BI en Salud: Registro de Salud Electrónico, Estado del Arte

Manuel Graña Romay

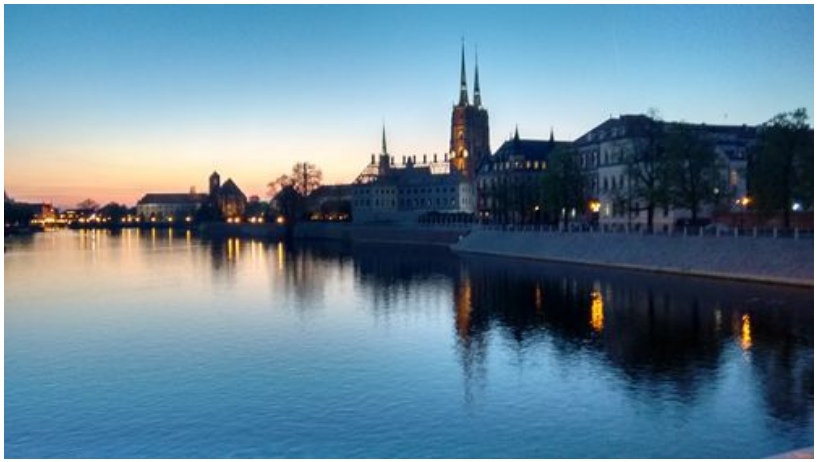
ENGINE Centre, Wrocław University of Technology

Grupo de Inteligencia Computacional (GIC); UPV/EHU; www.ehu.es/ccwintco

Localization

The GIC belongs to the University of the Basque Country (UPV/EHU)
Located in Donostia-San Sebastian, Spain, Europe
Summary of group works: www.ehu.es/ccwintco > Historial de grupo

- ENGINE center at the Wroclaw University of Techonology is funded by the EC for years 2014-2016

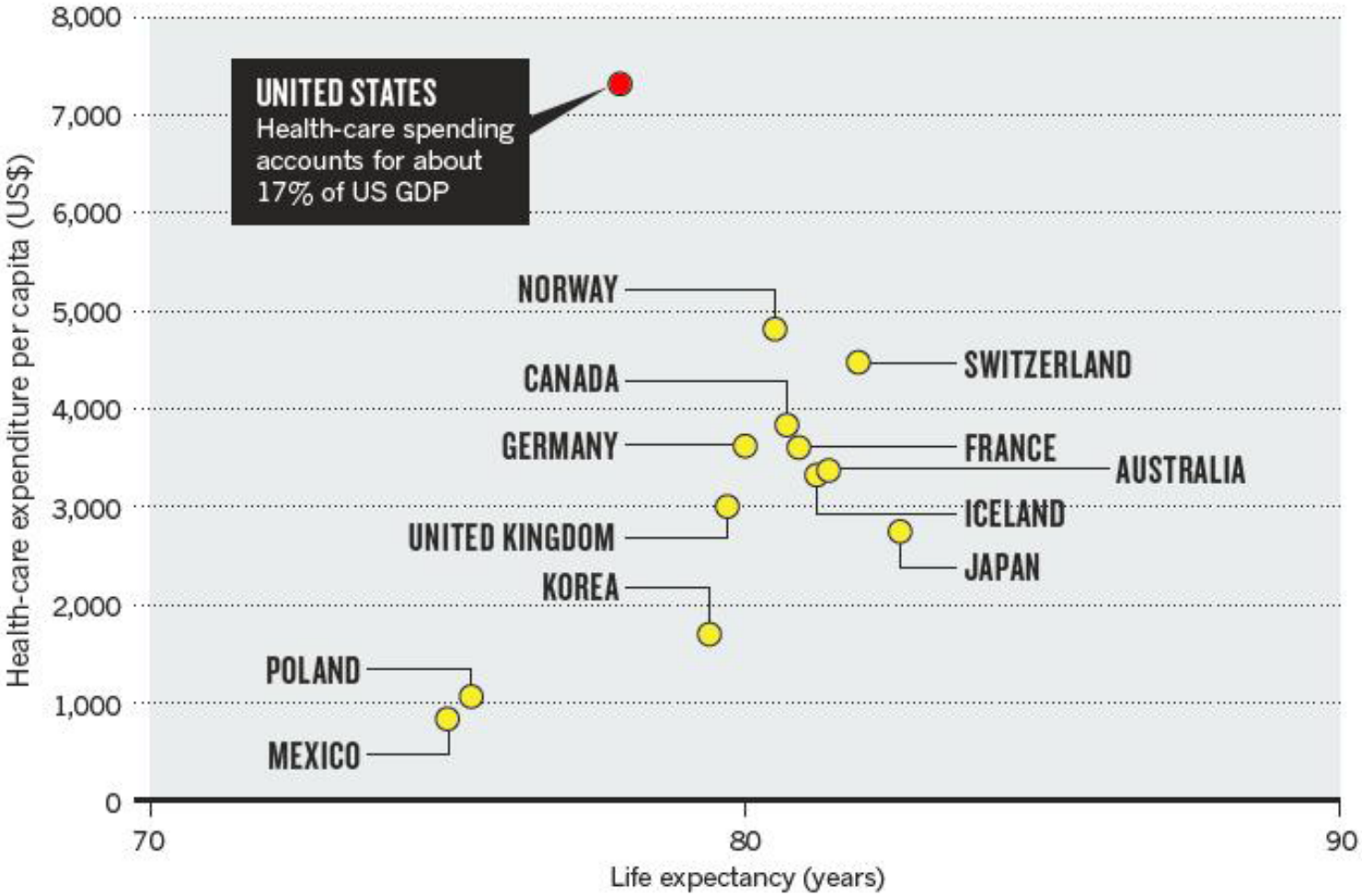


Contents

- Introductory ideas and history
  - Definition
- Vertical technological challenges
  - Privacy, security and authenticity
  - Interoperability
  - Workflow embedding
- Horizontal technological dimensions
  - Big Data
  - Machine Learning and Computational Intelligence
  - Social networks and personal health management
- Business opportunities

MONEY WELL SPENT?

The United States has not seen an increase in life expectancy to match its huge outlay on health care.



Hersh, W., Jacko, J. A., Greenes, R., Tan, J., Janies, D., Embi, P. J., & Payne, P. R. (2011). Health-care hit or miss? *Nature*, *470*(7334), 327.

