The Role of Analytics Governance to Promote Health Care Transformation

Bruce Levy[1] Karen Murphy[2] Greg Strevig[2] Ilia Decker[2] Kirk Hanson[2] Raghuveer Puttagunta[1] Shikha Sinha[1] Spencer Tavares[1] Leyla Warsame[1] David Vawdrey[2]

[1] Department of Medical Education, Geisinger Commonwealth School of Medicine, Scranton, Pennsylvania, United States
[2] Steele Institute for Health Innovation, Geisinger Health, Danville, Pennsylvania, United States

Address for correspondence Bruce Levy, MD, Geisinger Informatics, 100 North Academy Drive, MC 30-39, Danville, PA 17822, United States (e-mail: blevy2@geisinger.edu).

ACI Open 2022;6:e1–e10.

Abstract

Objectives Rapid digitization in health care during the 21st century has created significant data and analytics challenges for our providers and health systems. Just as information technology (IT) governance has helped manage exploding demand for IT services and increased efficiencies, analytics governance promises to bring these same benefits to data and analytics efforts. Potential governance models exist in other industries yet have not significantly penetrated health care.

Methods and Results Geisinger has implemented analytics governance throughout our enterprise. We identified and accomplished six core goals toward the establishment of analytics governance, including developing a vision; defining the organizational structure, roles, and responsibilities; managing our data assets; implementing robust data governance; establishing standardized analytics processes; and utilizing metrics to evaluate our progress. Early outcomes include improved tracking and intelligence around data/analytics requests, decreases in duplicative data/analytics efforts, the creation of the Enterprise Analytics Hub for employees to consume data, and initial steps toward self-service analytics.

Conclusion Our experiences support the proposition that analytics governance can provide meaningful benefits to health systems. It is clear from the experiences in other industries that health systems who can best manage their data and analytics will have a significant competitive advantage. Analytics governance will also provide a proper foundation for the use of advanced analytics, machine learning, and visualization tools, and prepare our workforce to utilize these tools for the benefit of patients.

Keywords
- medical informatics
- data science
- data visualization

Introduction

There has been a rapid digitalization of health care over the first two decades of the 21st century, driven by the widespread adoption of electronic health records (EHRs) and smart medical devices. This has resulted in an explosion of health data that are being collected and used to care for patients and to run the business of health care.[1] Not only has the volume of data increased tremendously, but there are also many additional types and forms of data compared with the past.[2] Health data exist in structured, unstructured, and semistructured forms, and much of the highly valuable data are difficult to meaningfully extract from unstructured data such as a physician's progress notes or a radiology image.[3,4] Health data are increasingly being shared between providers and with patients.[5] In addition to traditional patient health

received
April 13, 2021
accepted after revision
November 19, 2021

DOI https://doi.org/10.1055/s-0042-1744383.
ISSN 2566-9346.

Georg Thieme Verlag KG, Rüdigerstraße 14, 70469 Stuttgart, Germany

data, health systems are also collecting and attempting to incorporate disparate data sources such as insurance claims data, genomic data, socioeconomic data, and patient-generated health data.[6]

The growth of health data has placed an undue burden on our providers and health care workers. They are responsible for a large percentage of the data entered into the EHR.[7] The increasing amounts of time providers are spending entering data and placing orders in the EHR instead of with patients or off-hours has contributed to growing levels of provider dissatisfaction and burnout.[8,9] On the receiving end, providers are being cognitively overwhelmed by the ever-increasing data tsunami.[10] The volume and complexity of health data far exceeds the ability of the human brain to interpret. Analytics tools have been effectively utilized to help organize, evaluate, and visualize health data for providers and operations.[11,12]

This has fueled a growing demand for analytics to leverage clinically actionable data for patient care, which in addition to growing regulatory data reporting requirements and operational/financial data needs, has led to a growing recognition of and demand for big data analytics capabilities in health care. Providers, staff, managers, and leadership are demanding increasing amounts of analytics, overwhelming the capacity of the analysts to handle these requests. Big data was first defined in terms of its volume, velocity, and variety (the 3 V's)[13] but has been expanded to include veracity, variability, and visualization. Ultimately, big data are increasingly being defined by its value.[14]

While health systems spend tremendous amounts of money to collect, store, and use health data, most still consider this data a "by-product of health care delivery, rather than a central asset source for competitive advantage."[15] To obtain maximum value for the investment in electronic health data, systems need to treat data as the primary asset it represents and to integrate big data analytics, which will enable analysis of complex data and render valuable insights that would otherwise not have been identified.[16] It has been estimated that effective use of big data analytics could reduce U.S. health care expenditures by 8% and save more than $300 billion per year.[17] A survey of 26 big data implementations in health care identified the top five analytics capabilities: analytical capabilities for patterns of care, unstructured data analytical capabilities, decision support capability, predictive capability, and traceability.[18]

The response to this demand by hospitals and health systems has been inconsistent. Many systems currently have fragmented and siloed data and analytic efforts in which departments have their own teams that may not be aligned with similar units in other departments. This is like the early years of computerization in hospitals, where some departments had best-of-breed systems with minimal interfaces with other departmental systems. As the number and complexity of health information technology (HIT) systems used by hospitals increased, along with more sophisticated interfaces to share data and increasing user requests for new technical capabilities, it became obvious that health systems needed to develop some form of HIT governance to manage and prioritize these growing requests.[19] The same is now true for informatics, data, and analytics.

It has been recognized outside of health care that big data and analytics help companies achieve their business objectives through optimizing business processes and organizational performance.[20] One of the key findings from a survey of over 3,000 executives, managers, and analysts working across more than 30 industries and 100 countries is that "top-performing organizations use analytics more than five times more than lower performers."[21] Yet, another survey revealed that only 4% of over 400 large companies worldwide have the "right people, tools, data, and intent to draw meaningful insights from that data—and to act on them."[22] It has also been recognized that technical capability with big data analytics is not sufficient to drive innovation alone but requires analytics governance as a moderating influence.[23] Three areas of concern unique to analytics that require a governance function include integration, legal and regulatory issues, and human capital.[24] The examples from other industries appear translatable to health care.

