSM Health in St. Louis, MO. Nicholas Kockler, PhD, is the vice president for system ethics services at the Providence Center for Health Care Ethics in Portland, OR. Mark Repenshek, PhD, is the vice president for ethics and church relations for Ascension Health in St. Louis, MO. Christopher Ostertag, PhDc, is the director of mission services for OSF HealthCare in Peoria, IL. Kirsten Dempsey, PhDc, is the manager of mission integration at SSM Health in St. Louis, MO.

> *When a technology is still at an early stage of development, it is still possible to influence the direction of its development, but we do not know yet how it will affect society. Yet, when the technology has become societally embedded, we do know its implications, but it is very difficult to influence its development.*
>
> —Olya Kudina and Peter-Paul Verbeck, "Ethics from Within"

During the last five decades, health information has been rapidly digitized. With the development and integration of the electronic medical record (EMR), health data is being gathered and stored on an unprecedented scale. The ubiquitous adoption of smartphones and other personal devices affords wide-reaching opportunities to collect and exchange personal information and biometrics. Governmental agencies and private industries are also amassing digital information, such as demographic data, social media activity, direct-to-consumer genetic test results, and consumer purchasing practices.

We are well into the age of big data, which is enlivened by the assumption that all this information can enable predictive, actionable insights. Drawing on the accumulation of health data, there are a wide range of technologies—including artificial intelligence (AI)—that can be employed to improve clinical care, support population health strategies, and reduce waste. In general, advancing these goals can promote respect for human persons and advance the common good, which are central to the mission and ministry of Catholic health systems; however, ethical challenges arise when leveraging big data and deploying AI tools, prompting the need for data ethics. *Data ethics* is the study of the nature and impact of big data and whether the corresponding use thereof promotes human dignity and the common good. Data ethics involves an assessment of a broad spectrum of actions related to collecting, storing, sharing, and using big data.

In hopes of managing and using data responsibly, we argue that Catholic health systems must invest in cross-disciplinary data governance structures that facilitate *robust ethics integration* based on a shared moral framework; engaging in such work will lead to new data ethics questions, which require significant collaboration across disciplines. To make this argument in support of a commitment to data ethics, we proceed in five parts. First, we illustrate some of the ways that big data and AI are and will continue impacting health care, as well as the ethical challenges of collecting and using data. Second, we explain the meaning of data governance and the importance of establishing guiding principles within governance. Third, we ground our work in the principles of respect for persons and the common good and then draw on the *Rome Call for AI Ethics* (*Rome Call*) to highlight secondary ethical principles of data governance for Catholic health systems. Fourth, we identify key ethical questions stemming from these principles. Finally, we discuss the role of ethicists and modalities of ethics integration in data governance structures within Catholic health systems. To our knowledge, this is the first paper that has

Epigraph. Olya Kudina and Peter-Paul Verbeek, "Ethics from Within: Google Glass, the Collingridge Dilemma, and the Mediated Value of Privacy," *Science, Technology, & Human Values*, 44.2 (2019): 292, doi: 10.1177/0162243918793711.

attempted a comprehensive look at ethics integration within data governance of Catholic health systems.[1]

The Promises (and Perils) of Big Data and AI for Health care

Defining Big Data and AI

Big data and *AI* are commonly used terms, but their meaning can be elusive. In plain terms, *big data* refers to large and complex datasets obtained from new data sources. The size and complexity of the data exceeds what can be handled by traditional processing software.[2] Instead, advanced computational techniques can be used to analyze the data, understand patterns, and test hypotheses that would probably be imperceptible to a human who undertakes manual research.[3] These massive datasets are larger, more dynamic, and more complex than one human mind can review.

AI refers to the ability of a machine to conduct autonomous computational action through algorithms. In the most basic sense, AI involves a machine looking for patterns in a set of big data and sorting data fields into categories based on features that are often so subtle they are effectively imperceptible to humans. A machine uses AI when it applies a rule or algorithm (a set of instructions that tells a computer how to process data) to data in order to draw a conclusion.[4] In this sense, big data is indispensable for AI, and AI is likewise an essential technology to process and understand big data effectively.

AI has evolved rapidly from straightforward early accomplishments to far more complex domains. Early examples now seem simple and undaunting, such as the algorithm deployed in the 1990s by the US Postal Service to identify letters and spaces, allowing a computer to convert handwritten addresses to a digital format. This is an early example of natural language processing, which is a branch of AI concerned with a computer's ability to understand and respond to human text

1. See Jean Baric-Parker and Emily E. Anderson, "Patient Data-Sharing for AI: Ethical Challenges, Catholic Solutions," *The Linacre Quarterly* 87.4 (May 15, 2020): 471–481, doi: 10.1177/0024363920922690. Anderson and Baric-Parker offered a robust initial paper that provides a foundation for the ethical challenges of AI in Catholic health systems. Our project builds on the exploration begun in their paper, with deeper exploration of relevant ethical principles across the entire data lifecycle and with practical elements of accomplishing ethics integration within the applied context of Catholic health systems.
2. Kord Davis, *Ethics of Big Data: Balancing Risk and Innovation* (Sebastopol, CA: O'Reilly Media, Inc., 2012), 4.
3. One recent review identifies seven key characteristics of big data: volume, variety, velocity, veracity, value, variability, and valence. Sayantan Khanra and Amandeep Dhir, "Big Data Analytics in Health Care: A Systematic Literature Review," *Enterprise Information Systems* 14.7 (October 2020): 878–912, doi: 10.1080/17517575.2020.1812005.
4. A helpful and accessible summary can be found in Bernard Marr, *Artificial Intelligence in Practice: How 50 Successful Companies Used AI and Machine Learning to Solve Problems* (United Kingdom: Wiley, 2019).

and speech like humans do.[5] Today, AI is ubiquitous: smartphone and computer applications can correctly suggest what word is typed or spoken. Map applications offer not only accurate directions but also real-time insight into traffic flows. Credit card and mortgage companies make conclusions about whether to lend money to an individual. Targeted ads include personalized selections of products presumed desirable.

All of these technologies depend on massive amounts of data that are being processed through algorithms, which have become increasingly sophisticated. In a subset of AI called *machine learning*, algorithms have the capacity to parse, learn, and apply data to make conclusions. With machine learning, computers are able to refine an algorithm without the explicit direction of a human agent, independently adapting statistical modeling and drawing inferences, far beyond the rules given to it by a human.[6]

During the last decade, *deep learning* has emerged as a subset within machine learning. Deep learning can be defined as "neural network algorithms that learn the important features in data by themselves" and are "able to adapt through repetitive training to uncover hidden patterns and insights."[7] *Neural networks* consist of multiple layers of computations that work reciprocally to identify complex patterns in large datasets. Rather than refining a single formula, neural networks can break down an algorithm into a multistep process to accomplish more powerful predictive modeling. Deep learning (i.e., neural networks with three or more layers) can be used in supervised (utilizes labeled datasets) and unsupervised (uses unlabeled datasets) machine learning.[8]

With growing sophistication, AI draws on the massive amounts of data being accumulated in order to identify trends that have happened in the past, to make accurate predictions about the future based on those data, and, therefore and most importantly, to make actionable conclusions about the present moment. To put it mildly, the vast amounts of data that health systems hold or can obtain combined with evolving AI technologies hold vast promise for health and health care.

The Promises of Big Data and AI for Health Care

The US health care system faces many challenges, including troubling variations in quality, persistent health disparities, escalating costs, staffing shortages, record burnout, and the COVID-19 pandemic. Amidst these challenges, AI holds immense potential to create and sustain resolutions to these deeply rooted challenges. "The promise of artificial intelligence in medicine," Eric Topol writes, "is to provide composite, panoramic views of individuals' medical data; to improve decision making; to avoid errors such as misdiagnosis and unnecessary procedures; to help in the

5. IBM Cloud Education, "What Is Natural Language Processing?" IBM, July 2, 2020, https://www.ibm.com/cloud/learn/natural-language-processing.
6. Rodrigo Ceron, "AI, Machine Learning and Deep Learning: What's the Difference?," IBM, December 5, 2019, esp. fig. 1, https://www.ibm.com/blogs/systems/ai-machine-learning-and-deep-learning-whats-the-difference/.
7. Ceron, "AI, Machine Learning and Deep Learning," fig. 1.
8. Michael Matheny et al., *Artificial Intelligence in Health Care: The Hope, the Hype, the Promise, the Peril* (Washington, DC: National Academy of Medicine, 2019), 126–127.

ordering and interpretation of appropriate tests; and to recommend treatment."[9] Leveraging many kinds of data (see table 1), AI has already begun to add value in health care, and new applications are rapidly being envisioned, developed and integrated. While the possibilities are vast, we summarize five key areas of application.[10]

Diagnostics: Given the ability to analyze vast quantities of images at painstaking levels of detail, AI can dramatically augment both the speed and accuracy of diagnostics. In gastroenterology, AI has already been used to distinguish polyp types during colonoscopy, a task that can be exceedingly difficult and time-consuming for physicians but which AI accomplished rapidly and with strong accuracy.[11] In radiology, a deep learning algorithm (relying on a 121-layer neural network) detected pneumonia in chest X-rays more accurately than the actual work of radiologists against which it was measured.[12] In cardiology, deep learning techniques have diagnosed heart attacks with precision equal to that of cardiologists.[13] In the quest for better patient care, these kinds of tools will aid in diagnostics across all specialties, contributing to both speed and accuracy of diagnoses.

