

On visual Approaches towards Corpus Exploration

short presentations

egon stemle, verena lyding, lionel nicolas

{egon.stemle, verena.lyding, lionel.nicolas}@eurac.edu



European Academy of Bozen/Bolzano
(EURAC)

EURAC
research

May 4th, 2012

- 1 Visualisation - What it is 'nd ain't
- 2 From Data to LInfoVis
- 3 Example
- 4 References

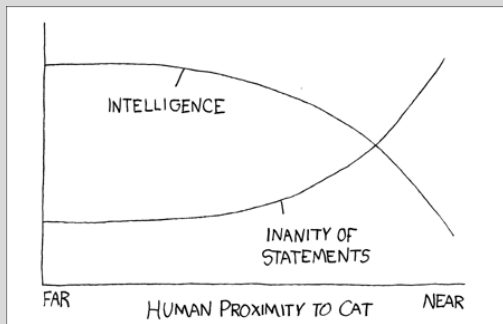
Maps, Charts, Graphs



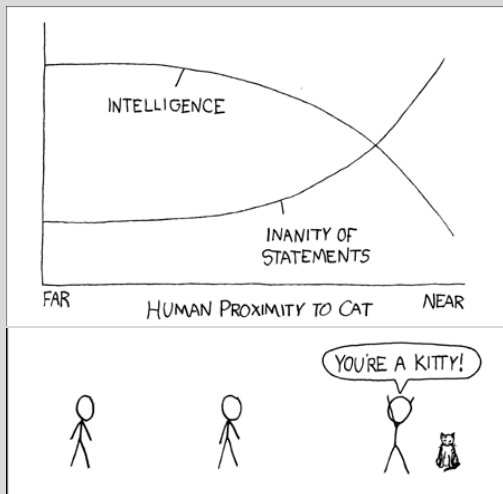
Maps, Charts, Graphs



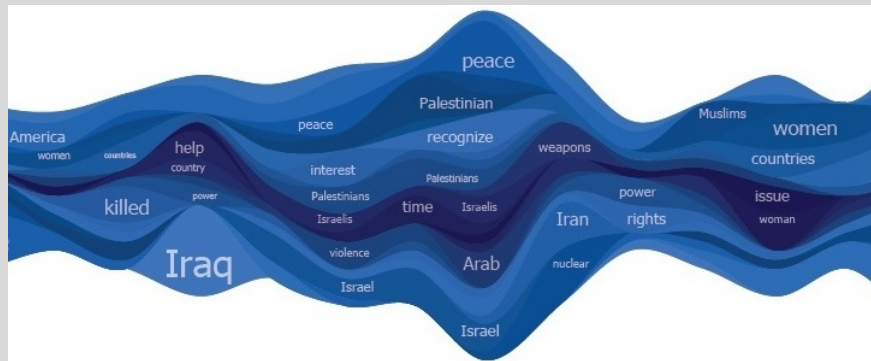
Maps, Charts, Graphs



Maps, Charts, Graphs - in Humor



Stream Graphs, Word Clouds, ...



Stream Graphs, Word Clouds, ...



...like an Elephant



...like an Elephant



- it's a pretty and colourful wild card - very tempting to put into papers, proposals & talks . . . but it is more than pretty pictures!

...like an Elephant



- it's a pretty and colourful wild card - very tempting to put into papers, proposals & talks . . . but it is more than pretty pictures!
- it's coded - a means of communication

...like an Elephant



- it's a pretty and colourful wild card - very tempting to put into papers, proposals & talks . . . but it is more than pretty pictures!
- it's coded - a means of communication
- it's an additional opportunity
 - to employ another communicative channel
 - to analyse data
 - to interact with data

Linguistic Information Visualisation (LInfoVis)

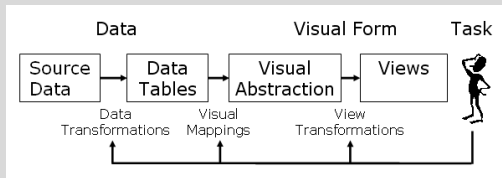
Employ techniques and principles of information visualisation in the context of linguistic data and derived data.

