

SEMANTICS ENABLED
PROACTIVE AND TARGETED
DISSEMINATION OF NEW
MEDICAL KNOWLEDGE

Lakshmith Ramaswamy & I. Budak Arpinar

Dept. of Computer Science, University of Georgia

laks@cs.uga.edu, budak@cs.uga.edu

MOTIVATION

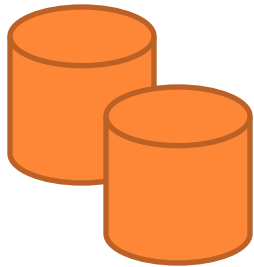


Prilosec + Plavix

Mr. Smith



Dr. Brown



Welcome | My Account

Home | Current Issue | Past Issues | Topic Collections | CME | Multimedia | Subscribe | Physician Jobs | For Authors

Original Contribution

EXPAND »

« Previous | Next Article
Table of Contents

Risk of Adverse Outcomes Associated With Concomitant Use of Clopidogrel and Proton Pump Inhibitors Following Acute Coronary Syndrome

P. Michael Ho, MD, PhD; Thomas M. Maddox, MD, MSc; Li Wang, MS; Stephan D. Fihn, MD, MPH; Robert L. Jesse, MD, PhD; Eric D. Peterson, MD, MPH; John S. Rumsfeld, MD, PhD

[+] Author Affiliations

ABSTRACT

This Article
JAMA. 2009;301(9):937-944
doi: 10.1001/jama.2009.26

- Abstract
- Full Text
- Full Text (PDF)
- *JAMA* Report Video

[–] SERVICES

- E-mail this article to a colleague
- Alert me when this article is published



News/Media

CardioSource Video News

Meeting Coverage

ACC Publications

— Journal of the American College of Cardiology

— JACC Cardiovascular Interventions

— JACC Cardiovascular Imaging

— Quality First SmartBrief

— CV News Digest

— Cardiology Magazine

Media Center

ACC in Touch

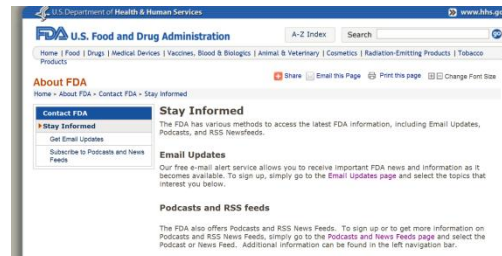
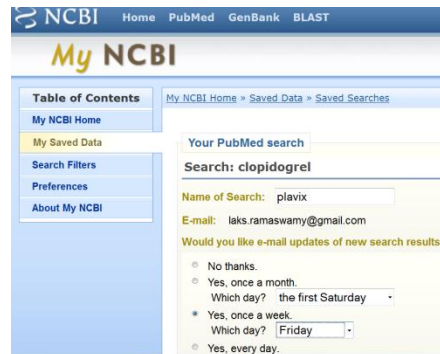
Home > News/Media > ACC Publications > Quality First SmartBrief

Quality First SmartBrief A Service of the PINNACLE Network

The Quality First SmartBrief is a free twice-weekly e-mail news briefing for health care stakeholders interested in quality care. It delivers must-read news about quality care directly to you each Tuesday and Thursday. Knowledgeable editors handpick key articles from hundreds of publications, summarize them and provide direct links to the original sources.

The Quality First SmartBrief features:

- Practice management tips that promote quality care
- E-health news and developments related to health IT, quality reporting and payment innovation
- Information on new products and technologies
- Exclusive updates from the American College of Cardiology and the

