



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE DELLA RICERCA

Alma Mater Studiorum Università di Bologna Archivio istituzionale della ricerca

PERCEIVE project - Deliverable D5.2 "Database of the topics and sentiments to be made available on-line for further research"

This is the submitted version (pre peer-review, preprint) of the following publication:

Published Version:

Availability:

This version is available at: <https://hdl.handle.net/11585/622227> since: 2019-07-17

Published:

DOI: <http://doi.org/10.6092/unibo/amsacta/5726>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)



PERCEIVE

Perception and Evaluation of Regional and Cohesion Policies by Europeans and Identification with the Values of Europe

Perception and Evaluation of Regional and Cohesion Policies by Europeans and Identification with the Values of Europe

PERCEIVE

GA nr. 693529

D 5.2 'Database of the topics and sentiments to be made available on-line for further research'

Table of contents

1. Introduction	3
2. Keywords translation.....	4
3. Data Collection	5
3.1 Communication of the EU.....	5
3.2 Financed projects' abstracts	5
3.3 Newspapers.....	5
3.4 Twitter	10
3.4.1 Institutional string	12
3.4.2 Discursive string	18
3.5 Facebook	18
4. Analysis: Topic Modeling.....	21
5. Conclusion.....	25
6. References.....	25
7. Appendix.....	26
7.1 Keywords in English.....	26
7.2 Keywords in Italian.....	29
7.3 Keywords in German.....	33
7.4 Keywords in Polish.....	38
7.5 Keywords in Romanian	42
7.6 Keywords in Swedish.....	46
7.7 Keywords in Spanish.....	50
7.8 English Stopword list	54

AUTHORS OF THIS REPORT AND EDITORS OF THE DELIVERABLE:

- Vitaliano Barberio, Research Institute for Urban Management and Governance, WU University of Economics and Business, Vienna, vitaliano.barberio@wu.ac.at
- Ines Kuric, Research Institute for Urban Management and Governance, WU University of Economics and Business, Vienna, ines.kuric@wu.ac.at
- Renate Meyer, Research Institute for Urban Management and Governance, WU University of Economics and Business, Vienna, Renate.Meyer@wu.ac.at
- Markus Höllerer, Institute for Public Management and Governance, WU University of Economics and Business, Vienna
- Edoardo Mollona, Università di Bologna, edoardo.mollona@unibo.it
- Luca Pareschi, Università di Bologna, luca.pareschi@unibo.it

LEADING PARTNER:

Alma Mater Studiorum – Università di Bologna [UNIBO, Italy]

ESTIMATED PERSON/MONTHS PER PARTNER (person/month):

Alma Mater Studiorum – Università di Bologna [UNIBO, Italy] : 6

Wirtschaftsuniversität Wien [WU, Austria]: 6

SUBMISSION DATE: *31th October 2017*

1. Introduction

This document describes the database we developed from data collection and organization activities as described in tasks 5.2 and 5.3 of the PERCEIVE project. The former task, which lasted from month 5 to month 10, consisted of an "Original data collection for exploratory topic modeling". The latter task, which lasted from month 10 to month 13, was named "Topic modeling and qualification". As stated in the Grant Agreement, this description of the database was expected around month 14 (October 2017), while the database itself will be made publically available on-line for further research only after the end of the PERCEIVE project. Indeed, task 5.4, which will last until month 15 is the "Actual Construction of a database of discursive topics". This document will thus focus on data collection and organization as well as provide an example of the topic models that we are developing coding discourses from different kinds of described datasets.

In order to conduct an exploratory documental historical analysis of EU Cohesion Policy and identity multiplicity, we planned to collect texts from several printed sources. During our data collection process, we improved our collection strategy by conducting several pilot tests. Table 1 presents the differences between the planned data collection process and the actual data that we collected.

Table 1 - Planned and realized data collection

	Planned data collection	Realized data collection
Level	Communication of the EU: documents constituting and explaining Cohesion Policy (periods 2007-2013 and 2014-2020)	We collected from the EC INFOREGIO web portal: <ul style="list-style-type: none"> • all issues of the PANORAMA magazine • all available brochures
	Financed projects' abstracts: from the portal of the EC (policy learning database)	We collected all descriptions of 'flagship' projects at the projects' pages of the EC INFOREGIO web portal
	Local press 8 national newspapers - 1 for each country in the consortium 8 regional newspapers - 1 for each region in the sample	We further elaborated and broadened the original objective and strategy. We collected 41 newspapers covering segments as indicated below: <ul style="list-style-type: none"> • 14 national 'quality' newspapers - 2 for each country represented in the consortium • 7 national 'tabloids' - one for each country represented in the consortium • 7 national 'business' newspapers - one for each country represented in the consortium • 13 regional newspapers - 2 for each country/region represented in the consortium, except for Austria/Burgenland, where only one relevant regional newspaper was

		found.
	Communication on social media (eventually add 3 months): Tweets: three to six theoretically sampled hashtags per country	We developed two research strings: <ul style="list-style-type: none"> • An 'institutional' one, based on institutional accounts and technical hashtags, was meant to capture the centrally-launched debate. • A 'discursive' one, composed of keywords translated in the different national languages, was aimed at capturing the debate launched by citizens.
	Communication on social media (eventually add 3 months) Blogs: two per country (left-wing vs. right-wing)	The debate in the public sphere moved from blogs to Facebook. We thus collected posts and blogs from 26 Facebook profiles, including LMAs and relevant politicians. For the analysis we focused on LMA profiles in order to ensure comparability.

We then conducted a semi-automatic analysis of collected textual data in order to model meanings produced through different media and by different actors. We are detecting topics, collections of frequently co-occurring words, and their qualification through sentiments analysis, automatic detections of emotional orientation of topics based on given vocabularies. The following paragraphs will detail the data collection process and provide an example of topic modeling.

2. Keywords translation

Data collection at national and regional level required a proper list of keywords in order to search for meaningful articles, tweets and posts. Therefore, our first task was to generate such a list and have it translated in all of the PERCEIVE project's national languages. The original list was developed in English and consists of seven sets:

- **Regional Policy (generic vocabulary).** This set includes generic words related to Regional and Cohesion Policy, such as development, strategy, investment, program, and infrastructure.
- **Thematic areas (2014-2020).** This sets contains words that define the thematic areas, such as "competitiveness of SMEs", and "Social Inclusion".
- **Additional themes RP** deal with urban and rural development, Northern Ireland and words referring to the enlargement of the EU.
- **EU Funds** specifically targets the official titles of EU funding programs.
- **Five targets 2020 strategy** identifies the five objectives of the 2020 strategy: employment, research and development, climate change and energy sustainability, education, and fighting poverty and social exclusion.
- **Cooperation** deals with the different cooperation programs and their vocabulary.

- **EU Identity** captures words used to define the idea of “being European”, or “europeanness”. In this set we find words referring to symbolic constructs, such as culture, nation, tradition, heritage, and identity.

All the words in each set were then translated thanks to PERCEIVE's partners. In the appendix we report the detailed original list, together with its translation in Polish, Romanian, Swedish, Spanish, German and Italian.

3. Data Collection

As we dealt with very different sources, each one needed a specific data collection strategy, which we describe in the following paragraphs.

3.1 Communication of the EU

We planned to sample communication materials directly produced by EC – DG REGIO in order to be able to explore the translation of content from central to local (i.e. national and regional) level. In line with initial task specifications we have collected information materials issued by the EC – DG REGIO and distributed through their web portal INFOREGIO¹ taking three main forms: first, and most important, all 66 issues of the PANORAMA magazine spanning over the 2006-2017 time period with irregular intervals (from 1 to 15 issues per year); second, all 62 brochures available (also 2006-17 period); third and finally, the EU communication plans/guidelines as both a source of textual data and conceptual content.

3.2 Financed projects' abstracts

We planned to collect communication materials reflecting how individual projects are communicated (possibly, but not necessarily by projects themselves). This objective has been accomplished to a satisfactory degree with a view to PERCEIVE's aim of inquiring the ways (i.e. language use) in which projects are communicated to the public. In fact, we have retrieved textual data – i.e. descriptions of 'flagship' projects (objectives and achievements) from the INFOREGIO portal under the Projects section². For the 2007-2013 programming period this set comprises 1250 projects and 752 major projects, while the 2000-2007 dataset comprises 715 projects and no major projects.

3.3 Newspapers

Newspapers have both empirical and conceptual importance for PERCEIVE as they constitute very active and relevant actors in the social-linguistic construction of the EU and its policies. Therefore, considerable effort was allocated by the involved research teams towards the collection of newspapers' contents, namely the body text of published articles. Our sampling strategy entailed several steps as well as the direct collaboration with all partners as described in more detail in the remaining part of this section.

A first step consisted of (re-)defining the **composition** of the 'desirable' sample in terms of newspapers' types or categories. In the initial drafting of the PERCEIVE project we planned this task in terms of minimal requirements: 7 national and 7 regional newspapers – one national and one

¹ At the publications section, the INFOREGIO portal provides a search mask where different communication genres (i.e. reports, brochures, magazines, guidelines, etc.) can be selected along with other relevant search criteria (i.e. programming period, language etc.). For more information visit the following page:

http://ec.europa.eu/regional_policy/en/information/publications/

² Available under the following address: http://ec.europa.eu/regional_policy/en/projects

regional newspaper for each country represented in the research consortium. However, early explorative sampling tests revealed that better coverage of relevant concepts (based on agreed-on keywords) could be obtained with a more extensive and structured sample. The final decision made on this matter was to include:

- Daily newspapers issued in the 7 countries and 9 NUTS2 regions of PERCEIVE case studies (see Deliverable 1.1).
- In terms of newspapers' categories, for each of the 7 national settings, we aimed at two national *quality* newspapers, one national *tabloid*, one national *business* and two *regional* newspapers per NUTS2 region case study (where available at a reasonable cost/effort to collect).

In a second step of the sampling process, we tried to **balance efficiency of data collection with the relevance of media sources** to mirror the discourse on EU Regional Policy and EU identity. Accordingly, we have decided to start our search from a highly comprehensive news database called FACTIVA³ which was available to one of the partners through an institutional license and therefore involved no additional cost. Then, all PERCEIVE consortium partners were asked to provide feedback regarding the relevance of available titles (i.e. newspapers) for their countries as well as to suggest eventual integrations. As a result of this integrated collective effort we were able to arrive to the selection described in table 2 below.

Table 2 - Newspapers' sample

Country	AT	UK	IT	ES	RO	SE	PL	TOTALS
number of newspapers	5	6	6	6	6	6	6	41
Sampled articles	4216	4669	9392	12064	11186	6667	8342	56536

A third and final step entailed operations aimed at 'cleaning' and organizing collected information (i.e. articles and related bibliographic information) into a dataset. In order to accomplish said objectives we started removing articles identified as 'duplicate' (i.e. same articles retrieved by different search keys on same newspapers) from the dataset. Then, we removed all unnecessary strings (i.e. copyright and bibliographic information, database export statistics from data providers and residual HTML lines) from the articles' text. Finally, we constructed a tabular (.csv) schema matching the textual content of each article with the corresponding structured data (i.e. authorship, date of publication etc.) through a unique identifier code. As the table shows, our objectives in terms of sample composition were fully achieved. In fact we managed to collect articles for all countries and regions listed in the categories and quantities above. In Austria (AT) we were able to only collect data for a single regional newspaper, which seems satisfactory nevertheless given the very small size of the NUTS 2 region. In fact, other sources we checked did not mention the respective keywords at all. More detailed information about which newspapers and keywords we matched in order to sample articles are reported in the following seven tables (one for each member country). Please notice that results for specific keywords and overall totals include repetitions (i.e. same articles sampled with different keywords) while totals per newspapers do not.

