School of Information Systems

SMU SINGAPORE MANAGEMENT UNIVERSITY

Lesson 9
Text Visualisation: Methods and Applications

Mentor: Dr. Kam Tin Seong
Associate Professor of Information Systems (Practice)
School of Information Systems, Singapore Management University

Content

- Introduction to text data and text visualisation
- Text Visualisation techniques and tools:
  - Tag Cloud
  - Wordle
  - Word Tree
  - Phrase Nets
- QA

School of Information Systems

SMU SINGAPORE MANAGEMENT UNIVERSITY

Motivation

- Textual data are readily available from various media.

By Channel NewsAsia, Updated: 08/10/2011

Market for new flats likely to stabilise in four years: PM Lee

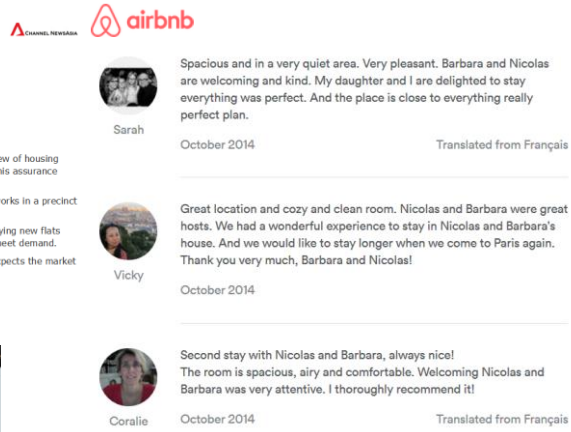
Market for new flats likely to stabilise in four years: PM Lee

SINGAPORE: Prime Minister Lee Hsien Loong said he expects the recent slew of housing measures to stabilise the market for new flats in four years. He also gave his assurance that the prices of new flats will remain affordable.

Mr Lee was speaking at a ceremony to mark the completion of upgrading works in a precinct in Ang Mo Kio.

In his speech, Mr Lee noted the concerns many Singaporeans have over buying new flats and he outlined several ways in which housing authorities are working to meet demand.

Some of these measures have started to take effect, and Mr Lee said he expects the market to stabilise in about four years' time.


Spacious and in a very quiet area. Very pleasant. Barbara and Nicolas are welcoming and kind. My daughter and I are delighted to stay everything was perfect. And the place is close to everything really perfect plan.

Sarah — October 2014 — Translated from Français

Great location and cozy and clean room. Nicolas and Barbara were great hosts. We had a wonderful experience to stay in Nicolas and Barbara's house. And we would like to stay longer when we come to Paris again. Thank you very much, Barbara and Nicolas!

Vicky — October 2014

Second stay with Nicolas and Barbara, always nice!
The room is spacious, airy and comfortable. Welcoming Nicolas and Barbara was very attentive. I thoroughly recommend it!

Coralie — October 2014 — Translated from Français

Why Visualise Text?

- Understanding – get the "gist" of a document
- Grouping – cluster for overview or classification
- Compare – compare document collections, or
- inspect evolution of collection over time
- Correlate – compare patterns in text to those in
- other data, e.g., correlate with social network

Levels of Text Representation

- Lexical level, transforming a string of characters into a sequence of atomic entities, called tokens.
- Syntactic level, identifying and tagging (anotating) each token's functions.
- Semantic level, extracting of meaning and relationships between pieces of knowledge derived from the structures identified in the syntactical level.

School of Information Systems

SMU SINGAPORE MANAGEMENT UNIVERSITY

Tag Cloud

- A tag cloud (word cloud, or weighted list in visual design) is a visual representation for text data, typically used to depict keyword metadata (tags) on websites, or to visualize free form text.
- 'Tags' are usually single words, normally listed alphabetically, and the importance of each tag is shown with font size or color.

Source: http://en.wikipedia.org/wiki/Tag_cloud

School of Information Systems

SMU SINGAPORE MANAGEMENT UNIVERSITY

Application of Tag Cloud I: Branding

- One-word tag cloud of DBS's corporate values statement created using Many Eyes.

