

# Cross Language search on the WEB and Digital Libraries

Luca Dini

Paolo Curtoni

[dini/curtoni@celi.it](mailto:dini/curtoni@celi.it)

CELI-Torino

[www.celi.it](http://www.celi.it)

Ad-hoc Clef – Ad-hoc ImageClef

# Challenges

## ■ Challenges

- Absence of a consolidated corpus
- Absence of a precise domain
- Presence of very technical query/docs
- Absence of style constraints and document formats (WEB only).

## ■ Approach (IT-EN):

- Query translation;
- Huge dictionaries;
- Query disambiguation mechanisms.

# Results

CELIdescNOEXPANSIONboost	NUM_RELEV	MEAN_AVE	R_PRECISION	PRECISION_AT_10	PRECISION_AT_20_R
CEL.titleNOEXPANSION	773	0,2397	0,2381	0,2400	0,1890
CELIdescNOEXPANSION	764	0,2268	0,2381	0,2320	0,1720
CEL.titleLisaExpansion	814	0,2238	0,2212	0,2160	0,1720
CEL.titleCw nCascadeExpansion	673	0,2390	0,2400	0,2020	0,1650
CEL.titleCw nExpansion	636	0,2110	0,2074	0,1980	0,1520
CEL.titleNOEXPANSIONboost	680	0,2035	0,2036	0,1900	0,1430
CELIdescLisaExpansion	793	0,1941	0,2016	0,1840	0,1470
CELIdescCw nExpansion	594	0,1792	0,1900	0,1760	0,1460
CELIdescCw nCascadeExpansion	602	0,1957	0,1982	0,1720	0,1380
CELIdescLisaExpansionboost	767	0,1908	0,1966	0,1660	0,1370
CELIdescNOEXPANSIONboost	733	0,1812	0,1811	0,1600	0,1300
CEL.titleLisaExpansionboost	712	0,1732	0,1795	0,1600	0,1370

System	REL_RETR	MAP	P10	P20
IT-EN-AUTO-NOFB-TXT-CELI-AND_CwnExpansion	1474	0,1392	0,2233	0,2000
IT-EN-AUTO-NOFB-TXT-CELI-AND_NOEXPANSION	1551	0,1486	0,2150	0,2092
IT-EN-AUTO-NOFB-TXT-CELI-OR_NOEXPANSION	1551	0,1441	0,2033	0,1975
IT-EN-AUTO-NOFB-TXT-CELI-AND_CwnCascadeExpansion	1590	0,1362	0,1917	0,1817
IT-EN-AUTO-NOFB-TXT-CELI-OR_CwnExpansion	1474	0,1284	0,1917	0,1775
IT-EN-AUTO-NOFB-TXT-CELI-AND_LisaExpansion	1655	0,1237	0,1867	0,1725
IT-EN-AUTO-NOFB-TXT-CELI-OR_LisaExpansion	1655	0,1189	0,1633	0,1558
IT-EN-AUTO-NOFB-TXT-CELI-OR_CwnCascadeExpansion	1590	0,1208	0,1583	0,1492