³ The collection of data for sources not available on FACTIVA was individually addressed case by case turning to other databases such as Lexis Nexis, APA DeFacto and individual content licensors.

PERCEIVE DEL. 5.2: 'DATABASE OF THE TOPICS AND SENTIMENTS TO BE MADE AVAILABLE ON-LINE FOR FURTHER RESEARCH'

Table 3 – Italy: newspapers and keywords

TITLE	TYPE	COVERAG	SEARCH KEYS	ARTICLES	TOTALS
La Repubblica	quality national	2005-2017	(unione europea or ue) and investment* and regional*	1064	3107
			(unione europea or ue) and fond* struttural*	503	
			(unione europea or ue) and politic* regional*	208	
			fondi europei and regional*	1332	
Corriere della sera	quality national	1997-2017	(unione europea or ue) and fond* struttural*	392	1291
			(unione europea or ue) and politic* regional*	98	
			(unione europea or ue) and (fond* eu or fond* europe*)	801	
Il sole 24 ore	business	2001-2017	(unione europea or ue) and fond* struttural*	1069	2353
			(unione europea or ue) and politic* regional*	250	
			(unione europea or ue) and (fond* eu or fond* europe*)	1034	
il giornale di calabria	regional	2008-2017	politiche regionali	425	1624
			fondi strutturali	296	
			fondi ue	231	
			fondi europei	672	
il resto del Carlino	regional	2005-2017	fond* (europe* or ue) and emilia* romagna	407	1180
			fondi strutturali	633	
			(politiche regionali or politica regionale or politiche di coesione or politica di coesione) and (ue or unione europea	140	
Leggo	tabloid*	2015-2017	(fond* europe*) or (fond* ue) or (fondi and ue) or (fondi and europ*)	213	598
			(politica europ* or politiche) and (ue or europ*)	296	
			fond* struttural* or invetiment* struttural* or (ue and struttural*) or (europ* and struttural*)	89	
			OVERALL		533

*Italy does not have tabloids in the common sense of the category, but this source (free press) gets close as of editorial format

Table 4 – Austria: newspapers and keywords

TITLE	CATEGORY	COVERAGE	SEARCH KEYS	ARTICLES	TOTALS
DerStandard	quality national	1992-2017	EU and Regionalpolitik*	326	1588
			(eu or europäische union) and investition* and regional*	440	
			EU and strukturfond*	385	
			strukturfond*	437	
DiePresse	quality national	1993-2017	EU and Regionalpolitik*	336	1382
			EU and strukturfond*	322	
			(EU or europ*) fond*	256	
			(eu or europäische union) and investition* and regional*	468	
Wirtschaftsblatt	business national	1998-2016	(eu or europäische union) and investition* and region*	1000	1339
			(eu or europäische union) and (regionalpolitik* or kohäsionspolitik*)	142	
			strukturfond*	197	
Kronen Zeitung	tabloid	2000-2017	eu and investition and regio*	255	439
		1994-2017	regionalpolitik* (eu OR europäische union)	104	
Bvz	regional	2003-2017	strukturfond*	80	
			euEuropäischeUnionFonds	50	158
			regionalpolitik	91	
OVERALL					4906

Table 5 – Poland: newspapers and keywords

TITLE	TYPE	COVERAGE	SEARCH KEYS	SAMPLED ARTICLES	TOTALS
Fakt	national tabloid	2009-2017	polityka regionalna ue	33	271
			polityk spójności	48	
			fundusz ue	190	
gazeta Wyborcza	national quality	2003-2017	polityk* spójno* or (polityk* regional* and europejsk*)	595	1910
			fundus* ue	385	
			fundus* struktural*	1124	
gazeta Olsztynska	regional	2013-2017	polityka spójności	13	1179
			Fundusze europejskie	737	
			regionalnych europejski	429	
gazeta Wroklawska	regional	2009-2017	polityka spójności	99	296
			fundusz ue	99	
			polityka regionalna ue	98	
gazeta Prawna	Business	2010-2017	(polityk* spójno*) OR (polityk* regional* AND europejsk*)	296	716
			fundus* ue	238	
			fundus* struktural*	294	
		2006-2010	(polityk* spójno*) OR (polityk* regional* AND europejsk*)	109	1011
			fundus* ue	444	
Rzeczpospolita	National quality	2001-2016	fundus* ue	940	2758
			fundus* struktural* and (ue or uni* europejsk*)	1302	
			(polityk* spójno*) OR (polityk* regional* AND uni* europejsk*)	794	
OVERALL					8141

PERCEIVE DEL. 5.2: 'DATABASE OF THE TOPICS AND SENTIMENTS TO BE MADE AVAILABLE ON-LINE FOR FURTHER RESEARCH'

Table 6 – Romania: newspapers and keywords

TITLE	TYPE	COVERAGE	SEARCH KEYS	SAMPLED ARTICLES	TOTALS
Adevarul	national quality	2012-2017	politic regional europe	134	3496
			fonduri structural	137	
			fonduri europene	3225	
ObiectivTulcea	regional	2010-2017	politic regional europe	74	870
			fond structural	62	
			fonduri europene	734	
Viata Libera Galati	regional	2009-2017	fonduri europene	148	255
			fonduri structural	81	
			politica regionala	16	
JournalulNational	national quality	2004-2017	politic regional ue	200	4028
			fonduri europene	3508	
			fondurile structurale	320	
ZiarulFinanciar	business	2000-2017	fondurile structurale	701	1991
			Fonduri europene	1000	
			politica regionala	290	
libertatea	tabloid	2016-2017	fonduri europene	430	546
			fodurile structurale / fond structural	24	
			politic regional europ	92	
OVERALL					11186

Table 7 – Sweden: newspapers and keywords

TITLE	CATEGORY	COVERAGE	SEARCH KEYS	SAMPLED ARTICLES	TOTALS
Svenska Dagbladet	national quality	2005-2017	regionalpolitik eu	102	2017
			strukturfond eu	102	
			regionalpolitiken	242	
			eu fonder	893	
			Sammanhållningspolitiken	7	
			Europeiska investerings	1887	
Aftonbladet	tabloid	2001-2017	regionapolitik eu	0	171
			strukturfonder	34	
			eu fonder	87	
			Sammanhållningspolitiken	1	
			regionalpolitiken	45	
			Europeiska investerings	4	
DagensNyether	regional	1994-2017	regionalpolitiken eu (loose term)	255	2330
			strukturfonder	435	
			Europeiska investerings (exact term)	720	
			eu fonder (loose term)	920	
gefleDagblad	regional	2004-2017	regionalpolitiken eu	39	104
			strukturfonder	24	
			eu fonder	41	
			Europeiska investerings	54	
dt	national quality	2002-2017	eu fonder	61	244
			strukturfonder	53	
			regionalpolitiken eu	39	
			Europeiska investerings	91	
Dagens Industri	business	2010-2017	fond* (eu or europeiska unionen)	990	1801
			investering and (eu or europeiska unionen)	262	
			strukturfond*	54	
			regional* and (eu or europeiska unionen)	495	
OVERALL					6667

PERCEIVE DEL. 5.2: 'DATABASE OF THE TOPICS AND SENTIMENTS TO BE MADE AVAILABLE ON-LINE FOR FURTHER RESEARCH'

Table 8 – Spain: newspapers and keywords

TITLE	TYPE	COVERAGE	SEARCH KEYS	SAMPLED ARTICLES	TOTALS
ElPais	quality national	2001-2017	(politica Regional) OR (politica* de cohesion) AND UE	716	1767
			fondos estructurales	948	
			fondos europeos and regional*	385	
ElMundo	quality national	2000-2017	fondos estructurales	1097	2827
			(politica Regional) OR (politica* de cohesion) AND UE	1116	
			fondos europeos and regional*	982	
Hoy (Extremadura)	regional	2006-2017	fondos estructurales	1257	3954
			fondos europeos regional	1168	
			politica de cohesion	1529	
ElPeriodico (Extremadura)	regional	2001-2017	fondos estructurales	242	1052
			(politica Regional) OR (politica* de cohesion) AND UE	346	
			fondos europeos and regional*	521	
Expansion	business	1995-2017	fondos estructurales	1080	1428
			(politica Regional) OR (politica* de cohesion) AND UE	329	
			fondos europeos and regional*	229	
20 Minutos	tabloid*	2013-2017	fondos estructurales	13	2557
			(politica Regional) OR (politica* de cohesion) AND UE	6	
			fondos europeos and regional*	100	
OVERALL					12064

*spain does not have tabloids in the common sense of the category, but this source (free press) gets close as of editorial format

Table 9 – United Kingdom: newspapers and keywords

TITLE	TYPE	COVERAGE	SEARCH KEYS	SAMPLED ARTICLES	TOTALS
The Times	national quality	2000-2017	(eu or eurpoean union) and fund* and regional	713	1200
			(eu or eurpoean union) and investment* and regional	555	
		1985-2017	eu and structural fund*	217	
the Sun	tabloid	1996-2017	(european union or EU) and (regional policy or regional policies)	163	
			(euro or eu or european) investment* or (euro or eu or european) fund*	368	614
			structural fund* or (structural and eu) or (structural and europ*)	123	
financial Times	business	2000-2017	eu policy or european policy or europolicy	160	
			(eu or eurpoean union) and funds and regional	813	1334
			(european or EU) and (regional policy or regional policies)	170	
			cohesion policy or european cohesion or eu cohesion* or cohesion fund* or (cohesion and eu polic*)	140	
the Echo	regional	2002-2017	structural fund*	410	
			(regional policy or regional policies) or ((policy or policies) and eu) or ((policy or policies) and europe*)	89	186
			(funds and eu) or (funds and europe*)	31	
the Brentwood Gazette	regional	2007-2017	(investment* and europ*) or (investment* and eu)	88	
			(regional policy or regional policies) or ((policy or policies) and eu) or ((policy or policies) and europe*)	197	444
			(funds and eu) or (funds and europe*)	151	
The Guardian	national quality	2000-2017	(investment* and europ*) or (investment* and eu)	146	
			(eu or european union) and (structural fund* or structural investment*)	163	891
			(regional polic* or cohesion polic*) and (eu or european union)	98	
			investment* and region* and (european union)	599	
OVERALL			eu funds or european funds	346	4669

3.4 Twitter

Regarding twitter we planned to sample tweets based on three to six hashtags per country. Indeed we did more, as we wanted to capture both the 'discursive' debate on Cohesion Policy, mostly enacted by citizens, and the 'institutional' debate, mostly generated by institutional profiles. A first choice that we made regards the time span to analyze: Twitter allows users to search and download tweets within a short time span of 7 days. In this vein, a search for tweets through hashtags (#), words and profiles (@) is possible. Older data is available by bundles comprising a maximum of 40.000 tweets or a maximum time span of 40 days at 1.000 \$ per bundle. In this case it is possible to use more logical operators for the research albeit this approach is not suited for a first exploratory analysis, as it could become very costly. We thus decided to keep downloading data in June and July 2017 based on keyword lists.