Among the major challenges to adopting analytics governance in health care are the absence of evidence of its practical benefits in health care,[25] and the absolute need for data and analytics accuracy due to medicine's life and death decision making.[26] In health care the energy and lifespan of big data also requires consideration. A 2016 article opines and presents data to support the concept that the energy of big data contains the holistic information content included in the data, the energy of an aggregated data set is greater than the sum of the energy of the individual data sets, and the lifespan of health data, as well as its value, decreases at an exponential rate.[27] Strategies to overcome some of these and other challenges to big data analytics include implementing governance, developing an information sharing culture, employing security measures, training key personnel to use big data analytics, incorporating cloud computing for analytics, and generating new business ideas from data analytics.[18,25]

The few published papers on analytics governance in health care are focused on ethical considerations. One proposed model for artificial intelligence (AI) governance in health care incorporates the four major ethical foci of fairness, transparency, trustworthiness, and accountability.[28] An article discussing the governance of automated image analysis and AI in radiology reinforces this viewpoint and indicates that radiologists have professional and ethical responsibilities around the use and minimization of implicit bias in use of analytics.[29] This ethical focus is important for analytics governance more broadly.

Geisinger, a large, rural, integrated health delivery system in central and northeastern Pennsylvania has both a culture of innovation and a long history in utilizing health data for patient care and operational efficiency. Geisinger implemented an EHR throughout its ambulatory care settings in 1995 and implemented the same EHR in its hospitals a short time later.[30] The clinical and operational leadership was intrigued by the RAND study published in the *New England Journal of Medicine* in 2003 that concluded that almost 45% of

care was suboptimal—including both too much, too little, and incorrect care[31]—and they set out to optimize processes that would improve care delivery. The enterprise, including its health system and health plan working in partnership, decided to apply process redesign methodology and reliability science to implement and consistently deliver evidence-based medical practices—a program that was dubbed "ProvenCare."

Geisinger's first value-reengineering project focused on coronary artery bypass graft (CABG), a relatively high-cost, high-volume, high-variability procedure. The team identified 40 measurable process elements and increased compliance with all 40 elements from 59% to 100%. This resulted in improved 30-day clinical outcomes in 8 of 9 measured areas.[32] The success of the initial ProvenCare CABG led to the growth of the ProvenCare portfolio to include dozens of acute and chronic health conditions and procedures,[31] each of which have provided both improved patient outcomes and financial savings.

One of the consequences of the ProvenCare initiative was the ever-increasing demand for high-quality data and analytics throughout the clinical and health insurance divisions of the Geisinger enterprise. Many departments and business units hired their own analysts to collect and analyze data. This resulted in fragmented and inefficient data and analytics environments despite strong central IT, informatics, and data teams. Leadership recognized that an enterprise-level initiative to establish analytics governance was necessary for efficient future growth. With the formation of the Steele Institute for Health Innovation, data, analytics, and informatics were directly connected with clinical redesign, digital transformation, AI, and health innovations.[33] The stage was now set for the development of analytics governance to address the identified deficiencies, leverage existing resources more effectively, and raise our analytics efforts to a higher level.

Methods and Results

Foundation for Analytics Governance

The existence of analytics governance structures in health care institutions can ensure not only that the institution's technology infrastructure is sufficient to support data and analytics needs, but also that issues related to people, process, and culture are sufficiently established; however, a necessary foundation first needs to be established.

Our review of the literature identified recommended elements of successful analytics governance structures. It was clearly important to align analytics efforts to overall organizational vision and business strategy.[21,34] The elimination of analytic silos and the development of common definitions, processes, and tools for analytics are necessary to the development of an analytics consuming and sharing culture.[18,21,24,34,35] The need for appropriate organizational structures, roles, responsibilities, and accountability for the analytics efforts are familiar to those already involved with other forms of governance.[18,24,34,36] Another common prerequisite was the effective management of data assets and employment of robust data governance (DG) structures.[18,24,35,37]

Table 1 Goals of analytics governance

Analytics governance goals
1. Develop the vision for data and analytics and connect it to the strategic priorities of the organization
2. Define the organizational structure, roles, and responsibilities
3. Manage the institution's data assets
4. Implement a robust data governance program
5. Establish analytics processes to standardize visualization and delivery of data
6. Promote the thoughtful implementation and rigorous evaluation of institutional programs and initiatives

We studied our current state and envisioned the desired future state for analytics. Based on what we learned from the literature and our own prior experiences with IT and other forms of institutional governance, we identified six goals for the initiative, which are described below. The sponsorship of executive leadership was obtained, including both dedicated resources and political support. We recognized the need for and set processes to ensure transparent and consistent communications with all stakeholders.

These are the six recommended goals for analytics governance (►Table 1).

1. Develop the vision for data and analytics and connect it to the strategic priorities of the organization: What is the future state being reached for? Why should everyone in your organization care? Without a clear vision this effort will quickly deteriorate into an exercise of governance for its own sake. Ideally, the vision should align well with your system's vision and strategic priorities. It should be something simple and easy to understand. We chose "*Use informatics to make better health easy*," as our vision, and aligned our vision with our health system's values of kindness, excellence, safety, learning, and innovation.

2. Define the organizational structure, roles, and responsibilities: This effort will reach into every corner of the enterprise and requires more than the typical organizational structure of an IT department. Dotted line connections and organized formal channels of communication are critical to ensure success.

Geisinger's Steele Institute for Health Innovation is led by the Executive Vice President and Chief Innovation Officer, who reports to the Chief Executive Officer. Informatics is one of the core functions within the Steele Institute, and it is headed by the Chief Data Informatics Officer (CDIO). Direct reports to the CDIO include leaders responsible for data management, enterprise analytics, informatics strategy, and clinical informatics.

