CSE512 :: 27 Feb 2014

Final Project Progress Presentation

@EEB037

CSE512 Students University of Washington
D5

Data Driven Documents Driving Development

Fahad Pervaiz & Trevor Perrier
Data Driven Development
Data Driven Development
## Immunization Numbers

<table>
<thead>
<tr>
<th>Period</th>
<th>Organisation unit</th>
<th>Fully Immunised</th>
<th>DPT3 doses given</th>
<th>Measles doses given</th>
<th>DPT1 doses given</th>
</tr>
</thead>
<tbody>
<tr>
<td>October 2010</td>
<td>Bambaa Medical Clinic</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>October 2010</td>
<td>Uzima MC (Kilifi)</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>October 2010</td>
<td>Mwanza Medical Clinic</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>October 2010</td>
<td>Menengai Health Care Serv</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>October 2010</td>
<td>Mtwaa Dispensary</td>
<td>153.0</td>
<td>127.0</td>
<td>153.0</td>
<td>117.0</td>
</tr>
<tr>
<td>October 2010</td>
<td>Vipingo Health Centre</td>
<td>75.0</td>
<td>102.0</td>
<td>75.0</td>
<td>107.0</td>
</tr>
<tr>
<td>October 2010</td>
<td>Union Medical</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>October 2010</td>
<td>Amari Family Medical Clin</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>October 2010</td>
<td>Michaela Denis MC</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>October 2010</td>
<td>Ganzo Dispensary</td>
<td>20.0</td>
<td>0.0</td>
<td>20.0</td>
<td>76.0</td>
</tr>
<tr>
<td>October 2010</td>
<td>Vutakaka Medical Clinic</td>
<td>2.0</td>
<td>8.0</td>
<td>2.0</td>
<td>9.0</td>
</tr>
<tr>
<td>October 2010</td>
<td>Chumasi Medical Clinic</td>
<td>2.0</td>
<td>8.0</td>
<td>2.0</td>
<td>9.0</td>
</tr>
</tbody>
</table>
DHIS2: District Health Information System 2

Current report visualizations are great for summary data analysis but not data exploration.

- No way to drill down into provincial level from national level.
- No interactive maps
- Complicated menu system for building reports.
DHIS2: District Health Information System 2

Current report visualizations are great for summary data analysis but not data exploration.

- No way to drill down into provincial level from national level.
- No interactive maps
- Complicated menu system for building reports.

Interactivity is key to data exploration, but this must be tied to the proper search queries to the underlying database.
Example: A3
What we plan to do.....

- Build a data exploration application to run on top of DHIS2
  - Focus on interactivity
    - Drill down from country ⇒ province ⇒ district ⇒ clinic
    - Easily add and remove data from the graph
    - Use brushing to filter graphed data.
  - Easily switch between views of data
  - Work with the DHIS2 app store so that any DHIS implementation can use our visualization tool.
Feedback?

1. What kind of data exploration are generally done by managers?
2. What data types are common for this exploration?
3. What are the most common ways to drill-down information on a visualization?
4. What drill-down operations works best on multidimensional dataset and are not complicated for our target population?
5. How to handle data density when user drills down on one dimension or branch and expands the visualization?

(Presented by Trevor Perrier and Fahad Pervaiz)
Dynamic News Timelines

November 2008 - February 2014

Michael Beswetherick & Joe Kohlmann
The New York Times Provides Aggregate News Pages for Certain Subjects (such as the POTUS)
Chronology of Coverage

FEB. 26, 2014
Pres Obama tells Afghan Pres Hamid Karzai that he has instructed the Pentagon to begin planning for a complete withdrawal of American troops from Afghanistan by the end of 2014; Obama also says the United States is still open to leaving a limited military force in the country, message aimed less at Karzai than at whoever will replace him. MORE »

FEB. 25, 2014
News analysis; Pres Obama has approached 2014 revolution in Ukraine with clinical detachment aimed at avoiding instability; approach is in stark contrast to Pres George W Bush's democracy-promoting response to Ukraine's Orange Revolution of 2004; while sympathetic to Ukraine's pro-Western protesters, Obama has not made global aspirations of democracy the animating force of his presidency. MORE »

