

**Component 11/Unit 8a**  
**Introduction to Data**

Infrastructure, architecture, data sets,  
structured and unstructured data, data  
warehouse and data repository

---

---

---

---

---

---

---

---

**Data Architecture Defined**

Two definitions will help us to define the meaning of the components making up the overall architecture that encompasses an electronic health record.

- “The configuration, structure, and relationships of hardware (the machinery of the computer including input/output devices, storage devices, and so on) in an information system” (LaTour, Eichenwald Maki, Glossary)
  - This definition address the complexity of hardware configuration related to the architecture
  - This may include older legacy systems based on mainframe with DOS-based systems
  - May be client/server architecture which is used by third party vendors to configure systems
- “A system that consists of individual databases contributing to a central data repository from which data may be either drawn directly to supply an EHR workstation or sent to a warehouse that performs sophisticated analysis on data to supply decision support.” (Amatayakul, Electronic Health Records, Glossary)
  - This definition relates more to the configuration of the interoperability of software that make up the data architecture

Component 11/Unit 8a Health IT Workforce Curriculum Version 1/Fall 2010 2

---

---

---

---

---

---

---

---

**Data Infrastructure**

The architecture provides a framework for hardware and software configuration. The data infrastructure, however, focuses on the data needs to operate a health care organization.

For the EHR we look not only at the operational end of the organization, but also the clinical needs and the concerns around the meaningful use of data collected

Component 11/Unit 8a Health IT Workforce Curriculum Version 1/Fall 2010 3

---

---

---

---

---

---

---

---

## Data Set Defined

- What is data? First of all, we need to understand that each fact collected is considered raw data
- Sometimes we create a group or set of facts (data) with uniform definitions to be used for a special use. This is called a data set

Examples of data sets in health care are:

1. Uniform Hospital Discharge Data Set (UHDDS). These are long standing data sets begun in the 1960s to provide information at a national level
2. Uniform Ambulatory Data Set (UACDS)
3. Minimum Data Set for Long-Term Care
4. Data Elements for Emergency Department Systems (DEEDS)
5. Health Plan Employer Data and Information Set (HEDIS)

Many other health care data sets are well established and provide helpful aggregate information

---

---

---

---

---

---

---

---

## Structured versus Unstructured Data

Structured data is readable by a computer. In an EHR, capturing data by individual fields provides structure. Common fields would be date of birth, gender, diagnosis/procedure code. When decision support systems are built, structured data can be used to provide information.

Unstructured data is not readable by the computer. It would include hand-written/transcribed narrative notes that are scanned in to a system. The notes are readable but not usable for clinical decision support systems

---

---

---

---

---

---

---

---

## Types of EHR Data

Here are some examples of structured/unstructured data you will find in EHR's

- Structured/discrete (e.g. CPOE, Registration data)
- Unstructured data:
  - Document image data
  - Transcribed reports (including voice recognition)
  - Video data (e.g. ultrasound)
  - Audio data (e.g. voice notations)
  - Vector graphic data (e.g. Fetal monitor, EKG)
  - Diagnostic Image Data (e.g. MRI, CT)

---

---

---

---

---

---

---

---

## Data Repository/ Data Warehouse

A part of the configuration of the architecture of an EHR consists of a data-base where all data from all systems can be collected and stored. Often this is called a repository or a warehouse. Further delineated, you may also find repositories that narrow the data more specifically, like the clinical data repository (CDR). It is essential to build the repositories in to the system so that data is protected and can be used for data analytics and decision-making.

- Here are a few related definitions
  - Data Repository: "An open-structure database that is not dedicated to the software of any particular vendor or data supplier, in which data from diverse sources are stored so that an integrated, multidisciplinary view of the data can be achieved; also called a central data repository or, when related specifically to healthcare data, a clinical data repository(CDR) (glossary, A)"
  - Data warehouse: " A database that makes it possible to access data from multiple databases and combine the results into a single query and reporting interface." (LaTour & Eichenwald Maki glossary)
  - Clinical repository: " a frequently updated database that provided users with direct access to detailed patient-level data as well as the ability to drill down into historical views of administrative, clinical , and financial data." (LaTour & Eichenwald Maki glossary)

---

---

---

---

---

---

---

---

## Conclusions

- Understanding terminology in data architecture is essential for someone who will be configuring an EHR
- This set of slides provided you with an introduction to some key terms that you will hear when working with information technology specialists.
- Keep in mind that the goal of developing an EHR is to collect meaningful data that can be used to provide quality health care

---

---

---

---

---

---

---

---

## Resources

- Amatayakul, Margaret K., Electronic Health Records: A Practical Guide for Professionals and Organizations, Fourth Edition, Chicago: AHIMA, 2009.
- LaTour & Eichenwald-Maki, Health Information Management. Concepts, Principles, and Practice, Third Edition, AHIMA Press

---

---

---

---

---

---

---

---