Clinical Decision-Making: AI will also be deployed to aid physician judgment in providing personalized predictions about what interventions will be most effective for a given patient. For example, machine learning was retrospectively deployed on patient datasets to make recommendations about the use of vasopressors, IV fluids, or medications to treat patients with sepsis. The study demonstrated that on average, AI identified which intervention would be most effective better than the clinicians who had cared for the patients.[14] Rather than rely on the recommendation of a treating physician, the vast amounts of available data about a patient and detailed algorithms will help us know how various interventions will impact this particular patient. These kinds of AI tools will improve treatment outcomes and reduce failed interventions, which can be dissatisfying and costly.

9. Eric Topol, *Deep Medicine: How Artificial Intelligence Can Make Health Care Human Again* (New York: Basic Books, 2019), 9.
10. For an overall review of AI applications, see Eric J. Topol, "High-performance Medicine: The Convergence of Human and Artificial Intelligence," *Nature Medicine* 25.1 (2019): 44–56, doi: 10.1038/s41591-018-0300-7.
11. Yuichi Mori et al., "Real-Time Use of Artificial Intelligence in Identification of Diminutive Polyps During Colonoscopy: A Prospective Study," *Annals of Internal Medicine* 169.6 (September 18, 2018): 357–366, doi: 10.7326/M18-0249; Pu Wang et al., "Development and Validation of a Deep-learning Algorithm for the Detection of Polyps during Colonoscopy," *Nature Biomedical Engineering* 2.10 (October 2018): 741–748, doi: 10.1038/s41551-018-0301-3.
12. Xiaosong Wang et al., "ChestX-ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases," *IEEE CVPR* (May 2017): 2097–2106, doi: 10.1109/CVPR.2017.369.
13. Nils Strodthoff and Claas Strodthoff, "Detecting and Interpreting Myocardial Infarction Using Fully Convolutional Neural Networks," *Physiological Measurement* 40.1 (January 2019): 015001, doi:10.1088/1361-6579/aaf34d.
14. Matthieu Komorowski et al., "The Artificial Intelligence Clinician Learns Optimal Treatment Strategies for Sepsis in Intensive Care," *Nature Medicine* 24.11 (October 2018): 1716–1720, doi: 10.1038/s41591-018-0213-5.

Mental Health: As a health care service and a public health concern, mental health has been grossly underprioritized and poorly resourced within the United States. AI has great potential to identify mental health patterns and crises from personal data. Mining the digitized world of personal communications and social media, AI has accurately identified depression and suicidal ideation.[15] It has even been demonstrated that mobile phone metadata could be effectively analyzed in real time to make inferences about the existence and severity of mood disturbances.[16] The monitoring of metadata could allow us to actively intervene before mental health challenges reach a crisis level. Ongoing surveillance of personal metadata could also pinpoint individuals who are experiencing acute mental health crises or suicidal ideation and intervene at critical moments, even when those individuals are not asking for help. Other AI models for mental health are being considered around speech, voice, and facial expressions, with similar aspirations for early intervention.

Population Health Management: On the same trajectory as mental health, the US health care model has disproportionately invested in acute and sick care and inadequately incentivized investments in services that keep people healthy, manage chronic diseases, and promote mental health. In a positive turn, health systems, insurers, governments, and other stakeholders are increasingly aligning around *population health* efforts that emphasize preventing illness and proactively managing chronic disease. AI can significantly bolster advanced population health efforts, which are enhanced by pinpointing in real time those patients who are at greatest risk and those for whom intervention will make the most significant impact. With risk stratification insights, health systems can offer care coordination, programmatic services, or urgent intervention at clinically pivotal times. Such interventions might result in a slowing of disease progression in outpatients. Risk stratification in real time might also result in anticipation and urgent intervention before a major medical event, both in the hospital and with outpatients. The impact of AI within population health strategies will enhance patient health and help to control the overall costs associated with utilization of acute services and rising chronic diseases.

Advancing Community Health: Outside the walls of a hospital, there are many opportunities to utilize additional datasets to gain insights and develop community level interventions. AI systems can enhance traditional public health practices, like disease surveillance. For instance, social media can be used to monitor pandemics, respond to instances of food poisoning, and predict tuberculosis transmission.[17]

15. Chiaming Tung and Wenhsiang Lu, "Analyzing Depression Tendency of Web Posts Using An Event-driven Depression Tendency Warning Model," *Artificial Intelligence in Medicine* 66 (January 2016): 53–62, doi: 10.1016/j.artmed.2015.10.003.
16. Bokai Cao et al., "DeepMood: Modeling Mobile Phone Typing Dynamics for Mood Detection," in *KDD '17: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (New York: Association for Computing Machinery, 2017), 747–755.
17. Mohammed Ali Al-Garadi et al., "Using Online Social Networks to Track a Pandemic: A Systematic Review," *Journal of Biomedical Informatics* 62 (August 2016): 1–11, doi: 10.1016/j.jbi.2016.05.005; Samir Bhatt et al., "The Global Distribution and Burden of Dengue," *Nature* 496.7446 (2013): 504–507, doi: 10.1038/nature12060; Jenine K. Harris et al., "Using Twitter to Identify and Respond to Food Poisoning: The Food Safety

Table 1

Areas of Relevant Data for Health Systems

Structured EMR data	Structured data has consistency and organization relative to pre-defined fields. Examples include utilization data, weight, blood pressure, blood type, diagnosis, disease stage, phenotype, lab values, genetics and genomics, payer type, demographics.
Unstructured EMR data	Unstructured data refers to data that is not contained in pre-defined fields, making it potentially more irregular and ambiguous. Examples include clinical notes (written by any discipline, such as medicine, nursing, social work, ethics consult, pastoral care), medical images, and structured data entered in an unstructured format (e.g., faxed copies of structured data).
Health system administrative data	Includes patient information, address/location, billing data, financial assistance applications, and provider claims data. AI applications include identification of inappropriate/fraudulent claims, improving coding processes, gaining insights from linking to social determinants of health and patient generated data.
Patient generated health data	Fitbit, phone trackers, health apps such as those found on the app store or offered by an employer or health care provider.
Public health/ socioeconomic data	Geospatial and aggregated datasets including community demographics, health surveys conducted by public health or governmental organizations, social determinants of health resource referral platform (e.g., UniteUs), and the Area Deprivation Index (ADI).
Consumer and behavioral data	Social media activity, location tracking, gym memberships and utilization, grocery store purchases and other purchasing histories, web browsing, etc.

Preventative measures such as identifying outlier air pollutants and predicting serious automobile injury can inform public health policy and community health education.[18] Advanced databases compile social vulnerability metrics for a block

STL Project," *Journal of Public Health Management and Practice: JPHMP* 23.6 (2017): 577–580, doi: 10.1097/PHH.0000000000000516; and Hiroshi Mamiya et al., "Towards Probabilistic Decision Support in Public Health Practice: Predicting Recent Transmission of Tuberculosis from Patient Attributes," *Journal of Biomedical Informatics* 53 (February 2015): 237–242, doi: 10.1016/j.jbi.2014.11.006.

18. Ilias Bougoudis et al., "Semi-Supervised Hybrid Modeling of Atmospheric Pollution in Urban Centers," in *Communications in Computer and Information Science*, vol. 629, *Engineering Applications of Neural Networks*, ed. Christina Jayne and Lazaros Iliadis (Cham, Switzerland: Springer International Publishing, 2016): 51–63, doi: 10.1007/978-3-319-44188-7_4; and Douglas W. Kononen, Carol A. C. Flannagan, and Stewart C. Wang, "Identification and Validation of a Logistic Regression Model for Predicting Serious Injuries Associated with Motor Vehicle Crashes," *Accident Analysis and Prevention* 43.1 (January 2011): 112–122, doi: 10.1016/j.aap.2010.07.018.

by block look at economic deprivation, which can offer insights for community outreach, public health, and public and private sector collaborations.