As for the discursive debate, we used the keyword list that we had developed to define relevant keywords used in all the languages specified. A first test included the translations of "Cohesion" and "European Funds":

Coesion OR Cohesion OR Kohäsionsfonds OR Coeziune OR Spójności OR Sammanhållningsfonden OR Cohesión OR "fondi europei" OR "Europejski Fundusz" OR "European funds" OR "europeiska fonder" OR "Fondul European" OR "Europäischer Fonds" OR "Fondos europeos"

These were the results:

Table 10 - Research string based on the translations of "Cohesion" and "European funds"

keyword	Occurrences (test on 5 th June 2017)
Coesion	95
Cohesion	2007
Kohäsionsfonds	0
Coeziune	4
Spójności	16
Sammanhållningsfonden	0
Cohesión	619
fondi europei	203
Europejski Fundusz	0
European funds	20
europeiska fonder	0
Fondul European	0
Europäischer Fonds	0
Fondos europeos	91

While this research string produced results in Italian, English and Spanish, it was not able to capture tweets in Swedish, Polish, Romanian and German. We then tried a combination of the translations of "European regional development fund", "Cohesion funds", "European social fund", "European structural and investment funds", "Regional Policy", and "convergence". Please note that i) we had to split searches in several steps, as there is a limit in the length of the research string that is accepted by Twitter, and ii) acronyms did not work as, out of context, they produce a large amount of results not related to Cohesion Policy.

Finally, upon analyzing results, we ended up compiling a research string that included the translation of "Regional Policy" as well. This is the final string that we employed for sampling tweets related to the discursive debate in the public sphere:

"politica regionale" OR "polityka regionalna" OR regionalpolitiken OR "politică regională" OR Regionalpolitik OR "Política Regional" OR "regional policy" OR Coesione OR Cohesion OR Kohäsionsfonds OR Coeziune OR Spójności OR Sammanhållningsfonden OR Cohesión OR "fondi europei" OR "Europejski Fundusz" OR "European funds" OR "europeiska fonder" OR "Fondul European" OR "Europäischer Fonds" OR "Fondos europeos"

As for the institutional debate, we started by compiling a list of relevant hashtags and accounts:

#interreg OR #CentralEurope OR #ce2020launch OR #CEcommunity OR @EU_Regional OR @EU_CoR OR @RegioEvaluation OR #EUinmyRegion OR #CohesionPolicy OR #euregions OR #EURegionsWeek OR #Regio

We discarded the idea of using words without the symbol for the hashtag, as the underscore was not properly recognized by the research tool, and several Portuguese tweets using the pronoun EU were added to the results. We obtained these results:

Table 11 - Relevant hashtags and accounts: results

keywords	occurrences
#interreg	895
#CentralEurope	26
#ce2020launch	0
#CEcommunity	0
@EU_Regional	1831
@EU_CoR	956
@RegioEvaluation	5
#EUinmyRegion	1344
#CohesionPolicy	1482
#euregions	331
#EURegionsWeek	50

We thus considered a larger list of both hashtags and profiles that we created both theoretically, by adding relevant keywords and institutions, and empirically, by performing a snowball addition to our research string of relevant items that were numerous in the sample originally downloaded. The following list reports hashtags and profiles considered.

Table 12 - Larger list of hashtags and profiles

hashtags	profiles
#SmartRegions	@RegioSweden
#RomaEU	@RegioPoland
#Regiostars	@RegioInterreg
#investEU	@RegioEvaluation
#InterregYouth	@OpenCoesione
#interregvolunteers	@InterregYouth
#interregRE	@INTERRECTweets
#ilovetheeu	@interregeurope
#FutureOfEurope	@europeanregions
#FutureCohesion	@EU_Social
#funduszeUE	@EU_Regional
#funds	@EU_Growth
#fondiUE	@CoR_President
#Fondieuropei	@AgenziaCoesione
#euregions	
#EUFunds	
#EUBudget	
#ESF	
#ERDF	
#efrd	
#DGregio	

Finally, due to the limits imposed to the length of the research string, we had to select only the most relevant. This is the final research string for the institutional debate, originated by institutional accounts or around technical keywords:

```
#InterregYouth OR #interreg OR #EUinmyRegion OR #CohesionPolicy OR
#euregions OR #ilovetheeu OR #FutureOfEurope OR #FutureCohesion OR
#euregions OR #EUFunds OR #ESF OR #ERDF OR #DGregio OR @EU_Regional OR
@EU_CoR OR @RegioEvaluation OR @RegioInterreg OR @INTERRECTweets OR
@interregeurope OR @europeanregions OR @CoR_President OR #NUTS2 OR
#interact OR #ESPON OR #ESPON2020 OR #URBACT
```

We used both research strings to download data twice a week during June and July 2017. Overall, we collected 42.778 tweets with the “technical” string, based on hashtags and profiles, and 104.895 tweets with the “discursive” string, composed of the keywords’ translations. Some descriptive analysis on these data follows.

3.4.1 Institutional string

Out of the 42.778 tweets captured, 13.341 were original tweets, while 29.387 were retweets. The average number of retweets for each tweet was 17,55, with a median of 5. Figures 1, 2, and 3 highlight the usernames most present in our sample - either as producer of a tweet or because they were mentioned by someone else - and the geographic distribution of the tweets collected.

Figure 1 - Usernames most present in the sample created through the institutional string

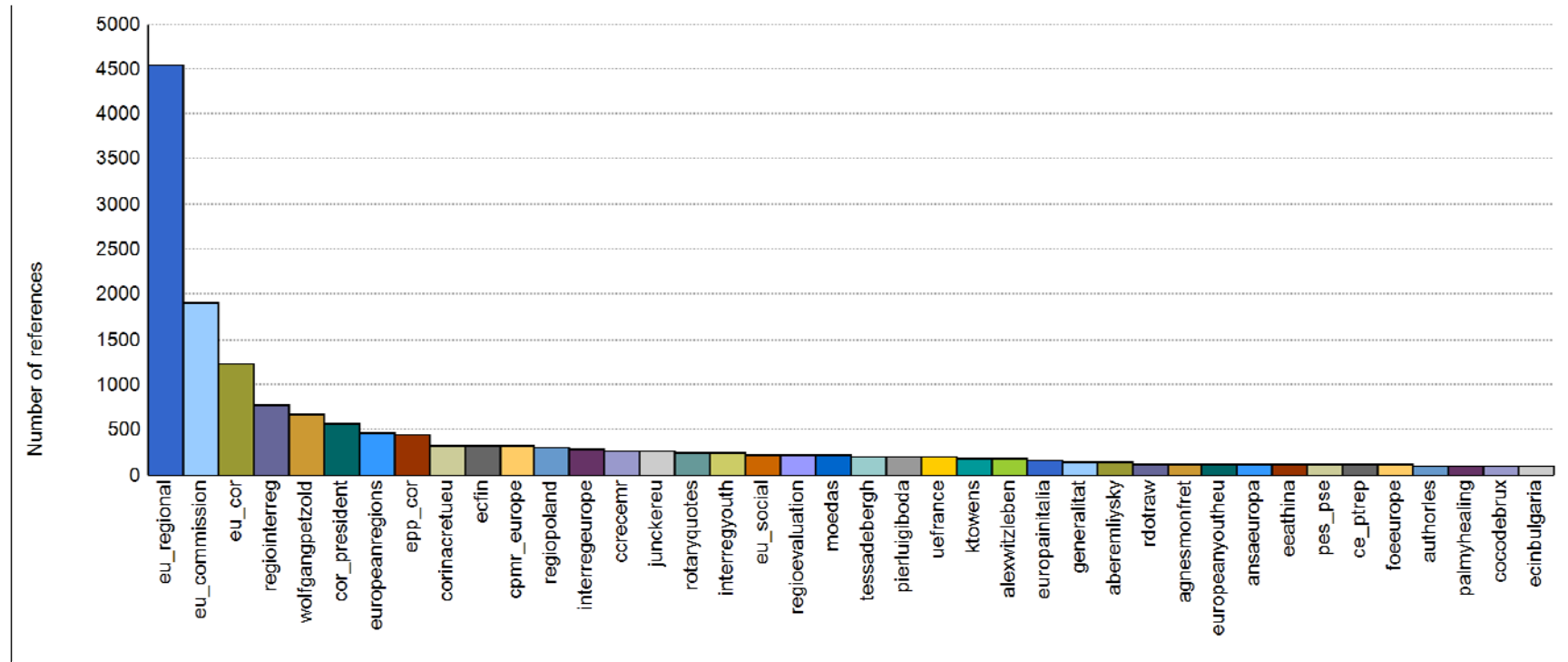


Figure 2 - Geographic distribution of the tweets created through the institutional string.

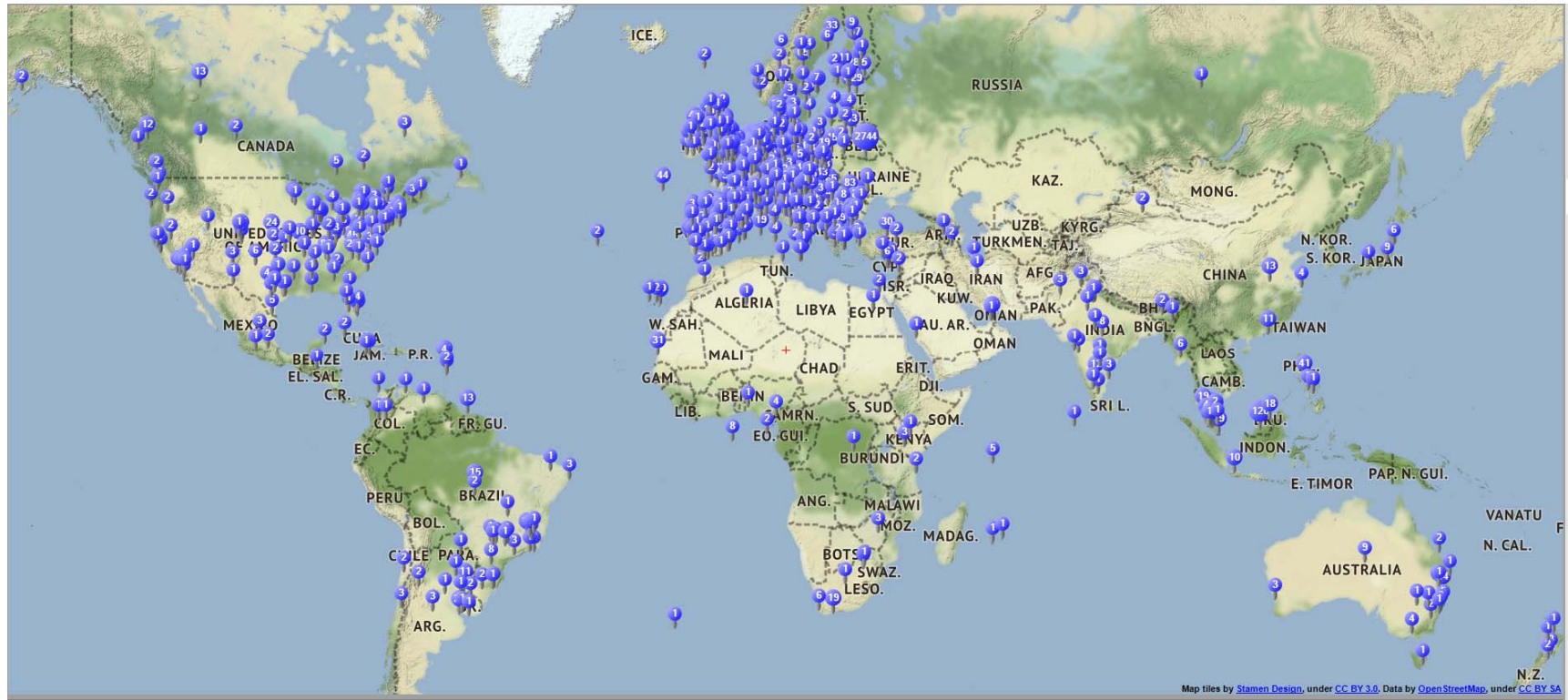


Figure 3 - Geographic distribution of the tweets created through the institutional string: detailed view.

