CHAPTER

5

HEALTH INFORMATION TECHNOLOGY INFRASTRUCTURE, STANDARDS, AND SECURITY

Learning Objectives

1. Define and use in context the technical terms related to information technology architecture and infrastructure.
2. Distinguish between the hardware and software components of an information system and provide illustrative examples.
3. Discuss basic telecommunication concepts.
4. Describe data storage options, discussing considerations, advantages, and concerns associated with each option.
5. Discuss data transaction types subject to electronic data interchange regulations.
6. Identify data standards organizations with influence in the healthcare sector.
7. Discuss provisions of the Health Information Portability and Accountability Act Privacy and Security Rules.

Overview

Infrastructure, broadly defined, consists of all components of an enterprise's information technology (IT) resources, including not only physical elements such as hardware and other equipment, networks, and data centers, but also software, operational and governance policies, and contractual relationships with vendors and partners. This superficial definition does not convey the scope of a typical system's components or the complexity of designing and managing a dynamic conglomerate that enables the enterprise to conduct its myriad business and clinical activities. These activities must be compliant with extensive government and industry regulations, incorporate numerous technologies, and ensure that clinical services are safe and effective. And all this has to happen under value-based business models driven by third-party payment regulations that constrain revenue enhancement opportunities.

While designing and managing the inherent complexities of information systems are mercifully the purview of a team of IT specialists, healthcare managers need conceptual understanding of system components, network structures, standards and regulations, security risks, and trending issues in health information technology (HIT). A basic level of knowledge is essential to effective participation in HIT budget development, negotiating system contracts, ensuring regulatory compliance, and assessing enterprise risk associated with information system policies and practices. As noted in chapter 1, managers, clinicians, and technicians possess differing bodies of knowledge relative to operational needs, system use, and technical design, and all viewpoints need to inform decisions about the enterprise information system.

Our world has become "connected" in the literal sense of the word. Approximately 95 percent of Americans own a cell phone, and most (77 percent) are smartphone users (Pew Research Center 2019). These smartphones provide constant personal communication opportunities through email, texts, and social media; instantaneous access to a seemingly infinite amount of information via the internet; and numerous apps to manage daily life activities. Things that once could be accomplished only at a hardwired desktop computer can now be done via a smartphone while riding a bus or standing on a beach. This ubiquitous connectivity exists in business computing as well. Information technology in healthcare enterprises, once deployed primarily as stand-alone applications in individual operating units such as the human resources department and the business office, is now conceived as a seamless integrated system of physical and virtual connections to devices that provide access to the full range of information needed for business and clinical operations—but a system that is secure and protected from unauthorized access. In fact, because of Bluetooth and other wireless technologies, very few data devices, even personal devices, used in a healthcare enterprise are not connected (or capable of connecting) to the enterprise information system.

The basic system components of first-generation computers (input, processing, then output) remain relevant, but the ways in which these actions occur have become more numerous and more sophisticated than were possible even five to ten years ago, as each generation of technology eclipses the last more quickly than previous generations evolved. This so-called law of accelerating returns (Kurzweil 2001) suggests that learning from one innovation informs future innovations for faster development, generating an exponential rate of return with regard to human-created technology.

Computing advancements are an excellent demonstration of this law. Computers entered our world in 1946 with the Electronic Numerical Integrator and Computer and its vacuum tube technology (Rosen 1969), moved through second and third generations (transistors and integrated circuity) in the 1950s and 1960s, and had evolved to a fourth generation (microchips) by

the early 1970s. Innovations in the second two generations decreased computer size, increased processing capacity and speed, and improved the user interface. Because these innovations also made computer technology more affordable, these advancements significantly increased the use of computers for business applications. However, the innovations of the fourth generation have literally integrated computing and its technology into our daily lives. The ability to network computers and devices, the emergence of the internet and World Wide Web, and the increasing cost-effectiveness of computers for personal use converged to spur innovation in technology and applications at an unprecedented pace.

One result of this rapid innovation and explosion of applications for healthcare is that the "schematic" of an enterprise information network has become more difficult to illustrate graphically as system components become increasingly distributed remotely and virtually, and third-party entities become major contributors to the governance of the information infrastructure. Thus, recognizing that actual infrastructure components will continue to evolve as new and improved technologies emerge at an increasing pace, this chapter will describe currently employed configurations, technology trends, and emerging HIT issues that healthcare managers should be conversant in and that should enable them to continue learning as their enterprise systems evolve in response to continued innovation, business needs, and regulatory guidelines.

