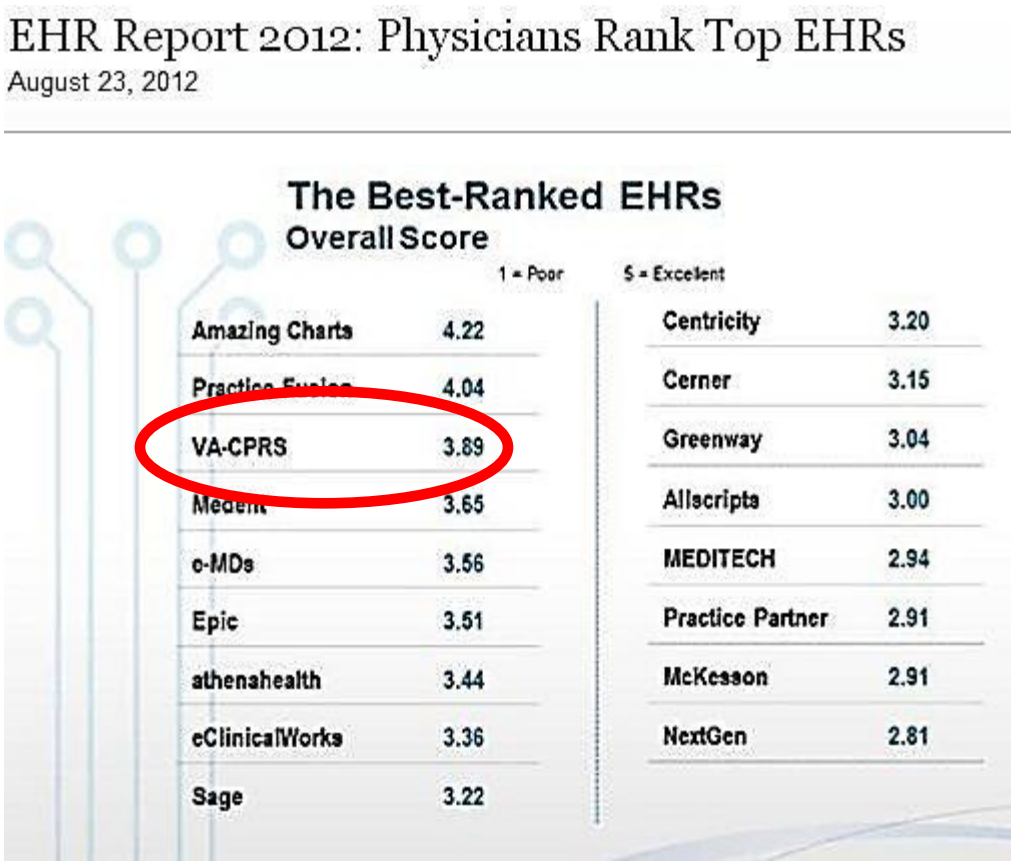
A Patient Centric Data Retrieval and Visualization Program

David Eibling MD
VA Pittsburgh Healthcare System
David.Eibling@va.gov

Augie Turano PhD
IT Director VINCI
Augie.Turano@va.gov

VISTA still Near Top of the Deck



EHR Report 2012: Physicians Rank Top EHRs

August 23, 2012

The Best-Ranked EHRs

Overall Score

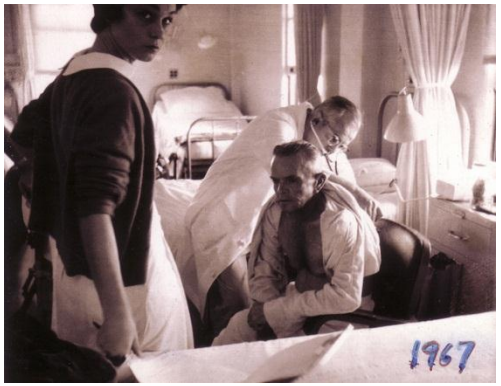
1 = Poor 5 = Excelent

Amazing Charts	4.22	Centricity	3.20
Practice Fusion	4.04	Cerner	3.15
VA-CPRS	3.89	Greenway	3.04
Medent	3.65	Allscripts	3.00
e-MDs	3.56	MEDITECH	2.94
Epic	3.51	Practice Partner	2.91
athenahealth	3.44	McKesson	2.91
eClinicalWorks	3.36	NextGen	2.81
Sage	3.22		

http://www.medscape.com/features/slideshow/EHR2012

Healthcare decisions have always been Information-Based

- Ancient writings
  - *"If a patient has this, then you should do that"*
- Traditional healers – until 1950's??
  - Apprenticeship – learn what to do by watching teachers
- Modern medicine
  - Scientific principles, data driven
  - Compensate for imperfections of human memory
  - Data necessary for decision-making
- Healthcare System
  - Decisions about resource allocation, oversight, etc

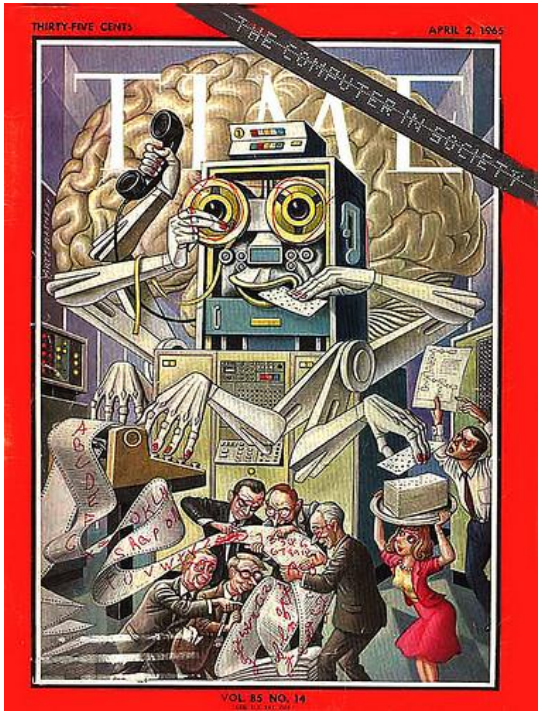


Bay Pines VA 1967

Photo from: http://www.flickr.com/photos/chris-warren-photos/2220257496/sizes/z/in/photostream/

What Information do **We** Need?

- Information about **This patient**
  - History of this disease – a narrative
    - *"Tell the story"*
  - Other facts about this patient
    - PMH, FH, SH, etc
- Information about **This disease**
  - Etiology
  - Natural history
  - Diagnostic tools
  - Ameliorating factors
  - Optimal Treatment
- Information about **This treatment**
  - Selection criteria
  - Probabilities of success, failure, other effects



Time April 1965

http://www.flickr.com/photos/79996062@N08/7344930088/sizes/m/in/photostream/

"If you install an EHR make sure you put a shredder by every printer" Dan O'Donnell MD

What Information do THEY* Need?