The Informatics Core is comprised of three overlapping functions: Data Management, Data Governance, and Data Delivery (►Fig. 1). Data Management is concerned with architecting, storage, and operations involving data. Data

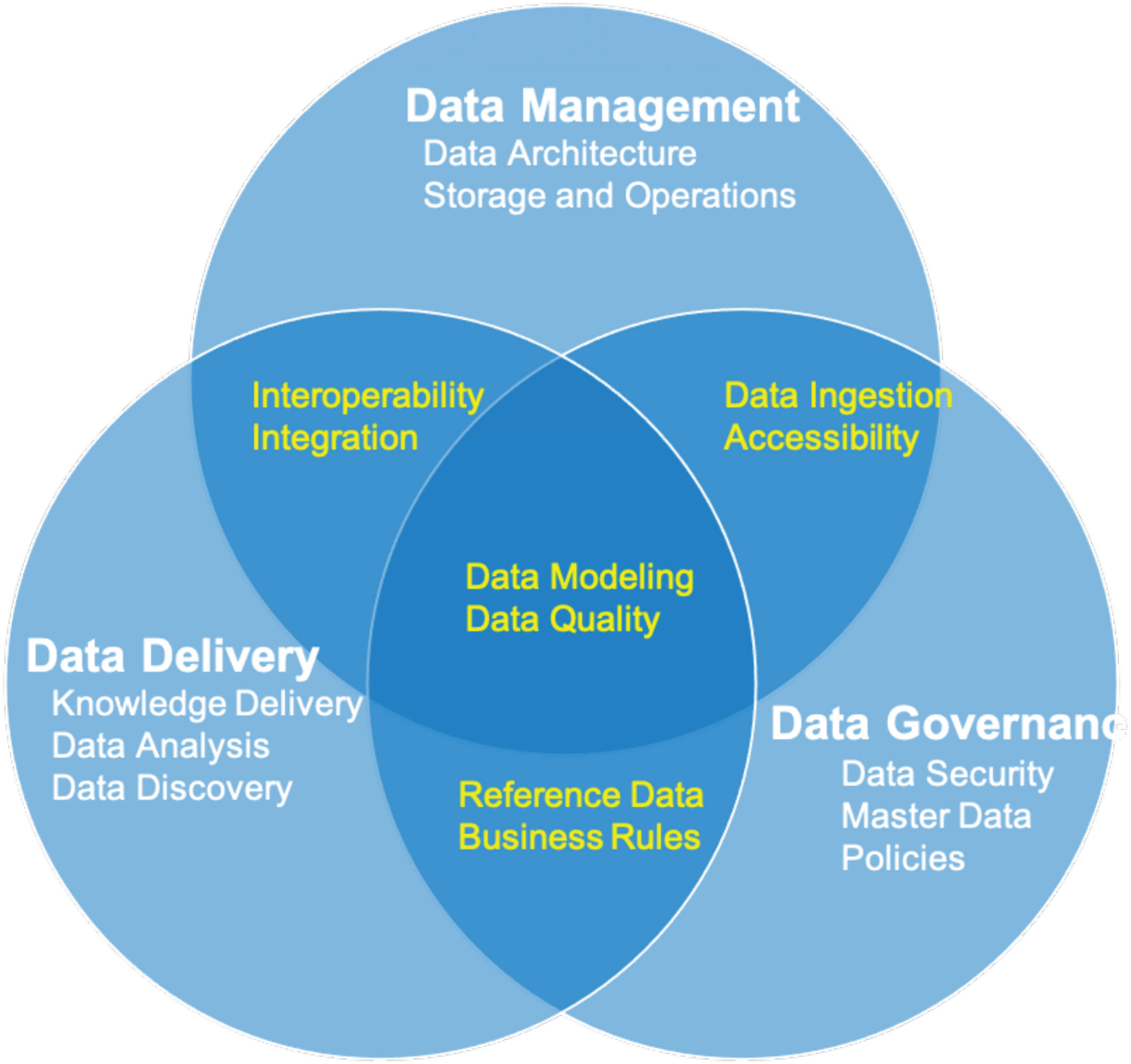

Fig. 1 The Informatics Core comprises three overlapping functions: Data Management, Data Governance, and Data Delivery.

Governance establishes the master data policies and works to ensure the security of our data and analytic reports. Data Delivery encompasses knowledge delivery, data analytics, and data discovery. Overlapping responsibilities across these functions include interoperability, integration, accessibility, business rules, and data ingestion, quality, and modeling.

3. Manage the institution's data assets: Many health systems have a data warehouse that is connected to the EHR and provides the primary source for data analytics. However, there are some shortcomings to the typical EHR data warehouse. It may be difficult or impossible to merge data from external data sources into reports. They are not well suited to the analysis of the massive quantities of unstructured data within the EHR or to feed that data into AI machine learning algorithms. The data may be a day or more old, insufficient in a world that demands real-time data.

Geisinger resolved this problem by creating a solution consisting of multiple different data storage formats and warehouses that collect data from numerous transactional systems (►Table 2). All data are imported into a large repository of big data (data lake) where the data are maintained in its native format without any normalization, processing, or manipulation. Several additional data warehouses with different subsets of the total data provide blended modeled data, EHR-based analytics, and population health analytics, among others. Provided by EHR and non-EHR vendors, these warehouses include both in-house and cloud-based solutions and have normalized/modeled data.

4. Implement a robust DG program: Effective governance of data are closely integrated with analytics governance, especially with increasing amounts of data sharing and analysis. Aspects of governing the data should include enforcing the use of terminology standards; master data

Table 2 Transactional systems that feed into data lake and warehouses

Transactional source systems for warehousing and analytics	
Electronic health record systems (multiple)	Customer relationship management system
Departmental systems (e.g., Laboratory Information System)	Socioeconomic data and social determinants of health
Picture archiving and communication systems (PACS) and other imaging systems	Facilities data (e.g., utilities, maintenance, construction, supply chain)
Imported health data (e.g., scanned documents, digital data, health information exchange)	Outside data (e.g., rankings and ratings, benchmarks, public health)
Patient-generated health data	Financial data, including available payor data
Research generated data	Student and trainee data

management; setting organizational policies and procedures (P&Ps) for information security, privacy, and data sharing, and monitoring compliance with those P&Ps; assessing and improving data quality; and ensuring data are available to measure health equity.

Geisinger recognized the need for an enterprise data strategy that included strong DG. A DG structure (►Fig. 2) was put in place with the vision to "*deliver the full value of our data*." The foundation of the DG structure is to govern our data to the minimum extent required to achieve our goals by defaulting to less governance and by expanding DG only when absolutely required.