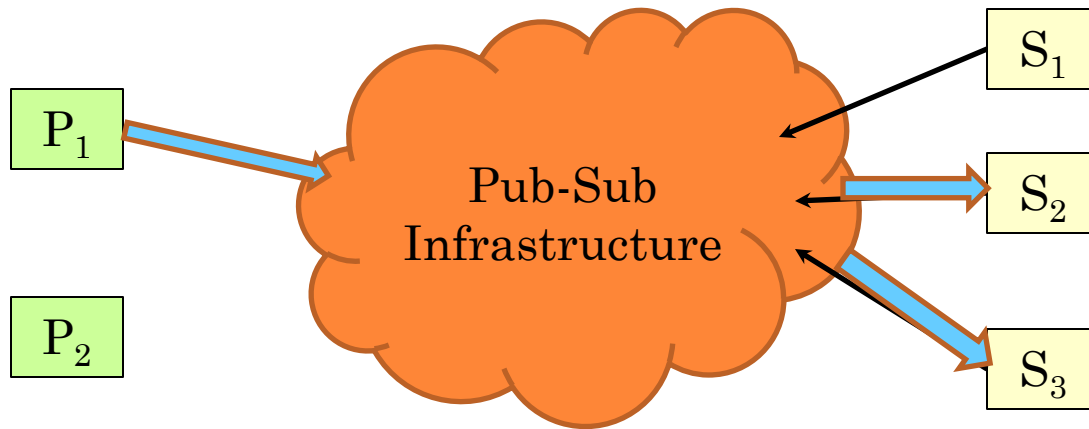


GOALS

- Cut down delays between medical discoveries and implementation via better knowledge dissemination
- Reduce information overload
- Minimal overheads on doctor establishments
- Medical information dissemination that is
 - Proactive – Push-based
 - Targeted
 - Timely



BACKGROUND: THE PUB-SUB PARADIGM



- Information filtering and propagation framework
- Subscriptions are continuous queries
- Used heavily in financial applications



QUERY TYPES

- Topic-based
 - Queries choose topics of interest from pre-specified hierarchy
- Content-based
 - Queries on content of published items
- Type-based
 - Queries specified on object types
- Distributed implementation to achieve scalability



PUB-SUB: PROS AND CONS

- Works well if
 - Know what information will be of interest
 - Anticipate the information that will be published
 - Published information is structured
 - Queries can be crisply specified
- Our scenario doesn't exhibit these characteristics
 - Scientific literature is free text
 - Cannot anticipate research
 - Too many queries or too coarse-grained queries



EHR AS BASIS FOR INFO DISSEMINATION

- Comprehensive and up-to-date (ideally !!) information about patient
 - Age, gender, physical activity, family support
 - Medical conditions, individual and family history
 - Drugs and medications, past reactions
- Can serve as basis for identifying relevant information
- Collected routinely as part of treatment process
- Fast increasing adoption