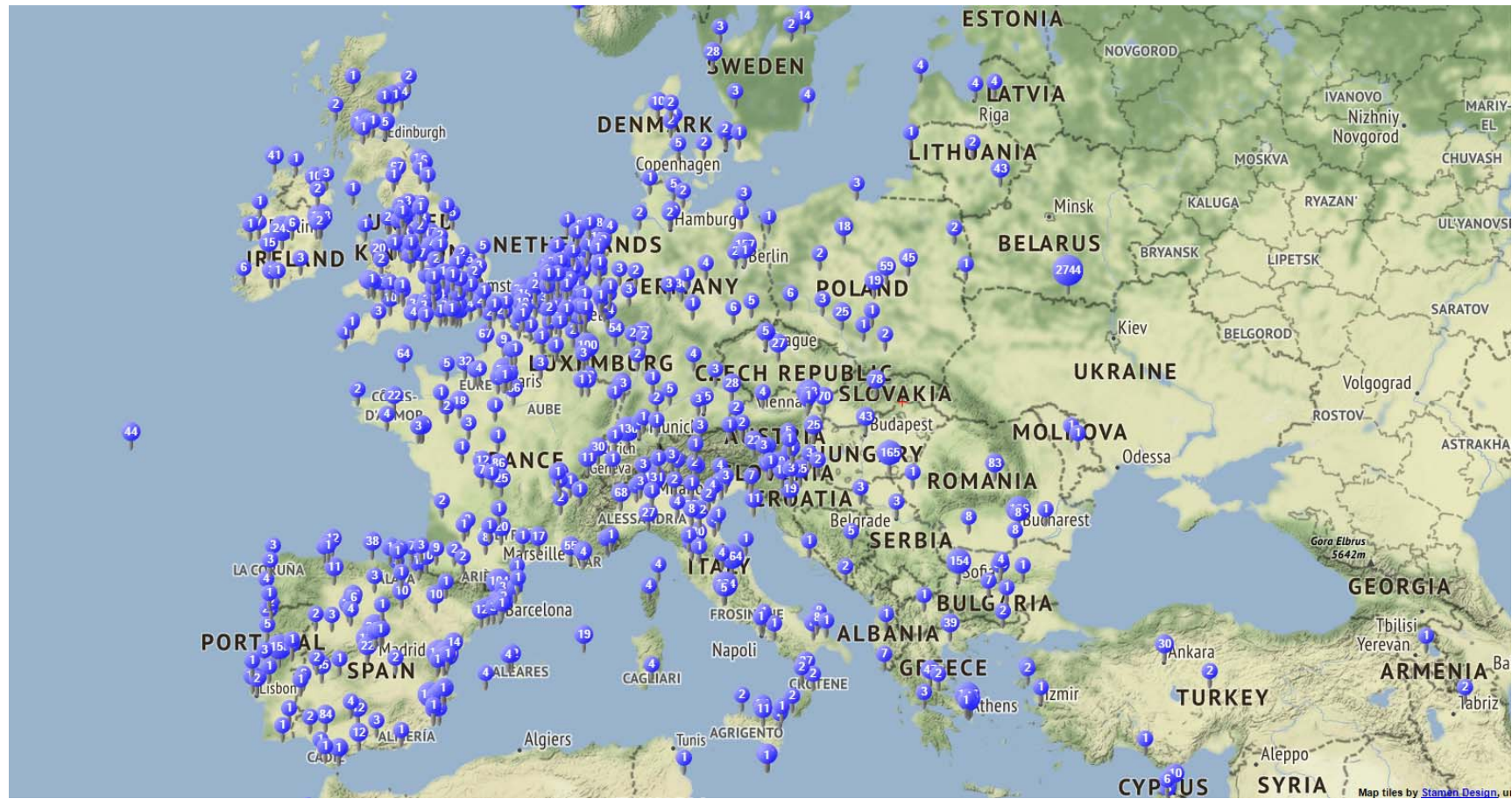


Figure 4 - Usernames most present in the sample created through the discursive string

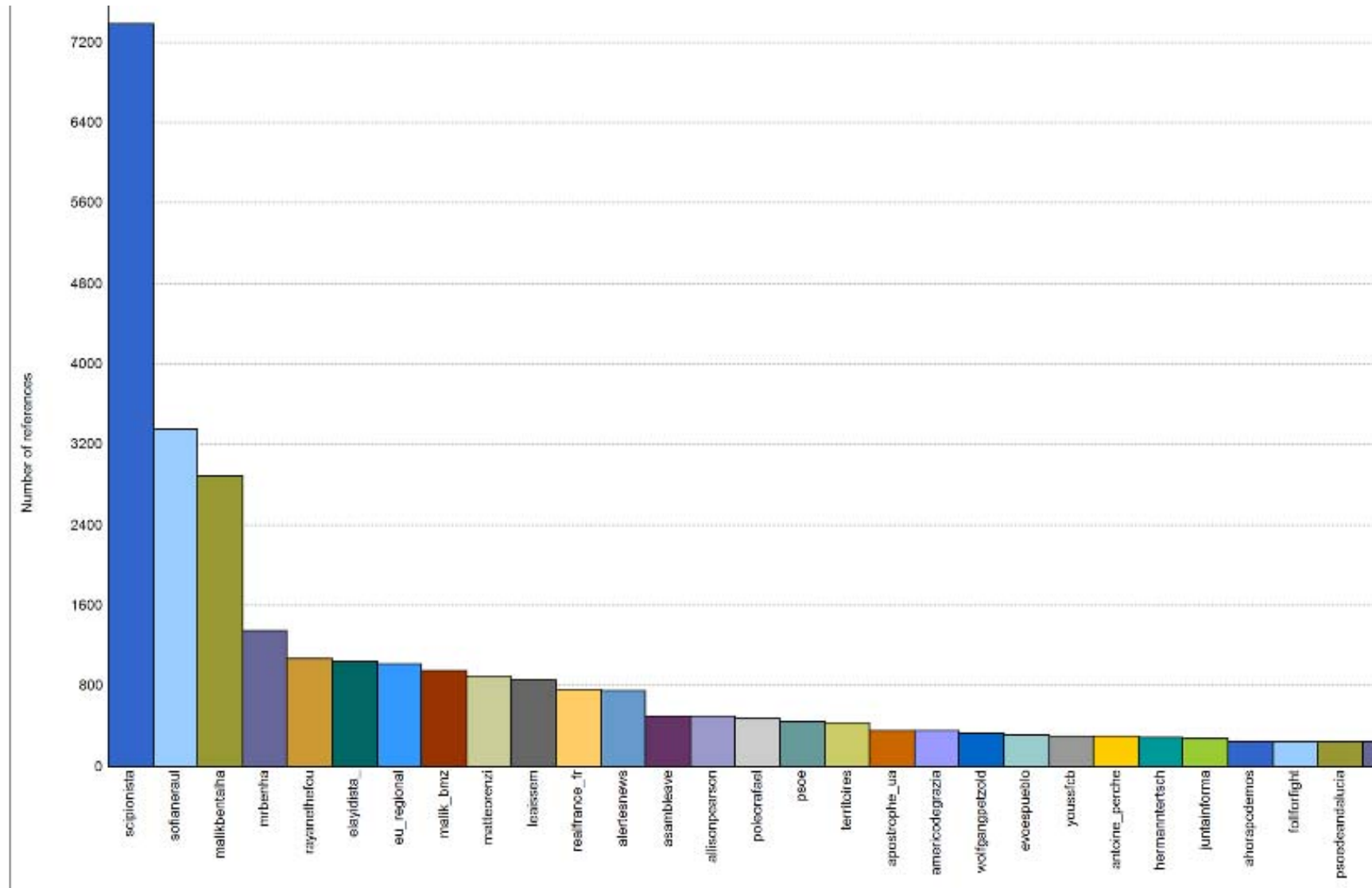


Figure 5 - Geographic distribution of the tweets created through the discursive string.



3.4.2 Discursive string

Out of the 104.895 tweets captured, 31.649 were original tweets, while 73.245 were retweets. The average number of retweets for each tweet was 1.060,8, with a median of 10. Figures 4 and 5 highlight the usernames most present in the sample created through the discursive string - either as producer of a tweet or because they were mentioned by someone else - and the geographic distribution of the tweets collected.

3.5 Facebook

The last set of data was aimed at capturing a debate in the public sphere that is constituted by pieces of text longer than tweets. Originally, we planned to download posts from two blogs per country: a right-wing and a left-wing one. Yet, when we searched for actual data, we noticed that this sort of debate had moved from blogs to Facebook in the last few years. Blogs' relevance for the online debate has indeed faded and it is not possible to find relevant up-to-date blogs dealing with Cohesion Policy. We thus developed a Facebook-based strategy of data collection: i) together with PERCEIVE's partners, we selected the Facebook profile for each Local Managing Authority; ii) then we asked each partner to identify at least two leading figures (e.g. politicians or journalists), who tackle the debate on CP in their country/region, in order to download posts and comments on their profile too; iii) finally we searched for relevant institutional Facebook profiles.

Table 13 specifies all the profiles we searched for and the details of posts and comments that we downloaded. Overall, we collected 56.167 posts and 504.744 comments from 26 Facebook profiles. While analyzing data, we noticed two issues:

- Facebook is not used in the same way by all the LMAs. While several LMAs have a Facebook profile devoted to Cohesion Policy especially, others only have a generic Facebook profile (as already noted in Deliverable 3.1), while the LMA in the UK has no Facebook profile at all.
- Facebook is not used in the same way by all politicians, and the variance in this case is greater when compared to the variance regarding LMAs. Some politicians, such as Matteo Salvini from the Italian party Northern League, massively use this channel of communication, whereas others, such as most of the Polish politicians, use Facebook for private reasons with their page unavailable for the public (and hence not downloadable). The case of Salvini is extreme as, with 299 posts, written in a couple of months, he collected almost 400.000 comments, thus reaching the maximum amount of downloadable data.

Consequently, in order to deal with data that are comparable both regarding the source and the dimension, we decided to focus only on posts and comments from LMAs' Facebook profiles.