Information System Components

The most basic tangible elements of an information system are hardware, software, data storage, and connections among the system components, each of which can be further categorized and described in terms of their functionality and integration. Simplistically, a computing system comprises these components:

- Input devices
- Processing unit
- Output devices
- Primary storage and secondary storage
- Communication devices

The communications devices create connections that enable the computer to interact with other computers or devices, either in or outside the organization. The ability to connect multiple devices that work collaboratively to complete a work process gives rise to the concepts of networking and telecommunications.

Computer hardware, the physical components and devices configured into an information system, comprises input and output devices, processing units, and storage media. Computer hardware spans incredibly broad spectrums of size and function. Some devices are small enough to be held in and manipulated with one hand. For example, a large segment of the general population enjoys personal computing with such devices as a tablet, notebook computer, or smartphone (or a smartwatch), and even these devices may be used for data capture, data processing, and output, and connect wirelessly to the internet or other devices.

A typical outpatient healthcare organization will have at least basic servers, which can store text files, and other servers may contain radiological or other diagnostic images. In a large integrated delivery network composed of multiple hospitals and outpatient facilities, the norm is more likely a dedicated floor or building filled with numerous servers with specified functions such as web servers that connect users to web pages, mail servers for storing email and email account information, and dozens of other file servers. Many diagnostic machines are computers that analyze clinical samples, process data, and produce reports. At the upper extreme of the computing spectra are large and powerful supercomputers.

Hardware is useless without programming or application software, which allow computers to perform specific functions. Software includes system operating instructions and applications that perform tasks, such as nursing documentation or billing. Applications to support healthcare enterprise computing are addressed later in chapters 8 and 9.

A key point for managers to remember is that software and hardware decisions must be made in tandem; one cannot be considered independently of the other. For example, a software application used by physicians may require a large amount of cache storage or a high processing speed to operate effectively, both of which are hardware factors. Healthcare managers need an understanding of basic software concepts to be knowledgeable participants in the complex processes of selecting, implementing, and testing software to maximize the value of their HIT investments. Knowledge needs include an understanding of the purpose and functionality of clinical, business, and communication application software; an awareness of the distinction between integrated and interfaced systems; a recognition of the role of system management software; and a general comprehension about programming languages and language translators.

Data storage options range from small independent devices (such as a thumb drive that you might use to store and transport a PowerPoint presentation) to large data warehouses that store millions of discreet clinical data elements, accessible by approved users in diverse locations simultaneously. Each storage option has unique security issues, some attributable to physical characteristics (e.g., small size that allows easy theft), and other risks that

emerge from how and by whom the data are accessed. Physical and virtual connections, which transform independent devices into an integrated network, are arguably the most complex and dynamic of these components. Let us consider these components by looking at their functionality.

Input: Capturing the Data

The power of an information system can be realized only when data and programs have been entered for processing and information is generated for the user. System designers can select from among multiple input options to meet the organization's needs for speed, accuracy, and cost-effectiveness for a given application. As technology advances, new modes of data input emerge, but few have become obsolete.

Although the keyboard remains a frequently used input device, healthcare organizations have found that other input devices are especially suitable for specific applications. For example, scanning devices provide an efficient and accurate means for tracking many types of inventory items, locating paper documents, and even identifying patients via bar code technology. Medical supplies, pharmaceuticals, and patient identification bands may be tagged with bar codes or graphical markings that perform several functions when scanned and recognized by the computer software. For example, a medical supply item or drug may be removed from current inventory, charged to a patient's account, and scheduled for inventory replacement with one simple scanning process. In addition, the computing skills and time constraints of the staff members who will enter the data may be a consideration in choosing the input approach. For commonly performed tasks such as medication orders, busy clinicians often respond better to data-entry options that are highly automated, such as prepopulated entries chosen from a drop-down menu, rather than to keyboard entry, which requires some skill, allows for more errors, and perhaps takes more time than is desirable.

Physicians may order diagnostic tests or medications simply by touching the monitor screen where a list of options is displayed. While discrete data are preferred for reporting purposes, scanning handwritten documents may make the information available to more users much sooner than if the document is audiorecorded and transcribed through keyboard entry. Scanned graphical material, such as electrocardiogram reports, can be accessed online by users in multiple locations, unlike hard copies stored in a single location, which requires users to travel to the storage location or the document to travel to the user.

Selection of the best input device for a given application should consider both efficiency and accuracy criteria. While speed of input provides user convenience, which is important to time-pressured clinicians, speed should not be gained at the expense of data quality, patient safety, and information confidentiality. Clinicians must be able to trust the accuracy of electronic

data, and the quality and resolution of captured images must be adequate for visual recognition and interpretation of clinical data.

