Crossreads towards a rhizomatic narrative

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.

- **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 Hopschotch (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 connections 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 approx. 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;
      - **I**: Tokenization
      - **II**: Standard stop word removal
      - **I**: Named Entity Recognition: Person, Location, Organization and Others (OpenNLP Named Entity recogniser)
      - **II**: Similarity 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 appropriate for crossreads: one or multiple authors, one or multiple genres, monolingual and/or multilingual collections, etc.