- Information about this **patient**
  - Nature of this disease
    - What is it, exactly?
    - *Coded information, not a story*
  - Other facts about this patient
    - Other illnesses
    - Who is paying for this care?
- Information about this **treatment**
  - What was done?
  - Where, by whom, with what resources?
- Information about the **outcomes** of the treatment
  - Did it work?
  - Were there complications?
  - What did the patient think?

* "They" = payers, govt agencies, patient advocacy groups, researchers and many others

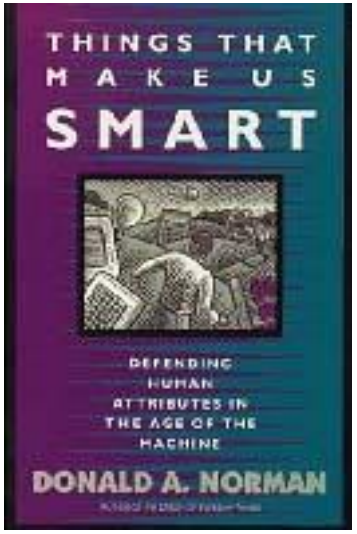
Why so much Information?

There is too much to know

- Increasing complexity of patients
  - Living longer
  - More illnesses
- Increasing complexity of care
    - And interactions between aspects of care
- More specialty care
  - More and more specialists know more and more about less and less
- More options – medications, treatment choices, etc
- Dramatic increase in literature
  - Up to 25,000 clinical trials published yearly
  - Thousands of Clinical Practice Guidelines
  - ***Even specialists can't know it all***

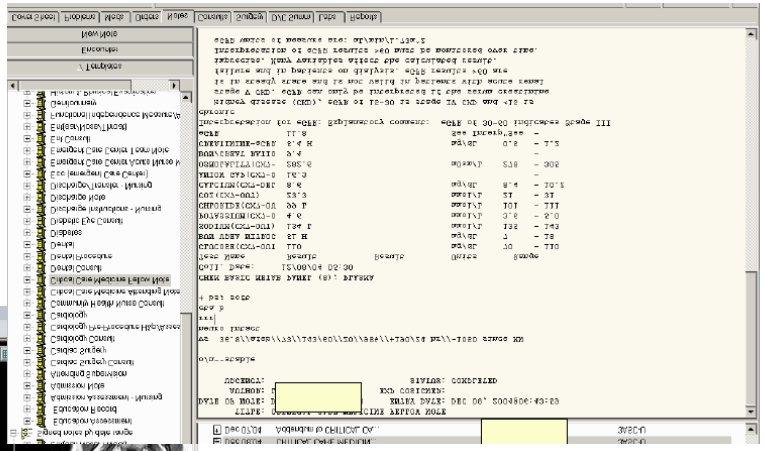
So we rely on External aids to acquire, organize, and access information

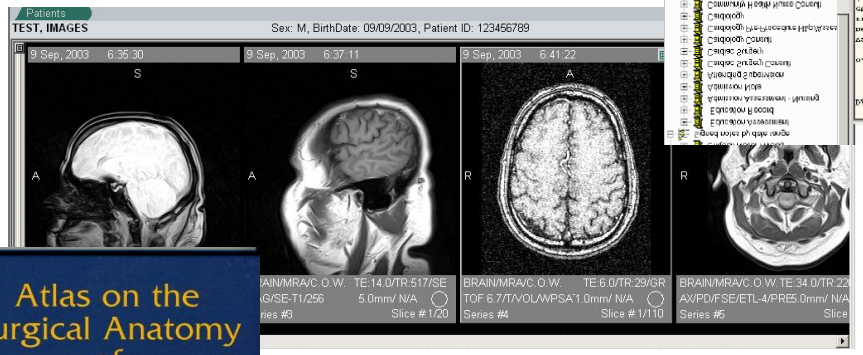
"The power of the unaided mind is overrated….(Its) real powers come from devising external aids that enhance (its) cognitive abilities."

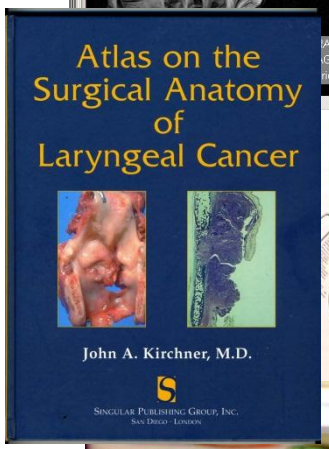


Norman. Things that Make us Smart 1993

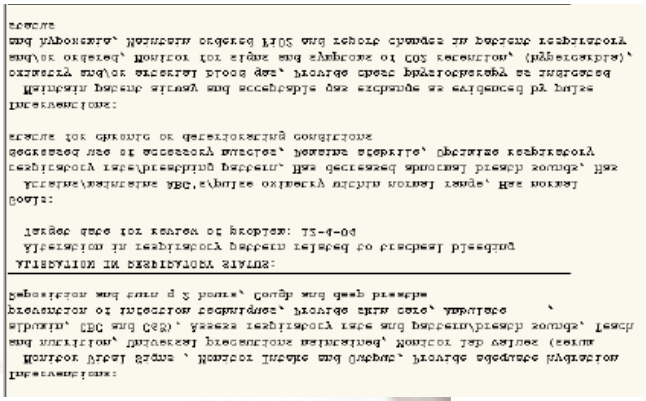
"External Aids" we use every day





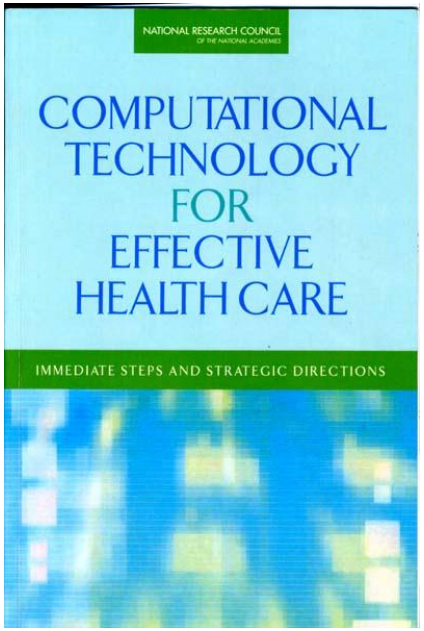


PubMed



EHR Systems Fail to support Physician Cognitive Tasks

"IT applications appear designed largely to automate tasks or business processes. They are often designed in ways that simply mimic existing paper-based forms and provide little support for the cognitive tasks of clinicians or the workflow of the people who must actually use the system"

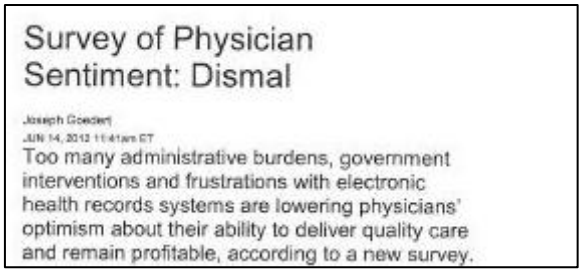


Report Released Jan 2009

Stead and Lin, eds. Computational Technology for Effective Health Care: Immediate Steps and Strategic Directions. National Research Council. National Academy Press Washington DC. 2009

What is the Effect on the Physician?

