Visualization & Visual Analytics 1

Angus Forbes

creativecoding.evl.uic.edu/courses/cs424

Aspects of the Project

Domain situation

Data/task abstraction

Visual encoding/interaction idiom

Algorithm

Text Visualization

Text Visualization Browser:
http://textvis.lnu.se/

Overview of *330* visualization techniques related to text, from publications released between 1976 and early 2016.

Text Visualization Data

Textual data?

novels, poems, newspapers, academic journals, conference proceedings, twitter, facebook, google searches, wikipedia, patents, chat room conversations, emails, coding comments, issue tracking, git repos, medical records / electronic health records, transcriptions, conversations between friends, official conversation (e.g., in courtroom), ... Others?

Text Visualization Questions

Interesting questions?

- what do people talk about in general?

- what do people talk about in different contexts? (e.g. presidential candidates in debates, fans at a sports game, friends at a party)

- how to the topics or events that people talk about change over time or in response to events?

Text Visualization Questions

Interesting questions?

- what do people write about?

- how can you compare how/what person or group X talks about to group Y?

- how can you extract facts from text?

- how can you extract metaphors from text?

- how can you extract opinion from text?

- how can you extract sentiment from text?

Text Visualization Questions

Interesting questions?

- how can you identify and describe relationships between characters or elements in text?

- what rhetorical strategies are most effective?

- how similar are two texts?

- where do the ideas in a text come from originally?

Text Visualization Questions

Interesting questions?

- where do ideas in a text come from originally?

- how do different types of textual communication differ?

- what is structure of a text? of conversation?

- which authors write texts together, or cite each other?

- which authors are the most influential?

Text Visualization Tasks

Understand – Summarizing, revealing, unique aspects and interesting features within a text

Group - Finding clusters of topics within a single text or across collections of documents

Compare - Exploring differences between different texts, or between a single collects at different times

Correlate - Finding patterns in a text that relate to other data, such as social networks or personal / world events

Not necessarily a comprehensive list of tasks!

Example

Domain: Health informatics

Data: Nurse reports, written and transcribed, historical and current; Patient's vital signs; Academic articles

Algorithm: Data mine patterns in historical records with known outcomes to active nurse reports for an individual patient

Visualization Tasks: Visualize similarity of patient with patients who had bad outcome (coma, death, etc); Visualize interpretations of the different nurses taking care of a patient

Text Visualization Issues

- Visualize the text itself or visualize facts about the text? (Word frequencies, word sentiment, topic clusters, metadata about speakers/writers, etc…)

- Do you *trust* that the extraction / interpretation of the text accurately represents all salient information for your task *without distortion*?

- Not all information is in the text – We bring knowledge/bias/expectations when reading
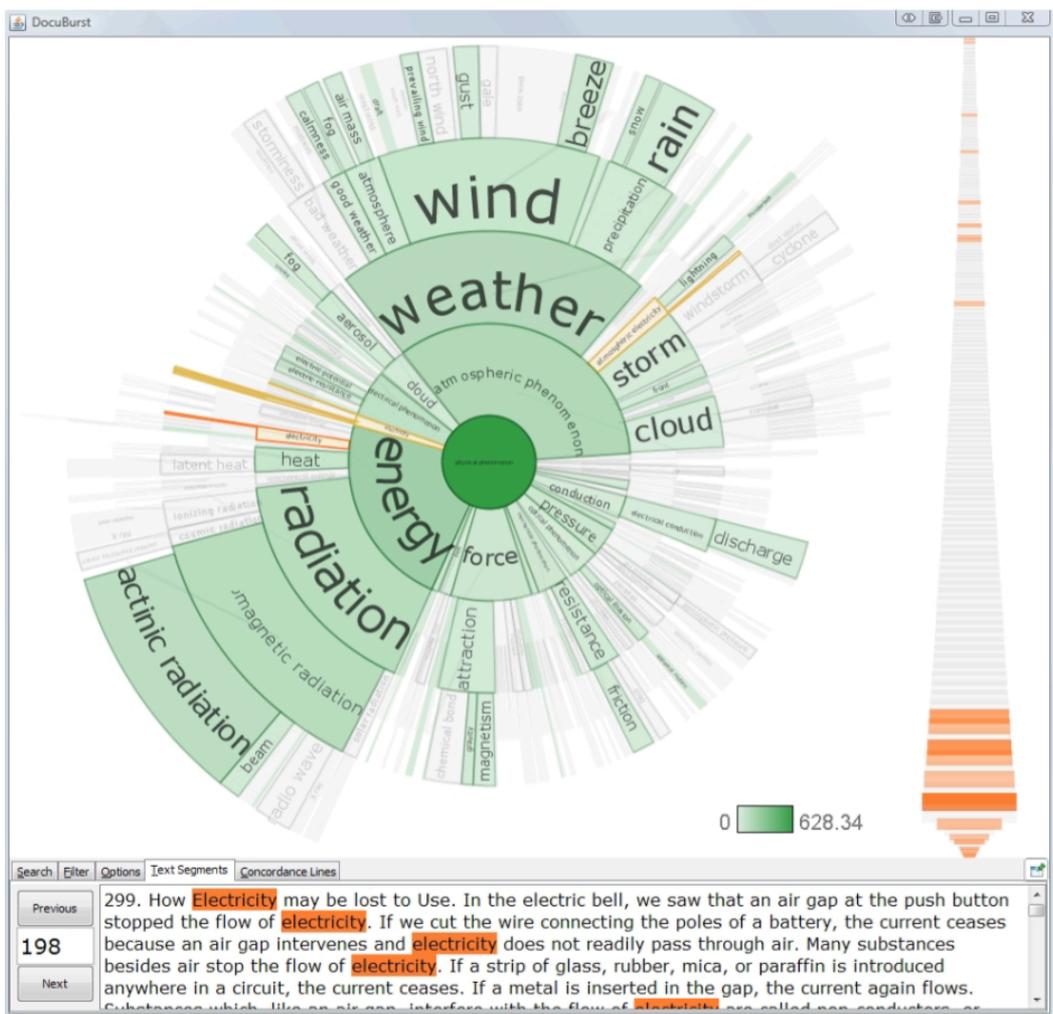
Example Text Visualization Tasks

- Show structure / repeated elements in a text

- Show evolution of popularity of names / terms in a text db

- Show patterns of discourse to help with computational linguistics tasks

- Show evolution of topics in academic papers or social media

- Extract and show quantitative sentiment and opinion measures from user comments

- Show patterns of sounds in poems or other text

Example papers

- DocuBurst (information visualization paper)
- Westeros (visual analytics paper)
- Morphable Word Cloud (algorithm paper)

Morphable Word Clouds, Chi et al.

DocuBurst, Collins et al.

Westeros Sentinel, Scharl et al.

Visual Analytics vs InfoVis

Information Visualization articles tend to present a single technique that could be applied to many different datasets in different domains

Visual Analytics articles tend to describe ways to use multiple techniques in order to make sense of a single domain

In-class group project

- Does it seem more like a visual analytics paper or an information visualization paper? (focus on a single technique that could apply to many domain, or on a single domain integrating many techniques?)

- What data does it discuss?

- What tasks is it trying to enable?

- What visual encodings & interaction idioms?

- How is/are the technique/s evaluated?

- How many people have cited this paper?

For Next Tuesday

- Each person in your group will take the lead on reading a paper related your domain, task, and/or visual coding.

- Be ready to answer questions about the paper and how it relates (or doesn't relate) to your proposed project.