Given these applications, AI elicits an optimism that will continue to inspire and sustain investments. AI holds vast potential to enhance population health strategies, improve outcomes, and reduce the total cost of care.

Ethical Concerns of Big Data and AI in Health Care

While AI warrants optimism for reasons we have just illustrated, excessive optimism can "substitute speculation about what AI might make possible for a realistic account of what it is likely to bring about."[19] The concerns, challenges, and questions of big data and AI should not be underestimated. While data itself may seem to be a neutral description, the ways in which humans build and interact with datasets are necessarily value laden. Care must be undertaken to understand the limits of datasets and make intentional choices about data and AI. Focusing on ethical issues should not temper the spirit of innovation or collective optimism but situate such sentiments within a mindfulness of the challenges AI is likely to bring about. In what follows, we summarize some of the ethical concerns with growing investments in big data and AI. These issues are examples of the kinds of concerns emerging and are certainly not exhaustive.[20]

Privacy and Security of Data and AI: Data gathered within health care contexts can reveal some of the most intimate aspects of our lives and behaviors. It has been clearly shown that de-identification of health data—which is itself an AI application—may not always impede re-identification. For example, hospitalized people mentioned in public news stories were linked to a specific hospitalization record in publicly available inpatient discharge data, which was purchased for fifty dollars by researchers.[21] Moreover, digitization opens new risks for malicious actors not only to breach public data but also to target the actual operations of AI. Adversarial techniques can intentionally create mistakes in the model or in the data being processed, with the intent to commit fraud or worse yet, to intentionally harm a person.[22] The term *killware* describes hacking with the explicit intention to

19. Robert Sparrow and Joshua Hatherley, "High Hopes for 'Deep Medicine'? AI, Economics, and the Future of Care," *Hastings Center Report* 50.1 (January 2020): 14, doi: 10.1002/hast.1079.
20. The issues we identify overlap considerably with those identified by Baric-Parker and Anderson, which include patient privacy and confidentiality, informed consent and transparency about collaborations with third parties, replicating bias and exacerbating inequalities, evidentiary standards for implementing AI in clinical care, and the clinical relationship.
21. Ishna Neamatullah et al., "Automated De-identification of Free-text Medical Records," *BMC Med Informatics and Decision Making* 8 (July 24, 2008): 32, doi:10.1186/1472-6947-8-32; Ji Su Yoo et al., "Risks to Patient Privacy: A Re-identification of Patients in Maine and Vermont Statewide Hospital Data," *Technology Science*, October 8, 2018, https://techscience.org/a/2018100901/.
22. Samuel G. Finlayson et al., "Adversarial Attacks on Medical Machine Learning," *Science* 363.6433 (March 22, 2019): 1287–1289, doi: 10.1126/science.aaw4399.

harm or kill people.[23] The first widely known patient death caused in part by hacking occurred in 2019.[24] While it was probably unintentional, it shows the depth of harm that can be caused by improper data security.

Perpetuating Bias in AI: AI depends on sets of assumptions, and these assumptions contain the (unintended) biases of human creators. Unintended bias can confound AI applications in multiple ways, including bias in datasets, bias built into algorithms, and bias in the implementation of AI-produced insights. These biases have the potential not only to reinforce existing social inequities and health disparities but also to worsen them. For example, in one instance, AI was used to create a risk score for patients. However, research uncovered that among black and white patients who had the same risk score, black patients were sicker than white patients. The algorithm used health costs as a proxy for health needs. Because black patients experience unequal access to services, they report fewer health costs. When care coordination resources were offered to patients based on the risk score, white patients with less severe illness would be offered services earlier in their disease process than black patients. Whereas 46.5 percent of black patients should have been offered additional care, only 17.7 percent were.[25] The disparity existing within the original dataset and a racial bias built into the algorithm exacerbated the existing disparity in accessing care and resulted in black patients receiving less access to services. Other examples of AI show bias across race, gender, and payor type.[26]

Provider Experience of AI: The introduction of complex predictive analytics may have a significant impact on clinicians' experience and sense of authority.[27] The clinician has historically had a responsibility to prudentially utilize generalized knowledge and experience to inform the care of a single patient, who always exceeds the sum of relevant data points about her. If not situated as an augmentation of the clinician, AI has the potential to dislodge the clinician's epistemic authority and diminish clinical reasoning skills.[28] A clinician may have reason to conclude

23. Hannah Mitchell, "6 Things to Know about 'Killware,' Cybersecurity's Next Big Threat," *Becker's Hospital Review*, October 12, 2021, www.beckershospitalreview.com/cybersecurity/6-things-to-know-about-killware-cybersecurity-s-next-big-threat.html.
24. Kevin Poulsen et al., "A Hospital Hit by Hackers, a Baby in Distress: The Case of the First Alleged Ransomware Death," *The Wall Street Journal*, September 30, 2021, www.wsj.com/articles/ransomware-hackers-hospital-first-alleged-death-11633008116.
25. Ziad Obermeyer et al., "Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations," *Science* 366.6464 (October 25, 2019): 447–453, doi: 10.1126/science.aax2342.
26. D. A. Vyas et al., "Challenging the Use of Race in the Vaginal Birth after Cesarean Section Calculator," *Women's Health Issues* 29.3 (May–June 2019): 201–204, doi: 10.1016/j.whi.2019.04.007; and Irene Y. Chen et al., "Can AI Help Reduce Disparities in General Medical and Mental Health Care?," *AMA Journal of Ethics* 21.2 (February 1, 2019): E167–179, doi: 10.1001/amajethics.2019.167.
27. Joshua James Hatherly, "Limits of Trust in Medical AI," *Journal of Medical Ethics* 46.7 (June 2020): 478–481, doi: 10.1136/medethics-2019-105935.
28. For a helpful commentary, see Michael Anderson and Susan Leigh Anderson, "How Should AI Be Developed, Validated, and Implemented in Patient Care?" *AMA Journal of Ethics* 21.2 (February 1, 2019): E125–130, doi: 10.1001/amajethics.2019.125.

that an AI-produced insight does not account for the full and unique situation of the patient in front of them. As algorithms become more reliable, clinicians may begin to feel obliged to defer to AI and hesitant to challenge the AI-produced recommendation. This dynamic could lead to dogmatism and gullibility and even reinforce inclinations towards defensive medicine.[29] While there is a clear place for AI to augment human intelligence, there also may be a tendency for AI to subsume prudential judgment by humans if it is not positioned carefully.[30] Further, while there is hope that AI will improve the provider experience, it also could add a new set of documentation burdens and source of dissatisfaction.

Patient Experience of AI: The patient's healing process is shaped not only by what services are delivered but also (and at times especially) by the patient's experience of being heard and receiving care. Mistrust could be created if patients begin to perceive that a nameless, faceless AI clinician is directing their care more than the human clinician with whom they are interacting.[31] Particularly if an AI system's inputs and operations are not visible to users—the notion of the *black box*—patients might find clinicians inadequately knowledgeable and potentially untrustworthy.[32] The impact of AI on the interpersonal encounter and the relationship of trust prompts questions about how to explain and disclose the use of AI. It may be appropriate to disclose and explain how AI is being used to inform critical clinical care decisions. There are probably situations wherein transparency around elements of AI could be seen as an essential element within informed consent. However, many patients will not understand what AI is or how it works. There is probably a sweet spot, as too much information or information that is too complicated may be unmanageable, but too little information may suggest secrecy and lead to suspicion.

Collaboration Across Organizations and Industry: Health systems that have privileged access to large health datasets are unlikely to accomplish major advances as independent systems. News stories have reported that US hospitals are regularly approached by tech companies who want to purchase de-identified health data, which "has become its own small economy."[33] Moreover, data-rich companies like Zillow or Visa might seek to sell their data to add to the predictive capabilities of

29. Thomas Grote and Philipp Berens, "On the Ethics of Algorithmic Decision-Making in Health Care," *Journal of Medical Ethics* 46.3 (March 2020): 205–211, doi:10.1136/medethics-2019-105586.
30. Stressing the assistive role of AI in enhancing but not replacing human intelligence, the American Medical Association uses the term *"augmented* intelligence." AMA, "Artificial Intelligence in Medicine," 2021, https://www.ama-assn.org/amaone/augmented-intelligence-ai.
31. This language has been used throughout the previously mentioned paper by Komorowski et al., "The Artificial Intelligence Clinician."
32. For a helpful summary, see David S. Watson et al., "Clinical Applications of Machine Learning Algorithms: Beyond the Black Box," *BMJ* 364 (March 2019): 1886, doi: 10.1136/bmj.1886.
33. Christina Farr, "Hospital Execs Say They Are Getting Flooded with Requests for Your Health Data," CNBC, December 12, 2018, https://www.cnbc.com/2019/12/18/hospital-execs-say-theyre-flooded-with-requests-for-your-health-data.html.

health systems and insurers. The exchange and sale of data prompts questions about what other stakeholders will do with the data, whose interests are being prioritized, what protections are in place, and whether collaborating organizations have shared values. HIPAA (Health Insurance Portability and Accountability Act) compliance is questionable within these kinds of partnerships, and there is great social interest in determining whether this data is considered protected health information and thus subject to relevant regulations.[34] Further, such partnerships again raise questions of liability and questions about who is responsible for AI design that results in direct or secondary harms. While new regulatory approaches are needed, regulation alone will not ensure public trust in health systems or that applications uplift human dignity and advance the common good.