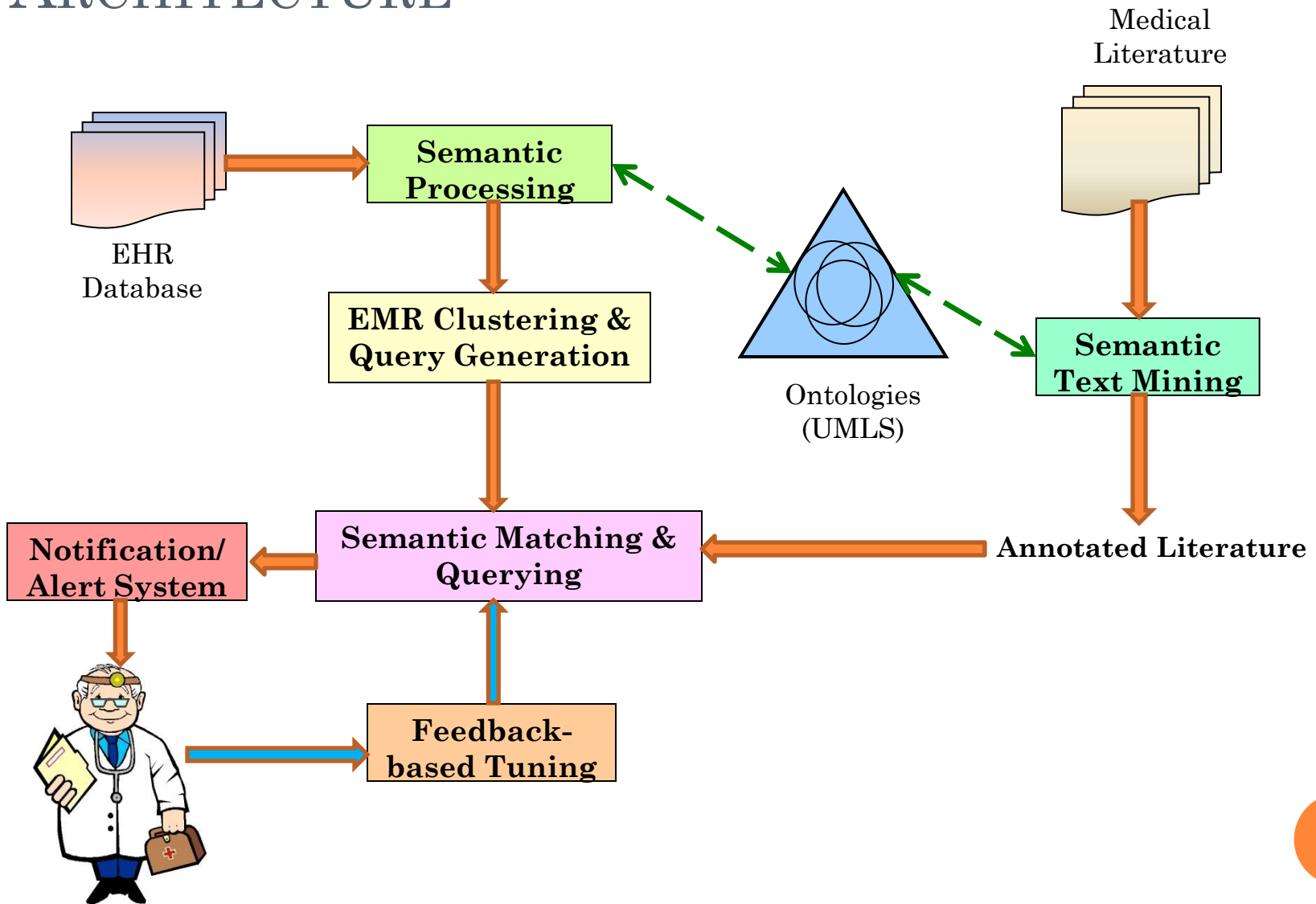


CHALLENGES

- Differences in terminologies and info representation
- Noisy EHR data
 - Incompleteness
 - Errors
- Personalization
- Scalability
 - Several thousand physicians, millions of EHRs