In the early phases of healthcare computing, data entry typically occurred at centralized locations, such as nursing stations or dictation rooms. Today, information systems are designed to facilitate data capture at the point of care, such as the patient's bedside or in other diagnostic or treatment areas. Often, data are captured concurrently with patient examination and treatment (point of care), through voice recorders, medical scribes, or digitally enhanced diagnostic devices. Areas such as the emergency department may use medical scribes, individuals who observe treatment and document clinician dictation and diagnostic test data in real time. Data also may be entered using computer workstations in or near the patient's room or by using a portable or handheld device that connects the user to the electronic health record (EHR) system. This ability to perform point-of-care data capture is mandated by regulations specific to reimbursement by Medicare that require hospitals to implement Certified Electronic Health Record Technology (CEHRT). Exhibit 5.1 summarizes commonly used data input devices.

Processing: Converting Data to Information

The hardware components of even the most powerful supercomputer cannot by themselves produce output that is of value to the healthcare manager, because they need a detailed set of instructions that describe, step by step, the tasks that must be performed to achieve a desired objective. This detailed set of instructions is known as a *program*, and programs are collectively referred to as *software*.

Although for many people software is equated with user applications (either general purpose or function specific), computer software also includes operating systems, utilities, programming languages, software development tools, and language translators. The *operating system* (OS) is the interface between the human user and the computer, managing the functioning of the software and hardware. Most people are familiar with the Microsoft Windows OS or the Apple iOS for personal computers, and Linux, which is open source. *Utilities* software performs general processing, computational functions, or system maintenance functions. Virus scanning and encryption are examples of utility software.

All software—application, operating system, or utility—consists of a detailed set of instructions describing the specific steps the computer is to perform. Processing instructions are communicated to the central processing unit (CPU) in a structured programming language, which has evolved over time from binary code (0, 1) to instructions resembling spoken language. Examples of programming languages include BASIC, COBOL, and Java, all of which have rules and context frameworks. Despite the number and type of programming languages in existence, the objective of all languages from the

EXHIBIT 5.1
Input Devices

Device	Description	Advantages	Disadvantages
Keyboard	Device containing a panel of "keys," including alphabetic and numeric characters and special function keys	Familiar, inexpensive, rapid entry (when done by skilled users on full-sized boards)	Poor keying skills result in data-entry errors; smaller boards on handheld devices may be difficult to use.
Pointing device (mouse, joystick, rollerball, light pen, touch screen)	Device that controls the screen cursor (locus of data entry); the pointer may be a finger or a special device	Easy to use, rapid data entry	Precision in pointing is required to avoid data-entry errors.
Scanning device (bar code reader, optical mark/character reader)	Device that captures data by reading differences in light reflection between the mark and the white space	Rapid data entry, good error control, useful in tracking systems	Limited amounts of data are captured; most data cannot be manipulated after entry.
Handwriting recognition device	A stylus or other device used to write data on touch-sensitive screen; may be optical scan of writing on paper	Familiar skill, no training required	Handwriting must be intelligible.
Image capture/video input (computed tomography, magnetic resonance imaging, webcam, digital camera)	Device that captures digital images, which are then stored in a system-defined format	Particularly useful in telehealth and diagnostic applications	Devices are relatively expensive; image resolution must be precise for diagnostic and other healthcare applications; file size can be very large.
Voice input technology	Microphone used to enter data and instructions; software program converts spoken language to machine language by digitizing sound waves	Technical skills not required	Devices are relatively expensive; machine must "learn" user's voice pattern and pronunciation; vocabulary must be built.

user's perspective is simple. The overarching goal is to communicate with the computer in some prescribed format so that useful output can be generated. Whereas skilled programmers may find reward in creating complex code, for the nonprogrammer user, the satisfaction of this communication process lies in the output created, not in the communication process itself.

The progression of programming languages can be tracked through successive generations, with each generation improving the computer–human interface. The evolutionary goal is to achieve *natural language* input, whereby the user is able to give verbal commands to a computer as easily as communicating with another—or as easily as you tell Alexa or Siri via your tablet or smartphone to add an item to your grocery list or play a specific song! A *language translator* program would convert natural language statements into the binary number commands a computer understands. The technology necessary to recognize spoken words, interpret their content, transform them into a set of procedures, and translate this sequence into machine commands is complex and has not been perfected. While various voice-recognition applications are available for business or personal use, acceptance for medical applications is not universal, although innovation continues apace.

Software issues are important for healthcare managers to understand for a number of reasons. First, although most healthcare organizations do little in-house development of software, the manager must be a knowledgeable participant in software acquisition. Managers must acknowledge that the quality of available software is variable, and in some cases software purchased at significant expense fails to meet expectations.