The three pillars of the DG program are to optimize data access, content, and quality. Our DG program has partnerships with our Information Security Office, Chief Privacy Officer, and Legal Counsel to provide a proper balance between appropriate access and necessary and required security. There is a continuous journey in the identification,

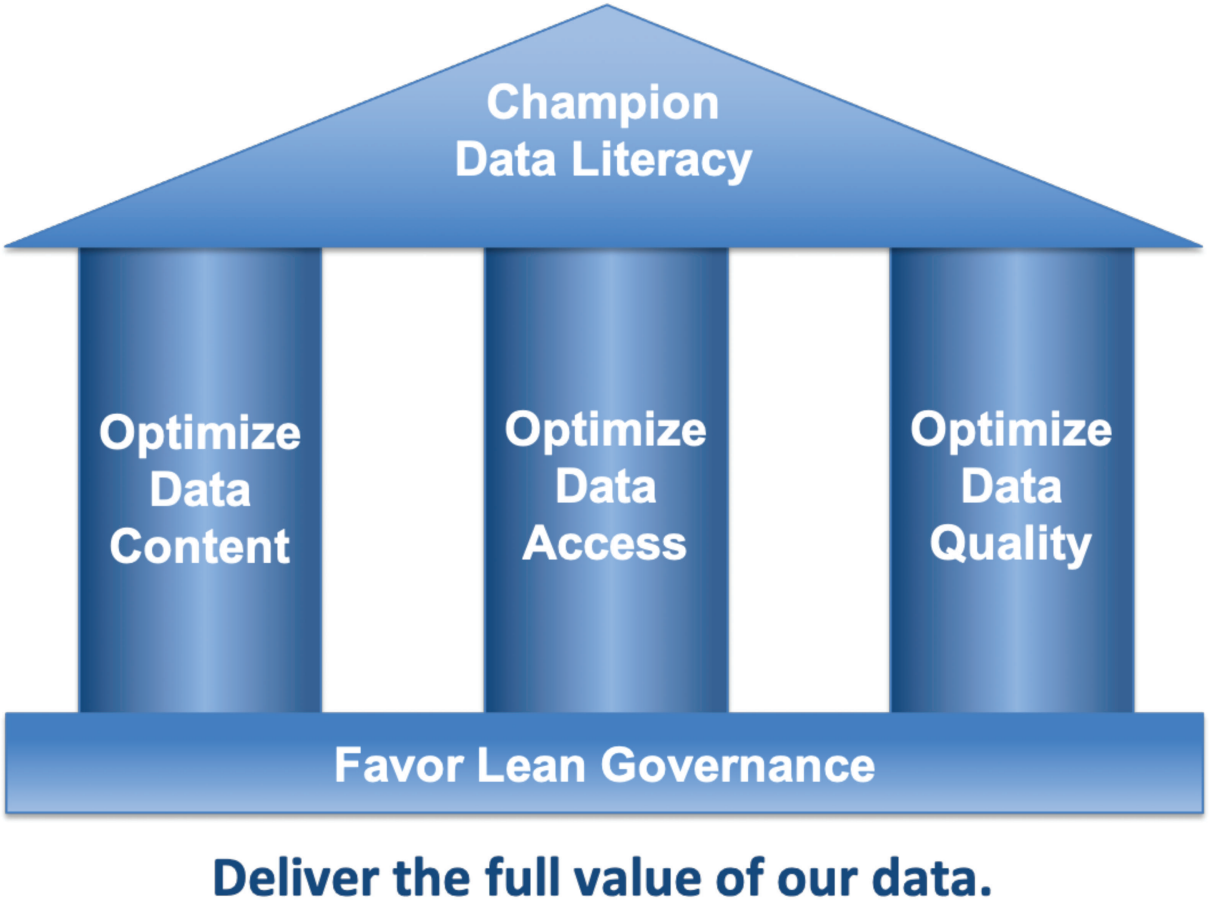

Fig. 2 Data Governance structure.

incorporation, and optimization of new data sources that match our enterprise and data acquisition strategies. Our DG program ensures that the completeness, validity, and timeliness of high value data are an enterprise priority and works to identify and enforce changes in source systems and workflows needed to improve data quality.

The overarching goal of our DG program is to champion data literacy as a core competency among our leadership, providers, nurses and other health professionals, managers, and staff. This will advance the causes of data-driven decision making and process improvement, as well as justify and encourage necessary discipline in data entry.

5. Establish analytics processes to standardize visualization and delivery of data: The selection of tools and establishment of processes includes adopting standards for tools and data visualization; establishing a request intake and prioritization process; monitoring the use evidence-based clinical decision support in the form of alerts, order sets, and templates; and creating a stewardship program for AI algorithms that includes assessing potential bias. It is important to align these tools and processes across the many different business units within a health system, especially in systems that have been utilizing siloed departmental best-of-breed analytics.

Like many health systems, Geisinger had analytics units with management and functionality dispersed throughout the entire enterprise. These siloed solutions obscured the visibility of available analytics; duplicated expert skillsets, analytics technology, and efforts; decentralized operational support for analytics; and resulted in disparate and disjointed reporting with negative impacts on analytics throughout the system.