Table 13 - Facebook profiles consulted and downloaded

	Region/Actor	link	download	captured post	captured comments
Italy	Emilia-Romagna	https://www.facebook.com/RegioneEmiliaRomagna/	yes	3.379	5.210
	Calabria	https://www.facebook.com/PorCalabria/	yes	428	339
	Fabrizio Barca - politician	https://www.facebook.com/fabriziobarca/	yes	2.195	7.948
	Matteo Salvini - politician	https://www.facebook.com/salviniofficial/	yes	299	391.451
Austria	Burgenland	https://www.facebook.com/rmbgmbh/	yes	578	68
	Christian Illedits – politician	https://www.facebook.com/illedits/	yes	1.105	605
	Hans Niessl – politician	https://www.facebook.com/hansniessl/	yes	4.897	14.602
Poland	Warmińsko-mazurskie	https://www.facebook.com/pg/RPO.Warmia.Mazury	yes	1.777	8.319
	Dolnośląskie	https://www.facebook.com/RPOWD	yes	831	117
	Tadeusz Iwiński – politician	https://www.facebook.com/tadeusziwinski/	yes	4.860	2.553
	Zbigniew Babalski – politician	https://www.facebook.com/zbigniew.babalski	not downloadable - private profile		
	Jerzy Szmit – politician	https://www.facebook.com/jerzy.szmit	not downloadable - private profile		
	Artur Chojecki - politician	https://www.facebook.com/artur.chojecki.568	not downloadable - private profile		
	Dawid Jackiewicz - politician	https://www.facebook.com/dawid.jackiewicz	not downloadable - private profile		
Romania	Regional level: Agentia pentru Dezvoltare Regionala Sud-Est	https://www.facebook.com/adrse.ro	yes	551	22
	National level: Ministerul Dezvoltarii Regionale, Administratiei Publice si Fondurilor Europene	https://www.facebook.com/MinisterulDezvoltarii	yes	4.687	1.339
	National level: Ministerul Fondurilor Europene	https://www.facebook.com/Fonduri-Europene-Structurale-%C8%99i-de-Investi%C8%9Bii-Rom%C3%A2nia-581612591868149/	yes	1.634	1.143
	Cristian GHINEA - Former Minister of EU funds	https://www.facebook.com/Cristian.Ghinea.Oficial/	yes	2.942	41.247
	Eugen TEODOROVICI - Former Minister of EU funds	https://www.facebook.com/EugenOrlandoTeodorovici/	yes	2.043	14.115
	Marian Jean MARINESCU - politician	https://www.facebook.com/marianjean.marinescu/	not downloadable - private profile		
Sweden	Tillväxtverket	https://sv-se.facebook.com/Tillvaxtverket/	yes	540	380
	Cecilia Malmström - politician	https://www.facebook.com/MalmstromEU/	yes	770	3.706
	Catarina Segersten Larsson – politician	https://www.facebook.com/public/Catarina-Segersten-Larsson	not downloadable - private profile		
	Sofia Jarl – politician	https://twitter.com/sofiajarl/media	not downloadable - private profile		

PERCEIVE DEL. 5.2: 'DATABASE OF THE TOPICS AND SENTIMENTS TO BE MADE AVAILABLE ON-LINE FOR FURTHER RESEARCH'

Spain	Junta de Extremadura	https://www.facebook.com/JuntaDeExtremadura	yes	16.134	4.265
	Juan José Ventura – journalist	http://es-la.facebook.com/juanjose.venturafernandez	not downloadable - private profile		
	Antonio Sáez – journalist	no facebook			
UK	No official facebook page for local and central LMA				
	Sajid Javid - Secretary of State for Communities & Local Government	https://www.facebook.com/sajidjavidbromsgrove/	yes	1.288	2.944
	Hazel Blears - politician	https://www.facebook.com/profile.php?id=100011395978393	not downloadable - private profile		
	John Denham - politician	https://www.facebook.com/johndenhammp	not downloadable - private profile		
	Eric Pickles - politician	https://www.facebook.com/eric.pickles/	yes	440	1.938
Institutional	EU Regional Policy	https://www.facebook.com/EUinmyregion	yes	970	1.027
	Interreg Europe	https://www.facebook.com/interregeurope/	yes	328	217
	Assembly of European Regions	https://www.facebook.com/EuropeanRegions/	yes	1.854	258
	European Committee of Regions	https://www.facebook.com/European.Committee.of.the.Regions/	yes	920	562
	Interreg Central Europe	https://www.facebook.com/InterregCE/#	yes	681	226
	Perceive Project	https://www.facebook.com/perceiveproject/	yes	36	143
				56.167	504.744

4. Analysis: Topic Modeling

We used topic modeling (Blei et al., 2003) to analyze our data. This technique provides an automated way to code the content of a corpus of texts into a set of 'topics' that are containers of meaningful words (Mohr and Bogdanov, 2013), with the latter co-occurring frequently. Topic modeling combines four important features. First, it can analyze bodies of texts that would be impossible for a human being to deal with because of their volume or extent. Second, once topics are automatically produced, they need to be interpreted – and topic modeling does not require the imposition of a priori-categories. The third relevant feature is that topic modeling categorizes words, not documents. It allows for variations in the meaning of terms in different contexts, and recognizes that the meaning of a word depends on the surrounding words. The fourth element is that topics are explicit and other researchers may reproduce the analysis, which improves reliability (DiMaggio et al., 2013). Overall, topic modeling complements systematic analysis with an inductive approach.

The most diffused implementation of topic modeling uses an algorithm called latent Dirichlet allocation (LDA) (Blei et al., 2003). LDA is based on Bayesian statistics and allows the development of topics in a completely automated way. Researchers make two decisions before running the model: the number of topics the model should produce, and whether topics should contain an equal number of words or not. Based on these parameters, the model provides the probabilities of words being used in a topic, as well as an account of the distribution of those topics across the corpus of texts. To put it more simply, the model places together terms that appear in the same texts more frequently than one would expect by chance. The idea is that each word of the corpus is assigned to a topic in an iterative process.

To perform topic modeling, we used Mallet - an open-source software developed by the University of Massachusetts Amherst working through command line in MS-DOS. The specific results of our analysis will be provided starting with Deliverable 5.3, but we present an example of the results here.

For each language we developed a so called stop-word list, which is a list of words the software will ignore, usually composed by articles, prepositions, adverbs, and other words with scarce substantive meaning. Then, it is possible to produce output grouping words in a number of topics requested by the researcher. Each model produces three main outputs:

1. *a list of words per topic* displaying the highest-ranked terms for each topic, where the prevalence of each word within a topic is adjusted for its prevalence within the corpus as a whole;
2. *a list describing how each word has been coded in each text analyzed.* Thanks to this list it is possible to distinguish the coding of each word within the text and therefore achieve a deep understanding of each topic; and
3. *a breakdown of the topics comprising each paper.* Thanks to this output it is possible to observe the composition of each source analyzed and calculate the prevalence of each topic throughout the sample.

The following table (14) shows a list of the 20 most important words for a 20 topic-solution that we developed by analyzing English newspapers' articles. In this case, our sample consisted of 4.669 articles; in particular, we analyzed:

- 1.200 articles published by *The Times*, and 891 articles by *The Guardian*, that are the “national quality” newspapers selected regarding UK.
- 1.334 articles published by *The Financial Times*, that is the “business” newspaper that we selected in the UK case.
- 614 articles published by *The Sun*, that is the “tabloid” newspaper selected in reference to the UK case.
- 186 articles published by *The Echo*, and 444 articles published by *The Brentwood Gazette*, that are the “regional” newspapers selected in reference to the UK case.

Newspapers differ according to the time span covered by their electronic database, as well as according to the searching tools available: for details regarding the time span covered by our sample, and the keywords used for the actual search, please refer to Table 9.

After removing all the words pertaining to the stop-word list, we ended up with a sample of 2.353.446 words. The longest article is composed by 7.091 words. Although this document is not aimed at analyzing topics in detail, we can depict few general characteristics of the model elicited. First of all, we can deal with the topics' meanings by analyzing the most important words. In the example at hand, we can highlight that topic 8 especially focuses on matters of budgets and funds, discussed within the European Commission. Topic 9, instead, is focused specifically on Brexit and effects of the vote of the connected referendum. We find names of the parties, here, together with those of the politicians that had a particularly prominent role during the referendum. By associating words such as “Greece”, “Greek”, “debt”, “financial” and “crisis”, topic 13 clearly evokes the debate on the Greek financial crisis. On the other hand, topic 12 deals with international diplomatic tensions. Here, words such as “security”, “war”, “border”, “Russia” and “Ukraine” delineate a debate on borders' issues. By focusing on words such as “transport”, “airport”, “city”, “car”, “water”, “rail”, “road”, “energy”, and “power”, topic 6 can clearly be connected to themes of energy and mobility.

In order to refine our analysis, we can then focus on articles that are mostly coded to a certain topic. For example, we can identify the article that is mostly coded to topic 9, that refers to Brexit, which is composed of 70,1% by topic 9. The title of the article is “Liam Fox 'nuts' jibe” and was published by *The Sun* on 17/8/2016. On the contrary, the article that is mostly coded at Topic 8, that refers to the EU budget, is composed of 91% by this topic. It is an article by George Parker, published in the *Financial Times* and titled “Brown gives EU rebate ultimatum”. These examples suggest how topic collection can efficiently guide discourse analysis.

Table 14 – Most important words for each topic

topic	0	1	2	3	4	5	6	7	8	9
20 most important words	council	trade	growth	party	call	Europe	energy	services	european	labour
	local	world	economy	minister	hall	European	city	director	commission	brexit
	essex	countries	markets	french	club	Policy	airport	education	budget	government
	school	china	year	government	members	Time	building	executive	states	cameron
	brentwood	global	market	france	church	Economic	transport	chief	funds	party
	people	international	bank	european	meeting	Union	project	community	countries	vote
	town	africa	rate	political	road	Political	car	john	member	referendum
	county	chinese	prices	prime	saturday	Make	water	service	brussels	britain
	residents	asia	inflation	president	brentwood	National	rail	professor	eu's	scotland
	southend	president	rates	europe	details	Future	road	london	europe	scottish
	years	india	month	election	group	Market	infrastructure	david	britain	david
	work	economic	brexit	german	information	Country	built	officer	aid	leave
	support	investment	data	germany	village	Made	years	health	regional	corbyn
	community	development	manufacturing	elections	visit	Government	year	university	union	secretary
	schools	foreign	economic	treaty	open	Single	industry	industry	money	minister
	children	asian	june	spain	meet	System	construction	school	euros	people
	councillor	nations	sector	parliament	ingatestone	Change	power	royal	structural	mps
make	south	european	leader	place	Clear	projects	head	year	election	
green	american	expected	voters	sunday	Power	investment	people	germany	speech	
high	japan	eurozone	poland	christmas	Important	land	chairman	enlargement	campaign	

topic	10	11	12	13	14	15	16	17	18	19
20 most important words	business	ireland	russia	greece	million	Lab	block-time	investment	tax	people
	funding	irish	turkey	greek	cent	Vote	published-time	companies	government	it's
	regional	food	security	eurozone	pounds	Ukip	bst	funds	public	time
	university	farmers	war	debt	group	Party	gmt	market	cent	years
	government	dublin	military	bank	billion	Total	updated-timeupdated	cent	year	don't
	research	ireland's	people	crisis	company	Election	hash	investors	people	back
	regions	oil	international	government	year	Majority	photograph	capital	pay	good
	development	northern	foreign	european	profits	Seats	itâ€™s	fund	spending	world
	london	fishing	police	banks	sales	Labour	today	financial	labour	work
	businesses	minister	president	cent	page	Electorate	morning	bank	years	life
	local	european	russian	euro	reported	Turnout	meeting	banks	work	day
	students	court	ukraine	finance	expected	Lib	april	year	budget	year
	education	government	border	financial	shares	Elected	thatâ€™s	business	jobs	lot

PERCEIVE DEL. 5.2: 'DATABASE OF THE TOPICS AND SENTIMENTS TO BE MADE AVAILABLE ON-LINE FOR FURTHER RESEARCH'

jobs	fish	western	athens	executive	Swing	deal	markets	cut	home
investment	newspapers	turkish	growth	euro	Member	march	european	rate	that's
universities	times	country	bailout	business	Votes	minister	private	britain	man
growth	fail	israel	economic	deal	Seat	with:block	europe	national	big
economic	fine	syria	fiscal	british	Hold	february	foreign	income	week
areas	pounds	iraq	monetary	bank	Maj	hereâ€™s	equity	local	things
city	olive	nato	economy	market	Tories	time	years	health	british