- **Increased time requirement**
  - More work shifted to physician
  - Estimated 30-40% reduction initially
  - Learning curve impact
- **Interference with Physician-Patient interaction**
  - Looking at computer rather than patient
    - More than 50% in my case
- **Impact on cognitive resources**
  - "Treating the computer" trumps treating the patient
- **Frustration, discontent with role, generalized unhappiness**
  - Burn-out symptoms in 50% !*
  - May be difficult to dissect out which is due to EHR itself or which is new administrative requirement

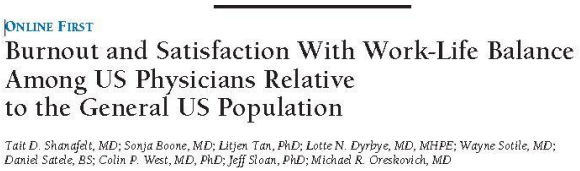

Survey of Physician Sentiment: Dismal

Joseph Goedert
JUN 14, 2012 11:41am ET

Too many administrative burdens, government interventions and frustrations with electronic health records systems are lowering physicians' optimism about their ability to deliver quality care and remain profitable, according to a new survey.

Health Data Management 23August2012

*Archives Int Med Online 20Aug2012

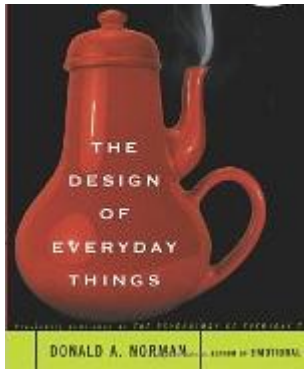

ONLINE FIRST

Burnout and Satisfaction With Work-Life Balance Among US Physicians Relative to the General US Population

Tait D. Shanafelt, MD; Sonja Boone, MD; Litjen Tan, PhD; Lotte N. Dyrbye, MD, MHPE; Wayne Sotile, MD; Daniel Satele, BS; Colin P. West, MD, PhD; Jeff Sloan, PhD; Michael R. Oreskovich, MD

How did this Happen?

- App Designers misinterpret intended use
- Poorly designed user interfaces
- Inadequate testing
- Too many administrative tasks



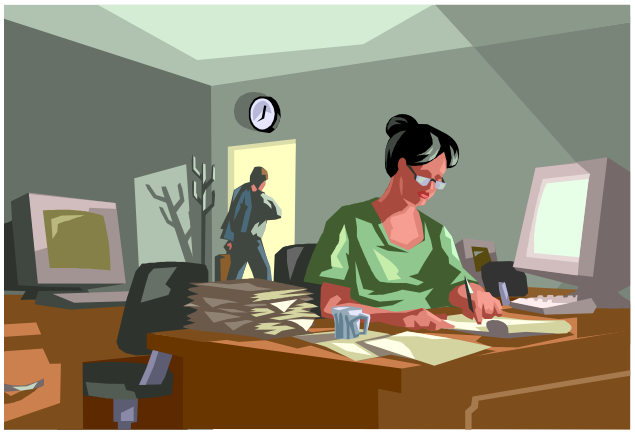


Image from Microsoft™

The most critical role of EHR Systems is

Clinical Decision Support

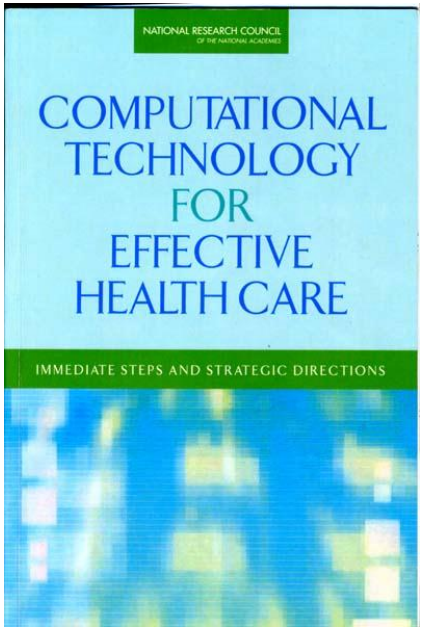
- Often defined as view alerts
  - But much more than that . . . . .

- "Global" definition – ***to provide all possible information to assist decision-making at the point of care***
  - Patient information
    - (this one)
  - Aggregated patient information
    - Information about all similar patients – predictive (pop stats) analysis
  - Evidence-based recommendations

Our Current Research

- **Web-based, secure, high speed, User-Centric Information Retrieval system for physician queries**
- **Accesses information within the VA VistA database**
- **Currently in prototype only**

"to exploit highly heterogeneous data effectively, users . . . need the ability to ask queries that span multiple data sources without requiring the data to be standardized . . . The user wants a single interface through which any query can be posed"



Stead Report Released Jan 2009

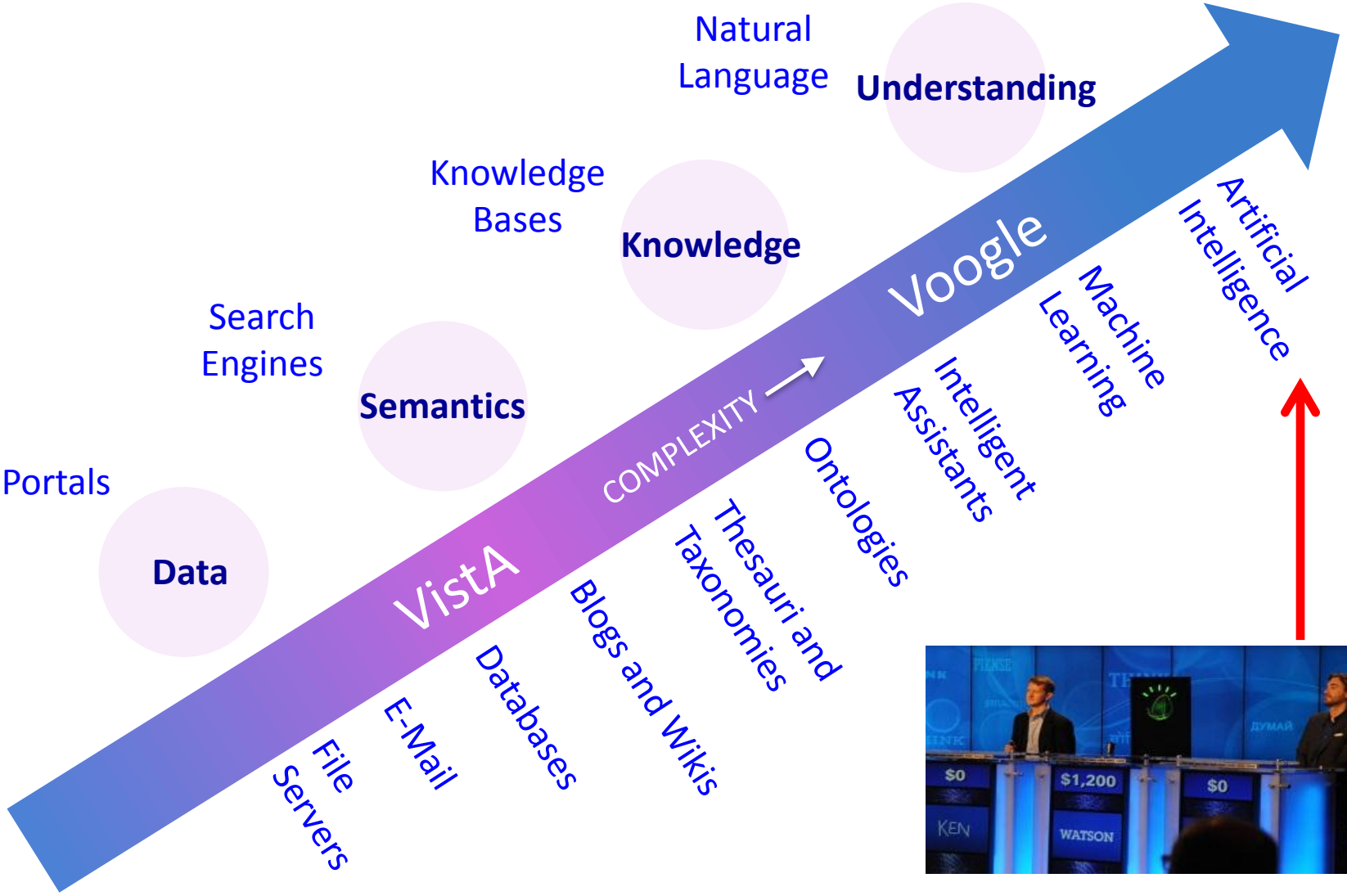
Stead and Lin, eds. Computational Technology for Effective Health Care: Immediate Steps and Strategic Directions. National Research Council. National Academy Press Washington DC. 2009