School of Information Systems

SMU Singapore Management University

Application of Tag Cloud I: Branding

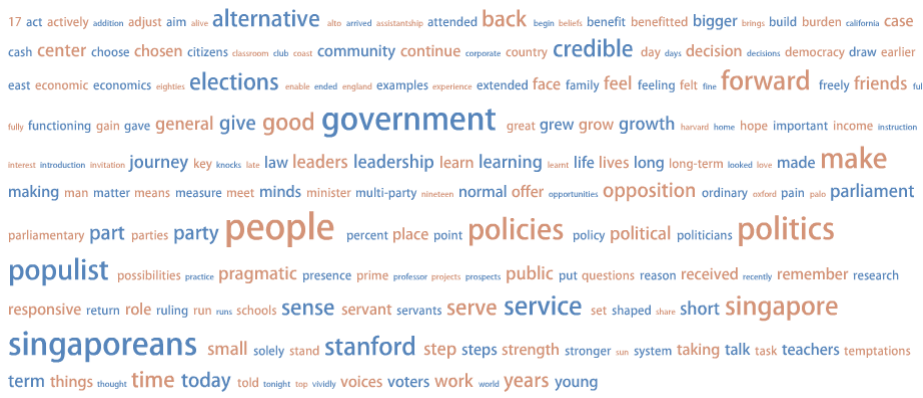
- Two-word tag cloud of DBS's corporate values statement created using Many Eyes.

School of Information Systems

SMU Singapore Management University

Applications of Tag Cloud II: Speeches

- Journey Back to Singapore by Chen Show Mao



School of Information Systems

SMU Singapore Management University

Applications of Tag Cloud III

- A blog tool or website analysis for search engine optimization

School of Information Systems

SMU Singapore Management University

Applications of Tag Cloud III

- Compress Yourself – Tag Cloud Your Resume

School of Information Systems

SMU Singapore Management University

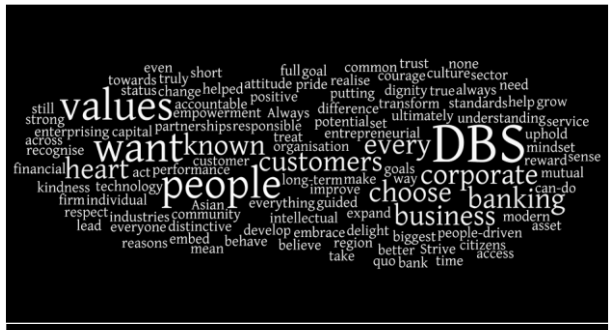
Wordle (http://www.wordle.net/)

- A toy for generating "word clouds" from text that you provide.