The Question of Excessive Connectivity: There are trillions of connected devices that are contributing more data points for powerful insights across our lives. But there is a tradeoff; as separate realms of our lives are encapsulated in what appears to be impersonal datasets, data-driven insights can become increasingly invasive. There is a data trail of very personal activities, such as where individual people travel, what websites they browse, what products they buy, what they post on social media, and how their wearable device has quantified daily activity. Perhaps we might feel uneasy about surveillance of our personal text messaging metadata, even if done for the laudable goal of suicide prevention. Perhaps we might feel that our private consumer purchasing habits should not impact our health care and coverage. Perhaps it makes us uneasy to know that one's credit score is a powerful predictor of employment success and is sometimes used to make hiring decisions. Substantive social dialogue is needed to consider whether there should be limits to the kinds of data that are shared, connected, and purchased by public health, health systems, insurers, and employers.

Addressing the Ethical Concerns of Big Data and AI

Although AI may conjure up images of a far-off future world, big data and AI have already transformed our world. AI systems are already offering actionable insights within health care delivery, health insurance, and direct-to-consumer health businesses, and more applications are on the horizon. The concerns raised by the use of big data and AI demonstrate that there are also widespread risks in maintaining patient and social trust. The interdependence of those investing in health-related AI means that every one of these stakeholders has a role in assuming moral responsibility and cultivating social trust. In their white paper *On Artificial Intelligence—A European Approach to Excellence and Trust,* the European Commission coins the phrase "ecosystem of trust."[35] Despite criticisms, this phrase helpfully encapsulates today's reality: building and maintaining trust requires actionable commitments to advance individual and common interests through data and AI endeavors within health systems as well as across all stakeholders and industry partners.

34. Baric-Parker and Anderson, "Patient Data-Sharing for AI."

35. European Commission, "On Artificial Intelligence—A European Approach to Excellence and Trust," (Brussels: EN COM, 2020), https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf.

Data ethics should not wait on a watershed moment when the time is finally right for critical engagement of values questions. Health systems are already incrementally and decisively investing in harnessing the power of big data and developing powerful AI tools. Whether they recognize them as value-laden or not, health systems are making value-laden decisions relevant to big data and AI every day. While the potential benefits of AI are evident, so too are value laden concerns with profound relevance to common human experience in matters of health. With structure and consistency, Catholic health systems have an opportunity to assume responsibility within a complex ecosystem and take on ethics integration work as they manage data, build algorithms, and establish interorganizational and cross-industry data partnerships. Catholic health systems must do so for two reasons. First, ethics integration pertains to the organizational integrity of Catholic health systems, which seek to operate in alignment with the Catholic moral tradition. Second, ethics integration is one necessary (although insufficient) factor to the critical work of cultivating social trust in health systems and contributing to an overall ecosystem of trust.

In the section that follows, we explore how Catholic health systems might kindle a growing commitment to data ethics, which we view as spanning a wide range of actions, including data collection, data quality, data access, data partnerships, data analytics, and AI design and use. These areas entail massive efforts involving thousands of employees and external partners. Considering this complexity, ethics integration will require thoughtful organization and alignment, starting within senior levels of leadership who establish clear expectations for work occurring across many data related domains. As such, we turn to the topic of data governance as an essential component to catalyze sustained work within data ethics in light of our Catholic moral tradition and shared second-order principles, which we unpack in the later sections of this paper.

Ethics Integration in Data Governance

Defining Data Governance

There are many definitions and models of data governance, which is applicable within any organization wherein data is a strategic asset. In general, data governance is the framework by which an organization identifies who has authority over its data, the rules by which data is to be utilized, how data utilization will be monitored and managed, and how enforcement of these rules and responsibilities is to be carried out.[36]

Of particular interest in this paper, data governance includes a primary responsibility for establishing a clear organizational philosophy around data and a set of principles to guide activities and decisions. John Ladley writes, "At the heart of effective governance are organizational principles. . . . Principles are statements of philosophy. . . . They are beliefs to be applied every day as guidance for procedures

36. Ibrahim Alhassan et al., "Data Governance Activities: An Analysis of the Literature," *Journal of Decision Systems* 25.supp1 (June 2016): 64–75, doi: 10.1080/12460125.2016.1187397.

and decision-making efforts."[37] Whereas policies or procedures apply to an already imagined context and thus are prone to declining relevance over time, Ladley notes that principles offer a consistent reference and guide, even and perhaps especially amidst questions that were not initially imagined when the principles were affirmed. In this paper, we focus on data governance because this is one central place wherein an organization can systematically catalyze actions and organizational norms that advance established values—a task that is of particular interest to health systems that seek to advance their mission and Catholic identity across all aspects of the data lifecycle.

Models of data governance vary, and topical focus within data governance has tended towards data security and data integrity. In the previous section, we illustrated the broad challenges health systems face as they integrate advanced AI technologies. The development and use of AI tools cannot be separated from the datasets on which they operate. Hence, data governance should not be compartmentalized to a few slices of the data lifecycle. Emerging views of data governance emphasize the need for an integrated approach, which "must increasingly include data pipelines as well as algorithms and their outcomes" as a recent *Wall Street Journal* article says. Looking forward, effective data governance must be holistic:

> The focus will shift from governing individual data elements to governing the entire data pipeline, with strategies to address the reliability of data sources and black box algorithms, whose inner workings might seem convoluted, and to monitor the output of these algorithms. … AI, machine learning, analytics, automation, and other modern technologies are only as valuable as the data they process and the business outcomes they support. Once an afterthought for many, data has become the lifeblood of an increasingly interconnected world. In order to survive and ensure their competitiveness into the future, organizations of all sizes must adapt and expand their data governance approaches accordingly.[38]

For visionary health systems, data governance must take a wide purview to be woven into the fabric of an organization's data-driven programming, services, and products. Data governance is the structure by which an organization drives the work of creating meaningful datasets and using, storing, sharing, and deploying data in responsible and strategically effective ways.

It is important here to draw a distinction between data governance and data management. The latter refers to the work of making decisions and the realm of operational implementation. The former, comprised of more diverse or multidisciplinary membership, clarifies who has the authority to make these decisions and by what principles, policies and procedures across the organization.[39] In short,

37. John Ladley, *Data Governance: How to Design, Deploy, and Sustain an Effective Data Governance Program* (San Diego, CA: Elsevier Inc., 2020), 27.

38. Tami Frankenfield et al., "The AI Era Is Here. Is Your Data Governance Ready?" *The Wall Street Journal*, February 24, 2021, https://deloitte.wsj.com/articles/the-ai-era-is-here-is-your-data-governance-ready-01614196928.

39. M.N. Kooper et al., "On the Governance of the Information: Introducing a New Concept of Governance to Support the Management of Information," *Interna-*

we see "data governance as the backbone of data management."[40] For example, data governance would involve identifying a foundational commitment (e.g., all partnerships should explore establishing a fundamental agreement to eliminate unfair bias and advance health equity). Data governance also involves establishing what leadership roles (e.g. general council, chief strategy officer, etc.) hold responsibility for ensuring this agreement is built into any data partnerships. Data management is then the work of strategically assessing potential partnerships, developing a shared vision with partners, and agreeing on how the partnership will conduct and report on relevant work.

This structure has essentially three interrelated components: administrative, operations, and technical.[41] The administrative component is made up of the governors (those who may make up a governance council, for example) who are responsible for the architecture—including scope of authority, policy and procedure, and defining data stewardship roles and responsibilities. The operations component includes the data stewards who are responsible for data standardization, definitions and ensuring compliance with the policies and procedures, rules and processes established by the administrative component. Finally, the technical component represents those persons who are responsible for applied realms that execute on the work of data governance. This component is the work of applying big data and AI to gain actionable insights—data integration, data modeling, etc.