VOOGLE Features

- **Clean “Google™-like” interface**
- **Data domains not limited to VistA – seamless internal and external data access**
- **No training required**
- **Natural language queries**
- **Direct data access, no menu navigation required**
- **Aggregated displays of multiple data fields**
- **User-directed display paradigms**
- **User-defined single-word groups (macros)**

Voogle off-loads the search task !

Digital Evolution



http://www.flickr.com/photos/ibm_research_zurich/5410545854/sizes/m/in/photostream/

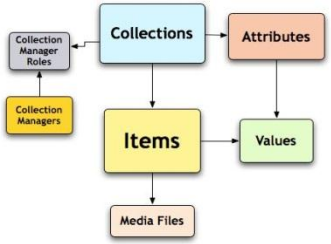
VOOGLE Knowledge Base

- **Knowledge base is key to Voogle functionality**
- **“Knows” where data located, or whether data present.**
- **Populated by a “crawler” that crawls through data structure and builds a “map” of data locations**
- **Can be configured to find information outside of VistA**
  - **“display expense form”**
  - **Could be programmed to access Pubmed**
- **Default display parameters can be overridden by user queries.**
  - **“ . . . as timeline , datagrid, chart”**

VOOGLE Methodology

Step 1:Crawl Data

Step 2: Create Metadata Knowledge Base



*

Step 3:Messaging and Visualization

*Image from http://blogs.law.harvard.edu/pkeane/2007/10/29/dase-metadata/

Getting Data

- FileMan Fields - FM API calls
  - Single value fields
  - Multiples
  - Pointers
  - Word Processing
- Documents – SF171, 9957 etc
- Table oriented data – SQL, Excel, etc
- RPC function calls
- Web Service calls – Bing, Google

FileMan Calls

- Set DIC=70000102,DIC(0)="QZM" D ^DIC
- D GETS^DIQ(DIC,DA,DR_"*","E","ANS", "ERR")
- S DIC=70000102,DR="4",DA=123 S DR(70000102.02)=".01:7",DA(70000102.02)=1 D EN^DIQ1

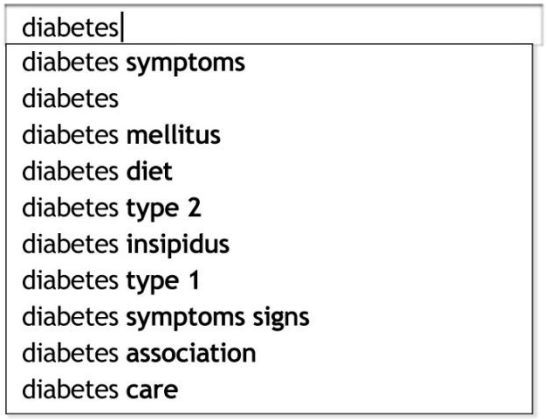
→Use indirection for dynamic program execution

VistA uses File Manager DB for data storage both structured and unstructured

Benefits

- Direct queries not a series of menus
- User adaption, retention, and satisfaction
- Decrease medical errors
- Decrease support and training costs
- Auto completion can help guide users

For example:



Mechanics

- Parse input - tokenize
- Isolate data fields call and quantifiers
- Consult knowledge base to get context and frames of reference
- Call web services, FileMan, RPC's calls from knowledge base index and search
- Choose best output format-table, chart

Architecture

- Commodity processors and mainstream operating systems
- Knowledge base to locate data and data concepts
- Fast retrieval implemented via new network (Infiniband) that is 30x faster than Ethernet
- SSD (solid state disks) that deliver 10x the data rates of SAN (storage area network)

Enhancing EHR Infrastructure

- Update infrastructure – new OS, New Filesystem, Fast IO
- Why is this important?
  - **Web Services**
  - **Commodity Processors, JBOD's – low cost**
  - **Fast dual networks (IB-Ethernet)**
  - **Objects**
  - **Fast performance**
  - **Encryption/security**
  - **Backups**
  - **High availability – paired db server mirrors**