Healthcare Continuum

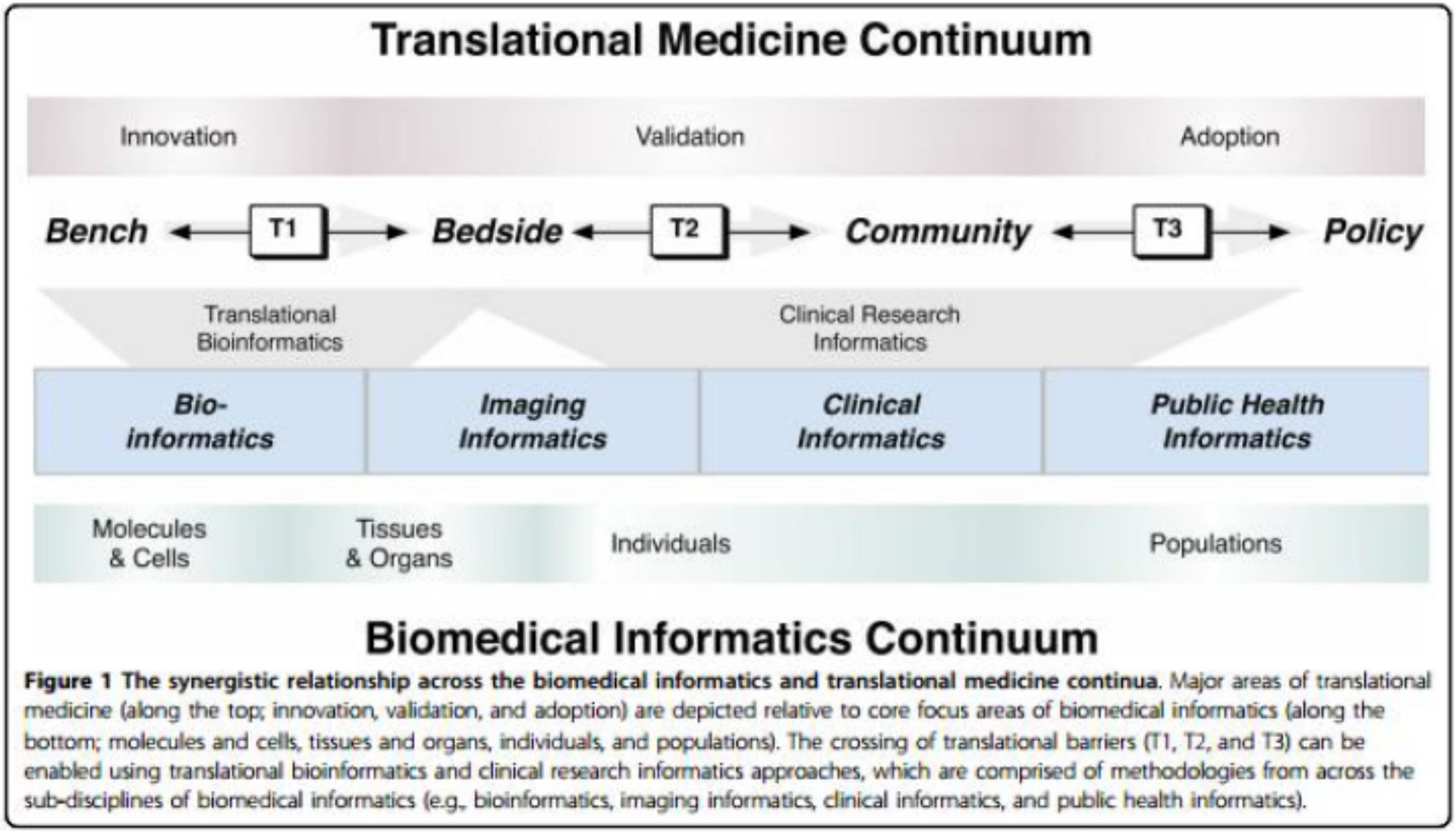


Figure 1 The synergistic relationship across the biomedical informatics and translational medicine continua. Major areas of translational medicine (along the top; innovation, validation, and adoption) are depicted relative to core focus areas of biomedical informatics (along the bottom; molecules and cells, tissues and organs, individuals, and populations). The crossing of translational barriers (T1, T2, and T3) can be enabled using translational bioinformatics and clinical research informatics approaches, which are comprised of methodologies from across the sub-disciplines of biomedical informatics (e.g., bioinformatics, imaging informatics, clinical informatics, and public health informatics).

Sarkar, Indra Neil. "Biomedical informatics and translational medicine." *Journal of Translational Medicine* 8.1 (2010): 22.

Introduction

- Healthcare is a big data problem

> In 2012, worldwide digital healthcare data was estimated to be equal to 500 petabytes and is expected to reach 25,000 petabytes in 2020.

Big Data Analytics for Healthcare

http://dmkd.cs.wayne.edu/TUTORIAL/Healthcare/

Introduction

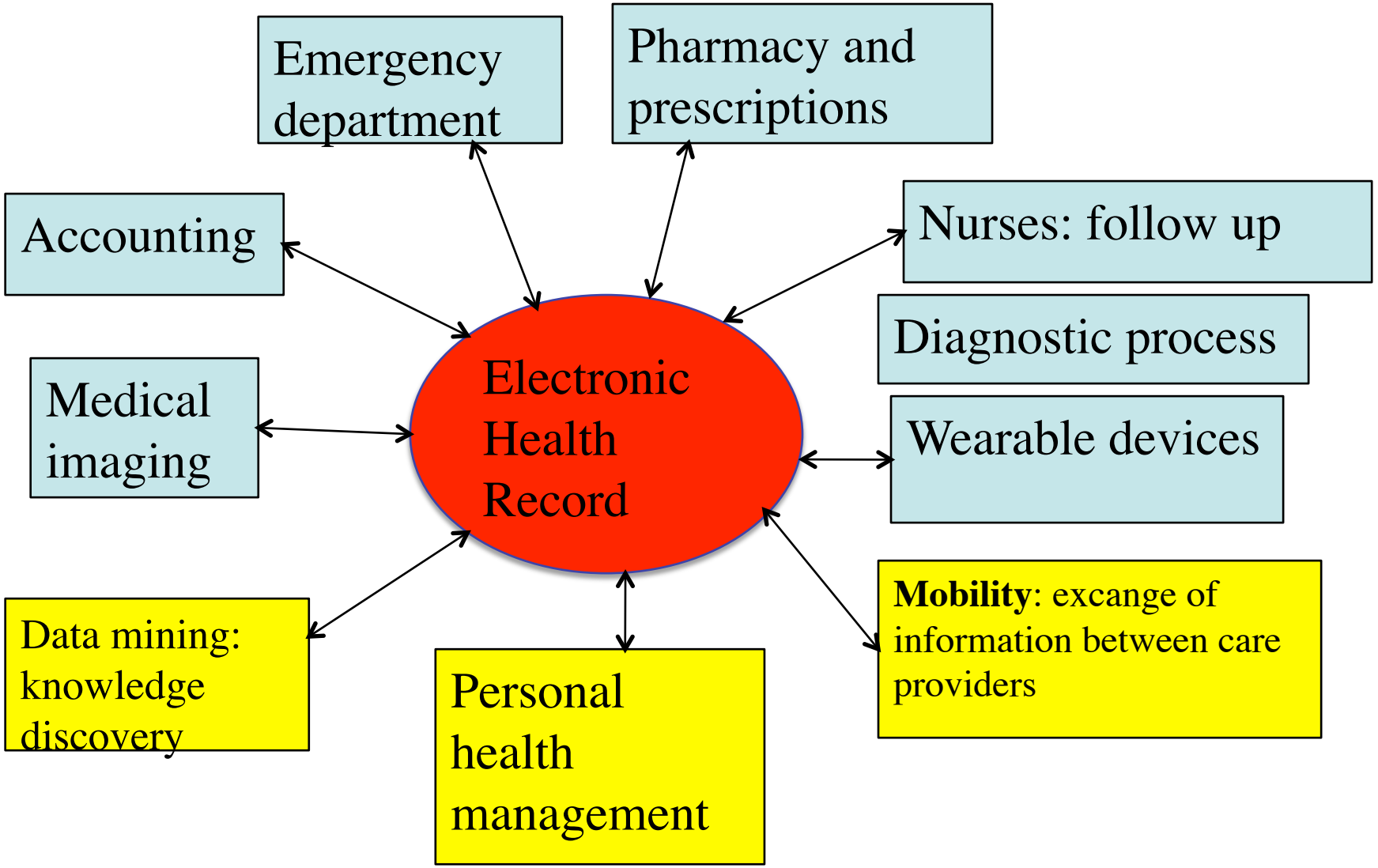
- Electronic Health Record
  - The Electronic Health Record (EHR) is a longitudinal electronic record of patient health information generated by one or more encounters in any care delivery setting.
  - Included in this information are patient demographics, progress notes, problems, medications, vital signs, past medical history, immunizations, laboratory data and radiology reports.
  - The EHR automates and streamlines the clinician's workflow. The EHR has the ability to generate a complete record of a clinical patient encounter - as well as supporting other care-related activities directly or indirectly via interface - including evidence-based decision support, quality management, and outcomes reporting.
- http://www.himss.org/library/ehr/?navItemNumber=13261

Introduction

- Many similar/related concepts
  - EMR : Electronic Medical Record
  - PHR : Personal Health Record
  - EPR: Electronic Patient Record
  - EHP: Electronic Health Profile

Introduction

- EHR: the central information object in health related information processes
  - Diagnosis
  - Data Mining
    - Public health
    - Industry
  - Accounting
  - Mobility and emergencies