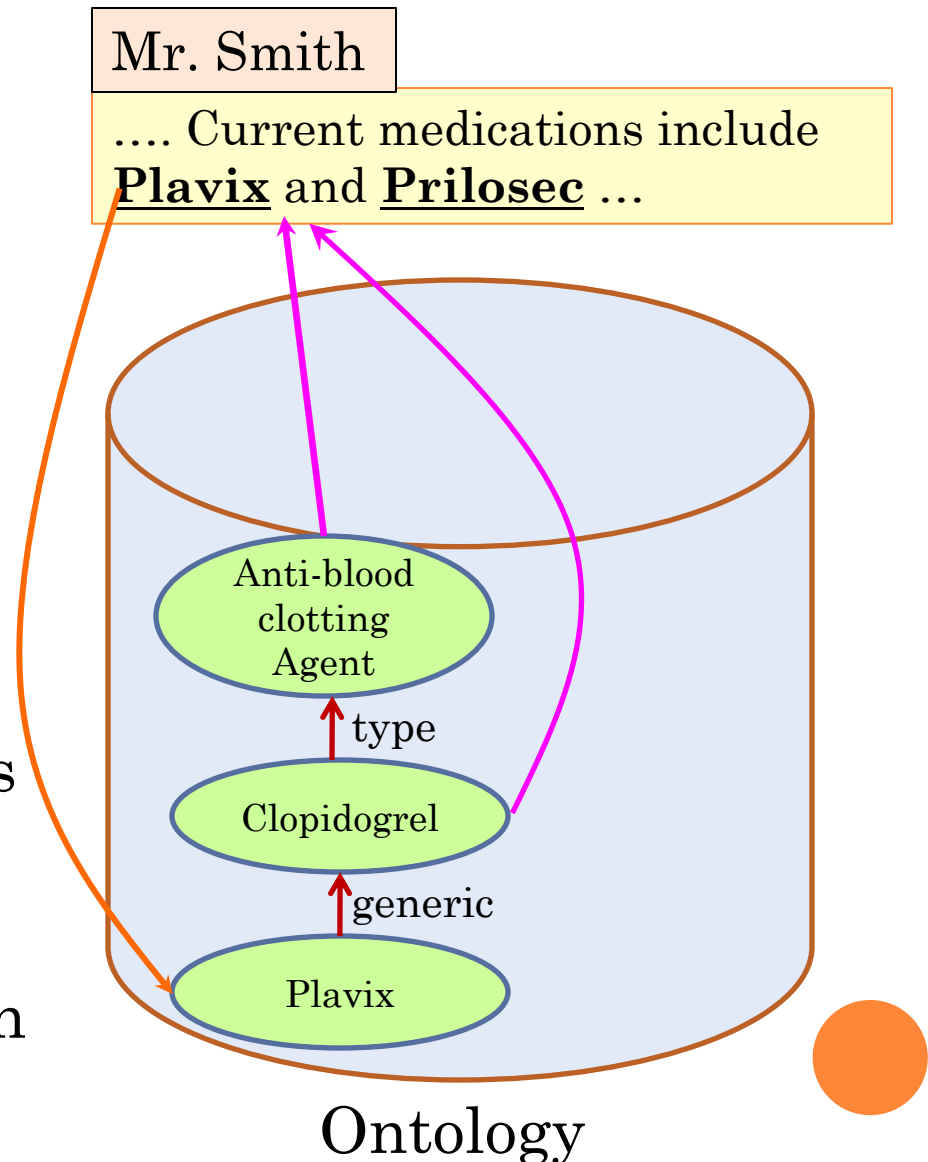


ARCHITECTURE



SEMANTIC ANNOTATION

- Foundation to support powerful query and relevance mechanisms
- Enhancing text with structured domain knowledge
- The result is a set of explicit assertions indicating named-entities within them
- Term, concept & relationship identification
- We use Stanford NLP



RELEVANCE DETERMINATION

- Content similarity b/w EHRs and medical articles
- Current approaches rely on syntactic similarity
 - Term vector approaches - Common words in documents and their relative importance
 - Importance measured in terms of relative frequencies (TF-IDF metric)
- Many shortcomings
 - Terminology differences
 - EHRs not likely to have significant word repetitions
 - Repetitions may not imply strong emphasis
 - Blind to relationships among words