FEB. 25, 2014
Louisiana Gov Bobby Jindal’s critical comments about Pres Obama in a news conference just after annual meeting of nation’s governors at the White House prompts immediate partisan arguing among dozen or so governors assembled just steps from the Oval Office. MORE »

FEB. 24, 2014
More than 500 leaders of a national network of young immigrants, frustrated that House Republicans say they will not move on immigration in 2014, have decided to turn their protests on Pres Obama in an effort to pressure him to act unilaterally to stop deportations. MORE »

FEB. 24, 2014
Bill Keller Op-Ed column examines how Pres Obama has lived up to the high expectations that he would reform the broken criminal justice system; contends that so far, the Obama administration has been unimpressive on the issue; expresses hope that the administration will do more in the remaining years of Obama's presidency. MORE »
Highlights From the Archives

NEWS ANALYSIS
Obama Wins a Clear Victory, but Balance of Power Is Unchanged in Washington
By PETER BAKER
After $6 billion, two dozen presidential primary days, four general election debates and more TV ads than anyone could watch, the two parties essentially fought to a standstill.
November 8, 2012 | US | NEWS ANALYSIS

MAN IN THE NEWS | BARACK HUSSEIN OBAMA
4 Years Later, Scarred but Still Confident
By PETER BAKER
President Obama is making the case that while progress is slow, he is taking America to a better place — and that he will be a better president over the next four years than the last.
September 6, 2012 | US | NEWS

Obama Elected President as Racial Barrier Falls
By ADAM NAGOURNEY
Barack Hussein Obama was elected the 44th president of the United States, as the country chose him as its first black chief executive.
November 6, 2008 | US | NEWS

ARTICLES ABOUT BARACK OBAMA

Newest First | Oldest First
Page: 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | Next >>

Shrugging Off Past Setbacks, Obama Plans Personal Role in Middle East Peace Bid
By MARK LANDLER
President Obama will press Prime Minister Benjamin Netanyahu of Israel to agree to a “framework” for a conclusive round of peace talks, officials
Articles

A (Paginated) Firehose, Starting with Most Recent Articles (Again)
Currently

The NYT Provides Aggregate News Pages for Certain Subjects (such as the POTUS)
A Dynamic Timeline for News Articles
With Interactive Exploration and
Linking of Related Subjects’ Timelines
A timeline with start/end point labels:

- Nov 2008
- Feb 2014

Bounded by starting node or date:

- Approximately concurrent events

Rects are a bit big here:

- Though: they might be top-sized now?

Again, what does position mean here? Might need to do API tests to see what data can be encoded vertically.
Multiple stories aggregated chronologically.

Aggregate title Jan-May 2010

Aggregate titles appear on hover or something?

Grouping threshold potentially user-defined (we'll see).
Transition to a new, linked timeline.

Healthcare Reform

Continue Exploring

Visual Link!

Click to Go Back

Starting Timeline Condenses

Nov 2008

Jan-May 2010

Feb 2014
The average browser window

Wasted area

Can we wrap the timeline?

Obvious inspiration from FCPX
Details on Demand

Efficient Use of Area

Lenses

Linked Displays

Narrative Visualization

Summarization
Interactions and Sample Usage

Figure 3: AP data from July - August 1990. A wide current in the river indicates heavy use of a topic, while changes in color distribution correlate to changes in themes. Figure 6: Parallel rivers let users compare AP data from Washington, D.C. and New York from the same time period.
Narratives Fisher et al, 2008
Front Row to Fashion Week

By MIKE BOSTOCK, SHAN CARTER, ERIK HINTON, CATHY HORYN and ERIC WILSON | September 12, 2013

Of the more than 300 collections shown during New York Fashion Week, here were the ones that created the most buzz and left the biggest impressions on fashion editors as they headed off to the next round of shows in London, Milan and Paris.

Calvin Klein

A beautiful, innovative collection in which Francisco Costa layered references to urban tribes, '80s art, handcraft and even, seemingly, radical chicks of the 1920s. It added up to a modern expression of fashion.