School of Information Systems

SMU Singapore Management University

Word Clouds of Corporate Values Statements



School of Information Systems

SMU SINGAPORE MANAGEMENT UNIVERSITY

Shaped Word Cloud

- Twitter Word Map for Android

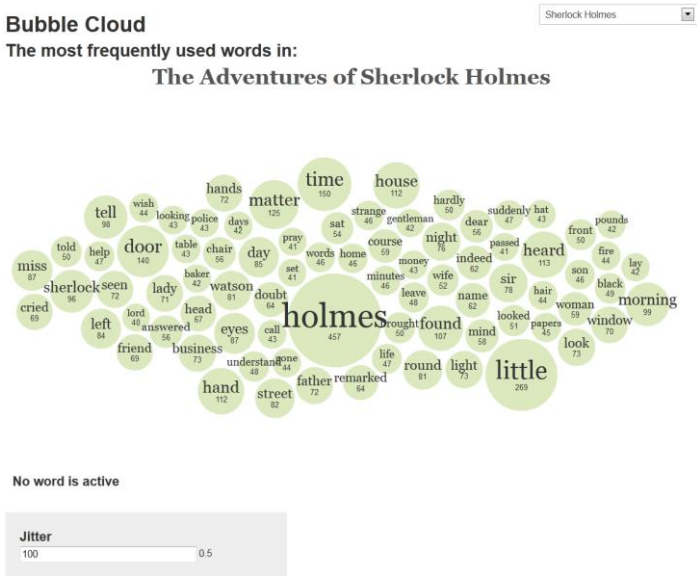
School of Information Systems

SMU SINGAPORE MANAGEMENT UNIVERSITY

Word Cloud in d3.js

cloud
RT
like
#cloud
Lol
#cloudcomputing
everyone
answer
live
morning
Computing
Cheltenham
sound
red
Extended
Judgment
falling
silver
ABS
woke
dark
fly
crazy
pretty
video
produces
Look
Atlas
nigga
think
great
gain
still
every
come
rain
sometimes
sad
something
literary
Master
outside
fuck
retweets
cover
theres
love
way
England
Ladies
First
murah
guy
Hope
house
just
Back
follow
get
little
audience
bitch
RETWEET
music
versa
Meanwhile
link
oh
vs
let
gonna
days
life
today
book
biar
kids
#TeamFollowBack
#commuting
festival
grey
talk
looks
omg
PIS
soon
good
never
even
one
real
HD
Cheltenham
morning
live
answer
pretty
crazy
woke
dark
fly
silver
ABS
falling
Extended
Judgment
everyone
never
good
talk
soon
video
produces
Sky
book
biar
kids
#TeamFollowBack
something
sad
Sometimes
rain
come
every
still
great
think
Atlas
grey
festival
#commuting
lining
literary
high
Master
fuck
outside
theres
retweets
cover
Ladies
First
murah
guy
England
way
love
go
house
just
Hope
9

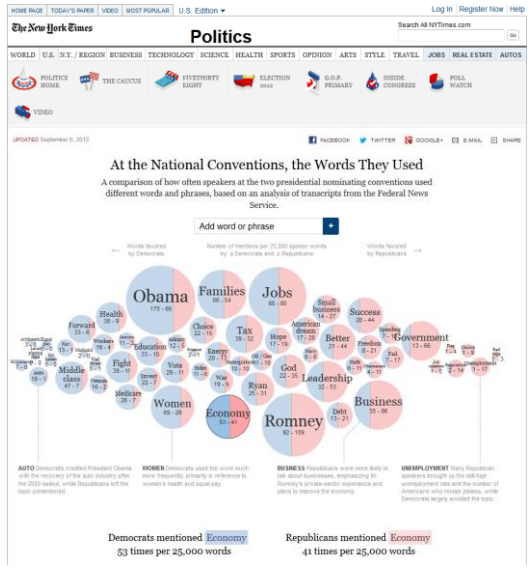
Bubble Cloud in d3.js



School of Information Systems

SMU SINGAPORE MANAGEMENT UNIVERSITY

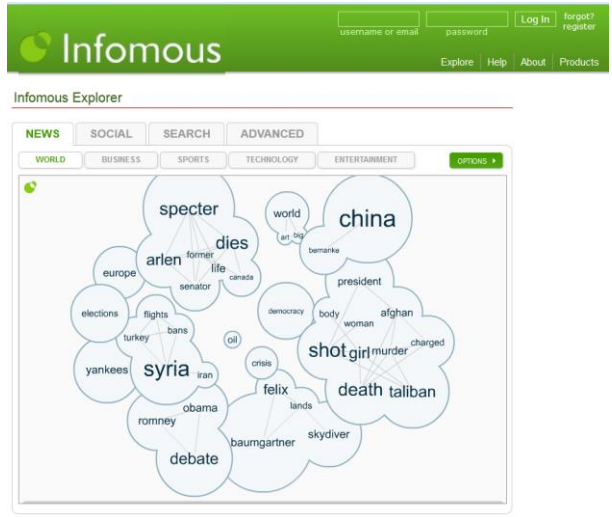
Bubble Cloud in Action



School of Information Systems

SMU SINGAPORE MANAGEMENT UNIVERSITY

Infomous (http://www.infomous.com/about)