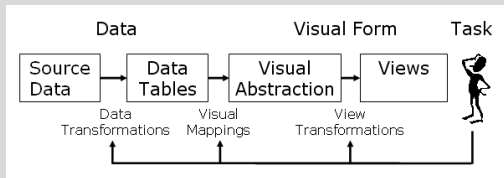
1 Visualisation - What it is 'nd ain't

2 From Data to LInfoVis

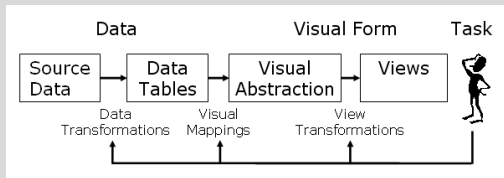
3 Example

4 References





- for LInfoVis: data reduction by abstraction - but which data & how?



- for LInfoVis: data reduction by abstraction - but which data & how?
- Visualise iteratively - discuss and refine
- use the feed-back to improve the transformations and views.

How to play the Wild Card

- Be relevant - Visualise things related to your goal
- Keep it simple - Do not make your visualisation more informative than is required
- Avoid obscurity, and ambiguity - Use known building blocks in a consistent manner

How to play the Wild Card

- Be relevant - Visualise things related to your goal
- Keep it simple - Do not make your visualisation more informative than is required
- Avoid obscurity, and ambiguity - Use known building blocks in a consistent manner
- Inventing new building blocks is a very late resort. . .

1 Visualisation - What it is 'nd ain't

2 From Data to LInfoVis

3 Example

4 References

A 'modified' KWIC View

The screenshot displays the Double Tree software interface. The main window shows a mind map centered on the word "benessere". The branches contain various related terms and phrases, such as "centri", "grad", "generano", "diffuso", "ecologico", "economico", "emotivo", "familiare", "globale", "irreale", "locale", "materiale", "paragonabile", "personale", "più", "psicologico", "quando", "reale", and "si".

On the right side, there is a control panel with the following elements:

- Term:**
- Context length:**
- Search:**
- Context as:** words POS
- Filter by preceding word:**
 - Filter by preceding word:
 - Position:
 - Minimum:
 - Maximum:
- Filter by following word:**
 - Filter by following word:
 - Position:
 - Minimum:
 - Maximum:
- Show only filtered
-
-

At the bottom of the control panel, there is a text box with the following content:

Double Tree is a compact representation of a KWIC for a single word.
Click on a word to expand the context by one level.
By Chris Culy and Verena Ludtke.
Done

-ta.

[Cah11] Cahaleng1102.wordpress.com.

LONDON CAHALLING - V for Vendetta, 2011.

Available from: <http://cahaleng1102.wordpress.com/2011/11/20/v-for-vendetta/trackback/>.

[CL10] Chris Culy and Verena Lyding.

Double Tree: An Advanced KWIC Visualization for Expert Users.

In *2010 14th International Conference Information Visualisation*, pages 98-103. IEEE, July 2010.

Available from:

<http://www.computer.org/portal/web/csdl/doi/10.1109/IV.2010.24>.

[Cla09] Jeff Clark.

Neoformix - Obama UN Speech StreamGraph, 2009.

Available from:

<http://neoformix.com/2009/ObamaUNSpeechStreamGraph.html>.

[CMS99] Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman.

Readings in Information Visualization: Using Vision to Think (Interactive Technologies).

Morgan Kaufmann, 1999.

Available from: <http://www.amazon.com/>

[Readings-Information-Visualization-Interactive-Technologies/dp/1558605339](http://www.amazon.com/Readings-Information-Visualization-Interactive-Technologies/dp/1558605339).

[Etz00] Erhard Etzlaub.

Kolorierte Erstaussgabe der Romwegkarte, 1500.

Available from: [http:](http://daten.digital-e-sammlungen.de/~db/0003/bsb00033752/image_1)

[//daten.digital-e-sammlungen.de/~db/0003/bsb00033752/image_1](http://daten.digital-e-sammlungen.de/~db/0003/bsb00033752/image_1).

[Han88] Itcho Hanabusa.

Blind men and an elephant, 1888.

Available from:

http://en.wikipedia.org/wiki/Blind_men_and_an_elephant.

- [Mun] **Randall Munroe.**
XKCD - Cat Proximity.
Available from: <http://xkcd.com/231/>.
- [Pri69] **Joseph Priestley.**
A New Chart of History, 1769.
Available from:
http://en.wikipedia.org/wiki/A_New_Chart_of_History.
- [SI10] **Julie Steele and Noah Iliinsky, editors.**
Beautiful Visualization: Looking at Data through the Eyes of Experts (Theory in Practice).
O'Reilly Media, 2010.
Available from: <http://shop.oreilly.com/product/0636920000617.do>.