Read more: Calvin Klein in Full Color

- Sand-colored, orange-accented canvas wrapped into a dress and suit
- A large emerald tweed coat with frayed, pronounced seams
- A boxy black jacket fringed with multicolored confetti strings
Project Plan

3/04  Prepare Initial Data Set
     Working demonstration of project

3/06  Finish View Experiments
     Figure out if we can use the live NYT APIs (or not)

3/11  Finalize Overall Interaction Design
     Have a full plan for user interactions and events

3/13  Final Project Poster Presentation
     Working demonstration of project

3/20  Final Project Paper and Software Due
     Finish the paper, write documentation, and get the final code on GitHub
Questions & Feedback

Michael Beswetherick & Joe Kohlmann
Visualization on Cosmological Halo Merger Trees

Laurel Orr and Jennifer Ortiz
CSE 512
Halo Merger Trees

**Halo:** dense clumps of particles that interact and merge

**Merger Tree:** shows the history of joining (merging) and splitting of halos over time

Scientists want to compare and analyze structures of different trees

![Diagram of Halo Merger Trees]

**PRESENT**

**PAST**
Previous Work

Our Task

320,589 halos with 286 present day roots across 27 timesteps
Create interactive visualization for astronomers to view and analyze halos

- Zooming
- Comparison between merger trees
- Path tracing of important halos
- Filtering
- Node/Path Annotations
Comparing Two Halos

Halo 87

Halo 25

Present Day

Subgraphs with same structure
Collapsible Tree
Force-Directed Graph

Using D3’s force directed layout with constraints to make it look like a tree

child has two parents so technically not a tree

present day halo
Help!

- How do we show a global view of a tree while allowing for zooming? How can we summarize the graph?
- Should the users be allowed to move the nodes?
- How can we allow for tree comparison?

(Presented by Laurel Orr, Jennifer Ortiz)
Voronoi Treemaps in D3

Paul Vines
Peter Henry
Treemaps

- **Squarified Treemaps**. Bruls et al., Data Visualization 2000.
- Already available in D3
- Real-time, interactive
Voronoï Diagrams

- Partition plane based on closest point
- Available in D3
Voronoi Treemaps

• *Voronoi Treemaps*. Balzer and Deussen, InfoVis 2005.
Voronoi Treemap Advantages

Original treemap:
Unbalanced aspect ratios

Squarified treemap:
Obscure heirarchy
Voronoi Treemaps Are Slow (Even on an 8 core Xeon 2.4GHz)

~7 minutes:

~6 minutes:

Figure 10: Enhanced AW Voronoi Treemap layout of 4075 nodes at 10 hierarchy levels (a brighter color indicates a lower hierarchy level)

Figure 11: Enhanced PW Voronoi Treemap layout of 16288 nodes at 7 hierarchy levels (a brighter color indicates a lower hierarchy level)
Tasks

• Literature for faster options (done)
  • *Fast Dynamic Voronoi Treemaps*. Sud et al., 2010.
    • GPU-based, ~50x speedup
    • CPU-based, core algorithm changes, beats above GPU solution

• Implement original core algorithm in D3

• Investigate Optimizations
  • Most promising is Nocaj and Brandes because it is CPU-based.
Questions

• How can we parallelize our tasks?
  • This project is somewhat serial in nature
• How can we parallelize our algorithm (LOL)?
  • Is WebGL or WebCL a good idea?
• Good ideas for optimization?
  • Plot only the first $n$ levels of the dataset
  • Plot first few levels as a standard treemap, then the remainder as a Voronoi Treemap
• What are some good available sample datasets?
• What interactivity can/should we support?

(Presented by Paul Vines & Peter Henry)
Direct-manipulation
font selection
Matthew Kay
Direct-manipulation font selection

Matthew Kay
nature and technology tell us that reality grows out of function (purpose), and out of how technic or technical), and out of how marvellous forms of nature and technology originated.

Elementar — a parametric type family
How about parametric, direct manipulation selection of existing fonts?
Stone Sans
Humanist sans-serif
Stone Sans
Humanist sans-serif
Angie Sans
Humanist semi-serif
Vollkorn
Serif
Apolline
Old-style serif
Sabon
Old-style serif
Garamond
Old-style serif
Garamond
Old-style serif
Thanks! Discussion—

Preventing jitter — optimal path through this space?

Handling multiple constraints?

Feature selection?

Matthew Kay / mjskay
Narrative Visualization

A story is worth a thousand pictures.