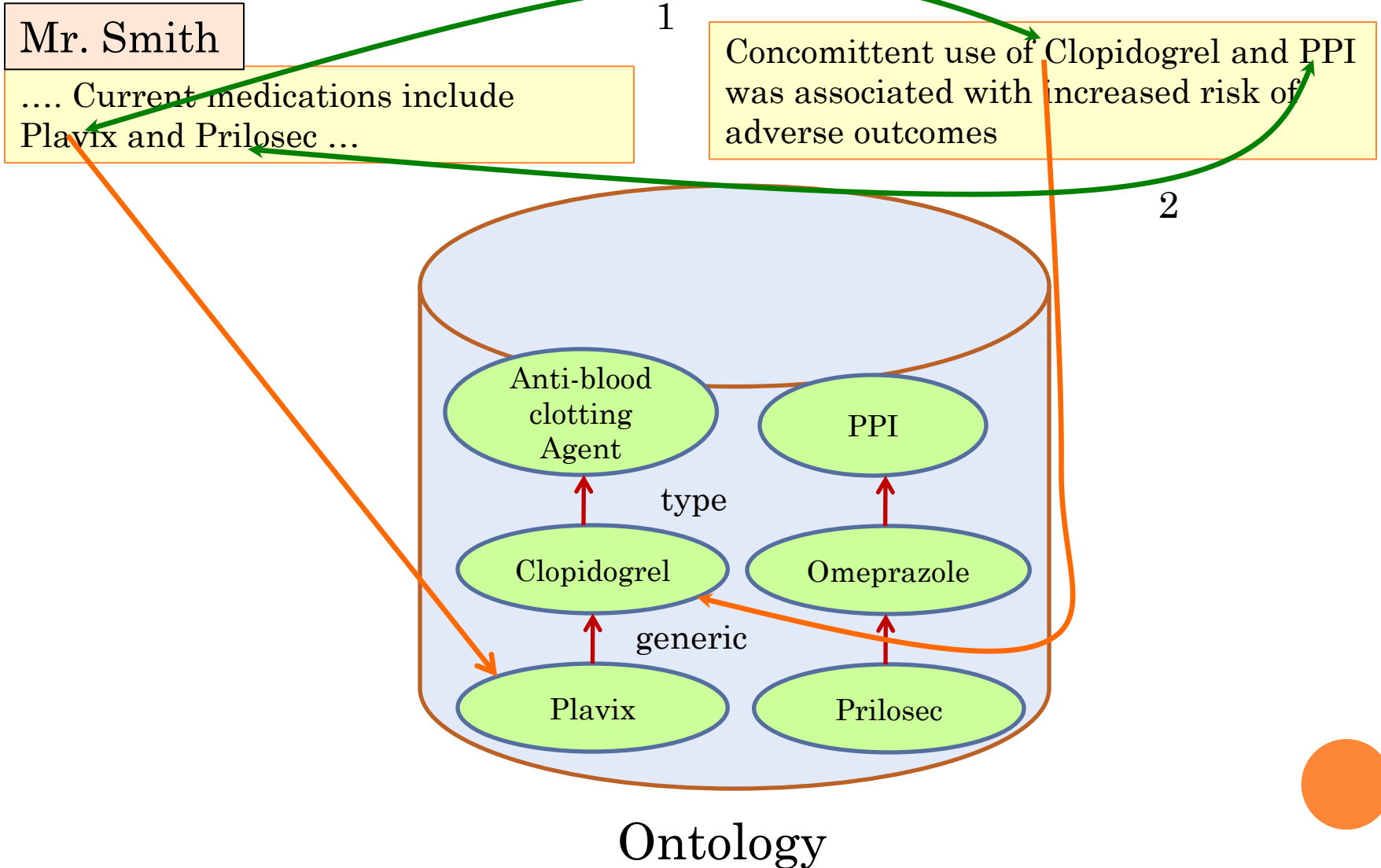


SEMANTICS DISTANCE-BASED RELEVANCE

- Quantifies relationship strength
- Based on structure of domain ontology
 - Min hops between concepts
 - # paths between concepts
 - Weighted hop distance
- Compute semantic distance b/w concept pairs
- Aggregate semantic distances
- Can be used for relevance determination, ranking etc.

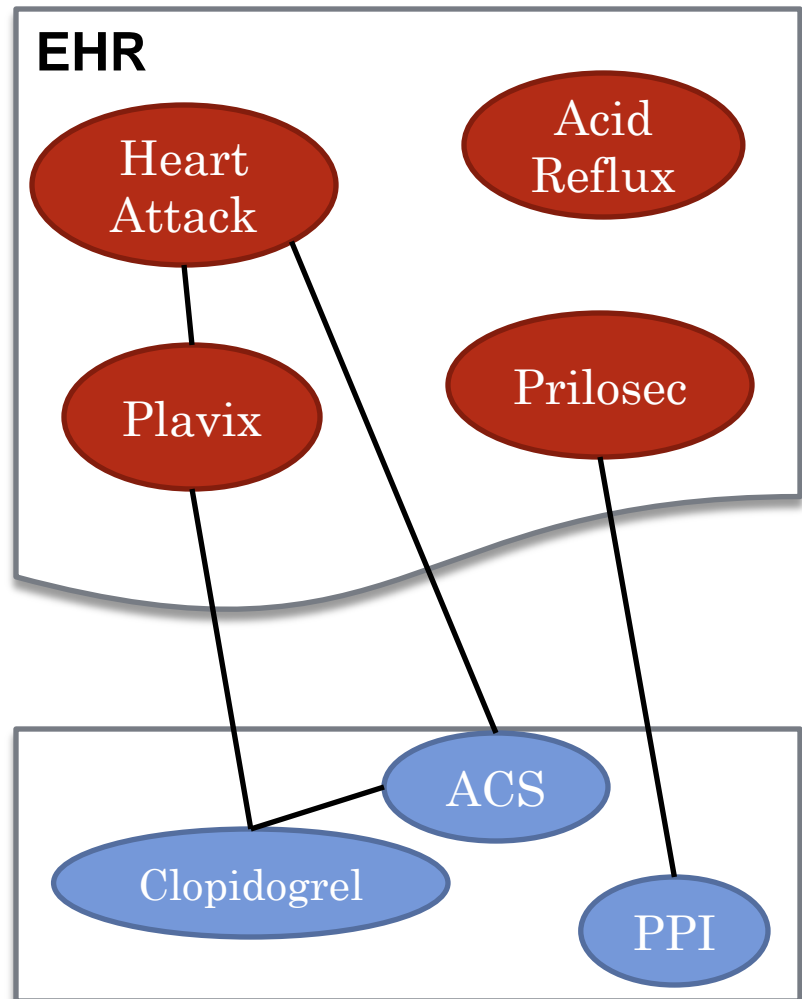


SEMANTICS DISTANCE



SEMANTICS GRAPHS-BASED RELEVANCE

- EHRs & articles mapped to semantic graphs
- Structural correlation of graphs
 - Identify links among concepts in EMR's and articles
 - Sub-graphs with dense cross-cutting paths signify higher degree of relevance
- More powerful but computation intensive



LOTS OF INTERESTING QUESTIONS

- Fuzzy matching
- How to associate weights with ontology links?
- How to incorporate user feedback?
 - Can it be used to strengthen/weaken relationships in ontology?
- How to scale the system?
 - Clustering EHRs
 - Distributed processing – Semantic overlays, Cloud
- Patient privacy issues
 - How much information can be exposed?
 - Where does the processing occur?