5. Conclusion

The document at hand has shown the sources employed to collect data and used to build our database of topics. We refined and improved the originally planned data collection strategy and collected texts from five main sources. First, at the level of EU communication, we collected all issues of the Panorama magazine and all available brochures. Second, at the level of the financed projects' abstracts, we collected all descriptions of the 'flagship' projects at the projects' pages of the EC INFOREGIO web portal. The third level regards newspapers and local press. Here, based on a keywords list that we developed and translated in all the case studies' languages thanks to the projects partners, we downloaded articles from 14 national 'quality' newspapers, 7 national 'tabloids', 7 national 'business' newspapers and 13 regional newspapers. Fourth, we collected communication on twitter based on two research strings that we developed. The first, that we called 'institutional', is aimed at capturing the centrally-launched debate, whereas the second, that we called 'discursive', is aimed at capturing the debate launched by citizens. Fifth, we collected posts and comments from 26 Facebook profiles, including LMAs and relevant politicians. Indeed, we collected texts of different kinds in order to analyze the public discourse on regional policies and structural funds: we collected texts from traditional sources, such as newspapers, but investigated new media as well. Here we combined means of communication such as Twitter, in which messages are conveyed in a compressed string of characters, and Facebook, which is populated by longer lines of text and provides more articulated accounts. The population of texts that we collected is appropriate to scrutinize the public debate under different angles, as we captured both technical and not necessarily technical texts. These data are now informing tasks 5.3 and 5.4, where we are analyzing them in order to construct a database of discursive topics. Finally, we gave a hint of what the output of topic modeling is. Technically, a topic is a cluster of frequently co-occurring words. However, these clusters provide rich insights on how the public debate is articulated, what the issues are that stimulate the debate and how these issues are defended or challenged. In the appendix we provide a list of relevant keywords translated in all the languages of the Perceive project.

6. References

- Blei, D., Ng, A., & Jordan, M (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- DiMaggio, P., Nag, M. & Blei, D (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of US government arts funding. *Poetics*, 41(6), 570-606.
- Mohr, J. & Bogdanov, P. (2013). Introduction, topic models: What they are and why they matter. *Poetics*, 41(6), 545-569.

7. Appendix

7.1 Keywords in English

KEYWORDS in ENGLISH
Regional Policy (generic vocabulary)
regional policy
cohesion policy
European solidarity
Competitiveness of SMEs
Sustainable development
2020 strategy
social exclusion
investment policy instrument
public investment
national investment
national-EU-budget
EU-investment
EU-budget
financing
Co-financing rate
national program
regional program
support
promotion
investment
restructuring
construction
building
infrastructure
development
innovation
start-up
sustainability

Thematic areas (2014-2020)
Research and Innovation
Information and Communication Technology
Competitiveness of SMEs
Low-Carbon Economy
Climate Change Adaptation & Risk Prevention
Environment Protection and Resource Efficiency
Network Infrastructures in Transport and Energy
Sustainable and Quality Employment
Social Inclusion
Educational and Vocational Training
Efficient Public Administration
Outermost and Sparsely Populated
Additional themes RP
Urban development
Rural development
Northern Ireland: the peace programme
Candidate countries (Enlargement of the EU)
EU Funds (see Regional Policy web portal)
European regional development fund (ERDF)
Cohesion fund (CF)
European social fund (ESF)
European agricultural fund for rural development (EAFRD)
European maritime and fisheries fund (EMFF)
European structural and investment (ESI) funds
5 targets 2020 strategy
Employment
Research and Development
Climate change and energy sustainability
Education
Fighting poverty and social exclusion
Cooperation

interregional cooperation program
territorial cooperation
transnational cooperation
macro-regional strategies
Danube region strategy
Baltic sea region strategy
alpine region strategy
Adriatic and Ionian sea strategy
URBACT
INTERACT (2013 is II)
INTERACT (2013 is II)
INTERREG (2013 is IV C)
INTERREG EUROPE
ESPO
European Spatial Planning Observation Network
Regions for economic change
REGIOSTARS award
EU Identity
being European
Europeans
"NUTS2ness"
"NATIONALness"
identity
identification
european identity
national identity
regional identity
NUTS 2
NATION
culture + NUTS2
national culture
regional culture
culture + NUTS2
culture + NATION
country

nation
region
history + NUTS2
symbol + NUTS2
tradition + NUTS2
landmark + NUTS2
heritage + NUTS2

7.2 Keywords in Italian

KEYWORDS in ENGLISH	KEYWORDS in ITALIAN	COMMENTS/SOURCE
Regional Policy (generic vocabulary)		
regional policy	Politica Regionale	Direct translation.
cohesion policy	Politica di Coesione	Direct translation.
European solidarity	solidarietà europea	Direct translation.
Competitiveness of SMEs	Competitività delle PMI	Direct translation.
Sustainable development	Sviluppo sostenibile	Direct translation.
2020 strategy	strategia Europa 2020	Direct translation with article. Also referred to as <i>Europa 2020</i>
social exclusion	esclusione sociale	Direct translation.
investment policy instrument	strumento della politica di investimento	"Politica di investimento" is the definition of Regional Policy in inforegio website
public investment	investimenti pubblici	Direct translation.
national investment	investimenti nazionali	Direct translation.
national-EU-budget	EU-budget nazionale	Direct translation.
EU-investment	EU investimenti	Direct translation.
EU-budget	EU-budget	Direct translation.
financing	finanziamento	Direct translation.
Co-financing rate	tasso di cofinanziamento	Direct translation.
national program	programma nazionale	Direct translation.
regional program	programma regionale	Direct translation.
support	supporto	Direct translation.
promotion	promozione	Direct translation.
investment	investimento	Direct translation.

PERCEIVE DEL. 5.2: 'DATABASE OF THE TOPICS AND SENTIMENTS TO BE MADE AVAILABLE ON-LINE FOR FURTHER RESEARCH'

restructuring	ristrutturazione	Direct translation.
construction	costruzione	Direct translation.
building	costruzione	Direct translation.
infrastructure	infrastruttura	Direct translation.
development	sviluppo	Direct translation.
innovation	innovazione	Direct translation.
start-up	startup start-up	
sustainability	sostenibilità	Direct translation.
Thematic areas (2014-2020)		Words in this set translated based on http://www.agenziacoesione.gov.it/it/politiche_e_attivita/programmazione_2014-2020/Politica_di_Coesione/Politica_di_Coesione.html#accept
Research and Innovation	ricerca, sviluppo tecnologico e innovazione	
Information and Communication Technology	tecnologie dell'informazione e della comunicazione	
Competitiveness of SMEs	competitività delle piccola e media imprese (PMI)	
Low-Carbon Economy	(transizione verso un) economia a basse emissioni di carbonio	
Climate Change Adaptation & Risk Prevention	adattamento a cambio climatico, prevenzione e gestione di rischi	
Environment Protection and Resource Efficiency	tutelare l'ambiente e promuovere l'uso efficiente delle risorse	
Network Infrastructures in Transport and Energy	promuovere sistemi di trasporto sostenibili e eliminare le strozzature nelle principali infrastrutture di rete	
Sustainable and Quality Employment	promuovere l'occupazione e sostenere la mobilità dei lavoratori	
Social Inclusion	promuovere l'inclusione sociale e combattere la povertà	
Educational and Vocational Training	Investire nelle competenze, nell'istruzione e nell'apprendimento permanente	
Efficient Public Administration	Rafforzare la capacità istituzionale e promuovere un'amministrazione pubblica efficiente	
Outermost and Sparsely Populated	le "Aree Interne" aree caratterizzate dal fenomeno del declino demografico, da	

	condizioni di disagio socio-economico, dall'esistenza di una distanza significativa dai centri di offerta dei servizi pubblici essenziali	
Additional themes RP		
Urban development	sviluppo urbano	Direct translation.
Rural development	sviluppo rurale	Direct translation.
Northern Ireland: the peace programme	Irlanda del Nord: il programma Peace	http://ec.europa.eu/regional_policy/it/policy/themes/northern-ireland-peace-programme/
Candidate countries (Enlargement of the EU)	Paesi candidati (allargamento EU)	Direct translation.
EU Funds (see Regional Policy web portal)	Fondi EU	Words in this set translated based on http://www.agenziacoesione.gov.it/it/politiche_e_attivita/programmazione_2014-2020/Politica_di_Coesione/Politica_di_Coesione.html#accept
European regional development fund (ERDF)	Fondo Europeo di Sviluppo Regionale (FESR)	
Cohesion fund (CF)	Fondo di Coesione (FC)	
European social fund (ESF)	Fondo Sociale Europeo (FSE)	
European agricultural fund for rural development (EAFRD)	Fondo Europeo Agricolo per lo Sviluppo Rurale (FEASR)	
European maritime and fisheries fund (EMFF)	Fondo Europeo per gli Affari Marittimi e la Pesca (FEAMP)	
European structural and investment (ESI) funds	Fondi strutturali e di investimento europei (SIE)	
5 targets 2020 strategy		Words in this set translated based on http://ec.europa.eu/europe2020/targets/eu-targets/index_it.htm
Employment	occupazione	
Research and Development	R&S / innovazione	-
Climate change and energy sustainability	Cambiamenti climatici /energia	
Education	Istruzione	
Fighting poverty and social exclusion	Povertà / emarginazione	
Cooperation	Cooperazione	
interregional cooperation program	programma di cooperazione	http://ec.europa.eu/regional_policy/sources/

PERCEIVE DEL. 5.2: 'DATABASE OF THE TOPICS AND SENTIMENTS TO BE MADE AVAILABLE ON-LINE FOR FURTHER RESEARCH'

	interregionale	cooperate/macro_region_strategy/pdf/report_implem_macro_region_strategy_it.pdf
territorial cooperation	cooperazione territoriale	same source as above
transnational cooperation	cooperazione transnazionale	same source as above
macro-regional strategies	strategie macroregionali	same source as above
Danube region strategy	strategia dell'UE per la regione del Danubio (EUSDR)	same source as above
Baltic sea region strategy	strategia dell'UE per la regione del Mar Baltico (EUSBSR)	same source as above
alpine region strategy	strategia dell'UE per la regione alpina (EUSALP)	same source as above
Adriatic and Ionian sea strategy	la strategia dell'UE per la regione adriatica e ionica (EUSAIR)	same source as above
URBACT	URBACT	http://urbact.eu/urbact-italia
INTERACT (2013 is II)	INTERACT II	
INTERACT (2013 is II)	INTERACT II	
INTERREG (2013 is IV C)	INTERREG IVC	
INTERREG EUROPE	INTERREG EUROPE	
ESPON	ORATE	http://ec.europa.eu/regional_policy/it/policy/what/glossary/e/espon
European Spatial Planning Observation Network	ORATE (Osservatorio in rete dell'assetto del territorio europeo)	http://ec.europa.eu/regional_policy/it/policy/what/glossary/e/espon
Regions for economic change	Regioni per il Cambiamento economico	https://enrd.ec.europa.eu/networks-and-networking/research-initiatives/research-initiatives-regions-economic-change_it
REGIOSTARS award	Premi REGIOSTARS	http://ec.europa.eu/regional_policy/it/newsroom/news/2017/02/15-02-2017-launch-of-the-regiostars-awards-2017
EU Identity	identità europea	
being European	essere europeo	Direct translation, does not sound good
Europeannes	europea	Direct translation, does not sound good
"NUTS2ness"	emilianità/calabresità	Direct translation.
"NATIONALness"	italianità	Direct translation.
identity	identità	Direct translation.
identification	identificazione	Direct translation.
european identity	identità europea	Direct translation.

national identity	identità italiana	Direct translation.
regional identity	identità regionale	Direct translation.
NUTS 2	NUTS 2	
NATION	Italia	Direct translation.
culture + NUTS2	cultura emiliana/calabrese	Direct translation.
national culture	cultura italiana	Direct translation.
regional culture	cultura regionale	Direct translation.
culture + NUTS2	cultura + NUTS2	Direct translation.
culture + NATION	cultura + NUTS 2	Direct translation.
country	Stato	Direct translation.
nation	Nazione	Direct translation.
region	Regione	Direct translation.
history + NUTS2	storia + NUTS2	Direct translation.
symbol + NUTS2	simbologia (?) + NUTS2	Direct translation.
tradition + NUTS2	tradizione + NUTS2	Direct translation.
landmark + NUTS2	luogo storico + NUTS2; monumento + NUTS2	Direct translation.
heritage + NUTS2	patrimonio/eredità culturale + NUTS2	Direct translation.

