

NLP per la Storia

Rachele Sprugnoli, Digital Humanities Group - Fondazione Bruno Kessler (Trento)
sprugnoli[AT]fbk[DOT]eu

Pipeline di strumenti NLP:

- EN (ma anche altre lingue): <https://stanfordnlp.github.io/CoreNLP/>
- IT: <http://tint.fbk.eu/>

Aggregatore strumenti NLP e linking:

- EN : <http://pikes.fbk.eu/>

Estrattore di topic/concetti-chiave/domani:

- MALLETT: <http://mallet.cs.umass.edu/>
- Topic-modeling-too:
<https://code.google.com/archive/p/topic-modeling-tool/downloads>
- KD: Keyphrase Digger
Download the script: <http://dh.fbk.eu/technologies/kd>
Use the online demo: http://celct.fbk.eu:8080/KD_KeyDigger/
Paper: Moretti, G., Sprugnoli, R., & Tonelli, S. (2015). Digging in the Dirt: Extracting Keyphrases from Texts with KD. *CLiC it*, 198.
https://iris.unito.it/retrieve/handle/2318/1532924/75495/Accademia_University_Press_978-88-99200-62-6.pdf#page=200
- L-KD:
Demo online: http://celct.fbk.eu:8080/L_KD/
Paper: Moretti, G., Sprugnoli, R., & Tonelli, S. (2016). KD Strikes Back: from Keyphrases to Labelled Domains Using External Knowledge Sources. In *Proceedings of the Third Italian Conference on Computational Linguistics (CLiC-it 2016)* (pp. 216-221).
https://www.researchgate.net/profile/Andrea_Abel/publication/311589135_CLIC-IT-2016_Proceedings/links/584fced808aecb6bd8d1dc55.pdf#page=218

Un paio di applicazioni.

- ALCIDE platform:
Paper principale: Giovanni Moretti, Rachele Sprugnoli, Stefano Menini and Sara Tonelli "ALCIDE: Extracting and visualising content from large document collections to support Humanities studies". *Knowledge-Based Systems*, vol. 111, pp. 100-112.
<http://www.sciencedirect.com/science/article/pii/S0950705116302635>
Demo online: http://celct.fbk.eu:8080/Alcide_Demo/
- Ramble On:
Info: <http://dh.fbk.eu/technologies/rambleon>
Paper: Menini, S., Sprugnoli, R., Moretti, G., Bignotti, E., Tonelli, S., & Lepri, B. (2017). RAMBLE ON: Tracing Movements of Popular Historical Figures. *EACL 2017*, 77. <http://www.aclweb.org/anthology/E/E17/E17-3.pdf#page=89>

Visualizzazioni:

- D3: <http://app.rawgraphs.io/>
- RAW: <https://d3js.org/>
- Gephi (per le network): <https://gephi.org/>

Tutorial su vari aspetti dell'utilizzo di NLP nella storia:

- GSCL 2013 Tutorial "Natural Language Processing for Historical Texts":
<http://dynalabs.de/mxp/gscl2013/>
- The Programming Historian: <http://programminghistorian.org/>

Libro:

- The Historian's Macroscopic: Exploring Big Historical Data:
<http://www.themacroscopic.org/2.0/>