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

<table>
<thead>
<tr>
<th>EHR</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazing Charts</td>
<td>4.22</td>
</tr>
<tr>
<td>Practice Fusion</td>
<td>4.04</td>
</tr>
<tr>
<td>VA-CPRS</td>
<td>3.89</td>
</tr>
<tr>
<td>Medent</td>
<td>3.65</td>
</tr>
<tr>
<td>e-MDs</td>
<td>3.56</td>
</tr>
<tr>
<td>Epic</td>
<td>3.51</td>
</tr>
<tr>
<td>athenahealth</td>
<td>3.44</td>
</tr>
<tr>
<td>eClinicalWorks</td>
<td>3.36</td>
</tr>
<tr>
<td>Sage</td>
<td>3.22</td>
</tr>
<tr>
<td>Centricity</td>
<td>3.20</td>
</tr>
<tr>
<td>Cerner</td>
<td>3.15</td>
</tr>
<tr>
<td>Greenway</td>
<td>3.04</td>
</tr>
<tr>
<td>Allscripts</td>
<td>3.00</td>
</tr>
<tr>
<td>MEDITECH</td>
<td>2.94</td>
</tr>
<tr>
<td>Practice Partner</td>
<td>2.91</td>
</tr>
<tr>
<td>McKesson</td>
<td>2.91</td>
</tr>
<tr>
<td>NextGen</td>
<td>2.81</td>
</tr>
</tbody>
</table>

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

Photo from: http://www.flickr.com/photos/chris-warren-photos/2220570496/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

“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
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"

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

*Archives Int Med Online 20Aug2012
How did this Happen?

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

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

Data Portals

Search Engines

Natural Language Understanding

Knowledge Bases

Knowledge

Semantics

Voogle

Artificial Intelligence

Machine Learning

Intelligent Assistants

Ontologies

Thesauri and Taxonomies

Blogs and Wikis

Vista

Complexity

File Servers

E-Mail

Databases

Portals

Data

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

Step 1: Crawl Data

Step 2: Create Metadata Knowledge Base

Step 3: Messaging and Visualization

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 goref/sec typical VistA system does 250,000 goref/sec)
- Storage: 60 terabytes
- Cost: $98,000
All commodity servers – no blades just rack mounts – swap them out if they break – no maintenance contract
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 searches are not always effective
Input: A Brown dog barked at the mailman.
Concept: Brown dog
Relation: Barked
Concept: Mailman

No rule creation or dictionary is necessary before indexing.
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
1. Split text into sentences.
2. Split sentences into **words**.
3. Use dictionary-based statistical models or syntactic parsing to **guess** which words might belong together.
4. Use a dictionary or ontology to **correct** and optionally enrich results.

### i.Know Approach
1. Split text into sentences.
2. Identify relational elements based on **compact** language models.
3. Relational elements **naturally** separate meaningful concept clusters.
4. Optionally use a dictionary or ontology to enrich results.
The iKnow approach to identifying concepts and relationships in unstructured text is only the starting point for more advanced analysis:

- Smart Indexing
- Smart Matching
- Smart Interpretation
Two patients are suffering from congestive heart failure.

Smart Indexing (concepts and relations):

<table>
<thead>
<tr>
<th>Concept</th>
<th>Two patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relation</td>
<td>Are suffering from</td>
</tr>
<tr>
<td>Concept</td>
<td>Congestive heart failure</td>
</tr>
</tbody>
</table>
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

Built on top of the iKnow query APIs, displaying a browsable overview of the semantic elements identified by the iKnow Smart Indexing API. Browse to the elements similar, related or containing the selected entity.

Use the filter button to filter the displayed results based on metadata criteria.

Top Concepts

<table>
<thead>
<tr>
<th>entity</th>
<th>frequency</th>
<th>spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>patient</td>
<td>36718</td>
<td>5564</td>
</tr>
<tr>
<td>veteran</td>
<td>29555</td>
<td>4770</td>
</tr>
<tr>
<td>history</td>
<td>26246</td>
<td>5412</td>
</tr>
<tr>
<td>yes</td>
<td>22571</td>
<td>4110</td>
</tr>
<tr>
<td>mouth</td>
<td>19415</td>
<td>3056</td>
</tr>
<tr>
<td>motion</td>
<td>19352</td>
<td>3171</td>
</tr>
<tr>
<td>pain</td>
<td>17518</td>
<td>4308</td>
</tr>
<tr>
<td>take</td>
<td>16862</td>
<td>2837</td>
</tr>
<tr>
<td>none</td>
<td>15805</td>
<td>3793</td>
</tr>
<tr>
<td>tablet</td>
<td>15202</td>
<td>2807</td>
</tr>
</tbody>
</table>