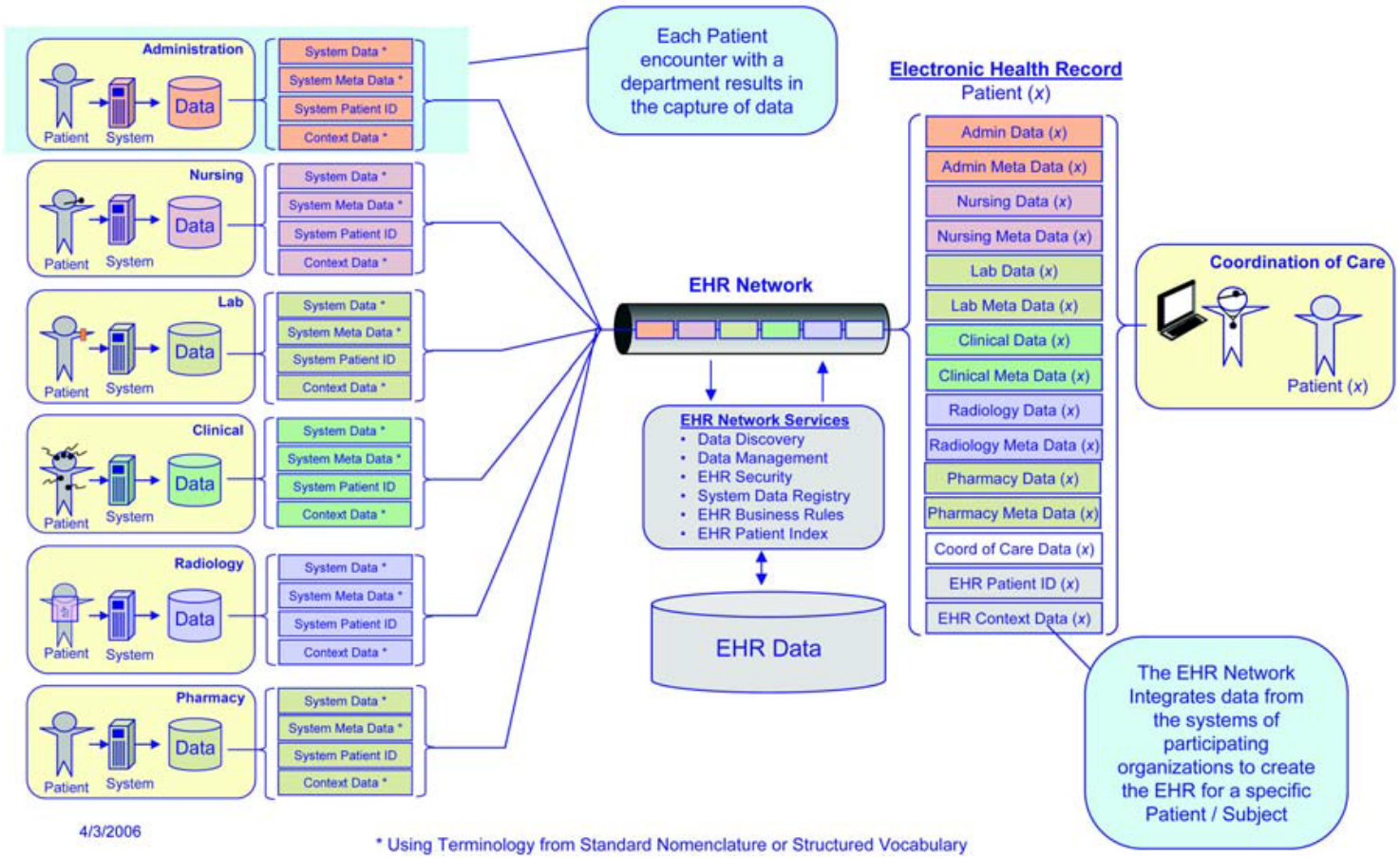

Emergency department
Pharmacy and prescriptions
Accounting
Nurses: follow up
Diagnostic process
Medical imaging
Electronic Health Record
Wearable devices
Data mining: knowledge discovery
Personal health management
Mobility: excange of information between care providers

Introduction

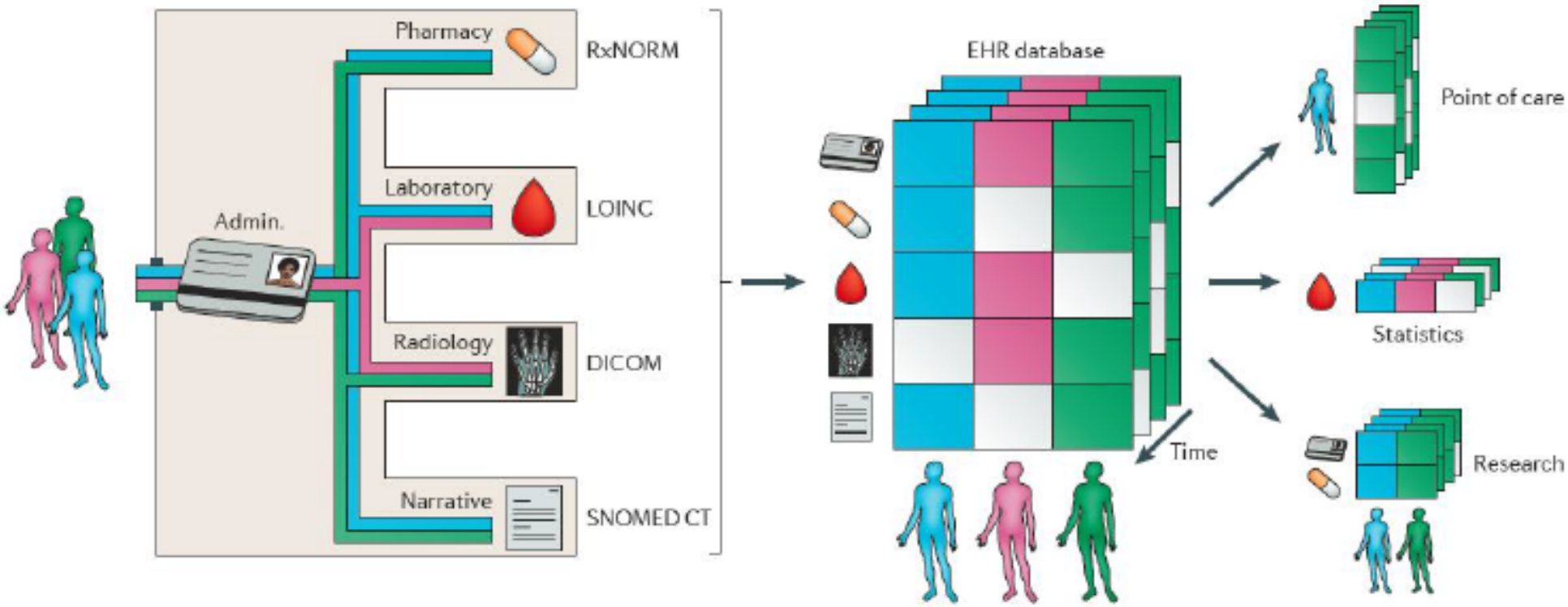
- Motivations and expectations:
- Health care process → Productive industrial process
    - Improvement of workflow; standardization;
    - Accountability : Quality control
        - Economical
        - Health care
    - Research of past events
        - Extraction of knowledge

Electronic Health Record – Concept Overview

The EHR represents the integration of healthcare data from a participating collection of Systems for a single patient.



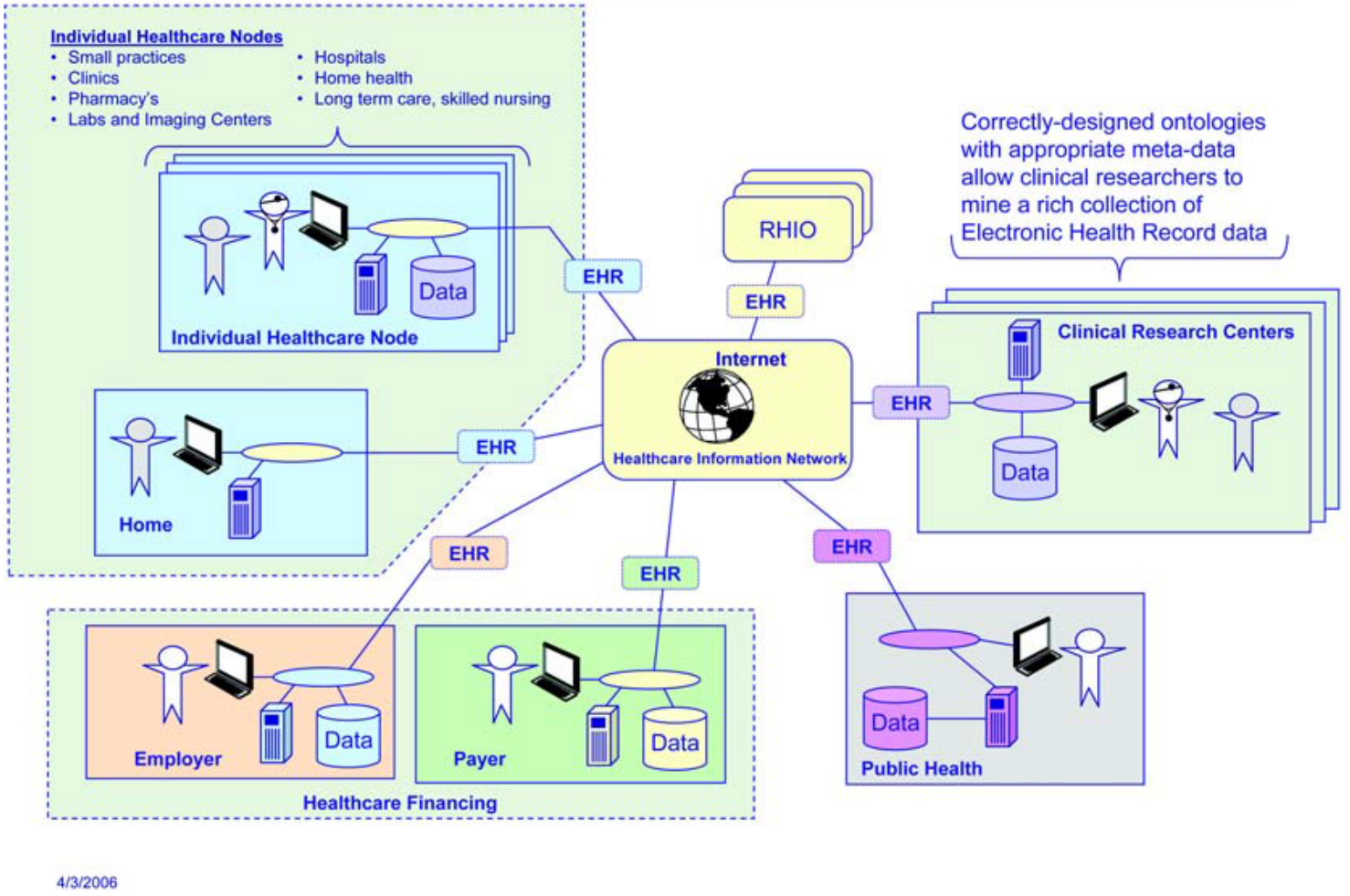
Data Collection and Analysis



Effectively integrating and efficiently analyzing various forms of healthcare data over a period of time can answer many of the impending healthcare problems.

