Lecture 10 Biomedical Information Systems and Knowledge Management

Andreas Holzinger
VO 709.049 Medical Informatics
18.01.2017 11:15:12:45

Science is to test crazy ideas – Engineering is to put these ideas into Business

Lecture 10 Biomedical Information Systems and Knowledge Management

a.holzinger@tugraz.at
http://hci-kdd.org/biomedical-informatics-big-data

Advance Organizer (1/2)

- Biomedical workflow management system = designed specifically to compose and execute a series of computational and/or data manipulation steps and/or workflows within the domain of bioinformatics;
- Business process re-engineering (BPR) = analysis and design of workflows and processes within an organization (hospitals). According to Davenport (1990) BPR is a set of logically related tools performed to achieve a defined outcome/Result;
- Clinical Pathway = aka care map, a tool used to manage the quality in healthcare concerning the standardization of care processes and promote organized and efficient patient care based on EBM;
- Digital Imaging and Communications in Medicine (DICOM) = a standard for handling, storing, printing, and transmitting data in medical imaging (also file format definition and a network communications protocol using TCP/IP);
- Evidence-based medicine (EBM) = aiming at developing mathematical estimates of benefit and harm from population-based research and apply these in the clinical routine, claiming that best research evidence on medical interventions come from experiments (e.g. randomized controlled trials);
- Health Level Seven (HL/7) = an International Organization accredited by the American National Standards Institute (ANSI) to push consensus-based standards representing healthcare stakeholders;

Advance Organizer (2/2)

- Hospital Information System (HIS) = integrated information system for administrative, financial, clinical etc. information management in a hospital;
- Integrating Healthcare Enterprise (IHE) = initiative by healthcare professionals and industry to improve the way computer systems in healthcare share information (i.e. promotes the coordinated use of established standards such as DICOM and HL/7);
- National Electrical Manufacturers Association (NEMA) = holds copyright of DICOM;
- Paradigm = according to Kuhn (1962) a shared view of a group of researchers, comprising 4 elements: concepts, theories, methods and instruments;
- Picture Archiving and Communication System (PACS) = system for handling images from various medical imaging instruments, including ultrasound (US), magnetic resonance (MR), positron emission tomography (PET), computed tomography (CT), endoscopy (ENDO), mammography (MG), digital radiography (DR), computed radiography (CR) ophthalmology, etc.;
- Workflow = consists of a sequence of connected steps, succeeding the flow paradigm, where each step follows the precedent;

Agenda for today

- 00 Reflection – follow-up from last lecture
- 01 Information Systems Challenges
- 02 Workflows – Clinical Pathways
- 03 Hospital Information Systems (HIS)
- 04 Multimedia: PACS
- 05 Data Standards for Communication
- 06 Open ...
- 07 Conclusion & Future Outlook

Warm-up Quiz

- A - View from above
- B - Profile coordinates view

Keywords

- Bioinformatics workflows
- Clinical workflow & management systems
- Cloud computing in healthcare
- Communication standards
- Digital Imaging and Communication in Medicine (DICOM)
- Formal methods & workflow modeling
- Health Level 7 (HL/7)
- Logical Observation Identifier Names and Codes (LOINC)
- Medical multimedia
- Mobile computing in medicine
- Personal Health Record (PHR)
- Picture Archiving and Communication System (PACS)
- Quality
- Software as a Service (SaaS)
- Systems architecture
- Unified Modeling Language (UML)

Learning Goals: At the end of this 10th lecture you ...

- have an overview about workflows and workflow modeling in health care;
- got an overview of typical architectures of hospital information systems for patient records as already discussed in lecture 4;
- have understood the principles of Picture Archiving and Communication PACS-Systems;
- know how important multimedia for medicine is;
- have a basic understanding of DICOM and HL/7;
- are aware of the constraints of open source software in the medical domain;
- have got an idea of possible future systems solutions;
01 Information Systems Challenges

Goal:
Unified View,
one single workflow,
Workflow continuity,
no media gaps
(DE: Medienbruch)


Not only a problem at cell level ...

Biomedical data
(e.g. clinical trial data)
*omics data

Clinical patient data
(e.g. EPR, images, lab results, medication, etc.)