Christine(Yenting) Liu, Nina(Zhuxiaona) Wei
Question

Which is the more effective way to present statistical data graphics, **synchronously** or **asynchronously**?
Question

Synchronous

Asynchronous

Synch-
Cognitive Load
Attention Limitation
Top-Down processing

Asynch-
Bottom-up processing

Output/klbs

Month

April
May
June
July

Month

April
May
June
July

Month

April
May
June
July
The role of Storytelling in visualization, and the importance of narrative visualization.
Gershon & Page, 2001; Wojtkowski & Wojtkowski, 2002; Ma et.al, 2012; Kosara & Mackinlay, 2013;

Design category, design space analysis, case study
Segel & Heer, 2010;

Visualization rhetoric, framing effects
Hullman & Diakopoulos, 2011

Animation transitions, Sequence of narrative visualization
Heer & Robertson, 2007; Hullman et al, 2013
Progress

**Done:** Literature Review

**Doing:** Experiment Design, Pilot Test

**To do:** Experiment, Data Analysis, Paper Write-up
Within-Subjects Design

**Independent Variables:**
1. Synchronous, Asynchronous

**Dependent Variables:**
1. Accuracy (describe the level of interpretation), objective
2. Likert Attitude Scale (based on the Design Criteria), subjective

**Experiment Stimulus:**
Basic Charts (bar) or Typical Narrative Visualization
Synchronous

Asynchronous

500ms
Synchronous

Asynchronous

500ms
Synchronous

Asynchronous

500ms
Synchronous

Asynchronous

500ms
Synchronous

500ms

Asynchronous
Synchronous

Asynchronous

500ms
Synchronous

Asynchronous

800ms
Appreciate any feedback

Research question specific enough, meaningful?

Suggestions for experiment design?
- Within-Subjects or Between-Subjects design
- Should study different types of chart? Charts with different complexity?
- Independent or Dependent variables?

Suggestions for experiment materials?
- Professional statistical data charts/figures from papers, unfamiliar domain Biology, Chemistry etc.
- Online narrative visualization, Human Development Trend etc.

Others?
Hope you enjoy our presentation.

Special thanks to Jeff, Steven(MSR), Ham
Visualization of Lattice Structure

Shengjie Wang
A Simple Lattice

sea

hawk

title
Seattle

sea

httle

hawk
Common Edge: “sea”
Visual Encoding: Graph Cluster

https://wisonets.wordpress.com/2010/11/22/
Visual Encoding: Common Edge

- sea
- hawk
- ttle
Visual Encoding: Compelling Edges

sea

ttle

hawk
Interaction: Drill Down Cluster

Cluster Simple

sea

hawk

ttle
Interaction: Query Edge

- sea
- hawk
- whatever
- title
Interaction: Query Edge

sea → ttle

hawk → whatever
Interaction: Add/Remove C.E.
Related Works

Feedback

• Should we cluster the graph based on nodes? Can we effectively show the information of conveyed in each path if we apply the clustering?

• My name: Shengjie Wang
Decoding the Text Encoding

Hamid Izadinia and Fereshteh Sadeghi

Final project progress presentation
Data Visualization
(CSE512)
From Text Visualization to Data

Input

Output

(People, 100)
(new, 83)
(jobs, 78)
(years, 77)
(make, 69)
(care, 8)
(child, 7)
Related Works

ReVision: automated classification, analysis and redesign of chart images.
Savva, M and Kong, N and Chhajta, A and , Fei-Fei L. and Agrawala, M and Heer, J.
Letter extraction & OCR
Word extraction
Word extraction & Size estimation
Challenges

• Dense arrangement of letters and words
• Different Fonts
• No color clue
Feedback

• Is there any data set available for this task?
• What kind of evaluation is appropriate?
• Any idea for redesigning the text visualization?
• Do you know of any scaling mechanism in standard text visualization softwares? (linear, log, ranking, …)

Hamid Izadinia, Fereshteh Sadeghi
Visualize NLP Annotations for Crowdsourcing

Hanchuan Li, Haichen Shen, Shengliang Xu and Congle Zhang
Statistical Nature Language Processing

• NLP: interaction between computer and human languages

• Statistical approaches have made great success
But ...

• Need labeled training data, which is very expensive
Even worse...

• Structured prediction, expert only

“I voted for Nader because he was most aligned with my values,” she said.
Even worse...