7.3 Keywords in German

KEYWORDS in ENGLISH	KEYWORDS in NATIONAL LANGUAGE	COMMENTS
Regional Policy (generic vocabulary)		no proper equivalent found on Inforegio (switching from EN to DE and vice versa)
regional policy	Regionalpolitik	
cohesion policy	Kohäsionspolitik	
European solidarity	EU-Solidarität, europäische Solidarität	
Competitiveness of SMEs	Wettbewerbsfähigkeit von KMU, Wettbewerbsfähigkeit der KMU	English version: +SME competitiveness
Sustainable development	nachhaltige Entwicklung	
2020 strategy	Strategie Europa 2020	English version: Europe 2020 strategy
social exclusion	soziale Ausgrenzung	

investment policy instrument	Investitionsinstrument	
public investment	öffentliche Investitionen	
national investment	nationale Investitionen	
national-EU-budget	nationales EU-Budget	
EU-investment	EU-Investition, EU-Investitionen	probably used in the plural in German
EU-budget	Haushalt der EU, EU-Haushalt, EU-Gelder	
financing	Finanzierung, finanzieren	English version: funding or financing
Co-financing rate	Kofinanzierungssatz, Kofinanzierungssätze, Kofinanzierungsanteil	
national program	nationales Programm	English version: national programme
regional program	regionales Programm	English version: regional programme
support	Unterstützung, unterstützen, Förderung, fördern	'Förderung' might indicate more of a financial aid
promotion	Wertschätzung, Begünstigung, Förderung, wertschätzen, begünstigen, fördern	
investment	Investition, Investitionen	
restructuring	Restrukturierung, restrukturieren	
construction	Bau, Neubau, Ausbau	
building	Gebäude, bauen	
infrastructure	Infrastruktur	
development	Entwicklung	
innovation	Innovation	
start-up	Start-up	
sustainability	Nachhaltigkeit	
Thematic areas (2014-2020)		<i>found through switching from EN to DE on http://ec.europa.eu/regional_policy/de/policy/how/priorities, http://ec.europa.eu/regional_policy/de/policy/themes/</i>
Research and Innovation	Forschung und Innovation	
Information and Communication Technologies	Informations- und Kommunikationstechnologien	
Competitiveness of SMEs	Wettbewerbsfähigkeit von KMU, Wettbewerbsfähigkeit der KMU	
Low-Carbon Economy	CO2-arme Wirtschaft	

PERCEIVE DEL. 5.2: 'DATABASE OF THE TOPICS AND SENTIMENTS TO BE MADE AVAILABLE ON-LINE FOR FURTHER RESEARCH'

Climate Change Adaptation & Risk Prevention	Anpassung an den Klimawandel und Risikoprävention, Klimawandel und Risikoprävention	
Environment Protection and Resource Efficiency	Umweltschutz und Ressourceneffizienz, Umweltschutz und effiziente Nutzung von Ressourcen	English version: +protecting the environment
Network Infrastructures in Transport and Energy	Vekehrs- und Energienetze, Netzwerkinfrastrukturen	English version: +transport and energy networks
Sustainable and Quality Employment	nachhaltige und hochwertige Beschäftigung	
Social Inclusion	soziale Integration, soziale Eingliederung	
Educational and Vocational Training	Bildung und Ausbildung, Aus- und Fortbildung, vocational training = Berufsausbildung, (educational training = Bildungsarbeit)	'Bildung und Ausbildung' is used for 'Education and training', 'Berufsausbildung' is used for 'vocational training' as such, 'educational training' I have found only once as 'Bildungsarbeit' / I would not use it as equivalent in general
Efficient Public Administration	effiziente öffentliche Verwaltung, Effizienz der öffentlichen Verwaltung	
Outermost and Sparsely Populated	in äußerster Randlage und dünn besiedelt	
Additional themes RP		
Urban development	Stadtentwicklung	
Rural development	ländliche Entwicklung	
Northern Ireland: the peace programme	PEACE-III-Programm, operationelles Programm für Frieden und Versöhnung	
Candidate countries (Enlargement of the EU)	Bewerberländer, Kandidatenländer	
EU Funds (see Regional Policy web portal)		
European regional development fund (ERDF)	Europäischer Fonds für regionale Entwicklung (EFRE)	
Cohesion fund (CF)	Kohäsionsfonds	
European social fund (ESF)	Europäischer Sozialfonds (ESF)	
European agricultural fund for rural development (EAFRD)	Europäischer Landwirtschaftsfonds für die Entwicklung des ländlichen Raums (ELER)	
European maritime and fisheries fund (EMFF)	Europäischer Meeres- und Fischereifonds (EMFF)	

European structural and investment (ESI) funds	Europäische Struktur- und Investitionsfonds (ESI-Fonds, ESIF)	There is no distinction in the German singular and plural between 'fund' and 'funds'
5 targets 2020 strategy		http://ec.europa.eu/europe2020/targets/eu-targets/index_de.htm , http://ec.europa.eu/europe2020/europe-2020-in-a-nutshell/targets/index_de.htm
Employment	Beschäftigung	
Research and Development	FuE, FuE und Innovation, (Forschung und Entwicklung und Innovation)	
Climate change and energy sustainability	Klimawandel und nachhaltige Energiewirtschaft, Klimawandel und Energie	
Education	Bildung	
Fighting poverty and social exclusion	Bekämpfung von Armut und sozialer Ausgrenzung, Armut und soziale Ausgrenzung	
Cooperation		
interregional cooperation program	Programm für interregionale Zusammenarbeit, Programm der interregionalen Zusammenarbeit	I cannot find the English version as such on Inforegio, I've used a literal translation
territorial cooperation	territoriale Zusammenarbeit	
transnational cooperation	transnationale Zusammenarbeit, grenzüberschreitende Zusammenarbeit, länderübergreifende Zusammenarbeit	German equivalent sometimes was 'territoriale Zusammenarbeit' in this case too
macro-regional strategies	makroregionale Strategien	
Danube region strategy	Donauraumstrategie, Strategie für den Donauraum, Strategie für die Donauregion, Donaustrategie	
Baltic sea region strategy	Strategie für die Ostseeregion, Strategie für den Ostseeraum, Ostseestrategie	
alpine region strategy	Alpenraumstrategie, Strategie für den Alpenraum	
Adriatic and Ionian sea strategy	Strategie für die Adria und das Ionische Meer	
URBACT	URBACT	
INTERACT (2013 is II)	INTERACT	
INTERACT (2013 is II)	INTERACT	
INTERREG (2013 is IV C)	INTERREG	

INTERREG EUROPE	INTERREG EUROPE	
ESPO	ESPO	
European Spatial Planning Observation Network	europäisches Raumbewachtungsnetzwerk	
Regions for economic change	Regionen für den wirtschaftlichen Wandel	
REGIOSTARS award	Auszeichnung REGIOSTARS	
EU Identity		all of the key words in this category were translated literally (not using InfoRegio)
being European	europäisch sein, Europäer/Europäerin sein (europäischer Bürger sein)	'europäischer Bürger sein' = being a European citizen
Europeannes	Europäizität, europäisches Dasein, Europäischsein	
"NUTS2ness"	Burgenländertum (burgenländisch as adjective)	
"NATIONALness"	Österreichertum (österreichisch as adjective)	
identity	Identität	
identification	Identifikation, Identifizierung	
european identity	europäische Identität, europäisches Identitätsbewusstsein	
national identity	nationale Identität, Nationalidentität (nationale Zugehörigkeit)	
regional identity	regionale Identität, Regionalidentität	
NUTS 2	Burgenland	oder 'NUTS 2'
NATION	Österreich	
culture + NUTS2	burgenländische Kultur	
national culture	nationale Kultur, Nationalkultur	
regional culture	regionale Kultur, Regionalkultur	
culture + NUTS2	burgenländische Kultur	
culture + NATION	österreichische Kultur	
country	Land	
nation	Nation, Staat	
region	Region, Gegend, (Bundesland)	in the specific case of Austria/Burgenland, we could use 'Bundesland' too
history + NUTS2	burgenländische Geschichte, Geschichte des Burgenlandes	

symbol + NUTS2	burgenländisches Symbol, Symbol für (das) Burgenland	
tradition + NUTS2	burgenländische Tradition	
landmark + NUTS2	burgenländisches Wahrzeichen	
heritage + NUTS2	burgenländisches Erbe	'burgenländische Tradition' could be used in this instance too, 'kulturelles Erbe des Burgenlandes'/'burgenländisches Kulturerbe' would denote 'cultural heritage'

7.4 Keywords in Polish

KEYWORDS in ENGLISH	KEYWORDS in POLISH	COMMENTS
Regional Policy (generic vocabulary)		
regional policy	polityka regionalna	
cohesion policy	polityka spójności/ or: polityka kohezji	The first option is much more popular
European solidarity	euuropejska solidarność	
Competitiveness of SMEs	konkurencyjność MŚP	
Sustainable development	zrównoważony rozwój	
2020 strategy	strategia 2020	
social exclusion	wykluczenie społeczne	
investment policy instrument	inwestycyjny instrument wsparcia/instrument polityki inwestycyjnej	
public investment	inwestycje publiczne/inwestycja publiczne/inwestycja ze środków publicznych	
national investment	inwestycje krajowe/inwestycja krajowa	
national-EU-budget	krajowy-unijny-budżet	
EU-investment	inwestycje UE/inwestycje ze środków UE	
EU-budget	budżet UE	
financing	finansowanie	
Co-financing rate	stopa współfinansowania/poziom współfinansowania	
national program	program krajowy	
regional program	program regionalny	
support	wsparcie	
promotion	promocja	
investment	inwestycja	

