Announcement

Semantic Community: Build the Health Data Infrastructure in the Cloud for the Health Data Initiative Forum

Medicare Zombie Hunter

Live Data Mining Using the PearlDiver Technologies Data Engine and Semantic Community Data Visualization

Problem:

Dataset:

Problem Notes:

Bucket Language Solution:

What we found....;

Semantic Community: Build the Health Data Infrastructure in the Cloud for the Health Data Initiative Forum

Semantic Community has entered two health data challenges as follows:

• March 14, 2011, Build HealthyPeople.Gov in the Cloud: Email from Silje Lier (HHS): Your Healthy People in the Cloud project was among one of the top submissions for the challenge, and we appreciate your efforts. Again, thank you for your submission, and keep up the excellent work.

• May 17, 2010, Community Health Data Visualizations

Semantic Community has built Data.gov and Health.Data.gov in the Cloud following the expert's advice:

• "Medical statistics will be our standard of measurement: We will weigh life and see where the dead lie thicker, among the workers or among the privileged", Rudolf Virchow, in Edward R. Tufte, Beautiful Evidence (2006), Graphics Press LLC.

• Tufte Comment on iPhone interface design: Better to have users looking over material adjacent in space within our eyespan rather than stacked in time. This is especially the case for statistical data, where the fundamental analytical task is to make comparisons. Also see page 159 in the above book reference.

Semantic Community has written eight editorials for Federal Computer Week on these activities:

1. Making Individuals Into Information Architects and Preservationists
2. Data Services - What Data.gov and Many Other Things Should Be
3. Federal Cloud Computing: It can really happen if we can do our own IT!
4. Gov 2.0 Platform Data Services with Cloud Computing: OMB Earmarks Database
5. Gov 2.0 Platform Data Services with Cloud Computing: HealthDataGov

http://semanticommunity.info/Build_the_Health_Data_Infrastructure_in_the_Cloud/Announcement

Updated: Sat, 19 Sep 2015 10:41:40 GMT
Powered by mindtouch
What’s In a Name for Open Government Data Sets?: Everything!

Build Health Care Data Analytics in the Cloud: How patient and provider data can be used to promote economic growth, improved health care, and save taxpayer’s money.

The Open Government Research and Development Summit: Ed Tufte Should Have Been in the House and Sooner!

Semantic Community has done/is doing the following for this event:

1. Medicare Zombie Hunter (real-time demos)
2. Apps Inventory and Visualizations (building on learning from this event)
3. Assisted Reproductive Technology Success Rates: 2008 (use in teaching college students)
4. Show Me the Data and It’s Value: Chart of the Day (AOL Government)

in collaboration with others as follows, respectively:

1. PearlDiver HealthCare Data Mining Engine
2. Clear Government Solutions, a GSA IaaS BPA awardee
3. Center for Health Information Technology (HIT) at George Mason University (Mason)
4. AOL Government (expect launch around July 4th)

to address the judging criteria:

1. use of government data, (Medicare, Health Data Warehouse and Health.Data.gov)
2. real world impact, and (promote economic growth, improved health care, and save taxpayer’s money)
3. design of the application or solution (Tufte’s principles) (Get Spotfire for iPad App)

and to:

1. create an interoperability interface to the health data infrastructure results like Pearl Diver and from this event
2. put that in the certified and secured Government Cloud as it has done recently with GeoEye for the Geospatial Community Cloud.
3. apply advanced semantic technologies to the above content: Be Informed Video and Semantic Insights Research Assistant Video (Fact Sheet)

The term "interoperable interface" comes from the recent Report to the President and Congress "Designing a Digital Future: Federally Funded Research and Development in Networking and Information Technology", Executive Office of the President and the President’s Council of Advisors on Science and Technology, December 2010 (see excerpts in the wiki). This report defines the term as follows: the means by which components of the smart grid can talk to each other, for example, or by which electronic health records can be shared and added to by many parties - are an important stimulus to technology innovation and adoption. Optimally these interfaces would be open: anyone may create products that use the interfaces without paying fees; and a public, transparent process is used to establish, revise the standards that define the interfaces.