Similar Entities

<table>
<thead>
<tr>
<th>entity</th>
<th>frequency</th>
<th>spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Results</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Related Concepts

<table>
<thead>
<tr>
<th>entity</th>
<th>frequency</th>
<th>spread</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Results</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

i.know
Note Entities Column

This is a sample User Interface built on top of the iKnow query APIs, displaying a browsable overview of the semantic elements identified by the iKnow Smart Indexing API. Click on an entity in a list to browse to the elements similar, related or containing the selected entity.

Use the filter button to filter the displayed results based on metadata criteria.

<table>
<thead>
<tr>
<th>Top Concepts</th>
<th>Similar Entities</th>
<th>Related Concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>entity</td>
<td>entity</td>
<td>entity</td>
</tr>
<tr>
<td>frequency</td>
<td>frequency</td>
<td>frequency</td>
</tr>
<tr>
<td>spread</td>
<td>spread</td>
<td>spread</td>
</tr>
</tbody>
</table>

Example:
- Entity: patient
  - Frequency: 30716
  - Spread: 5584
- Entity: veteran
  - Frequency: 29556
  - Spread: 4770
- Entity: history
  - Frequency: 26246
  - Spread: 5412
- Entity: yes
  - Frequency: 22871
  - Spread: 4110
- Entity: mouth
  - Frequency: 10415
  - Spread: 3386
- Entity: million
  - Frequency: 10382
  - Spread: 3171
- Entity: pain
  - Frequency: 17518
  - Spread: 4388
- Entity: take
  - Frequency: 18862
  - Spread: 2827
- Entity: none
  - Frequency: 15906
  - Spread: 2792
- Entity: tablet
  - Frequency: 15208
  - Spread: 2687

Related Concepts:
- No Results
Examine Source Sentence

<table>
<thead>
<tr>
<th>Top Concepts</th>
<th>Similar Entities</th>
<th>Related Concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>entity</td>
<td>frequency</td>
<td>spread</td>
</tr>
<tr>
<td>patient</td>
<td>36716</td>
<td>5564</td>
</tr>
<tr>
<td>veteran</td>
<td>3955</td>
<td>4770</td>
</tr>
<tr>
<td>history</td>
<td>2624</td>
<td>6112</td>
</tr>
<tr>
<td>rc</td>
<td>2266</td>
<td>2110</td>
</tr>
<tr>
<td>mouth</td>
<td>19415</td>
<td>3366</td>
</tr>
<tr>
<td>motion</td>
<td>18352</td>
<td>3171</td>
</tr>
<tr>
<td>pain</td>
<td>17518</td>
<td>4389</td>
</tr>
<tr>
<td>none</td>
<td>15865</td>
<td>3793</td>
</tr>
<tr>
<td>tablet</td>
<td>15208</td>
<td>2887</td>
</tr>
<tr>
<td>patient</td>
<td>36716</td>
<td>5564</td>
</tr>
<tr>
<td>veteran</td>
<td>3955</td>
<td>4770</td>
</tr>
<tr>
<td>history</td>
<td>2624</td>
<td>6112</td>
</tr>
<tr>
<td>rc</td>
<td>2266</td>
<td>2110</td>
</tr>
<tr>
<td>mouth</td>
<td>19415</td>
<td>3366</td>
</tr>
<tr>
<td>motion</td>
<td>18352</td>
<td>3171</td>
</tr>
<tr>
<td>pain</td>
<td>17518</td>
<td>4389</td>
</tr>
<tr>
<td>none</td>
<td>15865</td>
<td>3793</td>
</tr>
<tr>
<td>tablet</td>
<td>15208</td>
<td>2887</td>
</tr>
<tr>
<td>entity</td>
<td>frequency</td>
<td>spread</td>
</tr>
<tr>
<td>hypertensive cardiovascular disease</td>
<td>391</td>
<td>150</td>
</tr>
<tr>
<td>hypertensive cardiovascular disease</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>hypertensive cardiovascular disease axis I</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>hypertensive cardiovascular disease axis II</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>hypertensive cardiovascular disease axis III</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>hypertension</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>heart template</td>
<td>33</td>
<td>33</td>
</tr>
<tr>
<td>symptom</td>
<td>108</td>
<td>108</td>
</tr>
<tr>
<td>current residual symptoms</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>likely</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>complication</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

