Crossreads
towards a rhizomatic narrative

Jaume Nualart [1,2,3] Gabriela Ferraro [1,4]

[1] NICTA / Locked Bag 8001, Canberra ACT 2601, Canberra, Australia
[2] University of Barcelona. Catalonia
[3] University of Canberra. Australia
[4] The Australian National University, Canberra, Australia

NICTA
UNIVERSITY OF CANBERRA
AUSTRALIA'S CAPITAL UNIVERSITY
Universitat de Barcelona
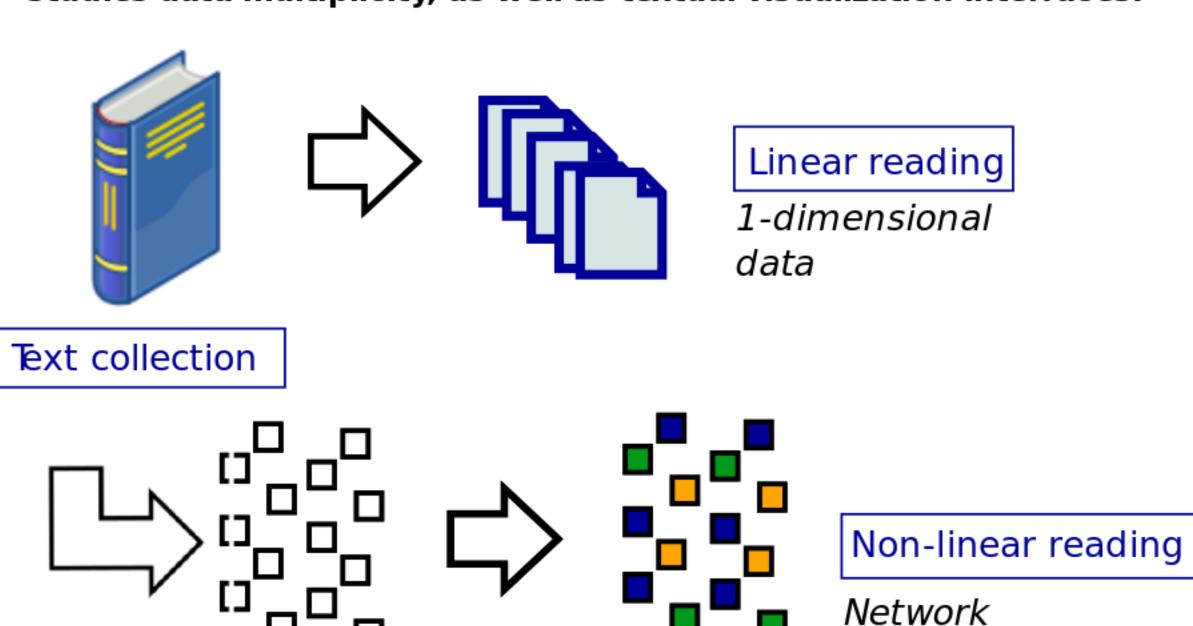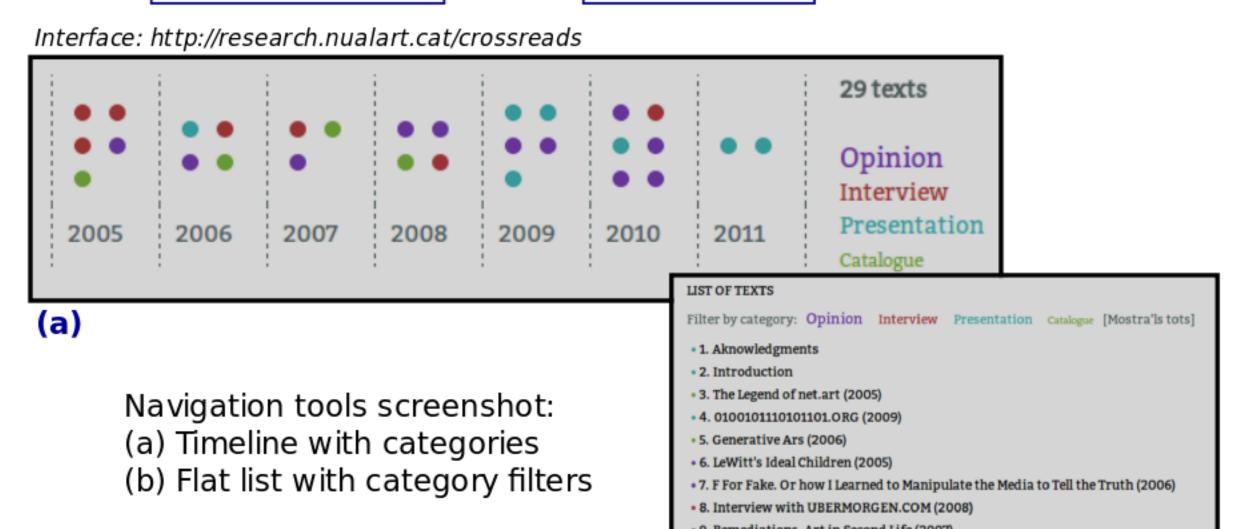ANU