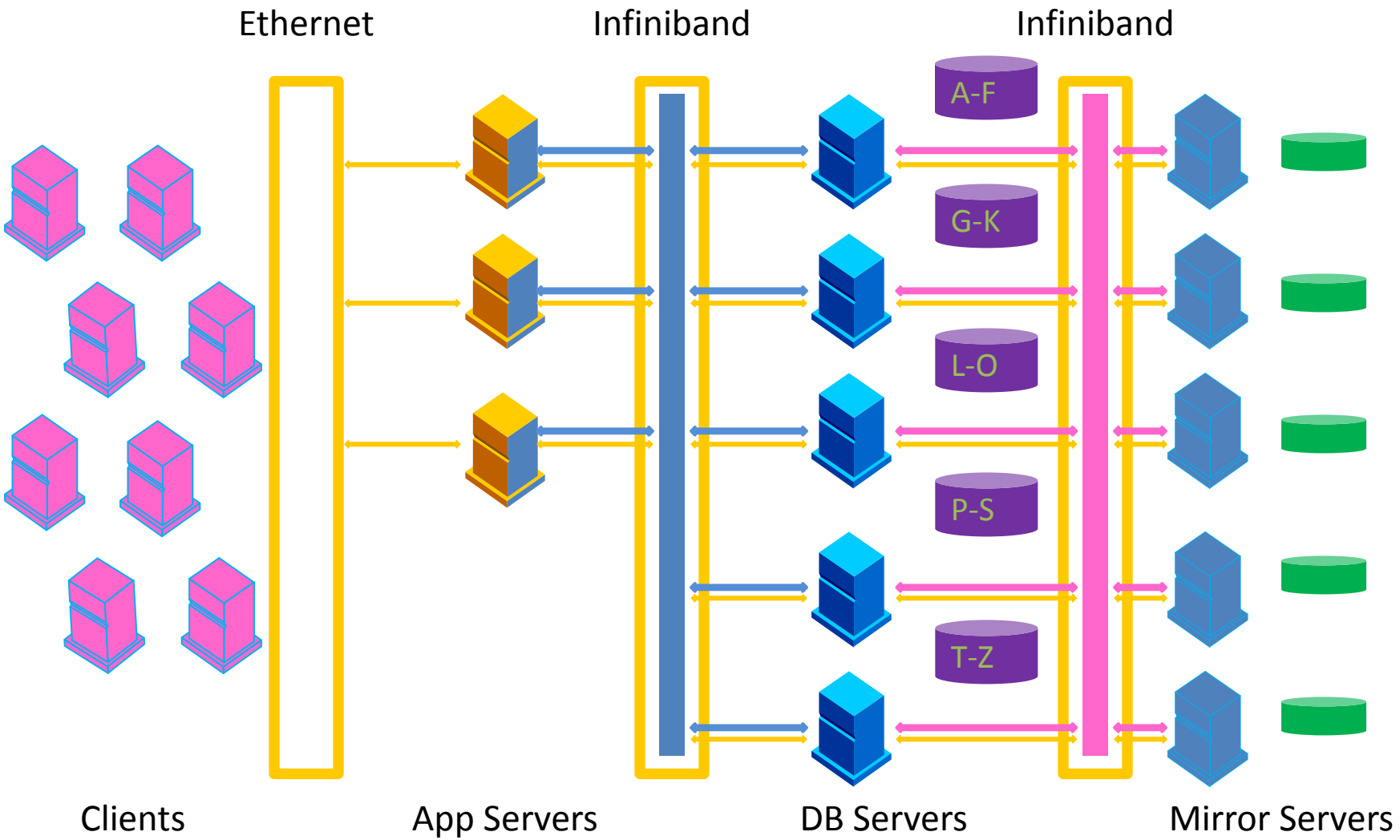
Hardware

- 12 Nehalem Servers (Quad)
- SATA JBODS with RAID Controller
- OS:Windows, Centos
- Networks:Infiniband (40gbs), Ethernet (1GeE)
- Horizontally scalable (2,400,000 gloref/sec typical VistA system does 250,000 gloref/sec)
- Storage:60 terabytes
- Cost:$98,000

Hardware

All commodity servers – no blades just rack mounts – swap them out if they break – no maintenance contract

VOOGLE Network Architecture



Security

- **VistA log in credentials**
- **Encrypted session identifier**
- **Knowledge base has authorization levels for each field**
- **User has authorization level – restricted by KB/User combination**
- **AES encryption (128 bit key)**

VOOGLE Summary

- **Powerful and simple**
- **High speed data retrieval**
- **Web-based**
- **Natural language interface**
- **Multiple data sources**
- **Adaptive knowledge base**
- **User-directed data displays**

Voogle Demo

Mining large amounts of Unstructured Data Using iKnow™ Software

- Natural “concept extraction”
- Enhancement to aid Decision Support
- iKnow ™ can “search” text on a single patient – or millions of patients
- Does NOT require structured data entry or rule building
- iKnow ™ will be incorporated into Intersystems Cache™ as of 2013

Unstructured Text

- **Far superior in "Telling the Story"**
  - **"All my patients look the same"**
- **Lots of information buried in free text**
- **Need a way to extract facts/concepts**
- **Discover correlations and relationships**
- **Can combine with structured data**
- **Classic word searchs are not always effective**

Image from Microsoft™

Concepts and Relations

Input: A Brown dog barked at the mailman.

Concept:Brown dog

Relation:Barked

Concept:Mailman

Barked

No rule creation or dictionary is necessary before indexing.

Images from Microsoft™

Ambiguity ????

BANK

- Of a river?
- Bank in the road
- Where you keep money
- Bank of computers

In Sentences --

- Breast cancer in women mushrooms.
- Teacher strikes idle students.

Process Breakdown

Classic Approach

Split text into sentences.

Split sentences into ***words***.

Use dictionary-based statistical models or syntactic parsing to ***guess*** which words might belong together.

Use a dictionary or ontology to ***correct*** and optionally enrich results.

i.Know Approach

Split text into sentences.

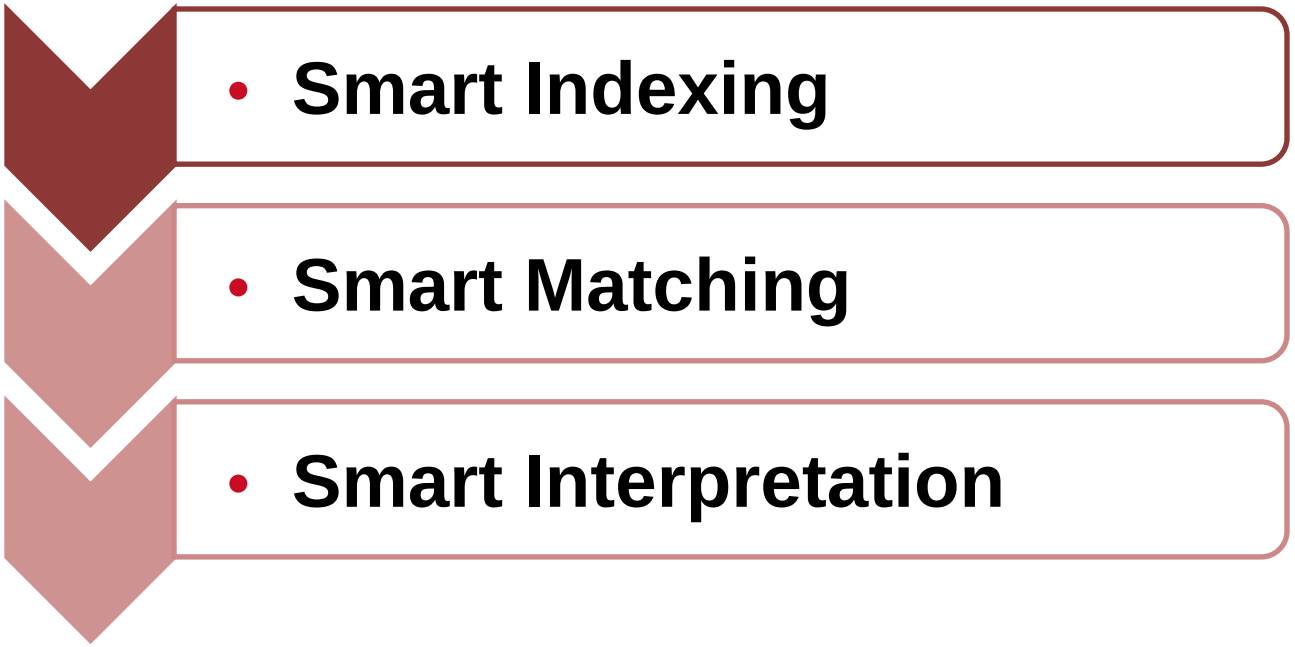
Identify relational elements based on ***compact*** language models.