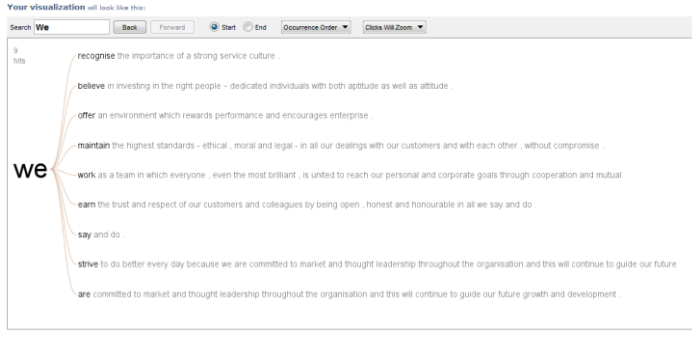
http://socialtimes.com/create-an-interactive-text-visualization-with-infomous_b95033

School of Information Systems

SMU Singapore Management University

Word Tree

- A visual search tool for unstructured text, such as a book, article, speech or poem. It lets you pick a word or phrase and shows you all the different contexts in which the word or phrase appears.
- The contexts are arranged in a tree-like branching structure to reveal recurrent themes and phrases.



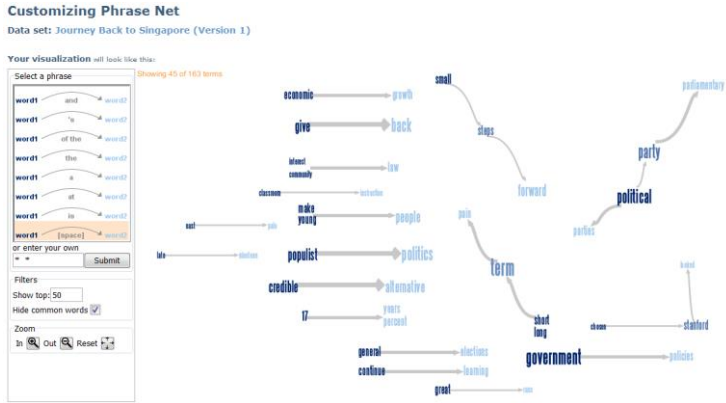
Link: https://www.jasondavies.com/wordtree/

School of Information Systems

SMU Singapore Management University

Phrase Net

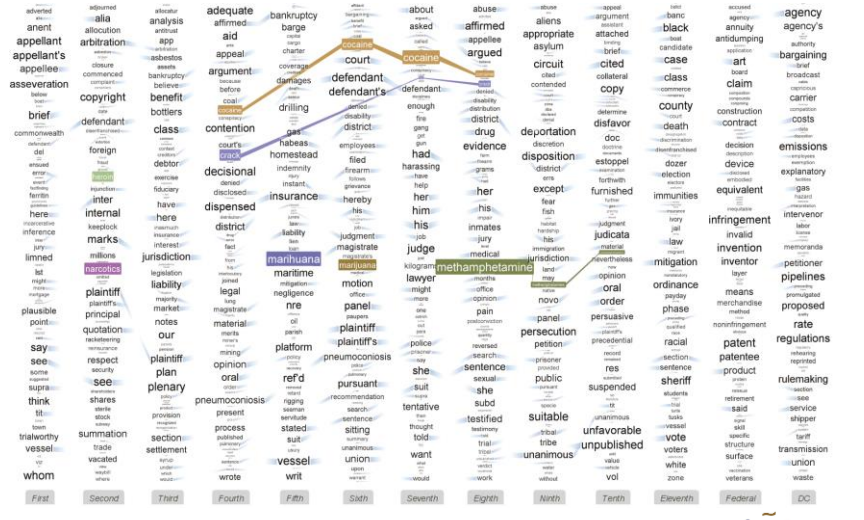
- A phrase net diagrams the relationships between different words used in a text. It uses a simple form of pattern matching to provide multiple views of the concepts contained in a book, speech, or poem.