Overall, the work of ethics integration in data governance is a foundational opportunity for many health systems. Mechanisms of accountability for integrating ethical commitments into the acquisition, use, commodification, and dissemination of health-related data best occurs at the governance level to ensure adoption throughout the organizational ecosystem. In this way, structures, roles, and decision rights are infused with a standard set of principles to which all those responsible for data management are held. Amidst innovative work that raises questions about the nature of big data, an essential foundation is the clear articulation of ethical principles. As such, we turn now to proposed principles for Catholic health systems.

Principles of Data Governance for Catholic Health Systems

Two Grounding Ethical Principles

While many moral principles are relevant to data governance, most described by authors are secondary (second-order) principles. For health systems rooted in the Catholic theological tradition, we believe grounding (first-order) principles are needed to provide a foundation and context for the secondary principles. With Catholic health systems being fundamentally committed to uplifting the dignity of the human person and addressing the needs of populations of people, we begin

tional Journal of Information Management 31.3 (June 2011): 195–200, doi: 10.1016/j.ijinfomgt.2010.05.009.

40. Peernova, "Data Management vs. Data Governance: What is the Difference?," *Peernova*, April 28, 2020, https://resources.peernova.com/data-management-vs-data-governance-what-is-the-difference/.

41. James C. Orr, *Data Governance for the Executive* (Colorado Springs: Senna Publishing, 2011).

with two grounding principles: *respect for persons* and the *common good*. A complete summary of Catholic thought on these two concepts is beyond the scope of this article. However, this brief overview provides an introduction and situates data ethics in relation to these two grounding principles.[42] As such, we hope that it will encourage future discussion on data ethics to acknowledge and further develop underlying metaphysical and moral assumptions.[43]

First, of paramount importance in the Catholic moral tradition is the sacredness of every human person and the subsequent duty of respect. Respect for persons is an obligation we all have to look upon our neighbor as "another self."[44] In order to respect persons, we must have a clear understanding of what a person is in the Catholic moral tradition. A person is an individual substance of a rational nature.[45] Every human being is a person, a unity of body and soul created in God's image and likeness.[46] As such, each human person has irrevocable worth and innate dignity, and many defining features stem from this divine image. As humans, we are capable of reason, we are subjects with free will and not merely objects, we are made for relationships with others and with God, and we are all equal by our very nature.[47] We have inherent rights, including the right to life, education, work, living wage, religious freedom, food, water, and shelter (*Compendium*, n. 155, 157). We also have responsibilities that correspond to those rights, including to love our neighbor, develop our own potential and that of others, respect the rights of others, seek and know the truth, respect and defend the dignity of all people, care for the environment, and promote the common good (*Compendium*, n. 156, 157). The principle of respecting persons means enabling each person to flourish as the kind of being they are and advancing the rights and goods that each person inherently deserves as a person of inherent dignity made in God's image and likeness.

---

42. In their seminal paper on AI in Catholic health care, Baric-Parker and Anderson put forth four principles: human dignity, common good, care for the poor, and stewardship. We affirm their approach, but highlight two for brevity, since stewardship can be viewed as an obligation stemming from the common good and care for the poor is wrapped up in both respect for persons and the common good.
43. Pontifical Council for Justice and Peace, *Compendium of Catholic Social Doctrine*, (Rome: Libreria Editrice Vaticana, 2004), chap. 3, subsequent references appear in the text; Anglican-Roman Catholic Consultation USA, *Images of God: Reflections on Christian Anthropology*, July 1983; and Valentina Napolitano et al., *The Anthropology of Catholicism: A Reader* (Oakland, CA: University of California Press, 2017).
44. *Catechism of the Catholic Church*, 2nd ed. (Washington, DC: United States Conference of Catholic Bishops/Libreria Editrice Vaticana, 2016 update), n. 1931. All subsequent citations appear in the text.
45. Thomas Aquinas, *Summa theologica*, trans. Fathers of the English Dominican Province (1920; New Advent, 2017), I.29.1 ad 1, https://www.newadvent.org/summa/1029.htm; and Fulvio Di Blasi, *From Aristotle to Thomas Aquinas: Natural Law, Practical Knowledge, and the Person* (South Bend, IN: St. Augustine's Press, 2021), 92–98.
46. This understanding can be compatible with a non-religious anthropology. Such a view could hold that human persons have intrinsic moral value or inherent moral worth without attributing it to being made in God's image.
47. Jacques Maritain, *The Person and the Common Good* (Notre Dame, IN: University of Notre Dame Press, 1966), 38–40, maritain.nd.edu/jmc/etext/CG03.HTM.

Second, Catholic health systems are called to advance the common good, which is "the sum total of social conditions which allow people, either as groups or as individuals, to reach their fulfillment more fully and more easily" (*Catechism*, n. 1906). This includes all people in a population; it does not leave anyone out because they are small in number or on the margins of society. To the contrary, we are called to express a preferential option for poor and vulnerable persons, given the dignity of each human person. Advancing the common good requires us to create social structures and organizations so that people can flourish in community with each other. It also requires a recognition that all resources were given to us by our Creator to be shared fairly by all people. They should be used to benefit everyone, rather than a particular group or subset of people, out of recognition for the universal destination of goods. Lastly, properly understood, the common good does not conflict with the good of persons, since "the common good, by its very essence directs itself to the persons as persons and directs the persons as individuals to itself" (see also *Compendium*, n. 277).[48] While the term *common ground* may be referenced in contemporary society or secular bioethics, the vision of the Catholic tradition is distinct.[49]

Six Secondary Ethical Principles

Many secondary principles stem from respect for persons and the common good, but not all are equally relevant. Fortunately, our robust theological foundations and anthropology lead us to secondary ethical principles that can—and have—been affirmed by diverse stakeholders. The Vatican has already taken the lead in facilitating groundbreaking consensus amongst such stakeholders. Jean Baric-Parker and Emily Anderson offer a helpful summary of initiatives led by the Vatican on AI.[50] Of particular note, in February 2020, leaders from the Pontifical Academy for Life, IBM, Microsoft, the Food and Agriculture Organization of the United Nations, and the Italian Ministry of Innovation became the first signatories of the *Rome Call*.[51]

Stemming from an awareness of both the positive potential and the risk of harmful or unfair uses of AI, the *Rome Call* seeks a shared framework. The statement asserts that "AI systems must be conceived, designed and implemented to serve and protect human beings and the environment in which they live," and "AI-based technology must never be used to exploit people in any way, especially

48. Maritain, *The Person and the Common Good,* 76–82; and Benedict XVI, *Caritas in veritate*, (June 29, 2009), n. 25.

49. Contemporary secular bioethics lacks such a substantive grounding about the human person or the common good. Its excessive focus on individuality and respect for autonomy, and consequent deference to informed consent as the determinative moral criterion makes it ill-equipped to protect persons who are unable to make their own decisions, have difficulty making decisions, or are otherwise vulnerable. Similarly, it has no substantive response to autonomous choices that harm the common good. For helpful background on the ascent of autonomy, see Therese Lysaught, "Respect: Or, How Respect for Persons Became Respect for Autonomy," *Journal of Medicine and Philosophy* 29.6 (January 2004): 665–680, doi: 10.1080/03605310490883028.

50. Baric-Parker and Anderson, "Patient Data-Sharing for AI," 476.

51. *Rome Call for AI Ethics*, (Vatican City: RenAIssance Foundation, February 28, 2020), https://www.romecall.org/wp-content/uploads/2022/03/RomeCall_Paper_web.pdf.

those who are most vulnerable. Instead, it must be used to help people develop their abilities (empowerment/enablement) and to support the planet." Both assertions are extensions of the principles of respect for persons and the common good. They presume an anthropology in which human beings are inherently valuable, even when vulnerable, and respect for persons entails respect for the environment. The *Rome Call* is a pledge to "set out from the very beginning of each algorithm's development with an 'algor-ethical' vision, i.e., an approach of ethics by design."[52]

The signatories encapsulate their commitment to integrate ethics into AI by affirming six shared principles: transparency, inclusion, responsibility, impartiality, reliability, and security and privacy. Again, while respect for persons and the common good are not specifically named, the six principles depict a moral reality grounded in both concepts. In what follows, we develop an expanded understanding of the six principles, which are explained with only a short phrase in the *Rome Call*. In table 2, we also provide a summary of the *Rome Call* principles and our proposed expanded meaning for Catholic health systems.

Reliability: Reliability means that systems and processes should work as expected. This principle applies to the technical realm of data collection and AI (e.g., AI should correctly identify results and avoid false positives and false negatives). Reliability also applies to organizational principles, policies, and procedures and the subsequent actions and processes that consistently lead to intended outcome: data collection and storage practices that follow stated values and policies. Audits lead to effective accountability and oversight. Ethical sensitivities established through principles and policies consistently result in leaders who prompt important questions and discernment processes at the appropriate times. Reliability ultimately means something is consistently true, accurate, and dependable.