Paths:

<table>
<thead>
<tr>
<th>pathid</th>
<th>path</th>
</tr>
</thead>
<tbody>
<tr>
<td>894338</td>
<td>ind is (history) of hypertensive cardiovascular disease</td>
</tr>
<tr>
<td>894805</td>
<td>ind is (history) of hypertensive cardiovascular disease</td>
</tr>
<tr>
<td>894031</td>
<td>hypertension with hypertensive cardiovascular disease</td>
</tr>
<tr>
<td>894030</td>
<td>hypertension and hypertensive cardiovascular disease</td>
</tr>
<tr>
<td>894031</td>
<td>hypertension and hypertensive cardiovascular disease</td>
</tr>
<tr>
<td>894031</td>
<td>hypertension and hypertensive cardiovascular disease</td>
</tr>
<tr>
<td>894031</td>
<td>hypertension and hypertensive cardiovascular disease</td>
</tr>
<tr>
<td>888342</td>
<td>2007 current residual symptoms of hypertensive cardiovascular disease</td>
</tr>
<tr>
<td>894072</td>
<td>2008 current residual symptoms of hypertensive cardiovascular disease</td>
</tr>
</tbody>
</table>
Can Drill to Original Documents

Paths

<table>
<thead>
<tr>
<th>pathid</th>
<th>path</th>
</tr>
</thead>
<tbody>
<tr>
<td>89333</td>
<td>[no is history of hypertensive cardiovascular disease]</td>
</tr>
<tr>
<td>894805</td>
<td>[no is history of hypertensive cardiovascular disease]</td>
</tr>
<tr>
<td>894031</td>
<td>[veteran's] diagnosed [hypertension] with [hypertensive cardiovascular disease] are [less likely] as not related or [complications] of [dm]</td>
</tr>
<tr>
<td>894030</td>
<td>[condition/disability ] whether or not [currently diagnosed hypertension] and [hypertensive cardiovascular disease] are as [likely] as not related or [complications] of [diabetes mellitus]</td>
</tr>
<tr>
<td>894024</td>
<td>[1] whether or not [currently diagnosed hypertension] and [hypertensive cardiovascular disease] are as [likely] as not related or [complications] of [diabetes mellitus]</td>
</tr>
<tr>
<td>890365</td>
<td>[hypertension uncontrolled] with [hypertensive cardiovascular disease]</td>
</tr>
<tr>
<td>886342</td>
<td>[1997 current residual symptoms] of [hypertensive cardiovascular disease]</td>
</tr>
<tr>
<td>886341</td>
<td>[yes date] of [diagnosis] of [hypertensive cardiovascular disease]</td>
</tr>
<tr>
<td>886340</td>
<td>[no is history of hypertensive cardiovascular disease]</td>
</tr>
<tr>
<td>884972</td>
<td>[2008 current residual symptoms] of [hypertensive cardiovascular disease]</td>
</tr>
</tbody>
</table>

Sources

<table>
<thead>
<tr>
<th>sourceld</th>
<th>contents</th>
</tr>
</thead>
<tbody>
<tr>
<td>9991</td>
<td>FILE: d:tiu_data\10k\358889383.txt</td>
</tr>
<tr>
<td>9991</td>
<td>... Yes DESCRIPTION OF THE EFFECTS OF THE PROBLEM ON USUAL DAILY ACTIVITIES: Mild/moderate restriction REMARKS: The hypertensive cardiovascular disease is at least not related to DM. ...</td>
</tr>
<tr>
<td>9666</td>
<td>FILE: d:tiu_data\10k\358887041.txt</td>
</tr>
</tbody>
</table>
| 9666 | ... SUMMARY OF PROBLEMS, DIAGNOSES, AND FUNCTIONAL EFFECTS --------------------------------------------- DIAGNOSIS AND ETIOLOGY: Hypertensive Cardiovascular disease, left ventricular 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: Hypertensive 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 [inadequate stress 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.