We decided to utilize the already existing analytics units dispersed throughout Geisinger while providing leadership, coordination, and education centrally (►Fig. 3). The hub serves to establish standards, policies, and procedures for analytics, and contains a centralized site for our analytics inventory, knowledge base, operational support, and

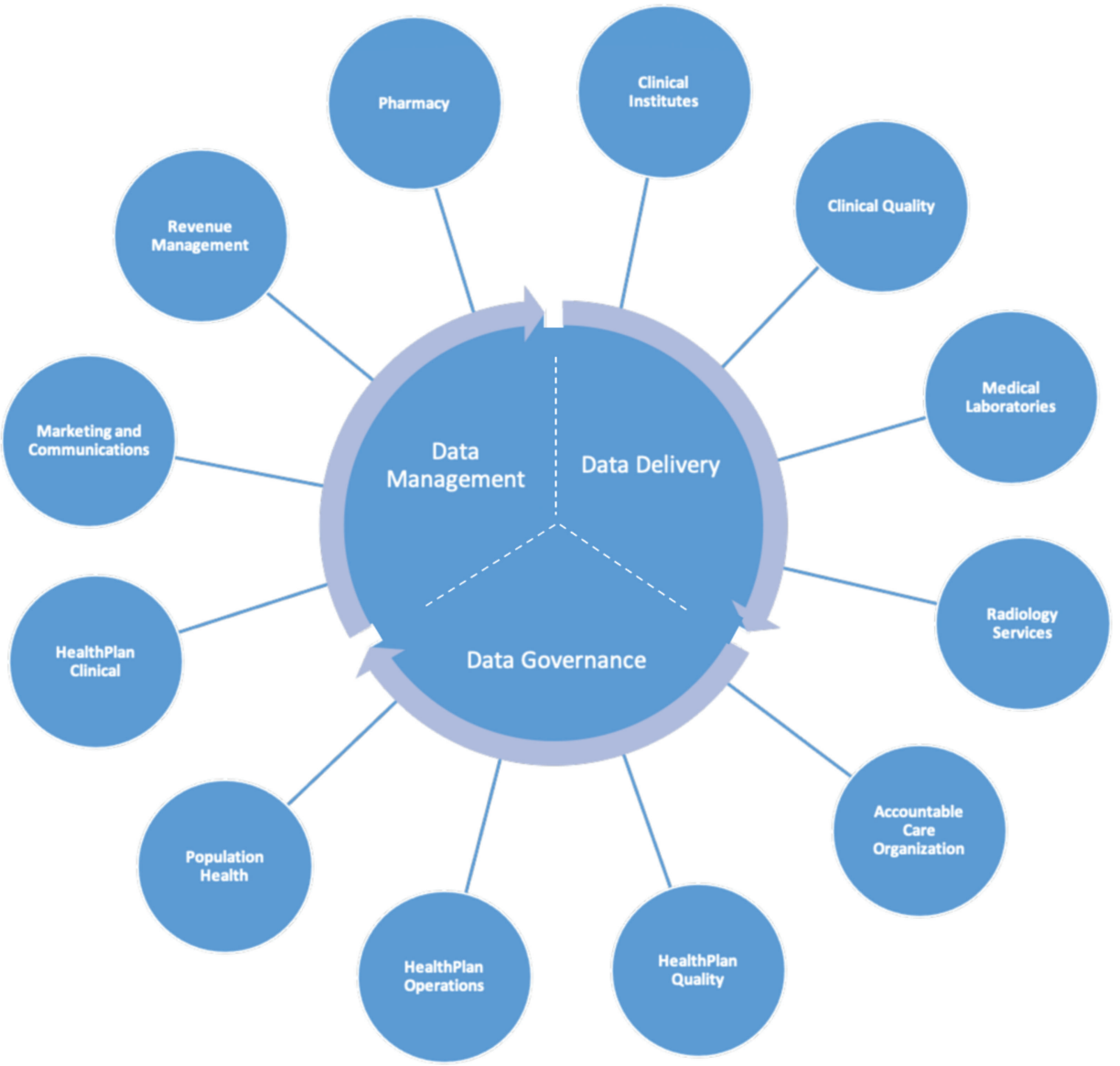

Fig. 3 Hub and spoke model of analytics governance.

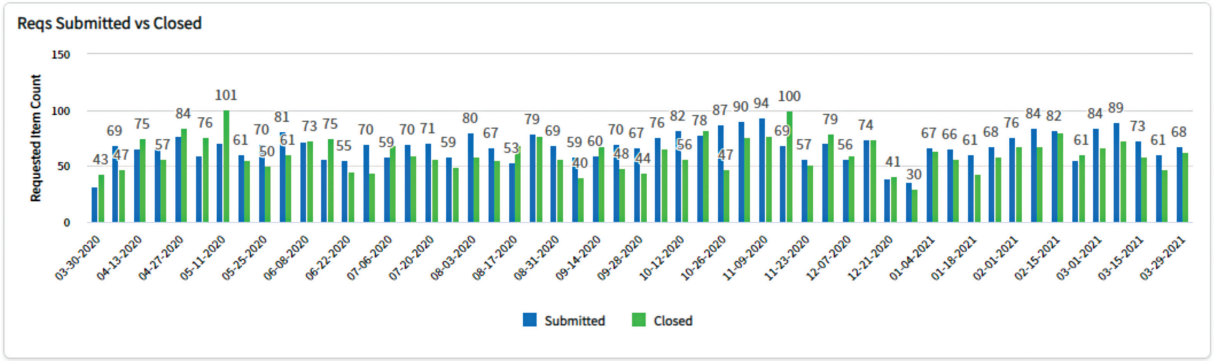

Fig. 4 Analytics requests submitted and closed per week, February 2020 through January 2021.

analytics requests. It provides centralized enterprise skillsets and employee training, education, and certification.

The spokes connect the already existing unit-level analytics functions to the hub. Job roles and the solutions supported are now centralized from the hub. However, each spoke physically remains within their respective business units to maintain effective communications between the spoke analytics teams and their customers. Master Project and Request Lists are maintained at the hub while each spoke has its own lists. Both common core and cross-training of data analysts ensure flexibility in personnel management to meet demand.

We implemented an enterprise IT service management tool to track analytics requests. Incoming requests are prioritized in collaboration with the clinical and operational leaders of the spokes and other stakeholder areas. Completion of requests typically involves an initial consultation and ongoing communications as requirements are formalized and implemented. Communication between analysts and requestors happens electronically, by phone, and in-person. Each step in the process is tracked through the service management tool and associated analytics.

6. Promote the thoughtful implementation and rigorous evaluation of institutional programs and initiatives: The governance process is a living process and the technology and capabilities within analytics are evolving rapidly. It is necessary to have the ability to quickly evaluate the effectiveness of existing and potential new tools, using quantifiable, measurable, previously defined outcomes; measuring their return on investment (ROI); and monitoring any sociotechnical impacts, especially unintended negative consequences, and to initiate rapid changes as needed.