Weakly structured, highly fragmented, with low integration

Health business data
(e.g. costs, utilisation, etc.)

Private patient data
(e.g. AAL, monitoring, sensor data, social network visits)


Patient-related Finding Data

Management of Clinical Trials (Research Network)

Chip-based genotypes

Mapping between Patient Sites and Clinics

Management of Chip-related and Annotation Data

RxNorm is a US-specific terminology that contains all medications available on the US market, part of UMLS.


How to combine these different data types together to obtain a unified view of the activity in the cell is one of the major challenges of systems biology.
02 Clinical Pathways and Workflows

Slide 10-3 Workflow > Interaction > Decision > Action

Slide 10-5 Workflow modeling in a nutshell
- Workflow modeling is the process of simplifying the real-world;
- Modeling is based on facts gathered during observations and we need to accept that this representation can never be perfect;
- Expectations from a model should be limited to the intentions with which it is designed for, be it problem solving or understanding of system intricacies (=elaborately complex details);


Slide 10-8: Example: Modeling in UML – Use Case Diagram

Slide 10-9: Modeling: Different views for different people

- Communication/Coordination diagrams show the message flow between objects and depend on the relationships between classes.
- Component diagrams are used to define the aspects of the run-time configuration of processing nodes and the components that run on those nodes. This shows the hardware & software requirements.
- Deployment diagrams are used to define the aspects of the run-time configuration of processing nodes and the components that run on those nodes. This shows the hardware & software requirements.
- Actors & Scenarios are used to define the aspects of the run-time configuration of processing nodes and the components that run on those nodes. This shows the hardware & software requirements.

Slide 10-10: Example Clinical Workflow

Slide 10-11: Example: WF-Optimization with mobile comp

Slide 10-12: Mobile Computing Project MoCoMed-Graz

Slide 10-13: Important: Microlevel – the view of the Manager

- The data model can also be viewed as a tool for decision-making.
- Output: improve productivity and patient care.

Slide 10-14: MoCoMed-Graz in operation

Slide 10-15: MoCoMedGraz Project – Benefit Summary

- This project was evaluated from a three level perspective: patients – doctors – managers.
- Theoretical Background: Bronfenbrenner model.
- Project demonstrates how both workflows and information quality can be optimized.
- Thereby raising both medical quality and patient empowerment.
- Success factors for the application within the hospital include: simplicity, usability, reliability.
Slide 10-16 Summary: Workflows in Bioinformatics

- A data-driven procedure consisting of one or more transformation processes -> nodes;
- Can be represented as a directed graph;
- Direction is time – i.e. the order of transformations;
- A set of transformation rules;
- The data flow originates from a source to a destination (or result) via a series of data manipulations;
- The specification is designed in a Workflow Design System (modeling component) and then run by a Workflow Management System (execution component).


Slide 10-17 Workflow Management: Definitions

- Def: WMS = a system that defines, creates and manages the execution of workflows. Its main components include:
  1. a graphical interface for composing workflows, entering data, watching execution, displaying results;
  2. an engine to store workflow descriptions, results of executions and related traces;
  3. a registry of available services, either local or remote;
  4. a scheduler able to invoke services included in the workflow at the appropriate time;
  5. a set of programming interfaces able to dialogue with remote services;
  6. a monitor tool for controlling the execution of the workflow;
  7. a set of visualization capabilities for displaying different types of results.


Slide 10-19 Architectures of Hospital Information Systems

1970s “Vertical Approach” – monopolistic mainframes
Central computer systems mainly for accounting, typical “data processing” (EDP)

1980s “Horizontal Approach” – evolutionary systems
Departmental clinical information systems, local area networks, distributed systems

2000s “Integrated Approach” – open, distributed systems
Hospital intranets, electronic patient/hospital record, mobile computing, “information quality focus”


