Information Governance for Analytics Support: Remember the Life Cycle Component

By Shannon H. Houser, PhD, MPH, RHIA, FAHIMA; Donna J. Slovensky, PhD, RHIA, FAHIMA; and Luona Wang

The volume and types of healthcare information created and captured grow constantly and exponentially, and the business, regulatory, and service drivers for information use continue to evolve as well. Both volume and importance of healthcare information compel robust and sophisticated information management practices and systems to ensure accountability, accuracy, and validity of information as a valuable organizational resource. Much of the value of this resource emerges from an organization's ability to optimize information through sophisticated analytics approaches to inform clinical and business decision support, and as a tool for organizational learning.

At the most basic level, the goal is to secure exactly the data needed—the best and most accurate version—in the most cost-effective manner possible, and to transform that data into the information needed to ground clinical and business decisions. Ideally and simply, data are captured once for use as many times as needed to generate numerous types of information. This requires storing the data in a format and location that allows accessibility on demand. Realistically there is nothing simple about securing, managing, and protecting the data and information resources of a healthcare organization, and transforming an organization's data into useful and reliable information increases in complexity with advances in informatics tools and methodologies. In fact, Big Data is big business, encompassing multiple professions and industries, and consuming vast amounts of money.

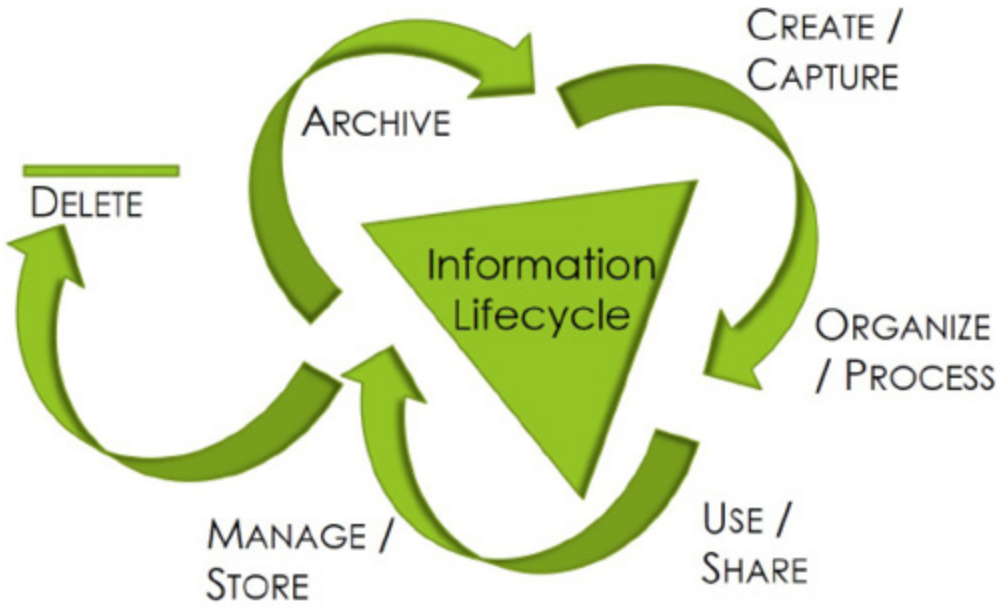
Current best practices in information management favor a governance approach that encompasses an organization's policies, business and clinical processes, and operational practices as well as the technology and infrastructure required to capture, use, and manage healthcare information over time.[1] An organization's information governance plans define a framework for oversight and accountability, and address the myriad problems that challenge information management, such as security and confidentiality, regulatory compliance, legal discovery, storage optimization, data migration initiatives, and others. However, many governance plans address one component insufficiently—the information life cycle. This omission is problematic because availability of current and relevant data is essential to ensuring analytical models produce reliable insights to guide decisions and strategies.