Geisinger's analytics and modeling toolchest consists of a variety of analytics and predictive modeling tools provided by multiple vendors that are available in both on-site and cloud-based environments. Each of these tools has different functionality, visualizations, and ease of user interfaces. By exposing all the data in our infrastructure to the full range of analytics and modeling tools, we can apply the right tool to the right data set for our needs.

Implementation and evaluation metrics are necessary to monitor the effectiveness of analytics governance and to allow adjustments to be made as needed. Categories of metrics that we are utilizing to evaluate the value achieved by the analytics core include clinical effectiveness, efficiency, satisfaction, financial ROI, equity, and dissemination.

Analytics Governance Early Outcomes

The establishment of analytics governance has shown early success but has also identified some challenges. For the first time we can track the initiation and progress of analytics requests at the hub and spoke levels (►Fig. 4). We receive approximately 900 new analytics requests per quarter yet are currently only completing an average of 874 requests per quarter. Roughly two-thirds of the currently active requests are >2 months old, and one-third are >6 months old. The top five requesting groups are the Medicine Institute, Population Health, Pharmacy, Marketing and Sales, and Steele Institute. Like many health systems, we struggled in 2020 with the rapidly changing data requests involving the coronavirus (Severe Acute Respiratory Syndrome Coronavirus 2) pandemic, but discovered that the relationship management functionality of analytics governance kept these data requests to a minimum due to the preemptive building of solutions as a result of open communications between the various stakeholders.

The demand for ever more sophisticated analytics from existing requestors continues to grow, and since our integration with Geisinger Commonwealth School of Medicine in 2017 our medical students and residents are increasingly requesting data as part of required educational projects or research. This observed growth has the capacity to overwhelm our available and recruitable data analyst resources. In response, we have taken initial steps to mitigate the situation. Each request is now reviewed to identify possible existing analytics resources that can meet most of their need without requiring a new report or tool. Preliminary data suggests that as many as 30% of the incoming data requests can be met with an already existing tool. Requests are being classified by type with the goal of identifying classes of analytics amenable for self-service by end users.

The need for self-service analytics aligns well with our already existing informatics educational initiative to increase the efficiency of our providers' use of the EHR for both data entry and knowledge retrieval. A series of EHR efficiency workshops that are specifically focused on specific specialty and subspecialty workflows have been created and delivered to groups of medical students, residents, and

faculty. An Introduction to Data Self-Service Analytics workshop was piloted in March 2021, where the Enterprise Analytics Hub, provider efficiency data and the self-service health data analytics tool contained within our EHR were demonstrated to a group of senior medical students. These students are expected to utilize these tools for self-service before placing an analytics request. In the coming academic year this training will be expanded to include all medical students, followed by residents and faculty in subsequent years. The rate of analytics requests by these learners will be studied to determine whether self-service can help to slow the increase in analytics demands, which should decrease the turnaround time for existing requests and free our data analysts to focus on the more sophisticated analytics for which we require their expertise.

The Enterprise Analytics Hub is another early success of analytics governance. The Hub brings together almost all the available clinical and operational dashboards and analytics reporting tools from across the enterprise and is available to all employees through Geisinger's employee portal (►Fig. 5). Users can browse the available dashboards and analytics reports and request access from a link on the home page of the Hub. Users can identify already existing resources, request a new analytics resource, or a consultation with an analyst. At the current time there are more than 450 unique resources available on the Hub. The creation of the Hub has also enabled us to combine multiple individual reports into single dashboards that enables a one-stop shopping solution containing governed and managed metrics.

Discussion

The development and use of analytics governance to enable and facilitate the effective and efficient use of data and analytics is a relatively new concept in general and has minimally penetrated the health care industry by the early 2020s. Appropriate governance for our data and analytics resources is as necessary as governance around IT.

In this article we have laid out how a large, integrated rural health system has embarked on the early parts of a journey toward the establishment of analytics governance, described some of the initial successes, and identified some initial challenges. We believe some form of analytics governance is necessary for health systems to fully realize the value locked within health data. The exact structure of this governance will vary between institutions and may look quite different from the model described. Analytics governance is not an independent and isolated undertaking but must be built upon supporting structures and integrated with other related governance activities, starting with the strategic vision of the organization and extending through governance of IT and data.

Analytics governance has realized early success and has delivered value to our system. We have made available robust, variable, and flexible sources of data and analytic