The centrality of truth and dependability hints at a deeper issue central to respecting persons and maximizing the value of data and AI efforts in service of the common good: trust. Certainly, accuracy is necessary for an ecosystem of trust, and it reflects our responsibility to seek and know the truth. Inaccurate or inconsistent results will lead to mistrust not just from the users of data or AI but also from patients. For health systems, this mistrust is of significant concern, especially as it will diminish the experience of vulnerable patients in need of services and providers who take pride in extending high quality care. There are no clear pathways to ensure trust with the public, a patient population, or subset of users, but innovative and collaborative ideas should be explored.[53] Nevertheless, building trust through

52. *Rome Call*.

53. See Mark Repenshek, "Liberating Big Data: Implications for a Traditional Consent Approach," *Health Care Ethics USA* 27.3 (July 2019): 35–40, https://www.chausa.org/publications/health-care-ethics-usa/archives/issues/summer-2019/liberating-big-data-implications-for-a-traditional-consent-approach. US Veteran Affairs principles eight and nine suggest user-friendly access to your own information and the ability to request amendments to correct inaccurate or incomplete facts. These features go a long way towards making up for the increasing irrelevance of consent given the power dynamics and incomprehensibility of terms of agreement. Moreover, the "use of innovative technologies, like blockchain, may be the very avenue through which trust is built."

reliability is essential for data use to succeed, because "without trust and love for what is true, there is no social conscience and responsibility."[54]

Transparency: Transparency contains two fundamental elements. The first is explainability. Computers augment human reasoning and support the good intentions of respecting persons and advancing the common good, not vice versa. When AI is not understood or explainable, we risk blind actions that lack intentionality and may perpetuate unknown but negative consequences. Health systems should be able to explain how an algorithm is working and what kinds of assumptions it makes. A health system cannot act with sufficient transparency while lacking knowledge of how an algorithm works, including what datasets were used to create the algorithm, what the shortcomings of those datasets might be, and the effects of bias throughout the dataset and algorithm design.

The second element of transparency involves actual communication via honest, open, and accessible dialogue with applicable individuals, including with data subjects or AI subjects. The subject of data is a person the data describes. The subject of AI is a person to whom an algorithm is applied. In general, respect for persons means people should have access to a reasonable amount of understandable information about what is happening with their data and how AI is impacting their care in significant ways and on an individual level.

The obligations stemming from the principle of transparency are proportionate to the gravity of potential impact to persons and the common good. For example, using AI to enhance the detection of health care fraud may not necessarily require deliberate disclosure with data subjects, as it is a process improvement to back-office operations that advance the common good by minimizing waste and unwarranted payments. In contrast, a life-impacting, individual clinical recommendation for an AI subject probably demands an accessible explanation of how the algorithm works—at least from the health system to provider, and in some cases some information might also be appropriately discussed between provider and patient. As AI systems are used more frequently, further consideration will need to be given to whether and when patients are informed that AI-produced insights are significantly impacting their care.

Transparency also harkens back to the topic of trust. Even if norms evolve wherein AI-produced insights are sometimes disclosed in impactful clinical scenarios, this transparency does not fully address the health system's responsibility. Regardless of what patients know (which is likely to have significant limitations), the organization must take accountability for the way that AI is working and have confidence that it has done due diligence in the reliability of the algorithm and the way that the algorithm aligns with other key principles, such as inclusion and impartiality.

Inclusion: The *Rome Call* describes inclusion as follows: "The needs of all human beings must be taken into consideration so that everyone can benefit and all individuals can be offered the best possible conditions to express themselves and develop." This stems from our duty to the common good and two aspects of respect for persons in particular: equality and diversity. The radical equality of all human beings calls us to account for the needs and concerns of all people so all

54. Benedict XVI, *Caritas in veritate* (June 29, 2009), n. 5.

can benefit. This goal of benefitting people creates an obligation to build a system that allows all people to flourish and reach their full potential (i.e., advances the common good).[55]

Inclusion also involves respect for human diversity. No one person or group of people can fully represent the breadth and depth of humanity as a whole. No one way of being human is perfect. Hence, we have a duty to respect the variety of religious and cultural values and customs throughout the world. This duty also extends to the poor and vulnerable since they are unlikely to flourish without assistance and likely to be inadvertently harmed without intentional safeguards. This special focus is sometimes called a preferential option for the poor.[56] Since respect for persons calls us to defend human dignity, we must pay special attention to protecting those whose dignity is most easily violated. The importance of considering the needs of the poor and vulnerable foreshadows the fourth principle.

Impartiality: The principle of impartiality requires promoting fairness and removing unintended or intended biases that perpetuate health disparities and social inequities. This is grounded in respect for persons since it means that individuals should not be unfairly disadvantaged. Yet, while human persons are called to live in right relationships and usually intend to do so, social injustices persist. Unintended biases are deeply engrained in human interactions and the systems humans build. Datasets and algorithms are prone to absorbing the biases and inequities of the world from which they are created, perpetuating or even exacerbating disparities. Any data driven project will reflect the biases present in prior decisions about data collection and the inequities present in the populations the data describes. Health systems must understand the limitations of their datasets and ensure that harmful biases or assumptions, such as the risk score previously described that disadvantaged black patients, are not built into AI. Health systems also have a responsibility to regularly review datasets and AI for new manifestations or new insights about how they may contain biases, result in unfair outcomes, or exacerbate health disparities. Such results undermine the common good as they limit people's ability to flourish.

Data governance that lives out inclusion and impartiality generates trust, and health systems can do this by paying particular attention to how diverse characteristics, broadly understood, are represented in membership and decision-making. Diverse characteristics (e.g., age, gender, race, discipline of expertise, physical ability, types of work experience, etc.) in leadership and governance contribute to richer dialogue and more robust ethical inquiry. Further, the principles of impartiality and inclusion prompt the question of whether and how physically, socially, and economically disadvantaged populations might be consulted for input or advice—especially with respect to populations with historical experiences that heighten mistrust. People have a right to have input into decisions that affect their lives, which is a value named subsidiarity in the Catholic tradition (*Compendium*, nn. 185–186). While individual patients obviously cannot be routinely involved in the development of AI, the principles of inclusion and impartiality prompt consideration of how the organization is facilitating participation in significant decisions.

---

55. US Conference of Catholic Bishops, *The Ethical and Religious Directives for Catholic Health Care Services*, 6th ed. (Washington, DC: USCCB, 2018), part one.
56. John Paul II, *Sollicitudo rei socialis* (December 30, 1987), n. 42.

Table 2

Six Data Ethics Principles for Catholic Health Care Systems and their Partners

Principle	The *Rome Call*'s Definition	Proposed Expanded Meaning for Data Governance
Reliability	**AI systems must be able to work reliably.**	Systems and processes will work as expected and consistently produce intended outcomes—in both the organizational and technical realms. The use of embedded feedback, regular review, and auditing of AI systems ensures accuracy and awareness of unresolvable limitations. Careful oversight of processes and procedures ensures consistency and accountability within governance and management.
Transparency	**In principle, AI systems must be explainable.**	The way that an AI system works can be clearly and comprehensively explained. Applied AI systems should be explained to physicians, data/AI subjects, and other users to a degree that is necessary and reasonable for showing respect, facilitating necessary understanding, and establishing/maintaining trust. This explainability requires organizational confidence in the quality of training data and a firm understanding of the sources and limitations of that data.
Inclusion	**The needs of all human beings must be taken into consideration so that everyone can benefit and all individuals can be offered the best possible conditions to express themselves and develop.**	The creation of datasets and development of AI pursues the goal of augmenting high-quality health care services, which advance the common good and benefit all people especially the poor and vulnerable. Investments need not be perfectly equal across populations, but strong consideration will be given to investments in data-driven projects and AI technologies that prioritize responses to the needs of the poor, vulnerable, and underserved.
Impartiality	**Do not create or act according to bias, thus safeguarding fairness and human dignity.**	The limitations of datasets will be understood and harmful biases or assumptions will not be built into AI. Regular review of datasets and AI will lead to corrections in newly identified or manifesting biases, unfair outcomes, or exacerbations of existing health disparities.
Security and Privacy	**AI systems must work securely and respect the privacy of users.**	The highest levels of security are maintained with sensitive data and similar standards are expected in data partnerships. Dialogue should occur in exploration of the limits of data collection and use across previously private realms of life.