Second, all software must be appropriately licensed and used only as specified by the license. Users may want to install personal applications on facility computers. While these issues can be controlled to a large degree by system security configurations, policies should be in place that emphasize the organization's strong stance on exclusive use of legally licensed software and facility ownership.

Third, managers should be aware of the rapid evolution of software versions. Operating systems and application software are constantly being revised. Generally, upgrades to major systems come with a cost, and sometimes that cost will exceed the value of the improved functionality, if the current version is meeting their needs. In other cases, the vendor might actually cease to support a given version, thereby forcing the user to upgrade. Again, knowledgeable participation by the manager is valuable in making upgrade decisions.

Finally, and perhaps most important in the current technology environment, the manager must understand the challenges created by the need for interfaces that link disparate software packages and system components.

Upgrading one module of an interfaced system may require extensive modification of the interface as well. Increasingly, there is a need to connect facility applications with those hosted by other providers in the continuum of care or with enterprise partners. A simple example is the electronic transfer of a patient prescription to an external pharmacy, a procedure that may contribute to patient satisfaction.

For stand-alone computers, the CPU is where the actual "computing" takes place. The speed and power of the CPU greatly influence the computer's capabilities, and each generation of computer technology has increased the processing capability while reducing the size of the processing component. An important by-product of processor evolution has been the increased speed of processing. Thus, processors are now smaller, faster, handle more volume, and cost less than their predecessors. You will recall we are now in the fourth computer generation (microprocessors embedded in microchips), enabling very small devices to have exceptional computing capacity.

Distributed processing, combining multiple processors either in one computer or across multiple networked computers, increases processing speed and computing power even further. A single computer may be configured to employ more than one CPU, but a more typical option is to connect multiple computers into a network. A local area network (LAN), sometimes referred to as an *intranet*, connects computers and peripheral devices, usually within a defined entity, such as a building or organization. The connections may be hardwired or wireless, or a combination. The intent of a LAN is to share resources such as software or output devices, facilitate data transfer among users, and provide internet connection.

A *wide area network* (WAN) extends this connectivity to a larger geographic region using multiple telecommunication networks. WANs may be private, such as those restricted to a business enterprise, or public, connecting many networks (LANs) together. The internet is a worldwide WAN. In today's world of "big data" and all-pervading connectivity, distributed processing (or distributed computing) takes on a complex reality. Simplistically, distributed computing systems' components cross multiple networks, and resources and information can be shared among an infinite number of users through communication linkages. The internet is embedded in our social and economic structures such that we take this extreme degree of connectivity for granted as the way we do business and manage personal affairs. A very large volume of healthcare business and information transfer is conducted via the internet. The relative ease with which data move across the internet and its widespread acceptability are undeniable. However, the inherent open access that supports this convenience is not without risk, as data security is a significant challenge that must be addressed when designing internet-based business processes.

Output: Making Information Available for Decisions

Accurate, comprehensive data are required to produce the information clinical and administrative decision-makers need, and the ability to access the information at the time it is needed is crucial. The actual work performed by the computer system is of little value until it is produced (output) in a usable format accessible to the user, such as in print or as a screen image, digitally for additional processing, or in audio or spoken form. An important goal of the IT industry is to make both data entry and retrieval as simple as possible.

Types of output of particular value to healthcare managers include visual displays, printed documents, and audio (including voice) output. The oldest and still most widely used form of displaying output from an information system is a video display screen, typically called a *monitor* for stand-alone devices or a *screen* on a handheld device. Technology has advanced from small monochrome screens into large (or very small), high-resolution liquid crystal displays (LCDs) that can be enabled for data entry by touch, thus serving as both an input device and an output device. These sophisticated monitors can display images at resolutions high enough to support clinical diagnosis and treatment. Where processing devices have evolved to smaller sizes, monitors have moved in the opposite direction for desktop and conference room use. LCD monitors come in more than a dozen sizes and vary in resolution and pixel density. Two or more monitors can be connected to a single computer to allow simultaneous viewing of data from multiple applications. Although most users prefer larger, higher-resolution monitors, purchasing decisions should be based in part on the applications to be used on the system and the data to be displayed on the monitor. For example, monitors to be used for image display need higher resolution than do monitors used for text processing. Individuals who are coding medical diagnoses and procedures may need multiple monitors, or very large split-screen monitors, to access multiple applications simultaneously. If the processing output is intended to meet the needs of a mobile user, the built-in screen in a smartphone or other handheld device serves that purpose. Again, the resolution and image quality needed must be considered when determining the type of device to host the application.