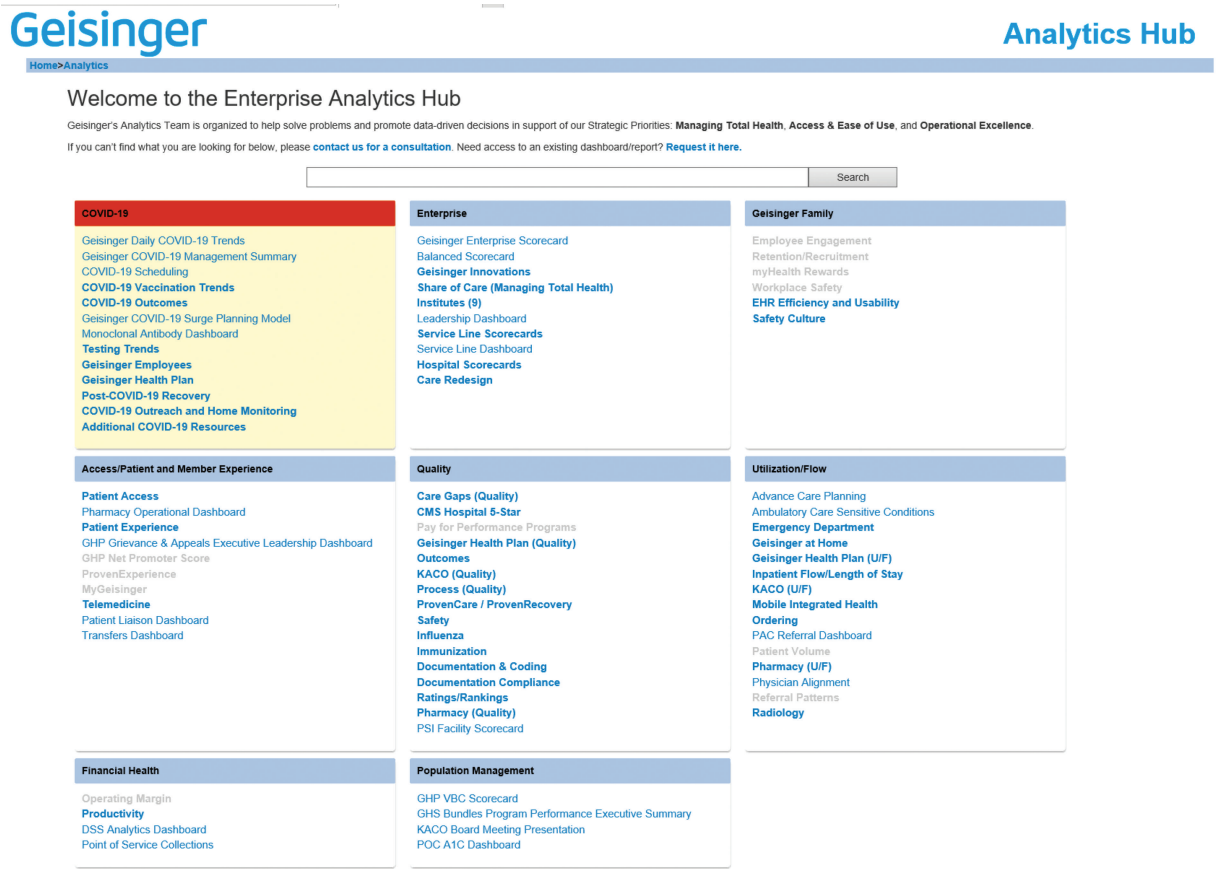

Fig. 5 Geisinger's Analytics Hub.

tools to meet the myriad needs of our integrated health system. Providers and staff are increasingly able to access and use available dashboards and reports, as well as request new analyses if existing sources do not meet their needs. Our culture has evolved to be more aware of and sensitive to data. It has become easier to integrate more complex sources of data and apply sophisticated analytic tools with reasonable expectations of valid results. Duplicate analytics efforts are decreasing, and analytic resources are more appropriately distributed across the enterprise level.

Several important lessons and enablers of success have been learned to date. First and perhaps most important is analytics governance is a long-term and complex undertaking. Without the strong support of leadership and a culture that is or can be aligned around extracting value from data, the likelihood of success is minimal. Having previous experiences with other forms of governance, such as IT, is necessary and good DG is a prerequisite for success with analytics governance. Excellent communications between the different groups of users and analysts are essential to effectively coordinate efforts. However, it is important to recognize that analytics is more a business function than technical competency, and this requires a rethinking of the typical IT analyst–requestor relationship to reflect this reality. It is also critical to identify, measure, and share clearly defined analytics outcomes to demonstrate value. These require flexibility, patience, and a willingness to adjust and adapt as the process unfolds and lessons are learned.

It is likely there are several distinct and isolated data and analytics efforts existing throughout most health systems. These teams can be your most effective allies in implementing analytics governance or can represent strong reservoirs of inertia that can be difficult to overcome. We found it was important to transparently communicate and reassure the existing teams that they would continue to focus on their existing responsibilities while gaining the advantages of a federation of similar teams across the enterprise. Trust in the new structure developed with time and experience.

The demand for data and analytics will continue to grow for the foreseeable future. Our experience has been that the availability of effective analytics dashboards and reports only drives demand for additional analytics of increasing sophistication and complexity. In addition, educational requirements for medical students and residents in the areas of patient safety and quality care are driving demand for analytics among these new consumers of health data. The current generation of trainees are the first to be raised to consume data and are clearly prepared for this challenge.

We have found many analytics requests are highly susceptible to a form of scope creep where a requestor appears to seek additional data or depth to an initial request. This is due to not only the growing recognition of the potential of analytics among end users, but also from a significant communication gap between our requestors and data analysts, especially involving clinical questions. We have found that trained provider informaticians and clinical informatics fellows can provide an invaluable resource in assisting fellow clinicians in crafting requests and helping data analysts in understanding the clinical question and identifying the correct data to query.

One necessary next step in our journey is the development and deployment of self-service analytics capacity. The initial educational foundation is already incorporated into the curricula of our medical school and many of our core residency programs, and needs to be expanded to include existing providers, nurses, management, and staff. Access will be granted to appropriate analytics tools that exist within many EHRs and as independent software. Analytics governance guides decisions regarding available tools and appropriate access.

Other issues to consider moving forward include processes to periodically and systematically review analytics tools to identify issues, flaws, or obsolescence; legal, regulatory, and process issues involving data sharing agreements; issues arising from antidata blocking legislation; addressing increasing cybersecurity threats while providing appropriate access to employees who may be working remotely; and the growing demands on our IT infrastructure to manage this increasing demand.

Conclusion

The demand for increasingly sophisticated data analytics throughout the health care enterprise continues to grow at a rate that will eventually overwhelm even the most highly resourced health systems. The experiences from other industries have clearly demonstrated that businesses who can best manage their data and analytics will have a significant competitive advantage. Our experiences support the proposition that analytics governance can help to manage this demand and provide meaningful benefits to health systems as well. Having a clear vision and plan for the effective use of governance for data and analytics is necessary. Leadership support, transparent communications, and specific measurable metrics are essential elements for success. Analytics governance will provide a proper foundation for the use of advanced analytics, machine learning, and visualization tools, and prepare our workforce to utilize these tools for the benefit of patients.

Clinical Relevance Statement

The article clearly lays out a case for and a pathway toward the establishment of analytics governance in health systems. We have identified six key goals leading to successful governance implementation. Benefits include greater efficiencies, an educated workforce, and ultimately improved patient care.

Author Contributions

All authors have significantly contributed to this work and this manuscript.

Protection of Human and Animal Subjects

No human or animal subjects were included in this study.

Funding
None.

Conflict of Interest
None declared.