We present Crossreads, a manner to deconstruct linear narratives and to read texts in multiple orders. This is an ongoing work that studies data multiplicity, as well as textual visualization interfaces.
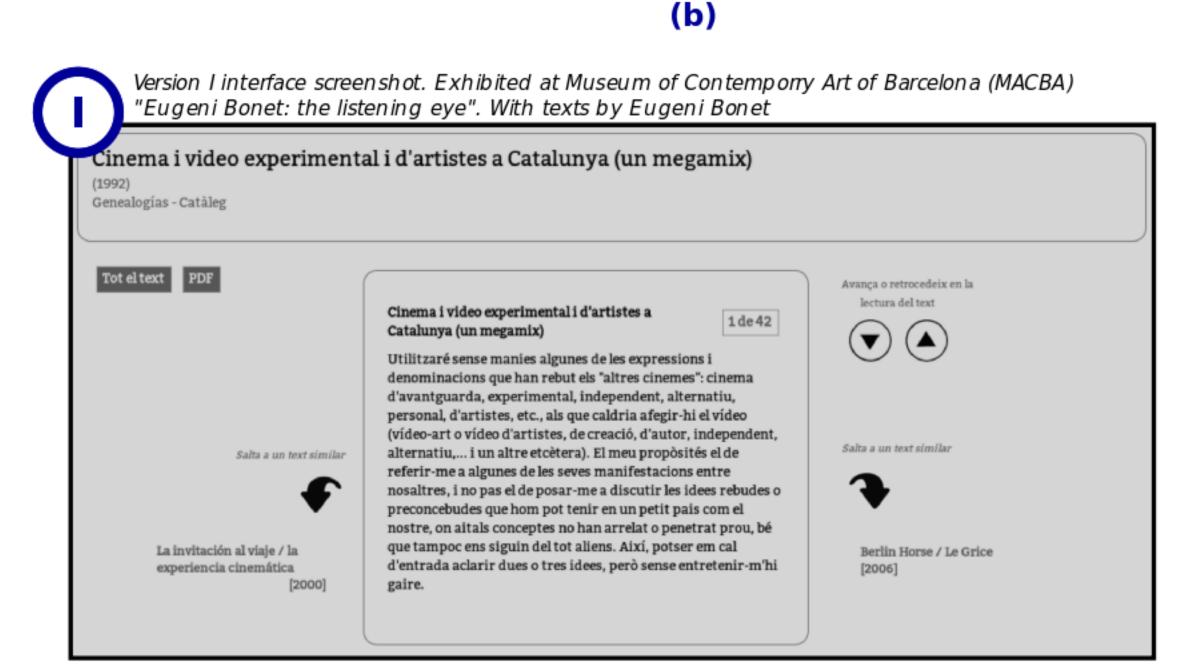
Text collection

Linear reading
1-dimensional data

Non-linear reading
Network data

Segmentation   Similarities

Interface: http://research.nualart.cat/crossreads

(a)

Navigation tools screenshot:
(a) Timeline with categories
(b) Flat list with category filters

(b)