Relational elements ***naturally*** separate meaningful concept clusters.

Optionally use a dictionary or ontology to enrich results.

Capabilities

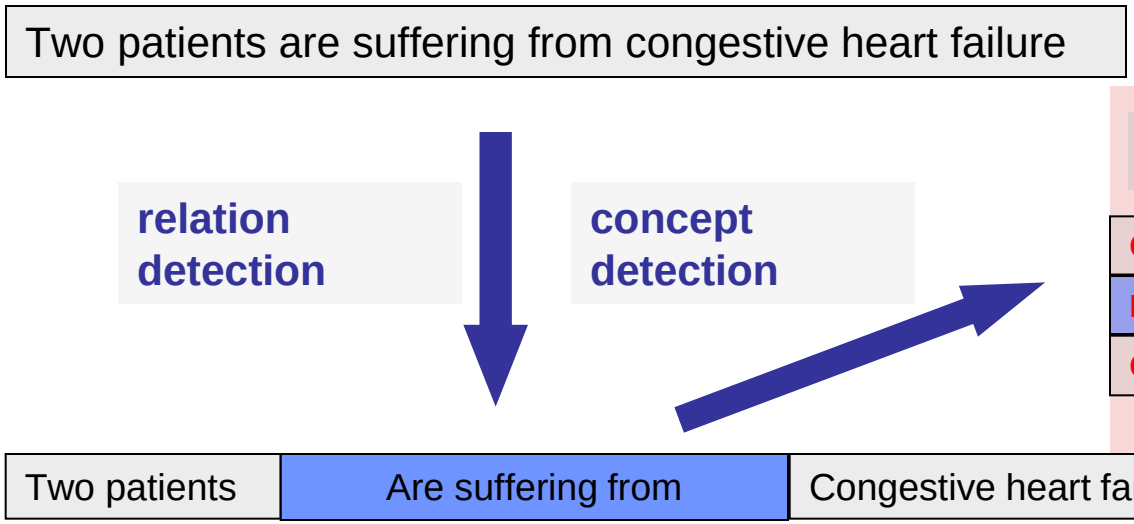
The iKnow approach to identifying concepts and relationships in unstructured text is only the starting point for more advanced analysis:



Smart Indexing

Two patients are suffering from congestive heart failure.

Smart Indexing (concepts and relations):



Smart Index

Concept	Two patients
Relation	Are suffering from
Concept	Congestive heart failure

Smart Mapping

The patient has « congestive heart failure »

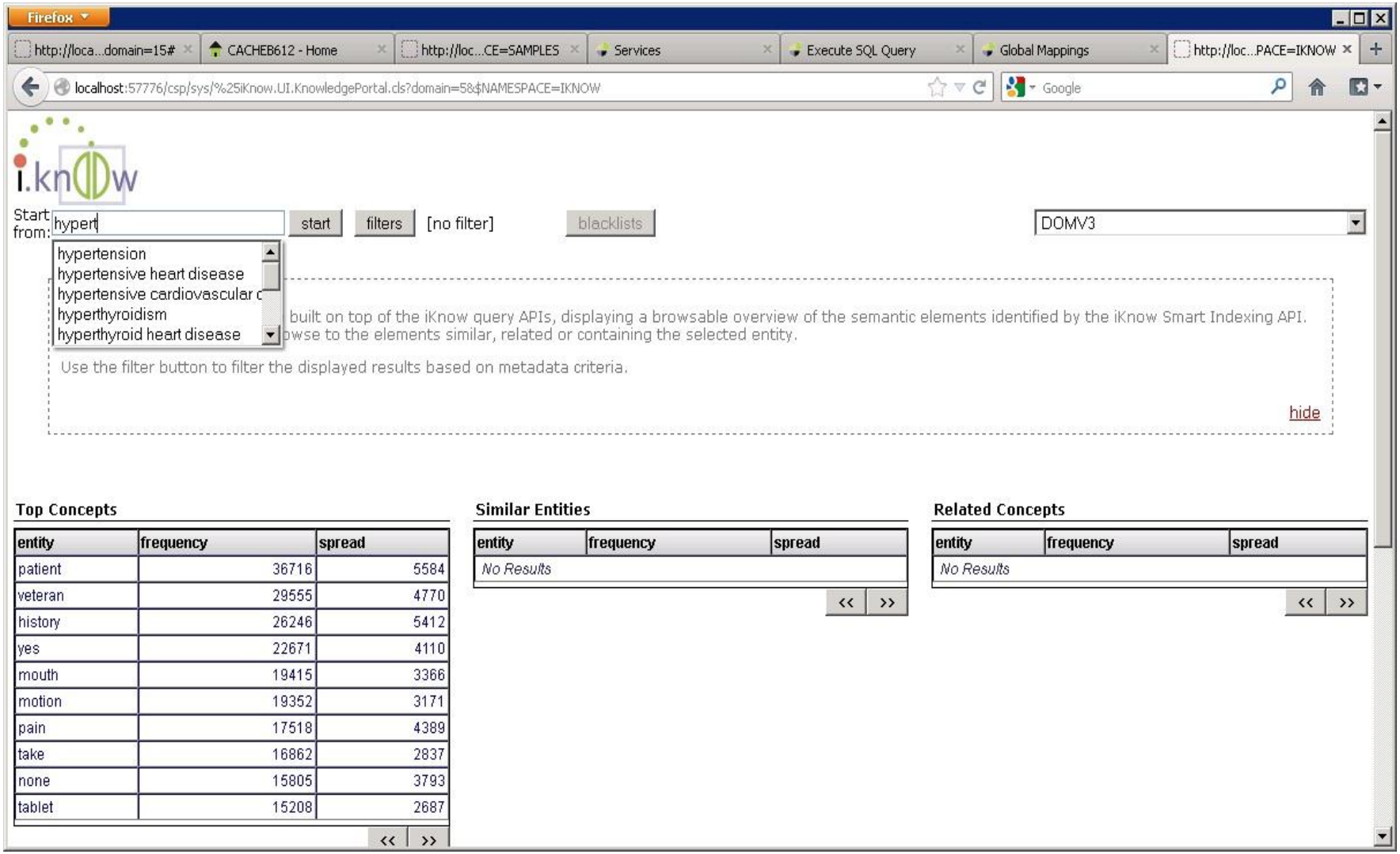
Congestive Heart Failure	**(Disorder)**	**SNOMED ID: 42343007**