Medicare Zombie Hunter

NOTE: Visualization cannot be shown publically because of restriction on Medicare data.
PearlDiver and the Semantic Community have joined forces to showcase the next generation of medical data mining technology.

Definition of a Medicare Zombie: A patient who has been reported as dead to Medicare who goes on to have further treatment by Medicare providers on Medicare’s dime.

The Medicare Zombie Hunter application performs live searches through ~3/4 of a billion Medicare records to find those pesky dead patients to keep coming back for more medical treatment. Based upon the user input, the Zombie Hunter will identify (to the extent allowed by law and CMS) the undead patients, the patient’s medical provider, the amount of money charged to Medicare, and how much money Medicare reimbursed for the treatment.

The data is then visualized by Semantic Community in Spotfire, a self-serve business intelligence tool.

PearlDiver and the Semantic Community are exploring opportunities to collaborate together to provide transparent health information to those decision makers who need fast answers to very complicated questions.

PearlDiver specializes in taking in very large patient datasets and organizing the data so that sophisticated questions can be answered in a very small amount time by domain/medical experts without needing the help of a data mining or database expert. PearlDiver is a small 5 year old technical company that is just now exploring how to share its next generation data mining tools with the Federal Government. PearlDiver specializes in sophisticated patient outcome analysis, econometric modeling, as well as providing the medical research community access to the largest public database of medical charging records (we currently have ~1.7 billion records from both private and public payers).

The Semantic Community is (1) working to create an interoperability interface to health data infrastructure results like Pearl Diver produces and (2) is working with Clear Government Solutions to get that in their certified and secured Government Cloud as it has done recently with GeoEye for the Geospatial Community Cloud.
Live Data Mining Using the PearlDiver Technologies Data Engine and Semantic Community Data Visualization

Handout

For Internet Explorer Users and Those Wanting Full Screen Display Use: Web PlayerGet Spotfire for iPad App

Problem:
Find all patients who had a diagnosis of DVT (blood clots in the legs) within 6 months of having a knee arthroscopy procedure.

Dataset:
Four years of Medicare Standard Analytical Files (~3/4's of a billion records) which includes:
•100% of the inpatient records over four years
•100% of the outpatient records over four years
•5% sample of physician records over four years

Problem Notes:
There are many different knee arthroscopy codes including the following CPT codes:

There are six different ICD-9 diagnosis codes for DVT:

http://semanticommunity.info/Build_the_Health_Data_Infrastructure_in_the_Cloud/Announcement
Updated: Sat, 19 Sep 2015 10:41:40 GMT
Powered by mindtouch™
Bucket Language Solution:

```
# # First collect all knee arthroscopy procedures
# knee =
{CPT-29881,CPT-29880,CPT-29875,CPT-29883,CPT-29886,CPT-29887,CPT-29885,CPT-29868,CPT-29870,CPT-29874,
CPT-29876,CPT-29877,CPT-29879,CPT-29884,CPT-29888,CPT-29889}

# # Now collect all the DVT folks
# dvt = {ICD-9-D-45340,ICD-9-D-45341,ICD-9-D-45342,ICD-9-D-45380,ICD-9-D-45111,ICD-9-D-45119}

# # Now let's see how many we have in common
# answer = knee AND dvt AFTER within 6 MONTHS

# # Dump out demographics spreadsheets of our answer bucket
# demographics answer

What we found…:

355,877 patients had a knee arthroscopy procedure
1,067,172 patients had a diagnosis of DVT
4961 patients had a DVT within 6 months of a knee arthroscopy procedure

This amounts to 1.4% of knee arthroscopy patients

This typically takes 3-6 minutes of real time to run this analysis on a laptop machine
```