School of Information Systems

SMU Singapore Management University

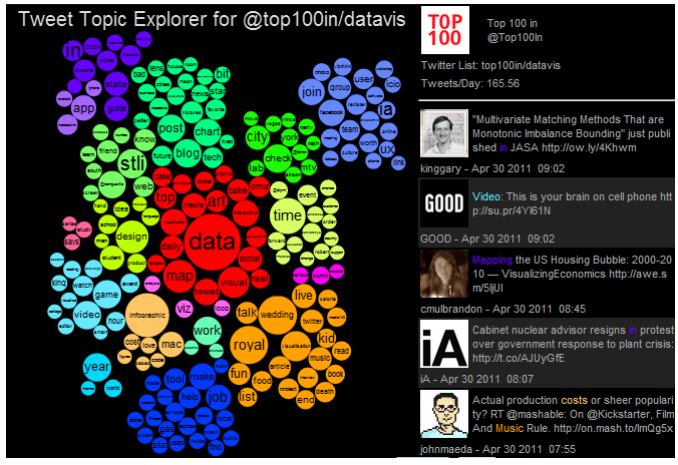
Parallel Tag Cloud



School of Information Systems

SMU Singapore Management University

Tweet Topic Explorer

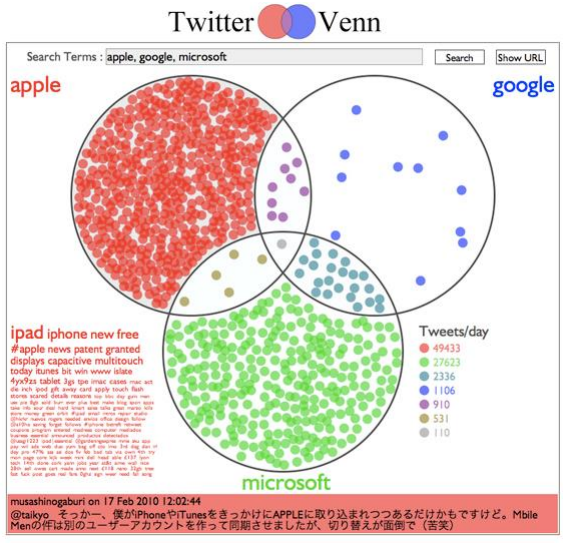


Source: http://www.neoformix.com/2011/ExploreTwitterLists.html

School of Information Systems

SMU Singapore Management University

Twitter Venn

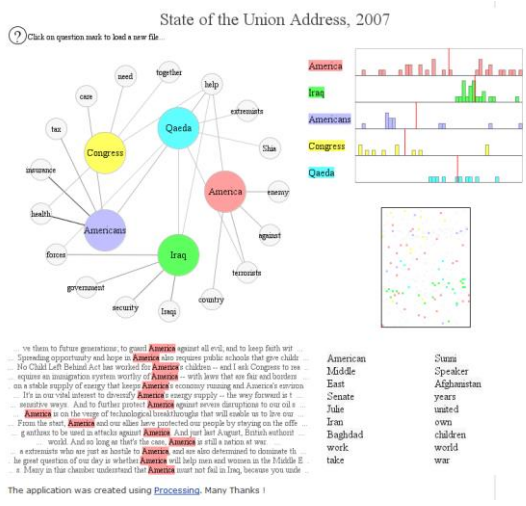


Source: http://www.neoformix.com/Projects/TwitterVenn/view.php

School of Information Systems

SMU Singapore Management University

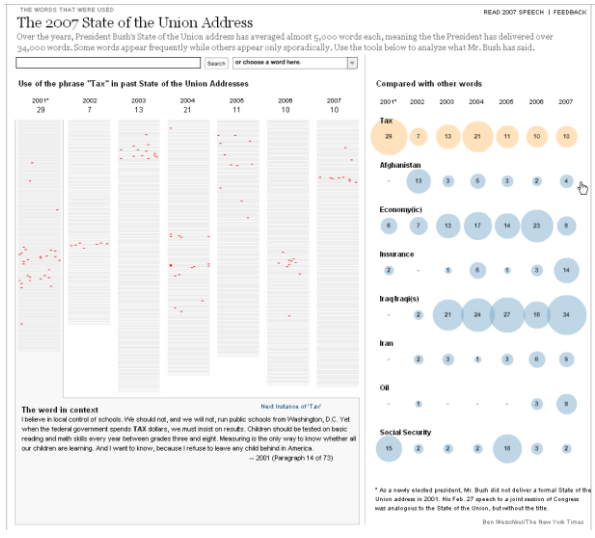
Text Explorer



School of Information Systems

SMU SINGAPORE MANAGEMENT UNIVERSITY

Text Visualisation in Data Journalism



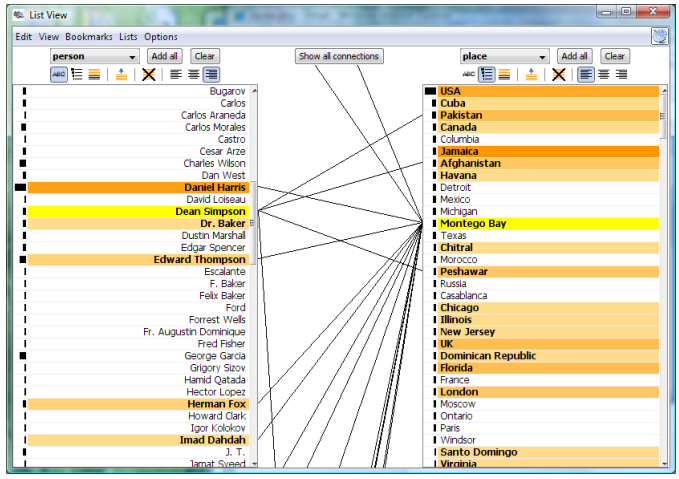
Source: http://www.nytimes.com/ref/washington/20070123_STATEOFUNION.html

School of Information Systems

SMU SINGAPORE MANAGEMENT UNIVERSITY