The patient has an « *acute* congestive heart failure »

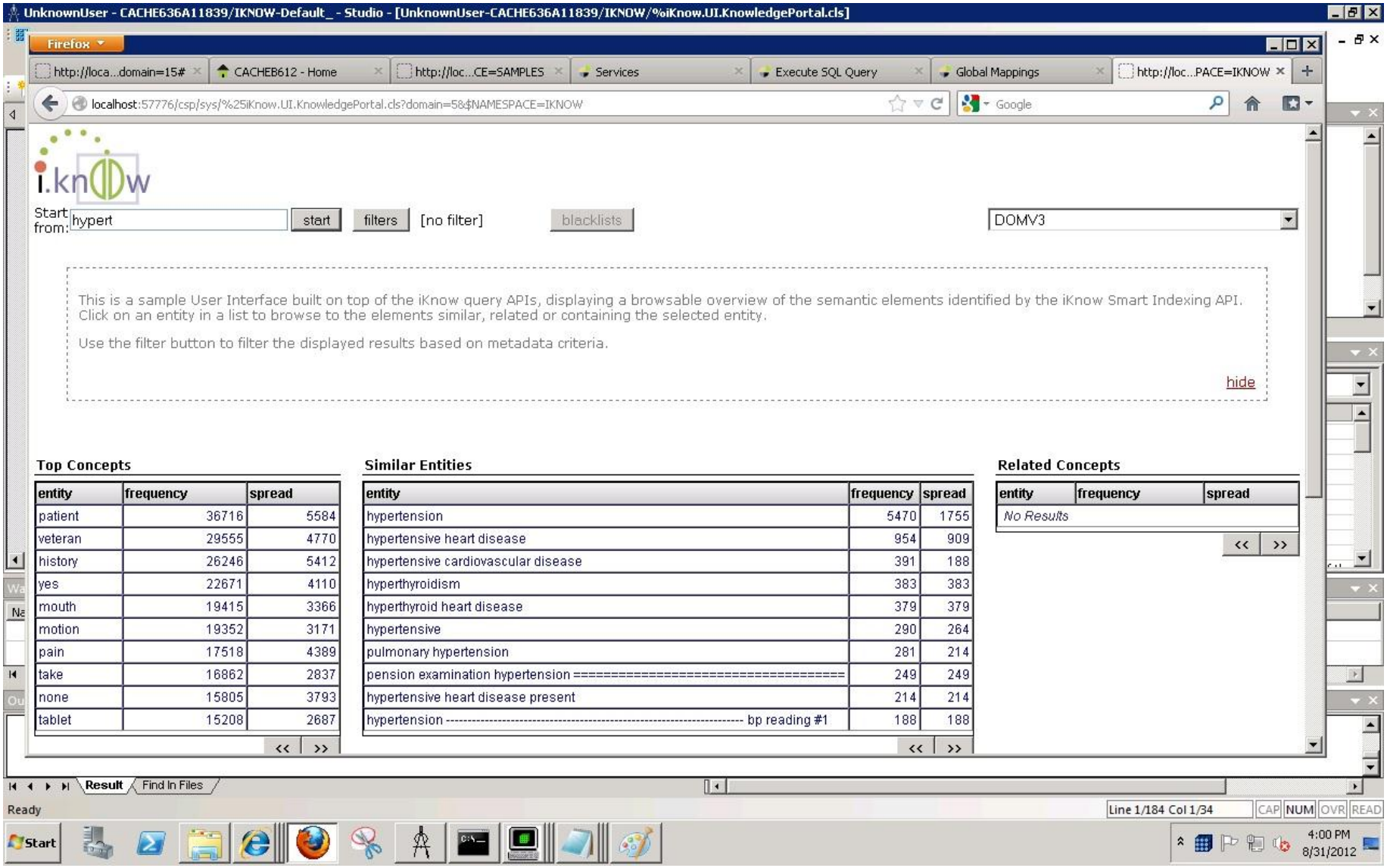
Congestive Heart Failure	**(Disorder)**	**SNOMED ID: 42343007**
Acute	**(Qualifier Value)**	**SNOMED ID: 373933003**