Table 2 (continued)

Six Data Ethics Principles for Catholic Health Care Systems and their Partners

Principle	The *Rome Call*'s Definition	Proposed Expanded Meaning for Data Governance
Responsibility	**Those who design and deploy the use of AI must proceed with responsibility and transparency.**	Bold leadership is needed to advance these principles and advocate for a sense of social responsibility across sectors. Sincere effort will be exerted listening to the public and specifically to marginalized populations. Building a stronger foundation of social trust in the practice of medicine and contributing to an ecosystem of trust across data and AI stakeholders and the public is the obligation of all major stakeholders working with big data and AI.

Security and Privacy: While security and privacy are not new to health care, their application in an era of big data and AI will require new regulatory processes and investments in social dialogue in light of changing social expectations, trends, and comfort levels. Breaches of security can lead to grave harms and death. Moreover, questions of privacy abound as big data connects previously private areas of our lives and allows for powerful, and sometimes invasive, predictions about our behavior, feelings, and health. Health systems have a particularly heightened responsibility to protect their data from unauthorized access. Health systems also must take leadership in facilitating social dialogue about emerging questions of privacy and security and building new standards that strengthen an ecosystem of trust as we approach more widespread AI applications.

Responsibility: All stakeholders involved in health-related big data and AI, and especially health systems, have an obligation to assume social and moral responsibility. The digitization of daily life and the delivery of health care mediates our habits of thinking, our patterns of knowing, and our interactions with one another. The impact is significant. Health systems hold a responsibility to ensure that further integration of big data and AI leads to better quality health care services, better experiences for patients and providers, a reduction in health disparities, and, ultimately, improvements in the human capacity to thrive in relationship with each other—in other words, a responsibility to create an ecosystem that respects persons and advances the common good. Health systems must step up as leaders across industries, calling technology companies and others to assume social responsibility in partnerships and more generally in serving the common good. Responsibility also means welcoming regulatory improvements and participating in their development, especially in the context where big data is rapidly shared across increasingly globalized spaces.

Key Ethical Questions

The principles of the *Rome Call* provide an ethical framework that can guide processes, decisions, and actions when established within the structure of data governance. In what follows, we provide examples of the kinds of key ethical questions that these principles might prompt. Rather than presenting an extended commentary on inevitably limited issues, we believe this approach offers an open-ended and though-provoking invitation to explore data ethics within governance.

In the tables that follow, we identify key questions that these six principles might prompt for data governance and management. These key questions were developed from our own assessment of the moral concerns of data governance and a review of the literature. They cover seven areas:

Data Governance Overall. (Table 3)

Data Collection: the creation, generation, purchase, or acquisition of data coming from internal or external sources, as previously summarized in Table 1. (Table 4)

Data Access: the ability to view, use, see, receive, or purchase data. (Table 5)

AI Design: the creation of an algorithm or model from a dataset in order to achieve a specific purpose or solve a specific problem. (Table 6)

Data Use: the application of AI or datasets to a person or group of people. (Table 7)

AI Refinement: the process of changing an AI system after deployment in order to improve accuracy by adding false positives and false negatives into the original dataset, tweaking the data to reflect more accurately the target population, or by other means. (Table 8)

As we proceed, we acknowledge that we are extending the *Rome Call* principles beyond the scope of algorithm design and use. The *Rome Call* principles provide a starting point and have applicability broadly within data ethics. The principles are not exhaustive of those needed within Catholic health systems, and the questions they prompt for us are certainly not comprehensively described. Additional sets of more specific principles may need to be established for specific areas of work within health systems and across new partnerships. Nevertheless, we must begin somewhere, and we encourage further thought from the Catholic tradition on data ethics.

These tables illustrate the kinds of questions that fall in the realm of data ethics. They also illustrate the way that guiding principles might serve to prompt ethical reflection or questioning. With respect to this kind of work, some organizations are already well on their way. For example, in the US, the Department of Veteran Affairs (VA) developed eight guiding principles that establish the foundation for actions of all stakeholders, guideposts for previously unimagined questions, and reference points to which actors can be held accountable.[57] Catholic health systems can learn from the VA's groundbreaking, secular thought leadership on ethics integration in

57. US Department of Veterans Affairs, *Enterprise Data Strategy: A Vision for the Future*, (VA: January 2021), esp. appendix D, "VA Data Guiding Principles," and appendix E, "VA Ethical Principles for Access to and Use of Veteran Data," 26–30, https://www.va.gov/oei/docs/VA_Data_Strategy.pdf.

Table 3

Data Governance Overall

Principles	Key Questions
Reliability	How mature are our organization's data governance efforts? What principles, policies, or roles and responsibilities should be prioritized for further development?
	How are leaders formed with an understanding of our ethical commitments and the philosophy and processes of data governance and management, as applicable to their level of responsibility?
Transparency	Do we make data governance and data management policies publicly available? What is the purpose of doing so and how might this impact the content of our policies?
Inclusion	How does the overall strategy around data collection efforts, AI tools, and programs/services reflect the value of serving all populations equally or, at minimum, fairly?
	How are diverse characteristics represented across data governance leadership and members?
Impartiality	What procedures are in place to ensure that unintended bias and unfair outcomes are being identified rather than enacted?
Security and Privacy	What level of control should we offer individuals regarding their data? Can the data subjects opt-out of the action? If so, how assertively must an opt-out option be presented? If not, what justifies requiring participation?
	How much does the organization rely on data de-identification, knowing re-identification is always a risk? Is HIPAA compliance sufficient? When might our mission call us to a higher standard than HIPAA compliance?
Responsibility	How does our overall strategy for data governance stem from our mission and established ethical principles? How does it shape our character as an organization?
	How does our organization build trust with our patients, physicians, employees, and community, as mistrust and misunderstanding about data are common? With respect to a particular action, how should we solicit feedback from the public or specially identified stakeholders?
	How does digitization and AI impact the creation of new jobs and the loss of existing jobs? What is the overall impact of employment opportunities for our lowest paid employees? Can job losses be mitigated?
	How are we participating in the development of more comprehensive regulatory improvements for the emerging world of data, AI, and partnerships?

data governance, although a detailed summary is beyond what we can describe here. However, many health systems have a significant opportunity to invest in the work of ethics integration. As we conclude this paper, we briefly explore potential roles of the ethicist within this work.

Table 4

Data Collection

Principles	Key Questions
Reliability	How is data integrity ensured, especially when data is received from external sources?
Transparency	What is the process for people to review their own data? How can people request a correction if they can show something is false?
Inclusion	Should we collect data on potential patients even if we have no relationship with them—for example, the entire population living in counties we serve? Should information collected before a clinical relationship is established be used to influence care?
	What subpopulations are not adequately represented in a dataset that is intended to describe a defined population?
Impartiality	Should all data elements be treated the same?
Security and Privacy	Was there a reasonable expectation of privacy when one of the datasets was collected? If so, what justifies violating that expectation of privacy?
Responsibility	Who owns the data being collected: us, the vendor, or the subjects? Does ownership mean the same thing as control? How do we ensure shared sense of social responsibility across partnerships?
	When merging datasets from different sources, how does the merger benefit the data's subjects? Will merging datasets lead to invasive connections between disparate or private areas of life that exceed a reasonable subject's comfort level?
	Do we require a specific purpose for collecting this data? If not, what reason justifies data collection for a general or unspecified purpose?

The Role of the Ethicist and Modalities of Ethics Integration

Catholic health systems have invested significantly in leadership within mission integration and ethics, prompting a question of what role such leaders might have in data governance. As implied by our attention to data governance, ethics integration is a shared responsibility that cannot be accomplished by one or several people alone. Nevertheless, ethicists or other ethics personnel (e.g., ethics champions or ethics liaisons) might be positioned to add value to the work of establishing shared principles, integrating and educating on those principles, and exploring ethical questions. We envision three such models for optimizing the impact of ethics personnel, which are not mutually exclusive and could be pursued in unison to support decision-making, processes, and policies.