Jensen, Peter B., Lars J. Jensen, and Søren Brunak. "Mining electronic health records: towards better research applications and clinical care." *Nature Reviews Genetics* (2012).

Future EHRs supporting Clinical Research



Contents

- Introductory ideas and history
  - Definition
- Vertical technological challenges
  - Privacy, security and authenticity
  - Interoperability
  - Workflow embedding
- Horizontal technological dimensions
  - Big Data
  - Machine Learning and Computational Intelligence
  - Social networks and personal health management
- Business opportunities

Privacy

- Informed consent: patient awareness of risks in EHR data sharing
- Anonimization of data for
  - Data management in the cloud
  - Research
    - Academic (data repositories on the web)
    - Industry

Security

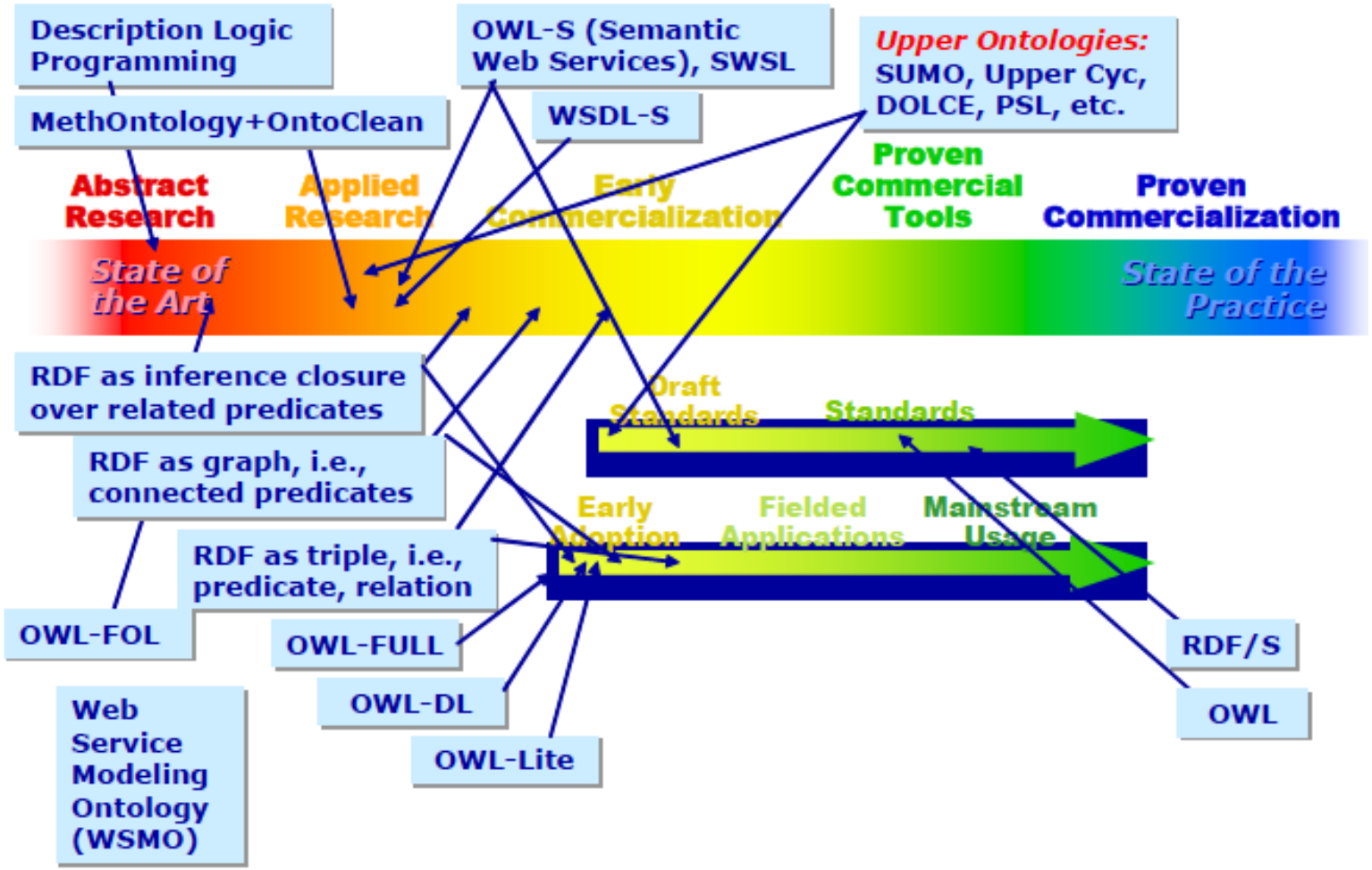
- Hard security based on cryptography
  - Data access protocols
  - Cloud computing: how to ensure that 3d parties manipulate your data without breaking it?
- Soft security based on trust assessment
  - Distributed systems with no central authority
  - Data exchanges

Authenticity

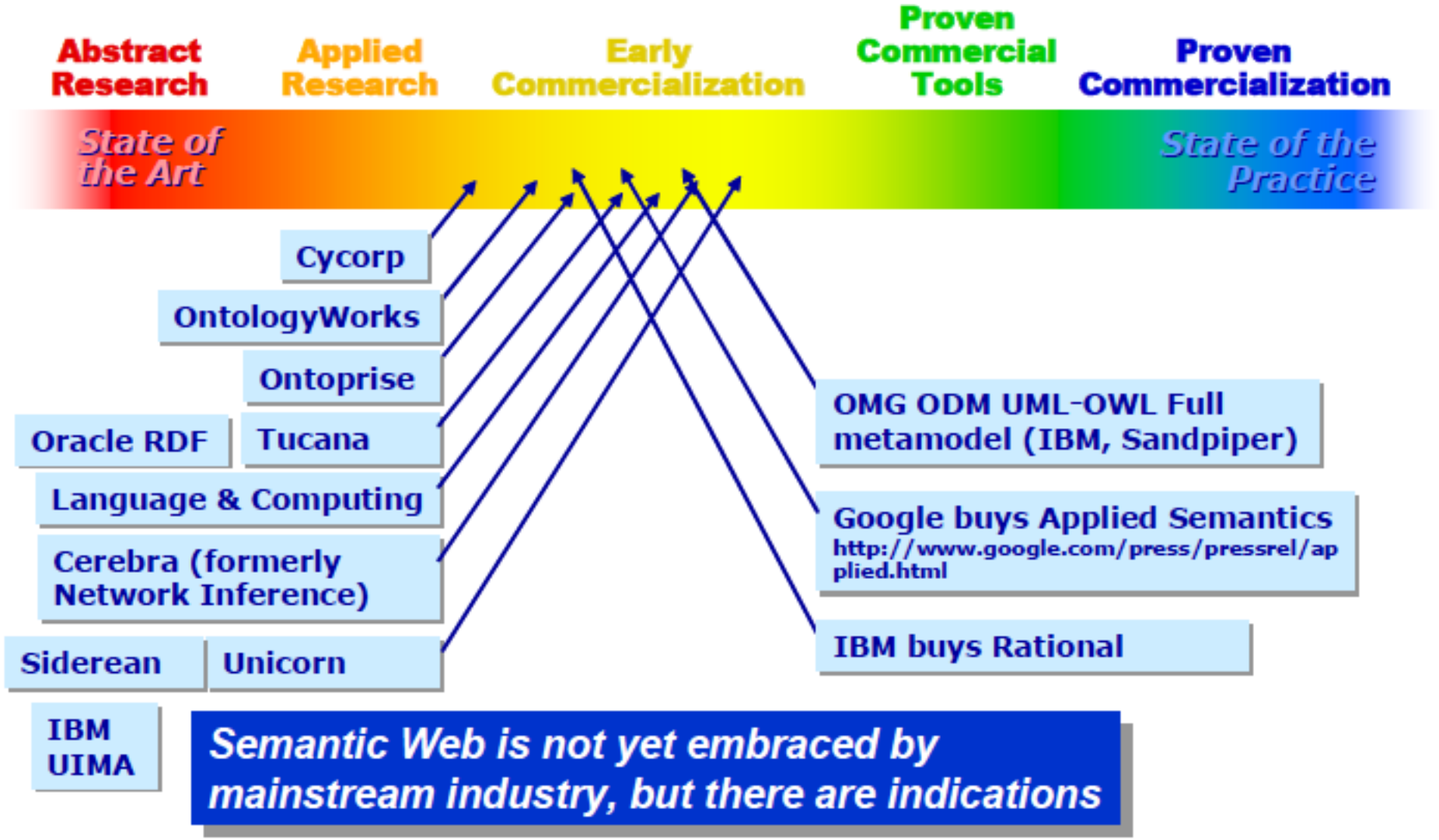
- EHR data must be univocally mapped to an identity
  - Data exchange needs ID certification
- EHR must be guaranteed to be free of manipulations
  - Image watermarking with demographic information

