Unstructured Data in Electronic Health Record (EHR) Systems: Challenges and Solutions

Healthcare Content Management White Paper

October 2013
www.datamark.net
The Electronic Future of Healthcare Records

With every year that passes, with every federal incentive payment to providers for adoption and meaningful use of certified electronic health record (EHR) technologies, patient care is destined for an electronic future.

More and more patients check in via tablets and kiosks, and more and more patients request their medical information in an easily transportable format—packaged in encrypted electronic files, rather than in manila folders stuffed with photocopies and faxes.

EHR systems are helping to provide one piece of the puzzle that makes up an entire patient medical history. However, it’s important to remember that “certified EHRs” (those designed to meet the standards and incentive payments of the Health Information Technology for Economic and Clinical Health (HITECH) Act) are systems designed primarily for structured data.

Structured data—information such as dates, patient names, identification numbers and diagnosis codes—is easier to collect and exchange between systems, because it is standardized, pre-defined, computer-readable and typically quickly accessible from a database. Structured data can pre-populate fields in electronic templates and be selected from pull-down menus, for example.

Unstructured Data—The Other 80 Percent of All Data

A significant piece of the puzzle that establishes the entire “patient narrative” for healthcare providers remains to be captured. This is all the unstructured data relevant to the patient that needs to be collected, captured and flowed into an organization’s EHR system.

Unstructured data is the information that typically requires a human touch to read, capture and interpret properly. It includes machine-written and handwritten information on unstructured paper forms, audio voice dictations, email messages and attachments, and typed transcriptions—to name a few.

Analysts from Merrill Lynch, Gartner and IBM have come to a general consensus that unstructured data accounts for 80 percent of the data in business organizations, and the same rule of thumb applies to the healthcare sector as well. In regard to documents used in healthcare, the Health Story Project estimates that some 1.2 billion clinical documents are produced in the U.S. each year, and about 60 percent of these contain valuable patient-care information “trapped” in an unstructured format.

Bottom line, much of the patient information that providers need for decision-making, and much of the data that could be used for trend analysis and research for the betterment of the public’s health is difficult to get to, difficult to standardize, and difficult to share between disparate computer systems.

Sources of Unstructured Data in Healthcare Organizations

• Medical claims
• Explanation of Benefits (EOB)
• Invoices and purchase orders
• Shipping documents
• Scanned medical reports
• Signed patient consent forms
• Handwritten notes
• Drawings
• Diagnostic images
• Voice dictation
• Email messages
• Email attachments
• Text messages
• Blog posts
• Tweets
• Online video
• Web pages
• Documents from other organizations
A recent Healthcare IT News survey found that healthcare organizations are still in a reactive mode as they develop or apply strategies to bring this vast amount of unstructured data in alignment with their EHR systems. According to the survey, 40 percent of respondents are implementing a solution for unstructured clinical data, 25 percent are creating a strategy or plan, and 18 percent are looking at what solutions are available in the marketplace.

Electronic Document Management Systems (EDMS)
The healthcare organization’s goal should be to create a complete EHR system, one that incorporates critical information from unstructured data sources as well as the structured data that defines a “certified EHR” in the eyes of the federal government’s HITECH incentive program.

This ideal EHR system should be built on the bedrock of sound Enterprise Content Management (ECM) principles, i.e., the tools, technologies, strategies and people that manage the entire lifecycle of information within an organization—from its arrival in the mailroom or through a secure server, to its flow through business processes, to its scheduled retention or destruction.

Unstructured data will be brought into the EHR flow using the technological components of Electronic Document Management Systems (EDMS), including:

- Automated forms classification and data capture systems.
- Document management systems for version control, check-in/check-out and security.
- eForms management for online data entry and data collection.
- Records management technology to control access to records, and retain or destroy records in accordance with state regulations.

State-of-the Art
A decade or more ago, typical enterprise PCs did not have sufficient processing power at a reasonable price point to drive effective automated forms classification or sophisticated optical character recognition (OCR) tools.

Fortunately, today represents a new era of high-performance computing at price points accessible to most enterprises. This vastly improved computing power has radically improved recognition accuracy of unstructured and semi-structured business forms.

Also, parallel processing power allows forms and data processing software engines to run algorithms concurrently and even send optical capture results to multiple OCR engines that “vote” to decide the best data result—achieving accuracy rates that would have once required human intervention.
EDMS as a Legal Archive

The importance of robust electronic document management systems in healthcare organizations cannot be overstated. The EDMS serves as a searchable legal archive of electronic documents—documents that may need to be presented because of a subpoena request, or to resolve an exception or dispute. Healthcare organization managers need to understand this distinction—a “certified EHR” will fall short when it comes to legal purposes because the certified EHR may only have access to information from a database, and not the original document in question.

Implementation Considerations

As can be seen by the sheer variety of content that healthcare organizations need to bring into an EHR system to create manageable, useful and actionable data, there is typically not one product that can do it all for an organization, although vendors’ marketing teams may do their best to convince buyers otherwise.

What is certain is that there are a number of very effective technologies and manual processing services that can be used to capture unstructured data and convert it into a formats that are easily searchable, transmittable, redactable when necessary, and secure.

However, it is important to remember that despite the best intentions and messages of product marketers, no forms-processing or workflow solution works flawlessly “out of the box.” Developing an overall enterprise content management strategy for a healthcare organization and integrating and implementing the necessary EDMS technologies is a highly complex endeavor.

Organizations would be well advised to seek the assistance of a consultant such as DATAMARK, with expertise in the fields of ECM, EDMS and business process analysis and re-engineering.

About DATAMARK

DATAMARK is a leading provider of multichannel customer contact center services, digital mailroom and mail center management, data entry, document processing services and business process re-engineering services for Fortune 500 companies, government agencies and other large enterprises.

Founded in 1989, the company is the strategic business process outsourcing (BPO) partner for companies across numerous industry sectors, including healthcare, insurance, banking and financial services, and transportation and logistics.

DATAMARK offers on-site, onshore and offshore processing facilities, delivering enterprise content management (ECM) and process-automation technologies and solutions to help organizations improve efficiency and profitability in all business functions.