References

1 Donovan F. Organizations see 878% health data growth rate since 2016. HIT Infrastructure. May 8, 2019. Accessed March 10, 2021 at: https://hitinfrastructure.com/news/organizations-see-878-health-data-growth-rate-since-2016
2 Dash S, Shakyawar SK, Sharma M, et al. Big data in healthcare: management, analysis and future prospects. J Big Data 2019;6:54
3 Malmasi S, Hosomura N, Chang LS, Brown CJ, Skentzos S, Turchin A. Extracting healthcare quality information from unstructured data. AMIA Annu Symp Proc 2018;2017:1243–1252
4 Lubner MG, Smith AD, Sandrasegaran K, Sahani DV, Pickhardt PJ. CT texture analysis: definitions, applications, biologic correlates, and challenges. Radiographics 2017;37(05):1483–1503
5 Villanueva AG, Cook-Deegan R, Koenig BA, et al. Characterizing the biomedical data-sharing landscape. J Law Med Ethics 2019;47 (01):21–30
6 Kakkanatt C, Benigno M, Jackson VM, Huang PL, Ng K. Curating and integrating user-generated health data from multiple sources to support healthcare analytics. IBM J Res Develop 2018;62(02): 1–2
7 Overhage JM, McCallie D Jr. Physician time spent using the electronic health record during outpatient encounters: a descriptive study. Ann Intern Med 2020;172(03):169–174
8 Aziz F, Talhelm L, Keefer J, Krawiec C. Vascular surgery residents spend one fifth of their time on electronic health records after duty hours. J Vasc Surg 2019;69(05):1574–1579
9 Shanafelt TD, West CP, Sinsky C, et al. Changes in burnout and satisfaction with work-life integration in physicians and the general US working population between 2011 and 2017. Mayo Clin Proc 2019;94(09):1681–1694
10 Furlow B. Information overload and unsustainable workloads in the era of electronic health records. Lancet Respir Med 2020;8 (03):243–244
11 Lin R, Ye Z, Wang H, Wu B. Chronic diseases and health monitoring big data: a survey. IEEE Rev Biomed Eng 2018;11:275–288
12 Weissgerber TL, Winham SJ, Heinzen EP, et al. Reveal, don't conceal: transforming data visualization to improve transparency. Circulation 2019;140(18):1506–1518
13 Laney D. 3D data management: controlling data volume, velocity and variety. META Group Research Note, 6; 2001
14 Gandomi A, Haider M. Beyond the hype: big data concepts, methods and analytics. Int J Inf Manage 2015;35:137–144
15 Murdoch TB, Detsky AS. The inevitable application of big data to health care. JAMA 2013;309(13):1351–1352
16 Insights SAS. Big data analytics: What is it and why it matters. 2021. Accessed March 11, 2021 at: https://www.sas.com/en_us/insights/analytics/big-data-analytics.html
17 Manyika J, Chui M, Brown B, et al. Big Data: The Next Frontier for Innovation, Competition, and Productivity. 2011. Accessed March 11, 2018 at: https://www.mckinsey.com/business-functions/mckinsey--digital/our-insights/big-data-the-next-frontier-for-innovation#
18 Wang Y, Kung L, Byrd TA. Big data analytics: understanding its capabilities and potential benefits for healthcare organizations. Technol Forecast Soc Change 2018;126:3–13
19 Šendelj R. Information technology and information management in healthcare. Stud Health Technol Inform 2020;274:139–158
20 Delen D, Demirkan H. Data, information and analytics as services. Decis Support Syst 2012;55:359–363
21 LaValle S, Lesser E, Shockley R, Hopkins M, Kruschwitz N. Big data, analytics and the path from insights to value. MIT Sloan Manag Rev 2011;52:21–31
22 Wegener S, Sinha V. The value of Big Data: How analytics differentiates winners. 2013. Accessed March 11, 2021 at: https://www.bain.com/insights/the-value-of-big-data/
23 Mikalef P, Boura M, Lekakos G, Krogstie J. The role of information governance in big data analytics driven innovation. Inf Manage 2020;57:103361
24 Avery AA, Cheek K. Analytics governance: towards a definition and framework. Emergent Research Forum. 21st Americas Conference on Information Systems. Puerto Rico; 2015
25 Mehta N, Pandit A. Concurrence of big data analytics and healthcare: a systematic review. Int J Med Inform 2018;114:57–65
26 Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst 2014;2:3
27 Dinov ID. Volume and value of big healthcare data. J Med Stat Inform 2016;4:3
28 Reddy S, Allan S, Coghlan S, Cooper P. A governance model for the application of AI in health care. J Am Med Inform Assoc 2020;27 (03):491–497
29 Ho CWL, Soon D, Caals K, Kapur J. Governance of automated image analysis and artificial intelligence analytics in healthcare. Clin Radiol 2019;74(05):329–337
30 Steele GD, Feinberg DT. ProvenCare: How to Deliver Value-Based Healthcare the Geisinger Way. New York: McGraw Hill Education; 2018
31 McGlynn EA, Asch SM, Adams J, et al. The quality of health care delivered to adults in the United States. N Engl J Med 2003;348 (26):2635–2645
32 Berry SA, Doll MC, McKinley KE, Casale AS, Bothe A Jr. ProvenCare: quality improvement model for designing highly reliable care in cardiac surgery. Qual Saf Health Care 2009;18(05):360–368
33 The Steele Institute for Health Innovation. 2021. Accessed March 11, 2021 at: https://www.geisinger.org/innovation
34 Grossman RL, Siegel KP. Organizational models for big data and analytics. J Org Design. 2014;3:20–25
35 Fadler M, Legner C. Toward big data and analytics governance: redefining structural governance mechanisms. Proceedings of the 54th Hawaii International Conference on System Sciences. 2021. Accessed September 28, 2021 at: https://scholarspace.manoa.hawaii.edu/handle/10125/71311
36 Elouazizi N. Critical factors in data governance for learning analytics. J Learn Anal 2014;1:211–222
37 Kamioka T, Luo X, Tapanainen T. An empirical investigation of data governance: the role of accountabilities. PACIS 2016 Proceedings, 2016;29. Accessed September 28, 2021 at: http://aisel.aisnet.org/pacis2016