Interoperability

- Semantic web (ontologies)
  - Mapping between vendor solutions
  - Standarization of terms



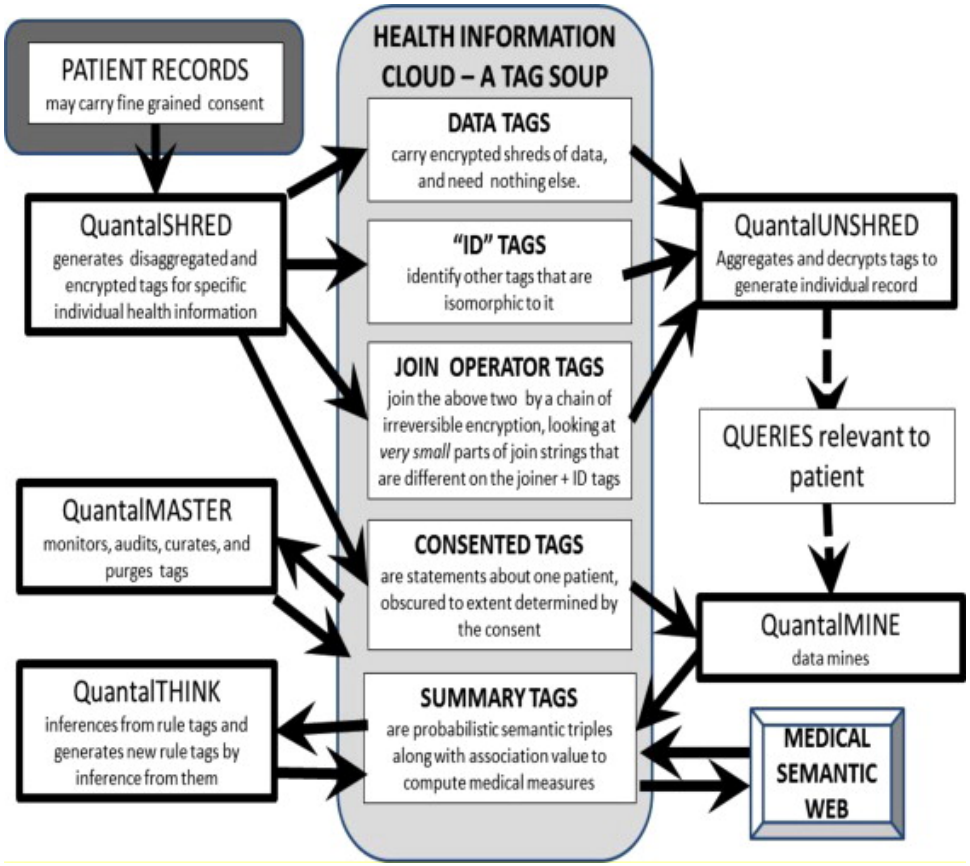
MITRE Corporation, report for the NIH (2006)


Abstract Research
Applied Research
Early Commercialization
Proven Commercial Tools
Proven Commercialization
State of the Art
State of the Practice
Cycorp
OntologyWorks
Ontoprise
Oracle RDF
Tucana
Language & Computing
Cerebra (formerly Network Inference)
Siderean
Unicorn
IBM UIMA
OMG ODM UML-OWL Full metamodel (IBM, Sandpiper)
Google buys Applied Semantics
http://www.google.com/press/pressrel/applied.html
IBM buys Rational
Semantic Web is not yet embraced by mainstream industry, but there are indications

Interoperability

- Universal translators between ontologies by probabilistic semantic reasoning
  - Barry Robson, Thomas P. Caruso, Ulysses G.J. Balis, Suggestions for a Web based universal exchange and inference language for medicine, Computers in Biology and Medicine, Volume 43, Issue 12, 1 December 2013, Pages 2297-2310

Secure interoperability



Workflow

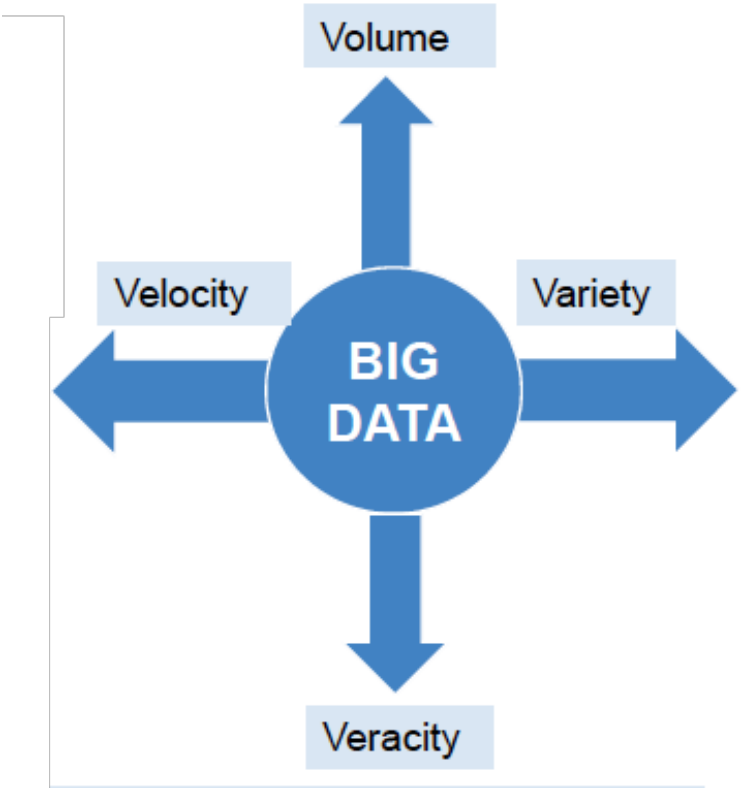
- Dealing with the EHR data completion hinders clinical processes.
  - 30-50% time in a visit devoted to filling information
- Clear improvements in some critical situations
  - Emergency departments