Ethics Leader as Integrated Collaborator: One model involves the integration of professional ethicists or other ethics personnel within data governance and operational infrastructure. Such integration can occur through membership of standing committees or councils that have as their purpose the execution and implementation of strategy and tactics in electronic medical records, population

Table 5

Data Access

Principles	Key Questions
Reliability	What processes exist to ensure anyone accessing our data is accountable for using it responsibly and will not use it in a way contrary to our ethical principles?
Transparency	If a breach occurs, under what conditions should disclosure occur? Should everyone in the data repository be informed, as opposed to only those affected by the breach?
Inclusion	Is the requested access for a specific purpose with a clear usage plan? If not, what reason justifies the requested general or unspecified access?
	Is it permitted to re-use data for something other than the originally approved purpose? If so, why? Note that this is illegal in some countries.
Impartiality	To what degree are external partners who are given privileged access to data aligned with our commitment not just to avoid exacerbating disparities but also to reduce health disparities?
Security and Privacy	How likely is re-identification? Could re-identification of this dataset lead to harm?
	Are security measures proportionate to the harms that could occur due to a data breach? If a breach could lead to killware, is the security the highest level possible?
	What steps are in place to regularly audit for unauthorized access?
	What process is in place so breaches or improper uses will be detected?
	Is there a clear, objective, and consistent process for granting access to data?
Responsibility	Are the data handling practices of people with data access in line with our values, practices, and industry standards?
	Is it ethically justified to sell data? Is there an ethical difference between selling data and monetizing the data by exchanging it for other goods such as an equity stake in a company, stocks, access to other data, etc.? How does the sale and purchase of data impact public trust?

health, human resources, and a variety of other institutional divisions that draw and use available data. This model could involve existing ethicists or evolve towards niche roles for ethicists who are explicitly focused on data governance as a chief data ethics officer. The challenge with this model is the large number of relevant groups with which the ethics leader needs to be in dialogue and the challenge of having sufficient technical knowledge of big data and AI.

Ethics Leader as Educator: Another model would focus on education as the key role for ethics integration. In light of the vast number of people involved in data and AI and the complexity of the relevant subject matters, the ethics leader might focus on building ethics competencies of specified leaders and groups within

Table 6

AI Design

Principles	Key Questions
Reliability	Does the AI system avoid using proxy data if at all possible? If not, is proxy data given more weight than directly relevant data? If so, what justifies this? Does the use of proxy data consistently lead to reliable and applicable insights?
Transparency	Do users know how the AI system was made and how it works?
Inclusion	How are inequities and disparities considered in the design process to ensure maximal benefit to all?
	If the AI system will be applied to a vulnerable population, should representatives of the population be involved or consulted in the design process?
Impartiality	What steps have been taken to minimize the blind spots in the AI system caused by the design team's own judgments and priorities?
	Do the results benefit or support the interests of most subjects? If not, are they neutral to most subjects? What recourse is available to those who are harmed or whose interests are undermined?
	Does the original dataset avoid proxy data if the outcome is potentially harmful?
	Does the original dataset used to train the AI system accurately represent the target population? Is the original dataset unbiased?
Security and Privacy	Will the results expose conclusions that a reasonable person would consider to be overly invasive or an invasion of private realms of their life?
Responsibility	What are the design team's guiding values and purpose for this particular project? When a partner is involved, have values and purpose been clearly identified and agreed upon?
	Do the design team members and the subjects share the same goals for the AI system's use?
	Does the design use data points that would be illegal, unethical, or rejected by experts in the relevant field for carrying inaccuracies, overgeneralizations, bias, or inequity?

the healthy system. This education might occur in standalone contexts or by integrating the ethics leader into other existing departments, meetings, or teams for just-in-time education.

Ethics Committee Model: Similarly, a third model establishes specific data ethics committees or subcommittees. This committee could be analogous to the functions and structures of hospital-based ethics committees but whose scope may apply to the horizon of data acquisition and application within the enterprise. In this sense, such entities would not be tied to specific facilities or lines of business but rather would be a conduit for interdepartmental coordination and collaboration in the space of ethical decision-making. Since ethics integration occurs widely

Table 7

Data Use

Principles	Key Questions
Reliability	Are the results reviewed and agreed upon by a human being before being applied to the target population? If not, are the results automatically applied to the target population without questions? What justifies bypassing manual review?
Transparency	Is the level of patient education or community engagement regarding the use of data proportionate to the depth of the impact on people's lives?
	If buying an AI system designed by a third party, are we allowed to see the original data to ensure a biased dataset did not create a biased AI system? If not, how do we know the designers share our purpose, goals, values, and commitment to unbiased results?
	Do subjects know data is being used? Do they understand the process, results, and intended outcome? What level of transparency do they have?
Inclusion	Are the AI system and related resources applied equally to all members of a population? If not, what justifies targeting a subset of the population?
	Does the AI system attempt to predict a person's future behavior on the behavior of other people? How is fairness ensured in light of this use of data?
Impartiality	Does the AI system exacerbate existing vulnerabilities, disparities, or inequities? Does it create a loop that does not worsen the disparity but perpetuates it?
Security and Privacy	What security or privacy concerns are raised by the application of this AI system, beyond issues that were prompted by collecting and storing related data?
Responsibility	What kinds of actions warrant higher level of scrutiny, given their potential to cause harm and/or affect a large number of people?
	What values, goals, and purposes are driving this data use?
	If the AI system uses proxy data or could cause harm, is there sufficient manual review for the application of results?

across data governance and management, ethics integration is not the work of a committee alone; however, the committee could provide a discussion forum for particularly critical or complex questions, and ethics personnel would probably have responsibility for managing the committee.

Serving in these roles, ethics personnel can support other modalities of ethics integration established through data governance. For example, processes should also be established for how decisions of high ethical impact will be made, with specific agreement on when leaders should employ a formal discernment process. Mechanisms for data ethics consultation could support the resolution of specific ethical questions, including basic questions or policy clarification (secondary data ethics consultation) and a formal ethics consultation from an ethics professional (tertiary

Table 8

AI Refinement

Principles	Key Questions
Reliability	AI refinement is a moral endeavor, not just an effort at accuracy. Does the process, frequency, and outcomes of refinement reflect this fact?
	How often should the AI system be audited, reviewed or refined? Is there an adequate process for an annual audit of AI systems?
	Is there a risk the AI system will cause or contribute to the thing it seeks to predict? Could it lead to a cycle that simply supports the AI system itself?
	What is the process for automatically collecting and including false positives and false negatives in revisions? How can users submit this information on their own?
Transparency	If applying an AI system has harmed subjects, how are the subjects notified of the harm? What guidance are they given to improve their results?
Inclusion	Are outcomes taken as unquestionably true, or are they seen as guideposts that must be correlated with the experience of people close to the issue?
Impartiality	Are outliers treated with respect and dignity? Is their treatment consistent with the fact that all datasets have outliers, so they are normal and expected?
Security and Privacy	Have any security or privacy complaints or concerns been received in the implementation of the AI system? How are they being addressed?
Responsibility	Do refinement processes include consideration of values to ensure alignment with ethical principles?
	Are the intensity and frequency of refinement proportionate to the seriousness of the AI system's implications for its subjects?

data ethics consultation). Robustly integrating values into policies, procedures, and charters assists people in engaging in ethical reflection as they think and act with respect to data. By positioning ethical principles, ethics leadership, and these modalities as a part of data governance, health systems will elevate the importance of values alignment and empower leaders to assume responsibility for advancing the interests of human persons, the earth, and the common good.

Taking Up the Opportunity to Lead in Data Ethics

There is no doubt that big data will grow and applications of AI will continue to pervade daily life and the health care system. Big data and AI hold vast promise for increasing the quality of health care services, enabling more effective population health strategies, and advancing mental and community health efforts. Several decades into the future, we will all look back on today and marvel at the profound

ways that big data and AI have transformed health care and modified our norms, practices, and interactions.

Catholic health systems have a timely opportunity to invest necessary resources in data ethics and building data governance structures that catalyze alignment between work and the values of Catholic health systems. Data governance is still in early stages of development across the nation for the majority of health systems. A 2017 survey suggested that more than half of US health care organizations lack a comprehensive data governance structures.[58] Even if this trend has improved during the last 5 years, the majority of health systems are probably at a pivotal moment wherein further development and investment will launch deeper maturation of data governance and management. As they do so, ethics integration is one essential piece of governance that warrants careful attention. Investing in building ethics throughout governance also offers Catholic health systems an occasion to consider Jean Baric-Parker and Emily Anderson's invitation to join the *Rome Call*, which to date has not been signed by any Catholic health system.

Catholic health systems are well acquainted with the experience of responding to health needs on a pioneering frontier of ever-changing contexts. Throughout their tradition of serving, Catholic health systems have fostered collaboration across disciplines to find innovative ways forward. They have embraced innovations that better meet the surrounding needs. And they have built organizational structures to promote values and support the organization in living into their mission and Catholic commitments. In this sense, the realm of data ethics is not unfamiliar territory. Data ethics offers another opportunity to faithfully accept the call to serve and lead with the tools and technologies of the day. We hope that Catholic health systems take up this invitation and seize the opportunity to act in faith and service in the era of big data.

58. Jennifer Bresnick, "56% of Hospitals Lack Big Data Governance, Analytics Plans," *Health IT Analytics: Quality & Governance News*, September 29, 2017, https://healthitanalytics.com/news/56-of-hospitals-lack-big-data-governance-analytics-plans.