Printers, too, have developed extensively from the early devices that were similar to typewriters, except they printed on track-fed continuous paper rolls. Today's color laser printers are capable of reproducing artwork, photographs, and detailed diagnostic images. These machines can print in a variety of sizes and on multiple grades of paper, cardstock, and other media. Most recently, the addition of three-dimensional printers, designed to add material layer by layer to create a three-dimensional object, are used in medicine. These "built" items may be models of organs created from a patient's imaging data, assistive devices such as hand splints, or other items from an

ever-growing list of possibilities. Photocopying machines now multitask as printers, and high-resolution printers are available for lease or purchase at acceptable costs. In fact, for many low-end printers, the cost of color ink cartridges compared with printer cost may cause a user to question whether the printer is the disposable item. Key printer characteristics to consider in lease and purchase decisions include memory capacity, print resolution, and print speed. Networking and wireless technology enable a single printer to service multiple computers and the many users who may share access to those computers. Thus, as with monitors, decisions about printer selection can be based on users and applications served rather than on cost alone.

As technology has enabled digitization of sound with good quality, audio output has become a more viable option in clinical technology applications. When digital text is converted to understandable speech by voice synthesis, an ordinary telephone can be used to access healthcare information. For example, a physician needing a patient's diagnostic test results could use a telephone to call the laboratory or radiology system and hear the results read by a voice synthesizer. Clinicians also can listen to body sounds, such as breathing or heartbeat, from distant locations using a telephone or other audio-transmitting devise. This ability allows expert consultation without patient travel or monitoring of homebound patients with chronic conditions. Collectively, these types of applications are referred to as telehealth, or mHealth.

Storage: Archiving for Active Use or Mandated Retention

Important decision factors for selecting or designing data and information storage for healthcare enterprises are volume, physical security, disaster recovery, and expansion planning. These points are discussed briefly here, but the savvy manager will use a just-in-time approach to explore relevant issues and available options in greater detail when the need to apply these concepts is relevant to the job.

Operational definitions for primary and secondary storage have evolved from early computing days when *primary storage* meant the data were stored on the computer's internal drive for access by the CPU, and *secondary* meant data were stored on external media. Currently, *primary storage* is the label attached to those repositories used for transactional data that are frequently accessed for business and clinical purposes, and *secondary storage* refers to repositories with an archival orientation for data accessed infrequently or not at all. Thus, the distinction between primary and secondary storage is based on data access or use frequency rather than the storage medium or the storage location. Data may be archived for anticipated future uses, such as clinical or business research or trend analysis. Data also may be archived to comply with state or federal mandates to retain business and

clinical records for specified periods. Most healthcare organizations have a record-retention schedule that specifies the length of time categories of records must be kept and the date such records can be destroyed as part of their data governance plan. Unfortunately, considering the costs and challenges of data storage and the risks of unauthorized access, many organizations opt to retain records indefinitely rather than implement the destruction schedule (Houser, Slovensky, and Wang 2017).

The goal for data utility, efficiency, and cost-effectiveness is to capture data once and store it in a single location, and to have the data from that location available as needed by any application or user. Replicating data for storage in multiple locations is undesirable for several reasons. First, capturing or inputting the data multiple times is an unnecessary expense. Each unique data capture or entry has an associated expense. If it costs $1 in personnel time to enter a birth date, and every patient's birth date is entered three times, with an annual inpatient census of 100,000, $200,000 would have been spent on the two unnecessary birthdate entries. Second, multiple captures or inputs create opportunities for increased data errors, as every entry poses independent risk of error. Third, if differing data formats are used, the data may not aggregate correctly when files are merged across applications. Data stored in multiple locations may change as data are manipulated, updated, or edited for differing purposes by the various users—which can pose legal risks if the appropriate data are not accessed in response to a specific inquiry. Thus, drawing data from one location may not inform decisions in the same way as drawing the "same" data from another location.

Key issues with regard to data storage include data classification, media used, location, cost, and security. Data classification assists in data management by categorizing data into types that have similar requirements for selected attributes, perhaps for security, regulatory compliance, or processing needs. Storage options, discussed in the following section, are numerous, but should be deliberately chosen to meet access and security needs. While decisions about type, location, and control of data storage are important from financial and data security perspectives, the real value of the repository—and thus the pivotal decision factor—lies in the accessibility and utility of the data housed inside.

Storage Options

The actual storage required for captured and archived data in a healthcare enterprise is massive, and the associated costs are a significant component of the total HIT cost equation. How much storage required for a given application is dependent on the type and volume of data captured, access and retrieval requirements, and retention requirements. For example, are data in text or nontext format? The size of image files and other nontext files is a significant contributor to the total volume of archival storage space

required, as nontext data require significantly more storage space (*Journal of AHIMA* 2011).