A Text Visualisation and Analysis System

- JIGSAW (http://www.cc.gatech.edu/gvu/ii/jigsaw/) for exploring and understanding document collection.

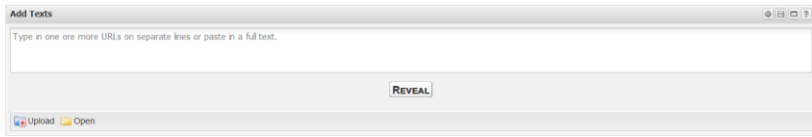


School of Information Systems

SMU Singapore Management University

Web-based Text Visualisation and Analysis Systems

VOYANT
see through your text



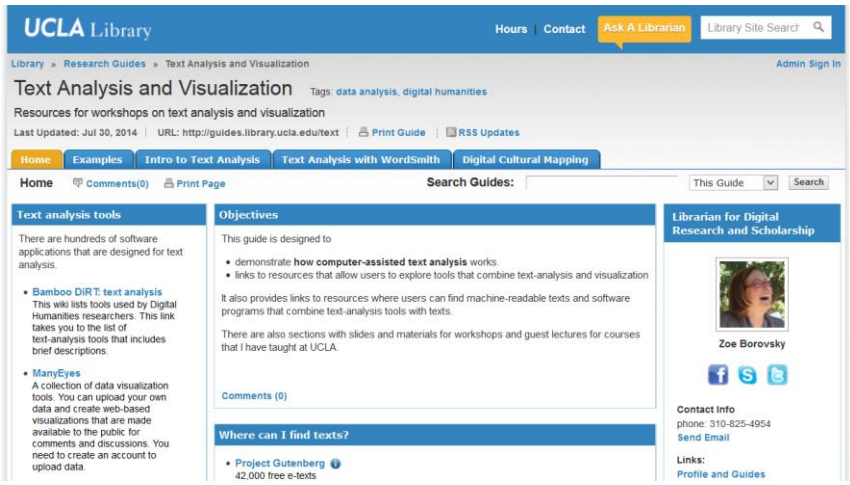
Voyant Tools is a web-based reading and analysis environment for digital texts. Find out more.

Source: http://voyeurtools.org/ and http://hermeneuti.ca/voyeur

School of Information Systems

SMU Singapore Management University

Reference



Source: http://guides.library.ucla.edu/content.php?pid=326428&sid=2670972

Reference

- Fun with Text Visualization http://campusguides.lib.utah.edu/content.php?pid=78288&sid=579667
- The pros and cons of word clouds as visualizations (https://www.visioncritical.com/pros-and-cons-word-clouds-visualizations/)