Contents

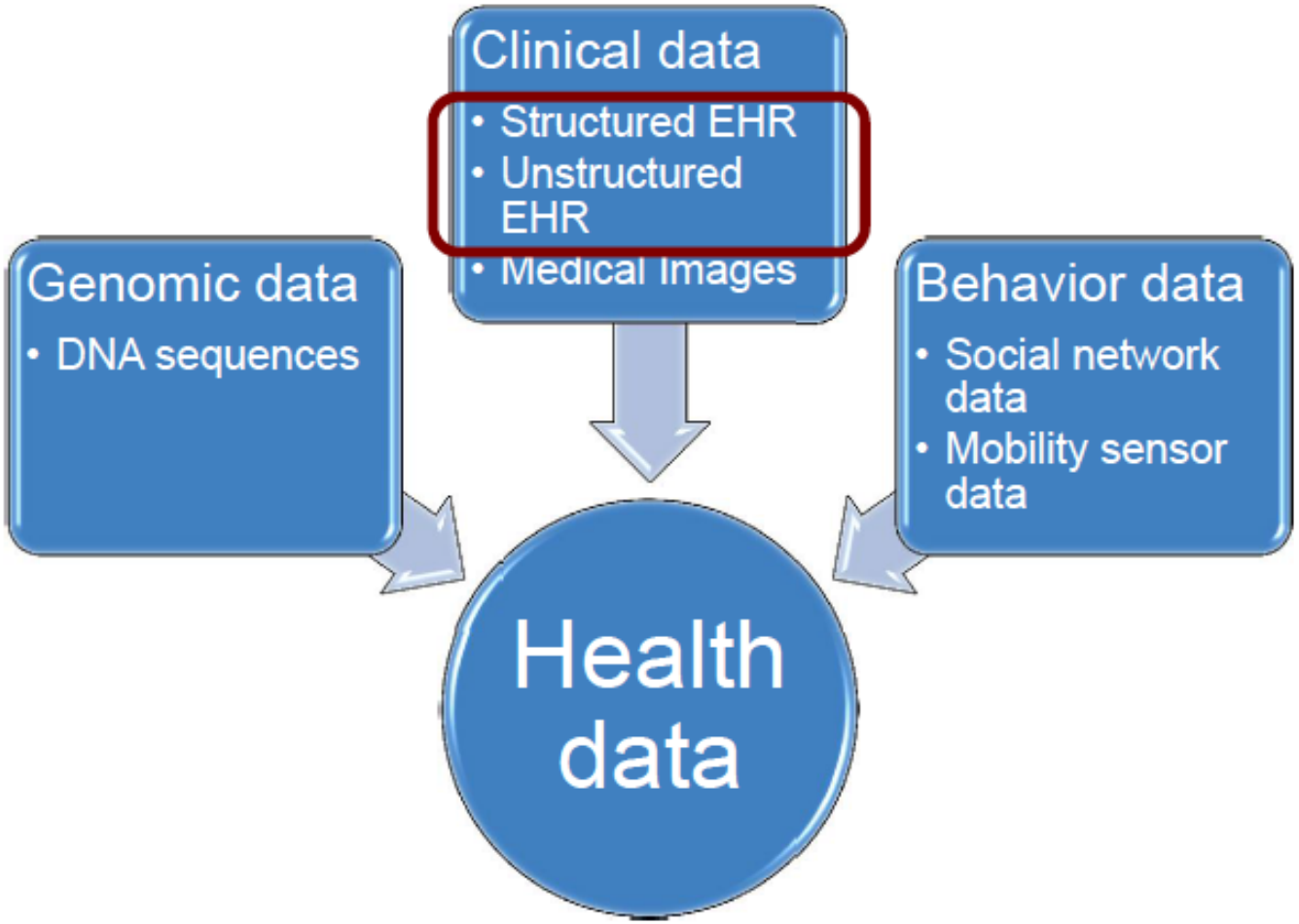
- Introductory ideas and history
  - Definition
- Vertical technological challenges
  - Privacy, security and authenticity
  - Interoperability
  - Workflow embedding
- Horizontal technological dimensions
  - Big Data
  - Machine Learning and Computational Intelligence
  - Social networks and personal health management
- Business opportunities