Many other questions must be posed to inform decisions about storage needs. For diagnostic images, how many procedures are performed in a year? How many years must original images be maintained? Must the diagnostic image and the clinical interpretation be maintained in the same file? Must previous diagnostic data be accessible for comparison with current diagnostic tests? Must the data be accessible in real time for an extended period, or can the data be archived quickly with minimal access requirements? These and other questions are paramount to establishing system data storage requirements.

In addition to nonvolatile storage options (those that retain data permanently), cache or active memory storage requirements for data that require rapid access and manipulation in real time can be extensive. When evaluating certain types of systems, cache capacity can become the determining factor in selecting one vendor product over another. For application-specific parameters, such as cache requirements for image viewing, including clinician users on the product evaluation and selection team is extremely important. If a technology solution does not produce quality data that can support clinical decision-making, the solution is insufficient. With issues of clinical adequacy, patient safety, and patient satisfaction, product cost is rarely the deciding criterion.

Perhaps the simplest and most controllable electronic data storage option is *on-premise hardware-based storage*, where data are housed on hard disks in arrays of network servers. While the technical points are beyond the intent of this discussion, the types of drives used in the array are determined by the type of data stored (structured or unstructured), with hybrid arrays addressing both cost and performance needs. A disk array is scalable, providing data storage and access for connected devices in the enterprise network. *Off-premise storage* may refer to a remote data center owned and managed by the enterprise, or a hosted solution outsourced to a vendor that provides data management services to meet enterprise needs. Generally, off-premise storage configurations will resemble those used on premise; the service is just in a location remote to users. However, this situation is changing as more enterprises adopt cloud storage or hybrid solutions of physical servers and cloud storage.

Cloud storage refers to an off-premise, distributed storage model where data are stored on the internet, generally through a contractual fee-for-service arrangement with an external vendor. The cloud service may be private or public. A *private cloud* is based on an IT infrastructure dedicated to a single enterprise, and offers greater security and control than a shared, public cloud service such as Amazon Web Services or Google Cloud. The technology may be owned and controlled by the enterprise, or a vendor may

provide dedicated network resources under a lease or contract arrangement. Healthcare enterprises, whose dynamic computing requirements are coupled with stringent security regulations, may find that private clouds meet those dual needs. A *public cloud* allows distribution of data over internet servers shared among multiple users. These virtual servers are accessed through an online interface. One advantage of public cloud service is that costs may be scaled by usage volume rather than a fixed price for the service. Data also may be distributed across multiple cloud providers for geographic benefits or to segregate data with differing access or security needs.

Storage Expansion Planning and Data Governance

It is a certainty, equally as profound as death and taxes, that the volume of data produced by healthcare enterprises will only increase. The corollary of that certainty is that data integrity and privacy and security regulations for archived health information will not lessen. Technology capabilities will continue to evolve, the types of data that can be captured will expand, and the storage media employed will change. All of these changes will occur rapidly and successively. A continuing challenge for information resource managers will be to ensure that previously archived data and information can be migrated to emerging storage media with no loss of data integrity, irrespective of current IT infrastructure.

The inevitable and constant acquisition and production of data in healthcare enterprises necessitates managing data purging and destruction as well as ensuring adequate storage capacity for archived data. As storage costs for many options have lessened, some managers have found it easier to expand storage capacity than to design and manage a data governance plan. This avoidance technique results in a circular information lifecycle model with no "death" (deletion or destruction) component (Houser, Slovensky, and Wang 2017). A robust approach to information governance is needed and should encompass organizational policies, business and clinical procedures, technology and infrastructure, and a well-defined accountability framework (Empel 2014).

From a financial perspective, one might consider that unimpeded growth will soon subsume more of an enterprise's IT budget than can be reasonably allocated to managing a vast amount of data that has no value to patient care or to business operations. Resources encumbered to manage data resources should be based on sound judgments that ensure accessibility of useful and reliable information to meet business, clinical, and analytical needs (Willig 2015). Thus, healthcare enterprises need to make deliberate distinctions between data that have ongoing utility or must be retained for regulatory compliance, and data that are retained as a result of insufficient data governance. However, selective archiving and destruction of data should be based on legal and regulatory guidelines to ensure defensible disposal (FTI Consulting 2015).

A well-documented data governance plan is important to ensure that data are maintained in accordance with business and clinical needs, securely protected to maintain patient privacy and meet regulatory requirements, and properly destroyed at the terminal point of their life cycle. The American Health Information Management Association (AHIMA) offers assessment tools and an information governance model created specifically for healthcare. Information about the model and assessment tools is available on their website at www.ahima.org.