This can aid with ICD-10, UMLS mapping

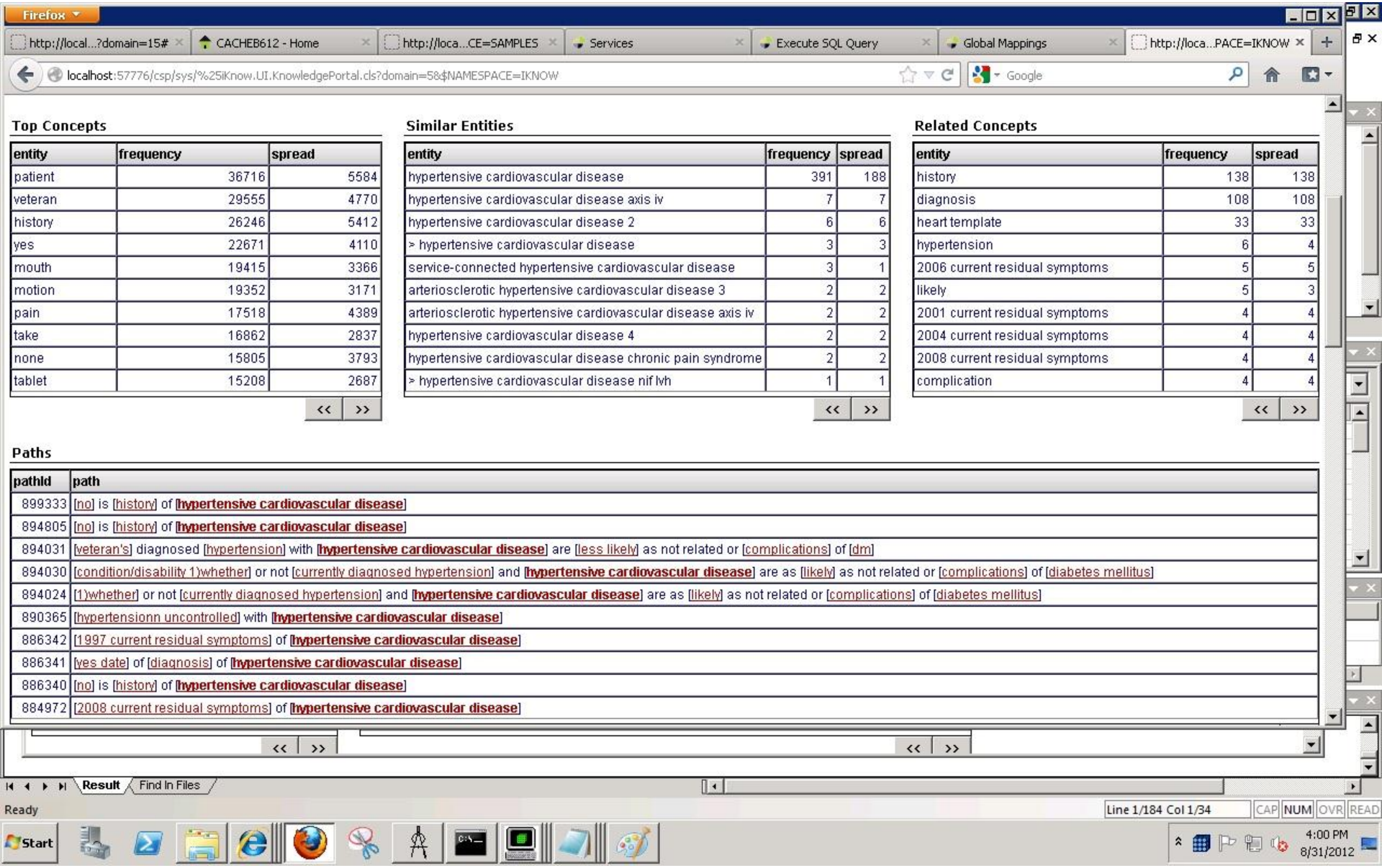
Choose Subject



Note Entities Column



Examine Source Sentence



Firefox

http://local...?domain=15# | CACHEB612 - Home | http://loca...CE=SAMPLES | Services | Execute SQL Query | Global Mappings | http://loca...PACE=IKNOW

localhost:57776/csp/sys/%25iKnow.UI.KnowledgePortal.cls?domain=5&$NAMESPACE=IKNOW

Google

Top Concepts

entity	frequency	spread
patient	36716	5584
veteran	29555	4770
history	26246	5412
yes	22671	4110
mouth	19415	3366
motion	19352	3171
pain	17518	4389
take	16862	2837
none	15805	3793
tablet	15208	2687

<< >>

Similar Entities

entity	frequency	spread
hypertensive cardiovascular disease	391	188
hypertensive cardiovascular disease axis iv	7	7
hypertensive cardiovascular disease 2	6	6
> hypertensive cardiovascular disease	3	3
service-connected hypertensive cardiovascular disease	3	1
arteriosclerotic hypertensive cardiovascular disease 3	2	2
arteriosclerotic hypertensive cardiovascular disease axis iv	2	2
hypertensive cardiovascular disease 4	2	2
hypertensive cardiovascular disease chronic pain syndrome	2	2
> hypertensive cardiovascular disease nif lvh	1	1

<< >>

Related Concepts

entity	frequency	spread
history	138	138
diagnosis	108	108
heart template	33	33
hypertension	6	4
2006 current residual symptoms	5	5
likely	5	3
2001 current residual symptoms	4	4
2004 current residual symptoms	4	4
2008 current residual symptoms	4	4
complication	4	4

<< >>

Paths

pathId	path
899333	[no] is [history] of [**hypertensive cardiovascular disease**]
894805	[no] is [history] of [**hypertensive cardiovascular disease**]
894031	[veteran's] diagnosed [hypertension] with [**hypertensive cardiovascular disease**] are [less likely] as not related or [complications] of [dm]
894030	[condition/disability 1)whether] or not [currently diagnosed hypertension] and [**hypertensive cardiovascular disease**] are as [likely] as not related or [complications] of [diabetes mellitus]
894024	[1)whether] or not [currently diagnosed hypertension] and [**hypertensive cardiovascular disease**] are as [likely] as not related or [complications] of [diabetes mellitus]
890365	[hypertensionn uncontrolled] with [**hypertensive cardiovascular disease**]
886342	[1997 current residual symptoms] of [**hypertensive cardiovascular disease**]
886341	[yes date] of [diagnosis] of [**hypertensive cardiovascular disease**]
886340	[no] is [history] of [**hypertensive cardiovascular disease**]
884972	[2008 current residual symptoms] of [**hypertensive cardiovascular disease**]

<< >> << >>

Result | Find In Files

Ready

Line 1/184 Col 1/34 CAP NUM OVR READ

Start

4:00 PM 8/31/2012

Can Drill to Original Documents

Paths

pathId	path
899333	[no] is [history] of [**hypertensive cardiovascular disease**]
894805	[no] is [history] of [**hypertensive cardiovascular disease**]
894031	[veteran's] diagnosed [hypertension] with [**hypertensive cardiovascular disease**] are [less likely] as not related or [complications] of [dm]
894030	[condition/disability 1)whether] or not [currently diagnosed hypertension] and [**hypertensive cardiovascular disease**] are as [likely] as not related or [complications] of [diabetes mellitus]
894024	[1)whether] or not [currently diagnosed hypertension] and [**hypertensive cardiovascular disease**] are as [likely] as not related or [complications] of [diabetes mellitus]
890365	[hypertensionn uncontrolled] with [**hypertensive cardiovascular disease**]
886342	[1997 current residual symptoms] of [**hypertensive cardiovascular disease**]
886341	[yes date] of [diagnosis] of [**hypertensive cardiovascular disease**]
886340	[no] is [history] of [**hypertensive cardiovascular disease**]
884972	[2008 current residual symptoms] of [**hypertensive cardiovascular disease**]

Sources

sourceId	contents
9991	:FILE:d:\tiu_data\10k\358989383.txt: ... Yes DESCRIPTION OF THE EFFECTS OF THE PROBLEM ON USUAL DAILY ACTIVITIES: Mild/moderate restriction REMARKS: The **hypertensive cardiovascular disease** is at leas not related to DM. ...
9856	:FILE:d:\tiu_data\10k\358987041.txt: ... SUMMARY OF PROBLEMS, DIAGNOSES, AND FUNCTIONAL EFFECTS ---------- DIAGNOSIS AND ETIOLOGY: **Hypertensive Cardiovascular diseas** hypertrophy, not in failure, Class II-B; ... Comments on records: VAE< March 2006: **Hypertensive Cardiovascular disease,** NIF, Class II-B, Mets 6 [Inadequate stress test] ... ********** SUMMARY OF PROBLEMS, DIAGNOSES, AND FUNCTIONAL EFFECTS ---------- DIAGNOSIS: **Hyperten Cardiovascular disease,** septal hypertrophy, not in failure, Class II-B; ... Comments on records: VAE< March 2006: **Hypertensive Cardiovascular disease,** NIF, Class II-B, Mets 6 [Ina test]

Gain Insight

- **Index** unstructured data to get to know what's in this other 85% without reading it
- **Match** the result to what you already knew to find out what you should have known
- **Interpret** the result to know what to do next

Application To Any Domain

- **iKnow ™ is not domain specific**
- **Input can be text from any domain and can feed a web service**
- **If the text size is reasonable the conversion of unstructured data to structured data can occur in real time**
- **Developers can design apps to use these features**

Conclusions

- **Healthcare is Information-rich, and the quantity and complexity of this information is increasing at exponential rates**
- **Information needs far exceed human cognitive abilities**
- **Current EHR designs are poor at matching physician patient care needs with tasks, overburdening physicians with system needs.**
- **Current emphasis should be on decreasing physician work by off-loading data entry and simplifying information access.**