DATA MINING HISTORICAL NEWSPAPERS METADATA

Jean-Philippe Moreux




DATA MINING HISTORICAL NEWSPAPERS METADATA

Old News Teaches History

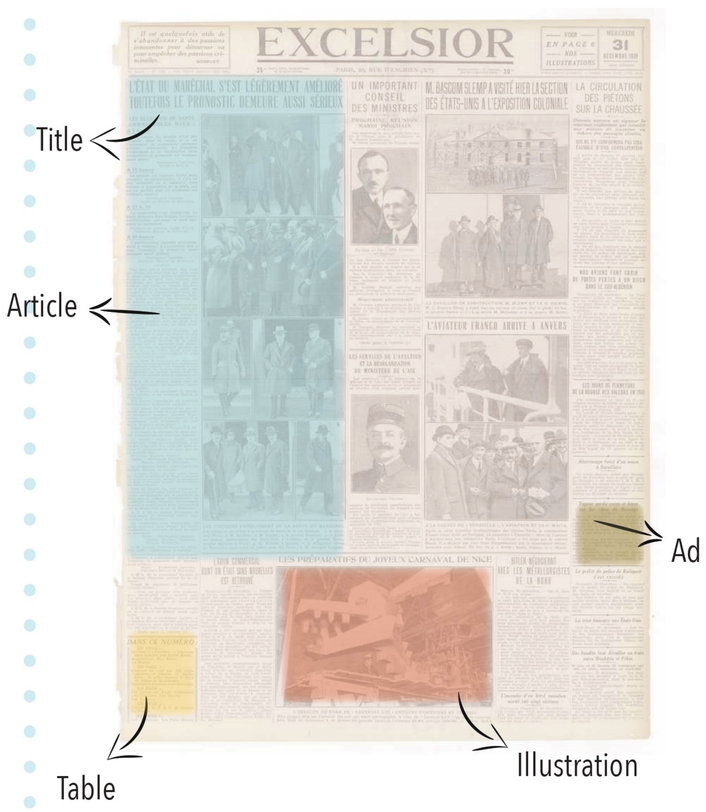
Would you like to provide:

- **Researchers** (DH, History of Press, Information Science) with quantitative medata on press layout and content?
- **Digital Curators** with insights on their collections?
- **Digitization Program Managers** with statistics?

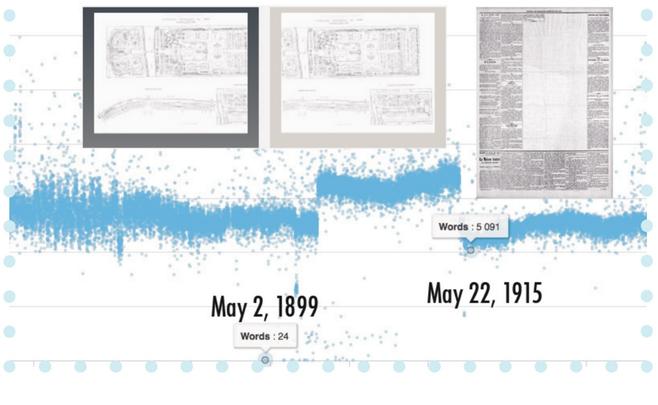
Just datamine and dataviz your METS/ALTO digitized press collection!

THE TEST BED

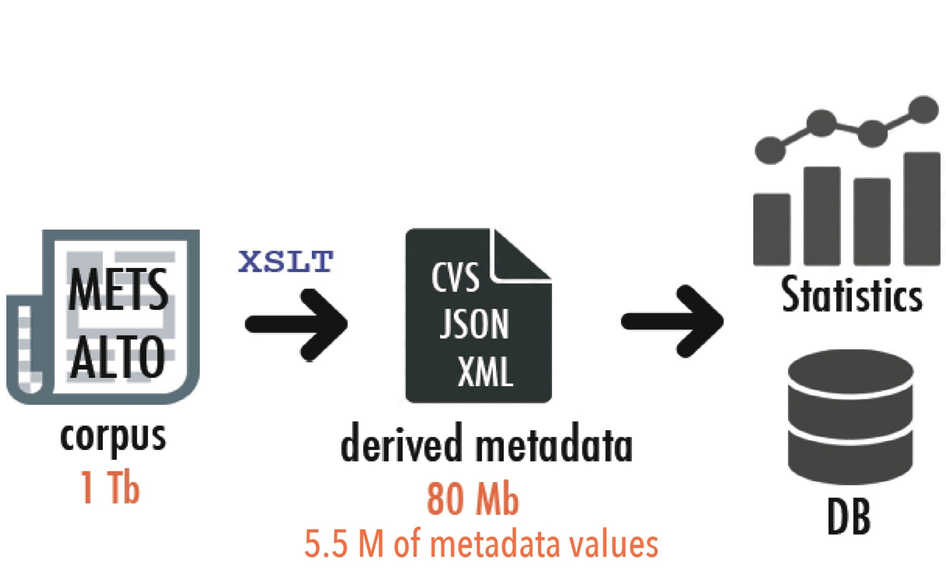
Europeana Newspapers project (2012-2015): 880K OLR'ed pages from BnF collection (1814-1944)