Slide 10-20 Basic architecture of a standard HIS

Slide 10-21 Model of Process Oriented Health Info Systems


Slide 10-22 Multimedia throughout the Hospital

03 Hospital Information Systems (HIS)

04 Multimedia Data in the Hospital: PACS

05

Data Standards for Communication and Representation

  1. A set of protocols for network communication;
  2. A syntax and semantics for commands and info;
  3. A set of media storage services (standard compliant);
- HL7 - Health Level 7
  1. HL 7 v2.3 messaging protocol to provide exchange of textual healthcare data between hospital information systems;
  2. Interface Implementation Model (IIM) contains data types, classes, state diagrams, use case models, and terminology to derive domain-specific information models;
  3. Clinical Document Architecture (CDA) is a document markup standard to specify structure and semantics of clinical documents in XML;
- IONIC: Logical Observation Identifiers Names and Codes
  1. Laboratory data (e.g., molecular pathology observations used for identification of genetic mutations, tumor genes, gene deletions, etc.);
  2. Clinical Observations (e.g., non-laboratory diagnostic studies, critical care, nursing measures, patient history, instrument surveys, etc.);
  3. Claims attachments (e.g., handles the definition of new LOINC terms and codes to manage claims-related data)


06 Towards Open Data

- Billions of biological data sets are openly available, here only some examples:
  - General Repositories:
    - GenBank, EMBL, HMCA, ...
  - Specialized by data types:
    - UniProt/SwissProt, MMMP, KEGG, PDB, ...
  - Specialized by organism:
    - WormBase, FlyBase, NeuroMorpho, ...
  - Details: http://hci-kdd.org/open-data-sets
Conclusion and Future Outlook

- Grand Challenges in this area:
  - Production of Open Data Sets
  - Synthetic data sets for learning algorithm testing
  - Privacy preserving machine learning
  - Data leak detection
  - Data citation
  - Differential privacy
  - Anonymization and pseudonymization
  - Evaluation and benchmarking

Please visit: http://hci-kdd.org/privacy-aware-machine-learning-for-data-science/

Attention: Medical certifications such as FDA in the US and CE marking in Europe do not apply to Open Source software;

- These certifications require a legal commercial entity to be identified as the owner of the product and warrant the legal liability of its distribution and commercial support.

- Open Source software being often developed outside commercial enterprises, such as academic groups or university research labs, do not have the proper legal structure to apply for such certifications.

Also, most Open Source products being distributed free of charge lack the legal binding between the provider and the user that is required for software distribution under FDA and CE certification.

Rahib et al. (2011)

1) The world (even the hospital world ;-) of tomorrow will be mobile ...
2) The data will be stored in the cloud as well as software-as-a-service will be used (Key problem: Privacy, Security, Safety and Data Protection).
3) Search tools will be integrated within the clinical workplace to provide a quick overview first, and then zoom and filter on demand ...
4) Integrated ML will enable knowledge discovery with the doctor-in-the-loop ...

Rahib et al. (2011)
Appendix: Electronic Health Records (1)

Appendix: Electronic Health Records (2)

Appendix: Electronic Health Records (3)

Appendix: OsiriX Open Source component architecture

http://www.osirix-viewer.com

OsiriX Imaging Software
Advanced Open-Source PACS Workstation
DICOM Viewer

for demanding users
for clinical use
for research
for 3D & PAWS
OsiriX 64-bit
OS (64-bit performance)
OsiriX MD
OsiriX HD
Open Source DICOM viewer


Example: Drug Target Identification


Bioinformatics Workflow Management System

http://i2creenasc.t2.hs-psu.edu/pacbio-demo


Quality – Quality – Quality

Experts consider health IT key for improving efficiency and quality of health care processes.

Please, carefully determine between:

Quality of Systems

and

Quality of Data

Quality of Information

Quality of Knowledge

Quality of Action


PACS R&D Progress and R&D Topics


PACS 2010: Multi-modality images on 6 CRTs

Pros/Cons of six PACS Implementation Models

**Pros**
- 1. More granular access control
- 2. Faster installation
- 3. Better scalability
- 4. Improved security

**Cons**
- 1. Less flexibility
- 2. Slower upgrades
- 3. More cost
- 4. Limited interoperability

Example: Images within the EPR

ISO/OSI Layer Model (ISO 7498) Communication Standards

Ex.: The Hong Kong Hospital Clinical Information System

**Big Fact: Personal Health Records: Poor Usability**
- 300+ different systems in the US
- Only 1% of the US population have access (2008)
- “Most EHR are 100% bad.”
  (Forrester Research, 2008)
- Biggest obstacles:
  - Cost factors
  - Poor usability
  - Need for UCD
  - Need for proactive wellness management