Disaster Planning and Data Recovery

Not only are healthcare enterprises accountable for protecting all medical and patient identification data maintained and used in the facility, they also must maintain a secure but accessible copy of these data in an off-site location in case information resources are damaged or destroyed by disaster. This obligation, required by the Health Information Portability and Accountability Act (HIPAA) Security Rule's Administrative Safeguards (Snell 2015), increases the secondary storage requirements imposed by the clinical and administrative operational needs of the enterprise. Information about IT disaster recovery plans and links to planning resources are available at www.ready.gov, an official website of the Department of Homeland Security.

Communication: Network Connectivity and Interoperability

Historically, two general approaches have been available for acquiring and implementing application software in a healthcare organization—integrated and interfaced. In the first approach, all modules required to satisfy the organization's computing needs are purchased from a single vendor. Typically, these modules have been designed to work with one another so that data transfer among modules proceeds smoothly. This type of system is known as an *integrated* information system. Epic (www.epic.com), Cerner (www.cerner.com), and Siemens Healthineers (www.siemens-healthineers.com) are well-known vendors of integrated healthcare system solutions for all types of organizations.

By contrast, each of the required modules could be purchased from the vendor thought to be the leader in that particular application area—or one that offers a unique feature valued by the enterprise. Historically, these high-performing applications were referred to as "best of breed," and one might argue that following this approach contributed significantly to the challenges of moving legacy systems toward interoperability. Although a given module might work well for its particular application area, connecting the module to other modules for data sharing could be problematic. For example, the data contained in one module could be incompatible with the

data format of other modules. The data formats or the vocabulary used could differ between the two systems. Something as simple as the way a date is recorded (e.g., 01–31–20 versus 2020–01–31) can prevent data from transferring or being matched correctly. Often, the solution is the development of an *interface*, which acts as a bridge between the two modules and which, for example, translates the data format into one that the receiving module can interpret, process, and store.

While these concepts remain conceptually relevant in an organization, the current computing environment in which healthcare enterprises operate requires not only connectivity among components of the internal enterprise information system but exchange of information between computers across industry networks with little intervention on the users' behalf (generally referred to as *interoperability*). Electronic health information must be transmitted by the collecting provider organization to third-party payers, other providers involved in the care of the patient, compliance and oversight agencies, the patient, and other business and clinical partners and stakeholders. The process of data transfer among various networks and systems, called *electronic data interchange* (EDI), requires that data are stored in standard formats, or are translated between sender and receiver, and that agreed-on communication protocols ensure data integrity after the transfer. The 21st Century Cures Act defines interoperability as "the ability to exchange and use electronic health information without special effort on the part of the user and as not constituting information blocking" (Office of the National Coordinator for Health Information Technology 2019).

In healthcare, EDI standards were mandated under HIPAA and include very specific data formats for exchanging billing information between covered entities. The EDI standards apply to all HIPAA-covered entities for the following transaction types (CMS 2017):

- Claims and encounter information
- Payment and remittance advice
- Claims status
- Eligibility
- Enrollment and disenrollment
- Referrals and authorizations
- Coordination of benefits
- Premium payment

Full interoperability has been difficult to achieve through collaborative efforts of consortia and interest groups, partly because early information systems were designed to be fully proprietary, ensuring more market share for vendors. There remains a lack of agreed-on standards that would ensure

a uniform exchange and processing of clinical and financial information between providers. Because interoperability between systems remains elusive, legislation and government regulations have been necessary to maintain forward momentum. In 2016, the Centers for Medicare & Medicaid Services (CMS) retitled its EHR Incentive Program as the Promoting Interoperability Program and began requiring use of CEHRT by eligible hospitals in 2019. These certified EHRs require easier exchange of certain elements of a patient's record between healthcare providers.

Recognizing that not all barriers to full interoperability are technical, the US Department of Health & Human Services (HHS) drafted the Trusted Exchange Framework and Common Agreement to support national network-to-network information exchange. In draft stage at the time this chapter was written, the final rule and current requirements will be available at www.hhs.gov.

Data Standards Organizations and Regulatory Bodies

EDI is more efficient and reliable if the data were created in accordance with a standard that makes data formats and definitions compatible. The American National Standards Institute, the National Information Standards Organization, the Organization for the Advancement of Structured Information Standards, and the Public Health Data Standards Consortium are examples of organizations and consortia working toward consensus standards for information products and services. Compliance with such accepted industry standards is an important criterion to apply when evaluating vendors' products. Current information about the standards organizations and their active initiatives is available on their websites, which are in the Web Resources section at the conclusion of this chapter.