- Structured prediction (generate trees/graphs), expert only

"I voted for Nader because he was most aligned with my values," she said.
Let Crowds do the job

• Visualize the NLP annotation

+ visualized toolkit

“I voted for Nader because he was most aligned with my values,” she said.
Goal

• Visualize two categories of annotations

• Undo/redo

Tree construction

Clustering

“I voted for Nader because he was most aligned with my values,” she said.
“I vote for Nadar because he most aligned with my values”, she said.
Question

• Flexibility: easily transferred to other tasks
  • Take HTML as input, with target objects tagged.
  • Let NLP experts to design the input visualization

• Interface

Presented by
Congle Zhang, Shengliang Xu, Haichen Shen, Hanchuan Li
Using Visualization to Optimize Text Classification

Jeff Snyder, Zorah Fung, Brian Walker, Zening Qu
Problem Description

- Text classification tasks are important in a wide variety of domains
- Create a model that assigns documents (tweets, papers) into classes ([important/not important], [positive/neutral/negative])
- We use visualization to
  - Help analysts to understand and tune the inner workings of their models,
  - Compare the performance of different models, at both a high level and on the level of individual documents, and
  - Understand the essential contours of their data that affect model performance.
Related Work


# Prototype and Storyboard

## Tweet Classification

<table>
<thead>
<tr>
<th>Actual Label</th>
<th>Predicted Label</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pos</td>
<td>Pos</td>
<td>546</td>
</tr>
<tr>
<td></td>
<td>Neu</td>
<td>213</td>
</tr>
<tr>
<td></td>
<td>Neg</td>
<td>401</td>
</tr>
<tr>
<td>Neu</td>
<td>Pos</td>
<td>235</td>
</tr>
<tr>
<td></td>
<td>Neu</td>
<td>497</td>
</tr>
<tr>
<td></td>
<td>Neg</td>
<td>102</td>
</tr>
<tr>
<td>Neg</td>
<td>Pos</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>Neu</td>
<td>195</td>
</tr>
<tr>
<td></td>
<td>Neg</td>
<td>783</td>
</tr>
</tbody>
</table>
Prototype and Storyboard

Tweet Classification

| Actual Label | Predicted Label | |
|--------------|-----------------|
| Pos          | Pos             | 546 |
|              | Neu             | 213 |
|              | Neg             | 401 |
| Neu          | 235             |     |
|              | 497             |     |
|              | 102             |     |
| Neg          | 89              |     |
|              | 195             |     |
|              | 783             |     |

Contributed against classification: [Color Scale]
Contributed to classification: [Color Scale]

Regularization: [Icons]

Model for predicted label: [Icon]
Model for actual label: [Icon]
Demo

http://cse512-14w.github.io/fp-jasnyder-zorahf-bdwalker-zqu/web

Jeff Snyder, Zorah Fung, Brian Walker, Zening Qu
CSE512 Final Project

Yang Chaoyu, Aniket Handa, Gregory Nelson, Alexander Conrad Nied
Motivation

Networks +

Time Varying Edge Data

Dangerous amounts of information, may not be meaningful.
Workspace

Using D3
Traffic flow is very popular
Not integrated to D3
Time-varying visualizations rare

http://bl.ocks.org/mbostock/4062045
http://boundary.com/blog/2012/05/14/visualizing-network-flow-data/
Methods

“PivotGraph” & roll-up

Featural Edges or Animated Packets

Sensors - Select a node
Get history/plots and traces
Storyboard
Storyboard

Brain + Activity + Focus

- Memory
- Speaking
- Hearing
- Context
- Semantics
- Feeling
- Moving
- Seeing
Questions

Did we miss any existing work?
Do you have a juicy dataset for this?
Do you have questions?

Yang Chaoyu, Aniket Handa, Gregory Nelson, Alexander Conrad Nied
Seattle Band Map
CSE 512 Final Project
Chase Wu | Susanne Hsu
Seattle Band Map

The current visualization is difficult to navigate and is not organized in a way that supports exploration or music discovery.
Relevant Work

Existing visualizations force users to define a starting point, rather than using the network structure to reveal areas of interest and guide discovery.
Challenges

How to best support exploration and discovery: what’s the most appropriate clustering and layout algorithm?

What information would you want to see/hear on the map and what would you want to do with it?

What information should be available in the global view vs. zoomed in view?

Susanne Hsu, Chase Wu