Information Life Cycle Models

Conceptually, the information life cycle is linear, beginning with data creation and ending when the data are deleted—a "birth" to "death" definition. However, the traditional healthcare approach has been to capture and keep all data, migrating it to newer systems and rarely purging any archived data. This practice by default results in a circular life cycle model (see Figure 1 below), with no exit ramp or endpoint to the model as data are never deleted or destroyed by plan. The model is deceptively simple visually, but living it actually is quite complex.

Figure 1: Circular Information Life Cycle Model

Figure 2: Expanded Information Life Cycle Model



The massive volumes of data in healthcare and the business and service drivers for data use compel organizations to employ sophisticated information management systems, most of which have evolved over many years. These systems rarely have been designed in their entirety in response to the organizations' planned information needs and intended use for analysis. Instead, they have emerged incrementally, with each new iteration layered onto existing systems, often migrating old data into newer systems. Sometimes older systems are retained and connections established to new systems, or older systems continue to operate independently of new ones. New information is constantly created and added to existing information, resulting in more information than is absolutely needed or will be used.

Organizations manage an overwhelming volume of data and experience unplanned data redundancy and data inconsistencies through multiple capture points and manipulation. While retention periods are legally defined for many categories of healthcare data, cost-effective digital storage options have allowed organizations to avoid making comprehensive data deletion decisions. It has been easier and affordable to migrate data to new formats rather than to define a data destruction plan and implement it, which also has associated costs.

The model in Figure 2 above is slightly more complex than the first model, and includes a destroy decision, but still omits nuances that are critical to robust information resource management, such as data validation, security backups, and value assessment prior to making the archive or destroy decision. Because current technological capability gives so many options for

cost-acceptable data capture and storage, many organizations haven't been adequately motivated to make structured business decisions from a life cycle perspective that are responsive to information expansion issues. They haven't distinguished among data that are "needed," data that are "wanted," and data that are "just there," and have no sound basis for resource investment decisions made to ensure optimal information accessibility to meet business and clinical needs.[2] Maintaining large volumes of data not verified for currency or relevance challenges the optimal deployment of analytics tools. Thus, skilled informatics and analytics expertise are needed to select the "best" data from among redundant sources to inform analytical models, and more time and financial resources are required to clean data prior to analysis.

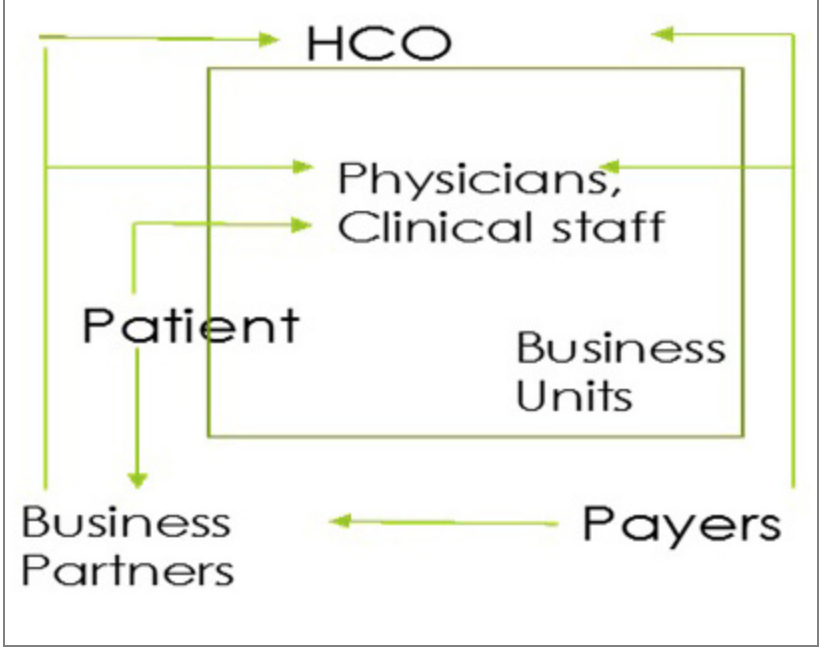
Information Life Cycle Governance

Considering the information life cycle from its definitional viewpoint as linear, incorporating components of the governance model, and involving the right stakeholders in planning can yield a robust enterprise data strategy to support clinical, business, and analytics needs. The groups below are among the important stakeholders to consider when defining the life cycle ending point for various data maintained for clinical, business, and analytical purposes:

- Business units (legal, human resources, finance)
- Research and analytics units
- Information security, privacy, and compliance
- Contract management
- Records management
- Information technology
- Patients/consumers
- Business partners

In addition to noting their primary environment as internal or external to the organization, relationships among the various stakeholders are important to consider as well. Mapping the key groups and relationships among the groups, such as the example in Figure 3 below, can be useful in setting priorities and making decisions about competing data requirements.

Figure 3: Stakeholder Model



Implementing an information governance model that incorporates and enforces the life cycle component can help organizations avoid two undesirable inverse data relationships:[3]

- Relevance and efficiency—the more information acquired and stored the less value it has if utility and access are not ensured.

- Time value of data—healthcare data have an important and complex time relationship. For many applications, the older the information the less value it has for decision support.

Some analysts suggest that information growth has reached a tipping point—information growth exceeds organizational budgets and processes for managing and governing that information.[4] Since it is unlikely that healthcare organizations will stop generating data, managing this imbalance will require more aggressive deletion and selective archiving practices as key components of the enterprise data strategy and governance plans.

Achieving optimal information resources to support an organization's clinical, business, and analytics needs requires commitment and collaboration throughout the organization, beginning with leadership and cascading down to the lowest information user level. At its core, the primary goal of such a program should be "defensible disposal" of data based on assessment of legal and regulatory guidelines, as well as the information's value to the organization in consideration of business objectives, clinical needs, and goals for organizational learning through analytics.[5]

While "defensible disposal" cues to legal incentives or mandates for retaining data, some business analysts propose that only one percent of data generated is actually subject to legal hold, and only 25 percent has ongoing business value.[6] While these numbers likely differ significantly when considering healthcare data, it's evident that great volumes of "digital debris" are currently stored and maintained at significant organizational cost in all industries. In addition to the financial costs, the loss of business efficiency and compromised data quality result in other costs that are less easily quantifiable. An attentive focus on defining information life cycle termination points will enhance the effectiveness of an organization's information governance programs.

Notes

1. Empel, Sofia. "Way Forward: AHIMA Develops Information Governance Principles to Lead Healthcare Toward Better Data Management." *Journal of AHIMA* 85, no. 10 (October 2014): 30-32.
2. Willig, James H. "The many lives of data." *Handbook of Healthcare Management*. Cheltenham, UK: Edward Elgar Publishing, 2015.
3. Kahn Consulting, Inc. "Information Lifecycle Governance." Information Governance Brief.
4. Paknad, Deidre and Rani Hublou. "Information Lifecycle Governance Leader Reference Guide: A Model for Improving Information and eDiscovery Economics with Information Lifecycle Governance." CGOC. 2012.
5. Ibid.
6. EDRM. "Disposing of Digital Debris." April 22, 2014.

Shannon H. Houser (shouser@uab.edu) is a health services administration/health informatics associate professor at the University of Alabama at Birmingham. Donna J. Slovensky (donnaslo@uab.edu) is a professor and senior associate dean for academic and faculty affairs at the University of Alabama at Birmingham, School of Health Professions. Luona Wang (luonaw@uab.edu) is a MSHA/MSHI graduate student at the University of Alabama at Birmingham's health informatics program.

Article citation:
Houser, Shannon H.; Slovensky, Donna J.; Wang, Luona. "Information Governance for Analytics Support: Remember the Life Cycle Component" *Journal of AHIMA* 88, no.6 (June 2017): 38-40.