There are several federal regulatory agencies that enforce HIT legislation and influence best practices in the sector. Some of the higher-profile agencies are presented in exhibit 5.2. Regulatory compliance generally, and information and data reporting specifically, is an extremely important and complex responsibility for healthcare organizations. Many organizations have delegated accountability for this responsibility to a chief compliance officer (CCO), generally an attorney with special training and experience in healthcare law and regulation. This person will have a staff who build and maintain the regulatory body of knowledge needed to protect the organization, develop and maintain relevant policies, prepare and submit regulatory reports, conduct risk assessments, and train clinicians and staff about compliance issues and appropriate practices. Some CCOs achieve recognition as a certified professional compliance officer, an exam-based designation by the American Academy of Professional Coders (www.aapc.com).

EXHIBIT 5.2
HIT Regulatory Agencies

Entity	Authority	Purpose
Health Information Technology Advisory Committee (HITAC)	21st Century Cures Act, P.L. 114-255 Federal Advisory Committee Act (FACA), P.L. 92-463, as amended, 5 U.S.C. App. 2	Recommends to the National Coordinator for Health Information Technology certain policies, standards, implementation specifications, and certification criteria relating to the implementation of a health information technology infrastructure, nationally and locally, that advances the electronic access, exchange, and use of health information.
Office of National Coordinator for Health Information Technology (ONCHIT)	HITECH Act of 2009 21st Century Cures Act of 2018	Responsible for advancing connectivity, interoperability, and usability of HIT. Oversees conditions of certification and trusted exchange framework.
ONCHIT HITECH Programs	American Recovery and Reinvestment Act (ARRA)	Supports nationwide implementation of HIT, including the State Health Information Exchange Cooperative Agreement Program, regional extension centers to assist primary care providers in adoption and meaningful use of EHRs, and the Workforce Development Program to train healthcare workers in new health information technologies.
Office for Civil Rights	Department of Health & Human Services	Enforces federal civil rights laws; conscience and religious freedom laws; HIPAA Privacy, Security, and Breach Notification Rules; and the Patient Safety Act and Rule, which together protect fundamental rights of nondiscrimination, conscience, religious freedom, and health information privacy.
Cybersecurity and Infrastructure Security Agency (CISA)	Cybersecurity and Infrastructure Security Agency Act of 2018	Communicates cybersecurity and infrastructure security knowledge and practices for federal network protection, cyberprotection, infrastructure resilience, and emergency communications.

Privacy, Physical Security, and Cybersecurity

The healthcare sector has experienced a major evolution—most health information and patient medical records are now online and connected across providers and health systems, and medical professionals enjoy significant advances in medical and information technology. However, as the healthcare field continues to expand its reliance on digital technologies, organizations face increasing concerns about safeguarding the privacy and security of information and the corollary of cybersecurity threats. Unfortunately, the technology advances that allow valuable, legitimate uses that improve business operations and clinical services also provide cybercriminals and recreational hackers with tools to breach systems and wreak havoc. EHRs, discussed in chapter 8, are a prime target for cyberattacks because of the extensive personal information they contain and the potential such data have for financial advantage.

Data breaches—generally, any unauthorized access to information—are a significant threat for healthcare enterprises, posing both reputational and financial risks. The HIPAA Breach Notification Rule, enacted under the Health Information Technology for Economic and Clinical Health (HITECH) Act of 2009, requires that breaches resulting in exposure of 500 or more individual records must be reported by the healthcare organization to HHS's Office for Civil Rights (OCR). Guidelines for submitting notice of a breach are available on the HHS website. For the year 2018, the OCR received notice of 351 data breaches involving exposure of more than 13 million health records—a significant increase over 2017, when 5.1 million records were exposed in 359 breaches (Donovan 2018a; *HIPAA Journal* 2018). The Ponemon Institute, an independent research firm that explores issues related to personal and business information, conducted interviews with more than 2,200 data protection and compliance professionals across 17 industries and identified 477 companies that experienced a data breach in the previous year (Ponemon 2018). With the caveat that healthcare organizations comprised only 1 percent of the sample, their findings calculate the average cost of a data breach to be $3.86 million, approximately $148 per record stolen. In addition, the financial penalties that covered entities could incur as a result of a data breach were outlined under the HIPAA Omnibus Rule, which became final in 2013. Because data breaches can be intentional (neglecting to manage known security risks) or unintentional (processes and technology were in place to prevent a network intrusion, but a sophisticated hacker gained access anyway), CMS defines a tiered approach to penalties that may be incurred by an offending covered entity.

Research reports such as these limited examples, coupled with high-profile media releases about individual security incidents, compel healthcare executives to consider information security a high priority in strategic planning and resource allocation. The reputational impact of a data breach that has been determined to stem from negligence and inadequate policies