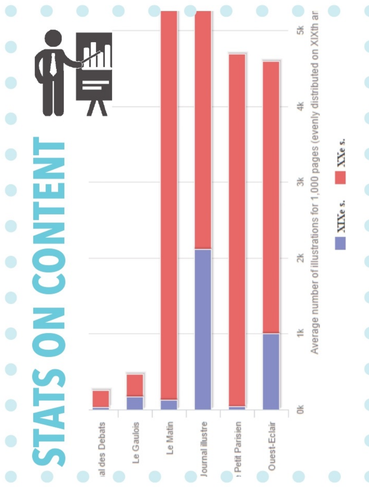
UNCOMMON ISSUES



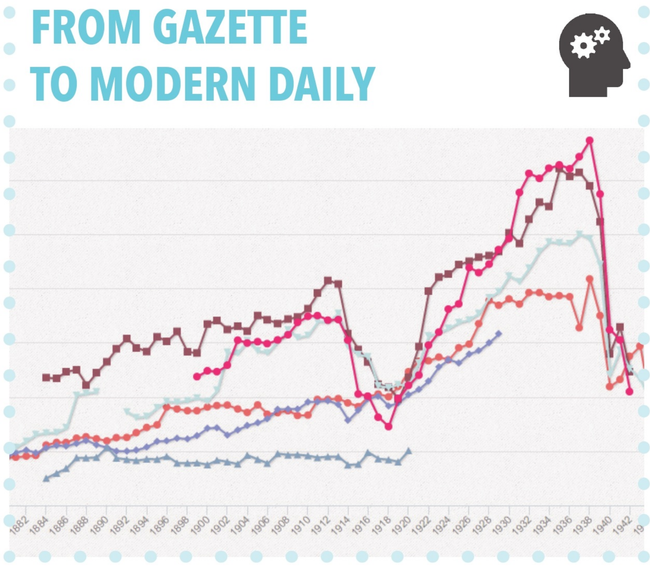
USE OF ILLUSTRATION



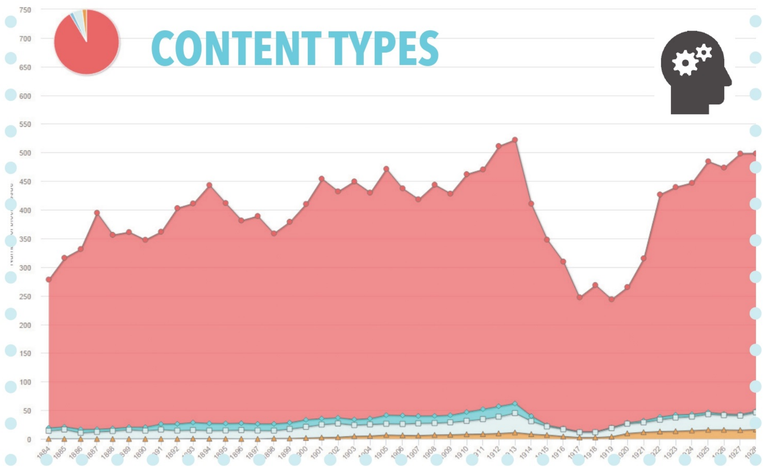
STATS ON CONTENT



FROM GAZETTE TO MODERN DAILY



CONTENT TYPES



europeana newspapers

Dive into the details of Europe's historic newspapers.

www.europeana.eu &
www.theeuropeanlibrary.org/tel4/newspapers

Jean-Philippe Moreux, BnF

Staatsbibliothek zu Berlin
Preußischer Kulturbesitz

This project runs from February 2012 to February 2015. It is led by the Staatsbibliothek zu Berlin and co-funded by the European Commission under the Competitiveness and Innovation framework Programme. http://ec.europa.eu/ict_psp