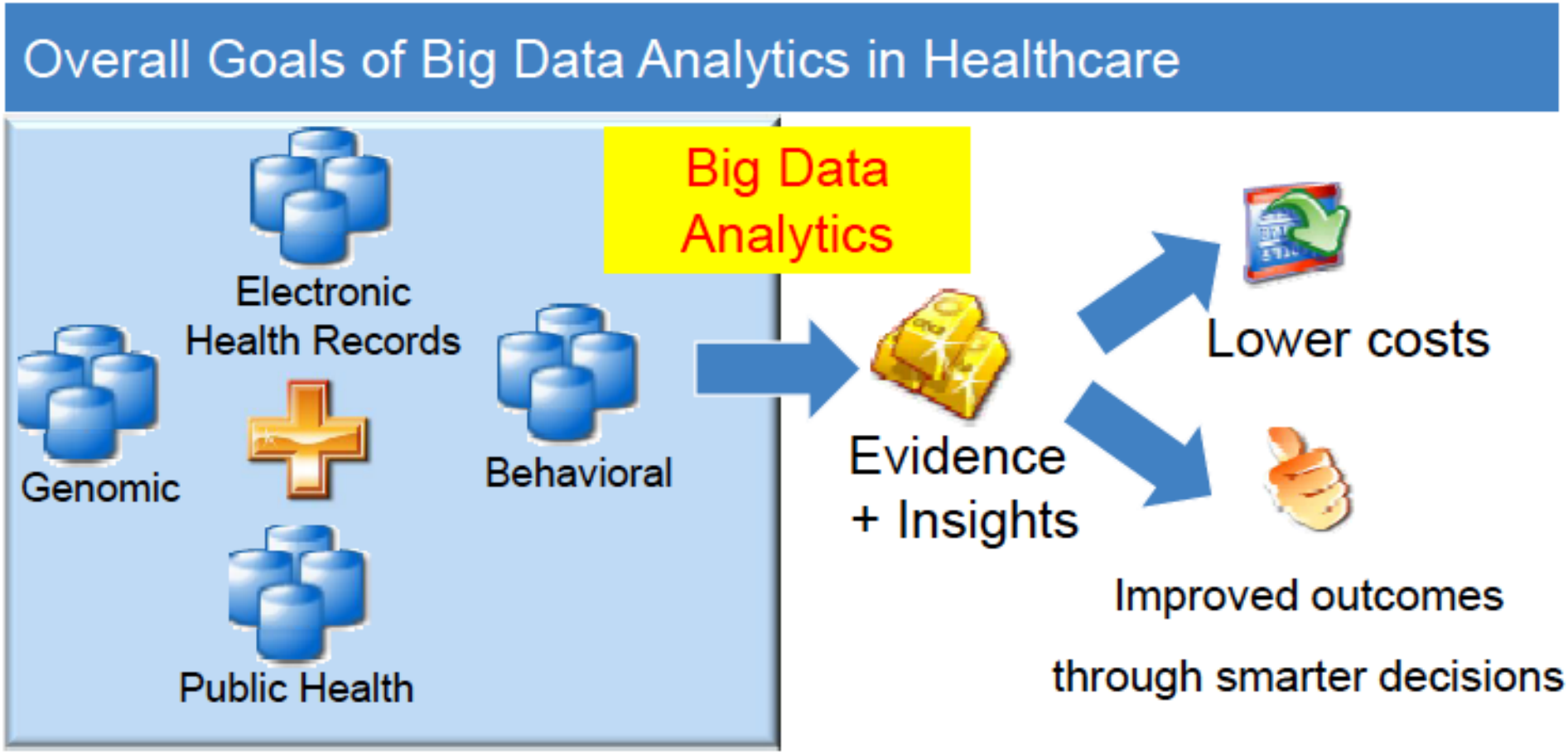
Big data

- Challenges:
  - Capturing
  - Storing
  - Searching
  - Sharing
  - Analyzing

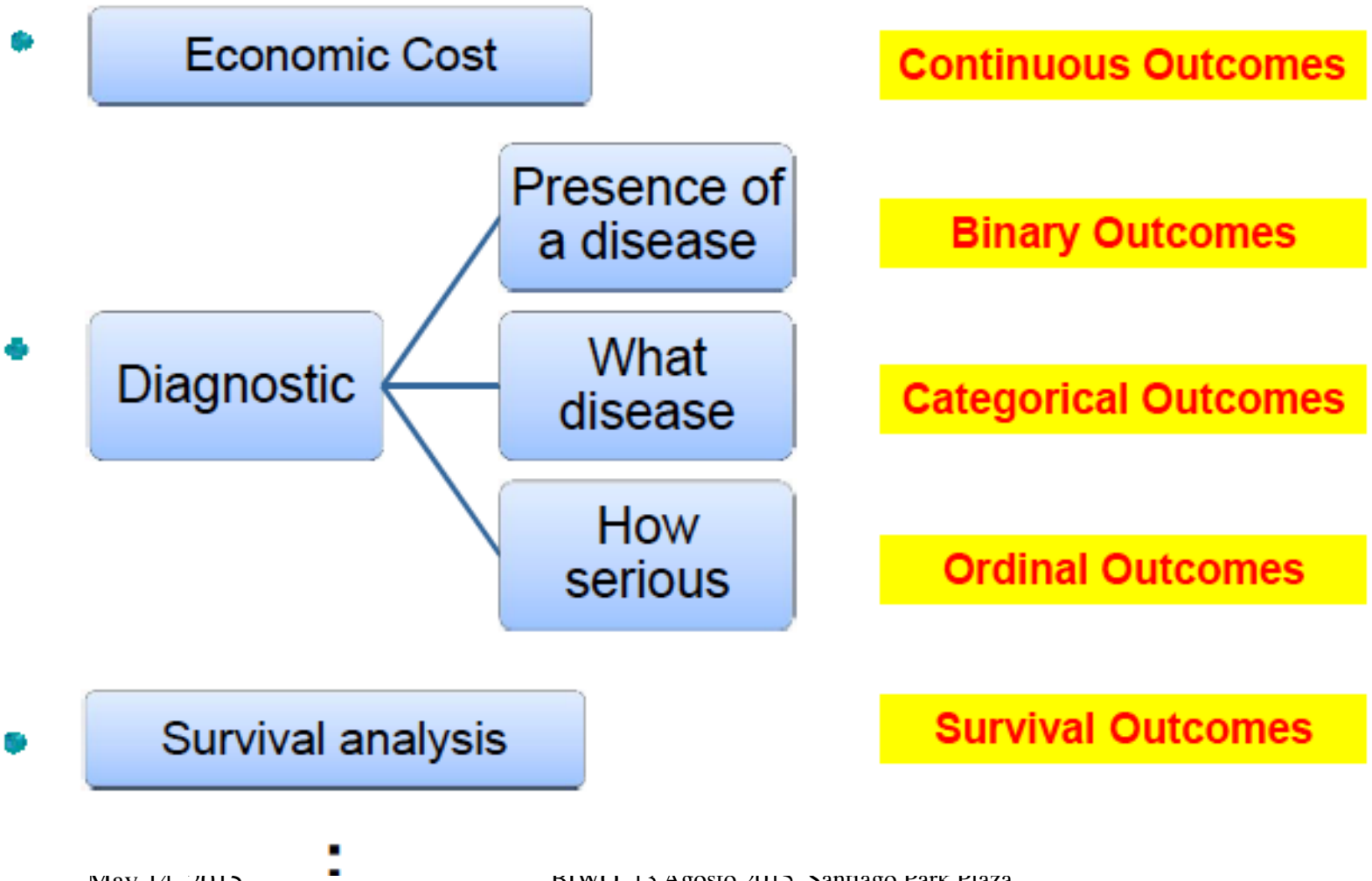


Big data


Clinical data
• Structured EHR
• Unstructured EHR
• Medical Images
Genomic data
• DNA sequences
Behavior data
• Social network data
• Mobility sensor data
Health data


Overall Goals of Big Data Analytics in Healthcare
Big Data Analytics
Electronic Health Records
Genomic
Behavioral
Public Health
Evidence + Insights
Lower costs
Improved outcomes through smarter decisions

Different Kinds of Outcomes



CASE STUDY: HEART FAILURE READMISSION PREDICTION

Big Data Analytics for Healthcare

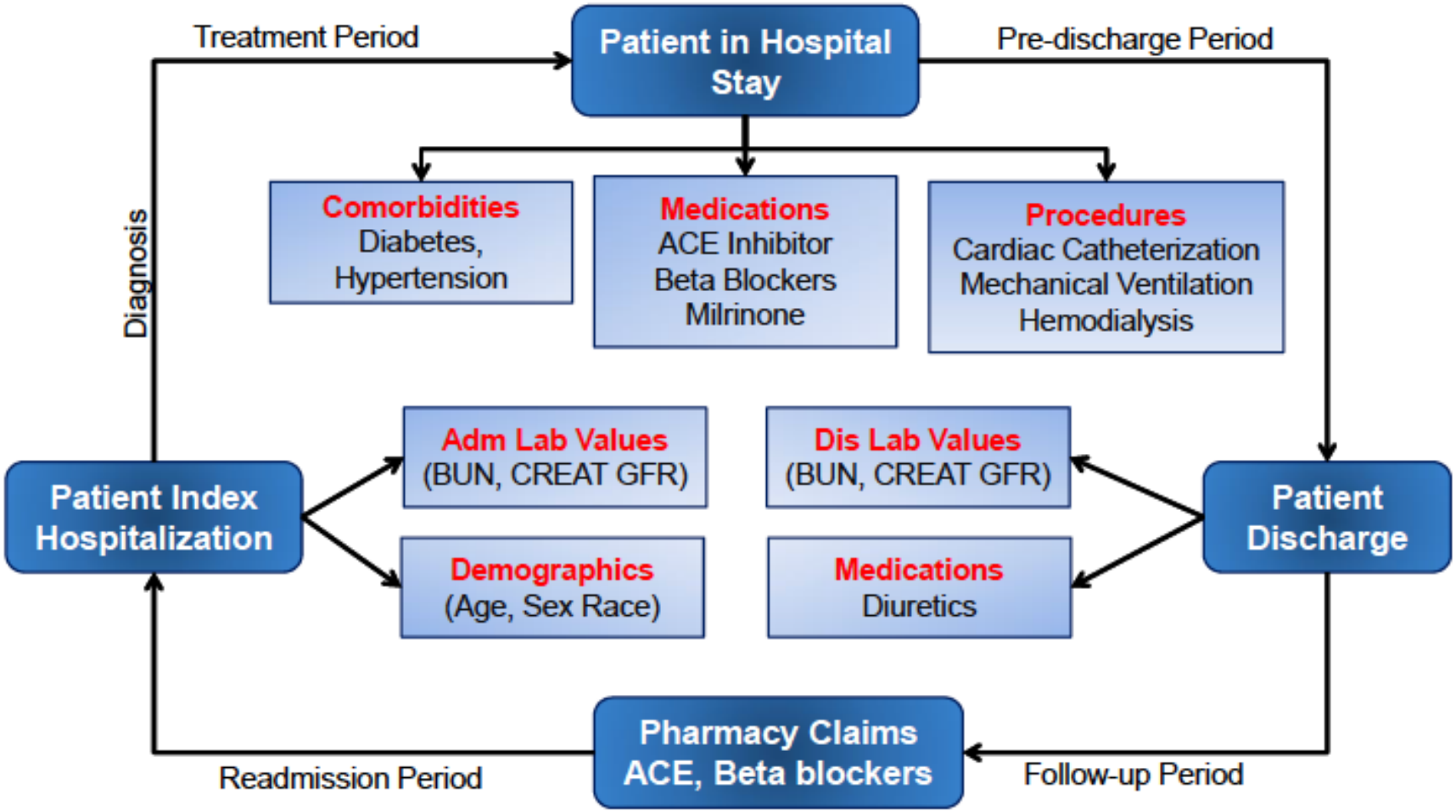
http://dmkd.cs.wayne.edu/TUTORIAL/Healthcare/

Penalties for Poor Care - 30-Day Readmissions

- Hospitalizations account for more than 30% of the 2 trillion annual cost of healthcare in the United States. Around 20% of all hospital admissions occur within 30 days of a previous discharge.
  - not only expensive but are also potentially harmful, and most importantly, they are often preventable.

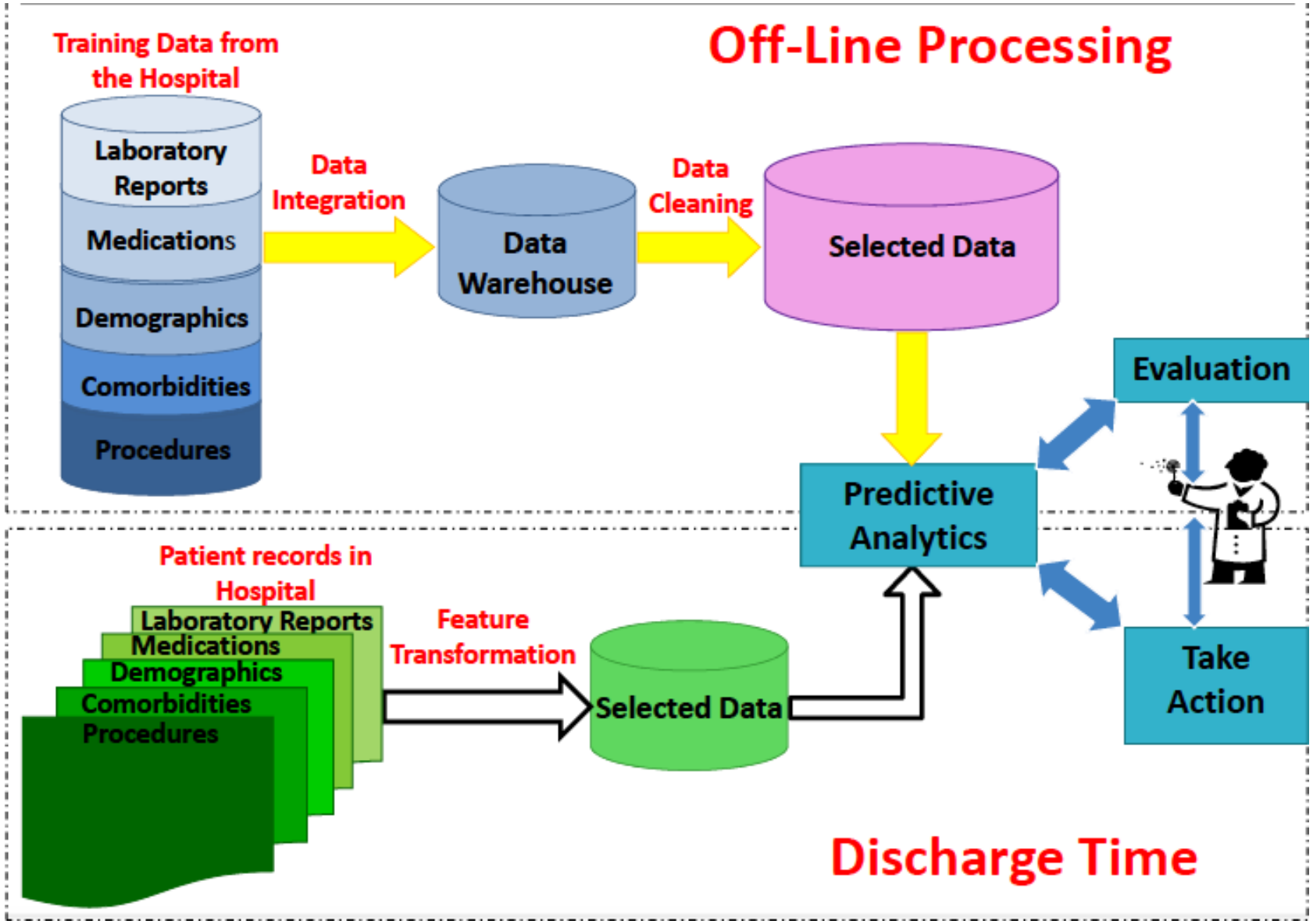
- Medicare penalties from FY12 - heart failure, heart attack, and pneumonia.
- Identifying patients at risk of readmission can guide efficient resource utilization and can potentially save millions of healthcare dollars each year.
- Effectively making predictions from such complex hospitalization data will require the development of novel advanced analytical models.