I

Version I interface screenshot. Exhibited at Museum of Contemporary Art of Barcelona (MACBA) "Eugeni Bonet: the listening eye". With texts by Eugeni Bonet
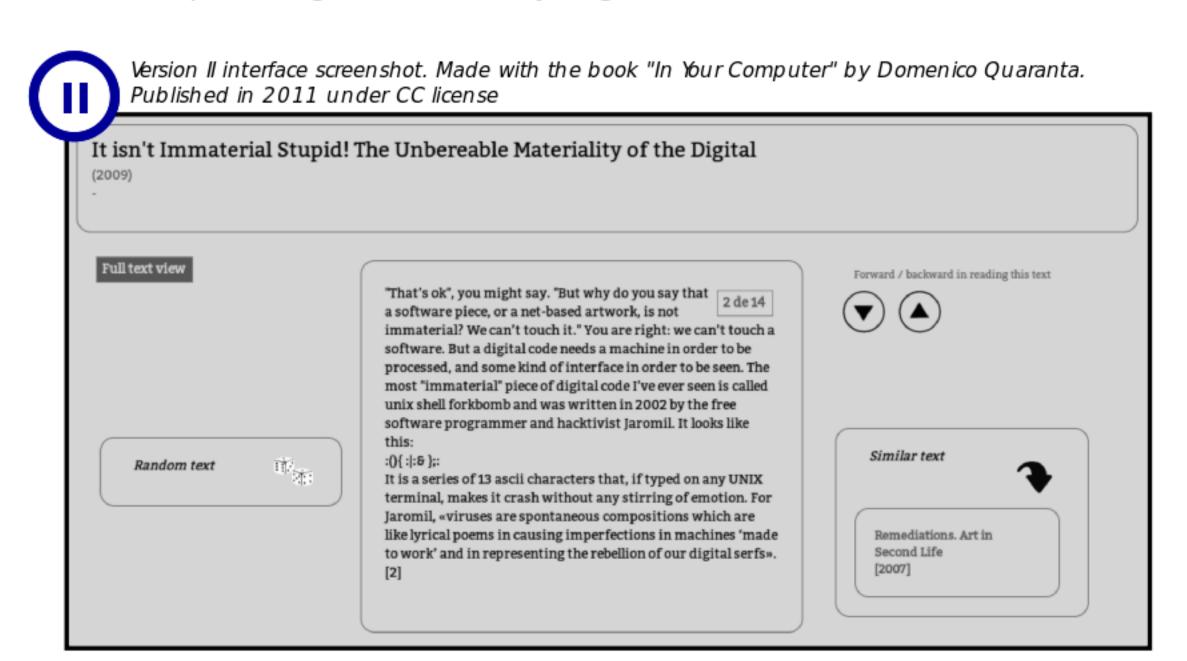
Each segment is linked to two similar segments.
Links have a wide range of similarity scores
No orphan segments = every segment offers a crossread

II

Version II interface screenshot. Made with the book "In Your Computer" by Domenico Quaranta. Published in 2011 under CC license

Each segment is linked to its most similar segment
To avoid repetition of pairs, segments that has been already set as a maximum similarity segment during ten iterations are skipped.
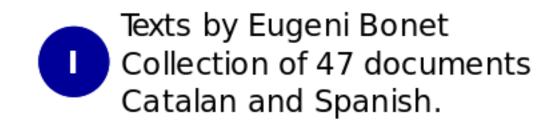
➥ Objectives

- Design an approach that allows reading a text in multiple ways or paths
- Exploit visualization techniques to support reader's exploration: interaction reader-text, text as a collection of segments, most popular reading paths, and so on.

➥ Related work

- Deleuze and Guattari have described the rhizomatic structure of knowledge. "A thousand plateaus: Capitalism and schizophrenia" 1987.
- In the novel Hopscotch (1966) by J. Cortázar, the author proposes two ways of reading the novel.
- The Project Xanadu (1960) is considered the first hypertext project in the digital era, and it was a visionary definition of WWW standards.
- Learners naturally make connections between pieces of knowledge, and they are better able to retrieve and apply their knowledge when those connection are accurate and meaningful.
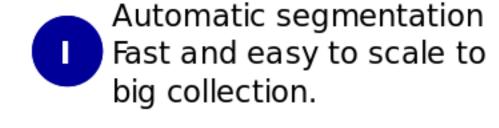
➥ Experiments

Data set selection

I Texts by Eugeni Bonet Collection of 47 documents Catalan and Spanish.

II Texts by Domenico Quaranta Collection of 29 documents English

Data segmentation

Each document is segmented into pieces. A segment contained aprox. 700 characters = one minute of reading

I Automatic segmentation Fast and easy to scale to big collection.

II Human segmentation Richer than the automatic one, but also subjective.

Data similarity

The computation of the similarity between segments comprises the following steps;
- Tokenization
- Standard stop word removal
- Named Entity Recognition: Person, Location, Organization and Others (OpenNLP Named Entity recogniser)
- Similarity Calculus between segments

➥ Conclusions & future work

Crossreads is an ongoing work that aims to explore the non-linear text reading.

Perform user studies to validate the proposed segmentations and the different versions of crossreads

Define which texts or text collections are appropiate for crossreads: one or multiple authors, one or multiple genres, monolingual and/or multilingual collections, etc.