

COMPARE-IT SELECT. Detailed corpus design

Here is a detailed description of the corpus COMPARE-IT SELECT. Information is given on the thematic sections of the articles collected, the number of words collected in each thematic section and the dates of publication of the articles.

COMPARE-IT SELECT_IT (Italy): 150'000 words

Politics (Politica)

corriere.it	20'000	11.2011 – 12.2011
lastampa.it	20'000	11.2011
repubblica.it	10'000	05.2011 – 11.2011

Economy (Economia, finanza)

corriere.it	30'000	09.2011 – 12.2011
repubblica.it	20'000	09.2011 – 10.2011

Sports (Sport)

corriere.it	20'000	01.2011 – 12.2011
lastampa.it	10'000	01.2011 – 12.2011
repubblica.it	20'000	03.2011 – 12.2011

COMPARE-IT SELECT_CH (Switzerland): 150'000 words

Politics (Politica)

cdt.ch	50'000	01.2013 – 12.2013
--------	--------	-------------------

Economy (Economia)

cdt.ch	45'000	01.2013 – 12.2013
gdp.ch	5'000	03.2013 – 09.2013

Sports (Sport)

cdt.ch	40'000	01.2013 – 12.2013
gdp.ch	10'000	08.2013 – 09.2013