Patient Readmission Cycle



Challenges with Clinical Data Analysis

- Integrating Multi-source Data – Data maintained over different sources such as procedures, medications, labs, demographics etc.
- Data Cleaning – Many labs and procedures have missing values.
- Non-uniformity in Columns – Certain lab values are measured in meq/DL and mmol/DL.
  - Solution: Apply the logarithmic transformation to standardize and normalize the data.
- Multiple instances for unique patients – Several records from different data sources are available for the same hospitalization of a patient.
  - Solution: Obtain summary statistics such as minimum, maximum and average for aggregating multiple instances into a single one.


Off-Line Processing
Training Data from the Hospital
Laboratory Reports
Medications
Demographics
Comorbidities
Procedures
Data Integration
Data Warehouse
Data Cleaning
Selected Data
Evaluation
Predictive Analytics
Take Action
Patient records in Hospital
Laboratory Reports
Medications
Demographics
Comorbidities
Procedures
Feature Transformation
Selected Data
Discharge Time

engine

UPV EHU

Interventions – Depending on the Risk scores at the Discharge

HOSPITAL RESOURCE UTILIZATION

Predictive Analytics

-  Staying back in hospital
- Home nurse visitation 
-  Telephone support
- Automated reminder 

HIGH

RISK

COST

LOW

THE COST SPENT FOR PATIENT CARE IS DETERMINED BY THE RISK

Social networks

- Social networks[1] allow
  - To share information
  - Counseling
  - Amateur epidemics
    - Analysis of google data uncovers flu epidemy

[1]https://www.patientslikeme.com

Social Media can Sense Public Health !!

During infectious disease outbreaks, data collected through health institutions and official reporting structures may not be available for weeks, hindering early epidemiologic assessment. Social media can get it in near real-time.

Twitter messaging correlated with cholera outbreak

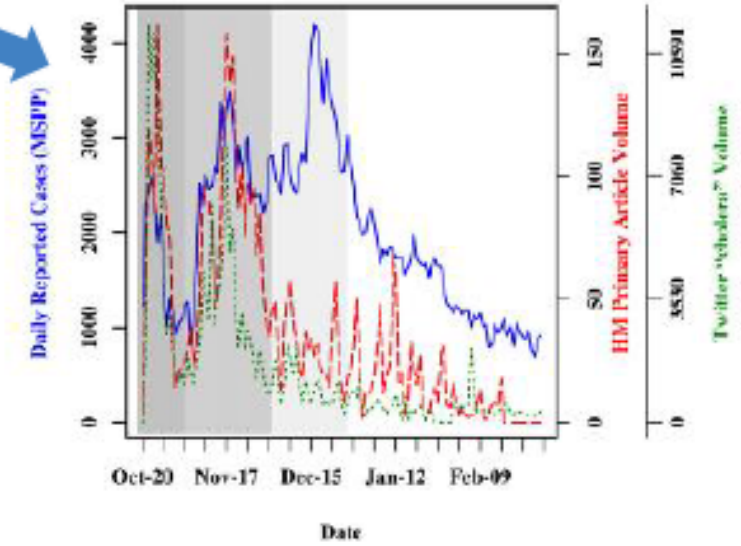
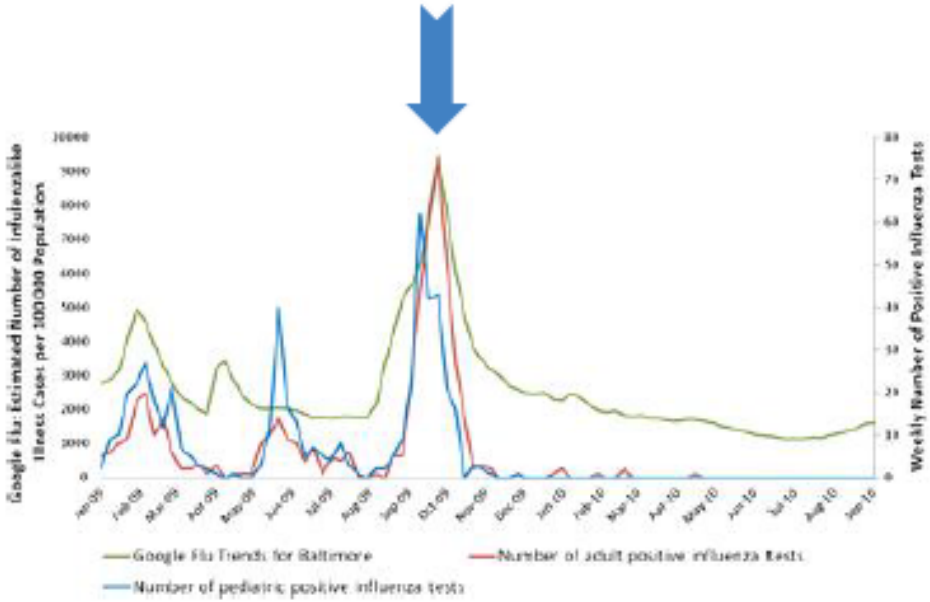


FIGURE 2. Daily reported case data for all departments from the Haiti Ministry of Health (solid), daily volume of primary HealthMap alerts (dashed), and daily volume of Twitter posts containing the word "cholera" or "#cholera" (dotted). Each curve has an initial peak at the onset of the outbreak (dark grey), and a peak during the time that Hurricane Tomas affected Haiti (medium grey). The first 100 days of the outbreak are shaded in light grey. Ministère de la Santé Publique et de la Population (MSPP) case counts peak again in late December, although HealthMap and Twitter volume only have daily variations during this time.

Chunara, Rumi, Jason R. Andrews, and John S. Brownstein. "Social and news media enable estimation of epidemiological patterns early in the 2010 Haitian cholera outbreak." *American Journal of Tropical Medicine and Hygiene* 86.1 (2012): 39.

Google Flu Trends correlated with Influenza outbreak



Dugas, Andrea Freyer, et al. "Google Flu Trends: correlation with emergency department influenza rates and crowding metrics." *Clinical infectious diseases* 54.4 (2012): 463-469.

Contents

- Introductory ideas and history
  - Definition
- Vertical technological challenges
  - Privacy, security and authenticity
  - Interoperability
  - Workflow embedding
- Horizontal technological dimensions
  - Big Data
  - Machine Learning and Computational Intelligence
  - Social networks and personal health management
- Business opportunities

Barriers to adoption of EHR

- Financial: is the system affordable and profitable?
- Technical: The EHR may require new computer skills to the physicians and staff members.
- time-related: learning and conversion of data to digital format for system initialization,
- psychological: changing habits and professional autonomy,
- social: the reaction of the social body to the system, adopting or rejecting it,
- organizational: how organizations and people assimilate changes in established processes. Is the system impeding or boosting the workflow?

Business oportunities

- Improving workflow:
  - Better interfaces for the human
    - Natural language processing
  - Better connectivity of EHR with devices
    - Body sensors
    - Wearable sensors
- Cloud implementations allow widespread deployment and optimal cost

Business

- Big data related
  - Data management and visualization
  - Data mining
    - Avoid prescription errors
    - Association discovery
- Intelligent anonymization processes

- Open standards: openEHR
  - Improve quality and confidence of th users
  - Validation and verification is open
- Social media specialization
  - Language
  - Endemic local diseases
  - Local demographic trends and tracks

Business

- Implementation of EHR and related techonologies is Government policy
  - In USA doctors get extra money for demonstrable use of EHR
  - Good: Government research funding programs
    - Good for start-ups
  - Bad: Government tend to trust big companies (trans-national)

Thank you for your attention

Electronic Health Record: implementation, data mining, security and user acceptance, EHRW2015

http://www.ehu.eus/ccwintco/index.php?title=EHR_Workshop_2015