PERCEIVE DEL. 5.2: 'DATABASE OF THE TOPICS AND SENTIMENTS TO BE MADE AVAILABLE ON-LINE FOR FURTHER RESEARCH'

restructuring	restrukturyzacja	
construction	budownictwo	
building	budynek	
infrastructure	infrastruktura	
development	rozwój	
innovation	innowacja/innowacje/innowacyjność	
start-up	start-up	
sustainability	zrównoważenie	
Thematic areas (2014-2020)		
Research and Innovation	badania i innowacje	
Information and Communication Technology	technologie informacyjno-komunikacyjne	
Competitiveness of SMEs	konkurencyjność MŚP	
Low-Carbon Economy	gospodarka niskoemisyjna	
Climate Change Adaptation & Risk Prevention	dostosowanie do zmian klimatu i zapobieganie ryzyku/adaptacja do zmian klimatu i zapobieganie ryzyku	this seems to be including two versions
Environment Protection and Resource Efficiency	ochrona środowiska i efektywne wykorzystanie zasobów naturalnych	
Network Infrastructures in Transport and Energy	infrastruktura sieciowa w transporcie i energetyce	
Sustainable and Quality Employment	zrównoważone i jakościowe zatrudnienie	
Social Inclusion	włączenie społeczne	
Educational and Vocational Training	szkolenia edukacyjne i zawodowe	
Efficient Public Administration	efektywna administracja publiczna	
Outermost and Sparsely Populated	najbardziej oddalone i słabo zaludnione	
Additional themes RP		
Urban development	rozwój obszarów miejskich	
Rural development	rozwój obszarów wiejskich	
Northern Ireland: the peace programme	Irlandia Północna: program dla pokoju	
Candidate countries (Enlargement of the EU)	kraje kandydackie (rozszerzenie UE)	
EU Funds (see Regional Policy web portal)		
European regional development fund (ERDF)	Europejski Fundusz Rozwoju Regionalnego (EFRR)	
Cohesion fund (CF)	Fundusz Spójności (FS)	
European social fund (ESF)	Europejski Fundusz Społeczny (EFS)	

PERCEIVE DEL. 5.2: 'DATABASE OF THE TOPICS AND SENTIMENTS TO BE MADE AVAILABLE ON-LINE FOR FURTHER RESEARCH'

European agricultural fund for rural development (EAFRD)	Europejski Fundusz Rolny na rzecz Rozwoju Obszarów Wiejskich (EFRROW)	
European maritime and fisheries fund (EMFF)	Europejski Fundusz Morski i Rybacki (EFMR)	
European structural and investment (ESI) funds	Europejskie Fundusze Strukturalne i Inwestycyjne (EFSI)	
5 targets 2020 strategy		
Employment	zatrudnienie	
Research and Development	badania i rozwój	
Climate change and energy sustainability	zmiany klimatu i zrównoważenie energetyczne/zmiany klimatyczne i zrównoważenie energetyczne	
Education	Edukacja	
Fighting poverty and social exclusion	zwalczanie ubóstwa i wykluczenia społecznego	
Cooperation		
interregional cooperation program	program współpracy międzyregionalnej	
territorial cooperation	współpraca terytorialna	
transnational cooperation	współpraca transnarodowa/współpraca transgraniczna	
macro-regional strategies	strategie makroregionalne	
Danube region strategy	strategia regionu Dunaju	
Baltic sea region strategy	strategia regionu Morza Bałtyckiego/strategia UE dla regionu Morza Bałtyckiego/strategia rozwoju regionu Morza Bałtyckiego	three versions here
alpine region strategy	strategia regionu Alp	
Adriatic and Ionian sea strategy	strategia Adriatyku i Morza Jońskiego	
URBACT	URBACT	
INTERACT (2013 is II)	INTERACT (2013 is II)	
INTERACT (2013 is II)	INTERACT (2013 is II)	
INTERREG (2013 is IV C)	INTERREG (2013 is IV C)	
INTERREG EUROPE	INTERREG EUROPA	
ESPO	ESPO	
European Spatial Planning Observation Network	Europejska Sieć Obserwacyjna Rozwoju Terytorialnego i Spójności Terytorialnej	
Regions for economic change	Regiony na rzecz zmian gospodarczych	
REGIOSTARS award	nagroda REGIONSTARS	

PERCEIVE DEL. 5.2: 'DATABASE OF THE TOPICS AND SENTIMENTS TO BE MADE AVAILABLE ON-LINE FOR FURTHER RESEARCH'

EU Identity		
being european	bycie Europejczykiem	
Europeannes	uropejskość	
"NUTS2ness"	????	sorry, no equivalent in Polish, you may use general: lokalność (as localness).
"NATIONALness"	polskość	
identity	tożsamość	
identification	identyfikacja	
European identity	tożsamość europejska	
national identity	tożsamość narodowa	
regional identity	tożsamość regionalna	
NUTS 2	region NUTS2/NUTS2	
NATION	naród	
culture + NUTS2	kultura + region NUTS2	
national culture	kultura narodowa	
regional culture	kultura regionalna	
culture + NUTS2	kultura + region NUTS2	
culture + NATION	kultura + naród	
country	kraj	
nation	naród	
region		
history + NUTS2	historia + region NUTS2	
symbol + NUTS2	symbol + region NUTS2	
tradition + NUTS2	tradycja + region NUTS2	
landmark + NUTS2	charakterystyczny obiekt + region NUTS 2	Sama nie wiedziałam co wybrać: punkt orientacyjny, charakterystyczny punkt; słup/kamień graniczny; kamień milowy, punkt przełomowy, punkt zwrotny
heritage + NUTS2	dziedzictwo + region NUTS2	

7.5 Keywords in Romanian

KEYWORDS in ENGLISH	KEYWORDS in ROMANIAN	KEYWORDS in ROMANIAN	COMMENTS
	without special characters	with special characters	
Regional Policy (generic vocabulary)			
European Union	Uniunea Europeana	Uniunea Europeană	<i>this concept is used and recognized in relation with EU construction</i>
regional policy	politica regionala	politică regională	
cohesion policy	politica de coeziune	politică de coeziune	
European solidarity	solidaritate europeana	solidaritate europeană	
Competitiveness of SMEs	Competitivitatea IMM-urilor		
Sustainable development	dezvoltare sustenabila / dezvoltare durabila	dezvoltare sustenabilă / dezvoltare durabilă	<i>we use both terms with the same common sense</i>
2020 strategy	strategia 2020		
social exclusion	excluziune sociala	excluziune socială	
investment policy instrument	instrument de politica de investitii	instrument de politică de investiții	
public investment	investitie publica	investiție publică	
national investment	investitie nationala	investiție națională	
national-EU-budget	fonduri UE		
EU-investment	investitii UE	investiții UE	
EU-budget	buget UE		
financing	finantare	finanțare	
Co-financing rate	rata de cofinantare	rată de cofinanțare	
national program	program national	program național	
regional program	program regional		
support	sprijin		
promotion	promovare		
investment	investitie	investiție	
restructuring	restructurare		
construction			
building			
infrastructure	infrastructura	infrastructură	
development	dezvoltare		
innovation	inovatie	inovatie	

PERCEIVE DEL. 5.2: 'DATABASE OF THE TOPICS AND SENTIMENTS TO BE MADE AVAILABLE ON-LINE FOR FURTHER RESEARCH'

start-up	start-up		
sustainability	sustenabilitate / durabilitate		<i>we use both terms with the same common sense</i>
Thematic areas (2014-2020)			
Research and Innovation	cercetare si inovare	cercetare și inovare	
Information and Communication Technology	Tehnologiile informatiei si comunicatiilor / TIC	Tehnologiile informației și comunicațiilor / TIC	
Competitiveness of SMEs	Competitivitatea IMM-urilor		
Low-Carbon Economy	Economie cu emisii reduse de dioxid de carbon		
Climate Change Adaptation & Risk Prevention	Adaptarea la schimbarile climatice si prevenirea riscurilor	Adaptarea la schimbările climatice și prevenirea riscurilor	
Environment Protection and Resource Efficiency	Mediu si eficienta resurselor	Mediu și eficiența resurselor	
Network Infrastructures in Transport and Energy	Rețele de transport si energie	Rețele de transport și energie	
Sustainable and Quality Employment	Sustenabilitatea si calitatea locurilor de munca	Sustenabilitatea și calitatea locurilor de muncă	
Social Inclusion	Incluziune sociala	Incluziune socială	
Educational and Vocational Training	Educatie si pregatire profesionala	Educație și pregătire profesională	
Efficient Public Administration	Administratie publica eficienta	Administrație publică eficientă	
Outermost and Sparsely Populated	Ultraperiferica si densitate a populatiei redusa	Ultraperică și densitate a populației redusă	
Additional themes RP			
Urban development	Dezvoltare urbana	Dezvoltare urbană	
Rural development	Dezvoltare rurala	Dezvoltare rurală	
Northern Ireland: the peace programme	Irlanda de Nord: programul Peace		
Candidate countries (Enlargement of the EU)	Tari candidate (Extinderea UE)	țări candidate (Extinderea UE)	
EU Funds (see Regional Policy web portal)			
European regional development fund (ERDF)	Fondul european de dezvoltare regionala (FEDR)	Fondul european de dezvoltare regională (FEDR)	

PERCEIVE DEL. 5.2: 'DATABASE OF THE TOPICS AND SENTIMENTS TO BE MADE AVAILABLE ON-LINE FOR FURTHER RESEARCH'

Cohesion fund (CF)	Fondul de coeziune (FC)		
European social fund (ESF)	Fondul social european (FSE)		
European agricultural fund for rural development (EAFRD)	Fondul european agricol pentru dezvoltare rurala (FEADR)	Fondul european agricol pentru dezvoltare rurală (FEADR)	
European maritime and fisheries fund (EMFF)	Fondul european pentru pescuit si afaceri maritime (EMFF)	Fondul european pentru pescuit și afaceri maritime (EMFF)	
European structural and investment (ESI) funds	Fondurile structurale si de investitii europene (ESI)	Fondurile structurale și de investiții europene (ESI)	
5 targets 2020 strategy			
Employment	Ocuparea fortei de munca	Ocuparea forței de muncă	
Research and Development	Cercetare si dezvoltare	Cercetare și dezvoltare	
Climate change and energy sustainability	Schimbari climatice si durabilitate energetica	Schimbări climatice și durabilitate energetică	
Education	Educatie	Educație	
Fighting poverty and social exclusion	Combaterea saraciei si excluziunii sociale	Combaterea sărăciei și excluziunii sociale	
Cooperation			
interregional cooperation program	program de cooperare interregionala	program de cooperare interregională	
territorial cooperation	cooperare teritoriala	cooperare teritorială	
transnational cooperation	cooperare transnationala	cooperare transnațională	
macro-regional strategies	strategii macro-regionale		
Danube region strategy	Strategia pentru regiunea Dunarii	Strategia pentru regiunea Dunării	
Baltic sea region strategy	Strategia pentru regiunea Marii Baltice	Strategia pentru regiunea Mării Baltice	
alpine region strategy	Strategia pentru regiunea alpina	Strategia pentru regiunea alpină	
Adriatic and Ionian sea strategy	Strategia pentru regiunea Marii Adriatice si Marii Ionice	Strategia pentru regiunea Mării Adriatice și Mării Ionice	
URBACT	URBACT		
INTERACT (2013 is II)	INTERACT (2013 is II)		
INTERACT (2013 is II)	INTERACT (2013 is II)		
INTERREG (2013 is IV C)	INTERREG (2013 is IV C)		
INTERREG EUROPE	INTERREG EUROPE		
ESPO	ESPO		
European Spatial Planning	Reteaua europeana de observare a	Rețeaua europeană de observare a	

PERCEIVE DEL. 5.2: 'DATABASE OF THE TOPICS AND SENTIMENTS TO BE MADE AVAILABLE ON-LINE FOR FURTHER RESEARCH'

Observation Network	dezvoltarii si coeziunii teritoriale	dezvoltării și coeziunii teritoriale	
Regions for economic change	Regiunile, actori in schimbarea economica	Regiunile, actori în schimbarea economică	
REGIOSTARS award	Premiile RegioStars		
EU Identity			
being european	a fi european / european		
Europeannes	europenism / europenitate		<i>we use both terms with the same common sense</i>
"NUTS2ness"	brailean / buzoian / constantean / galatean / tulcean / vrancean / dobrogean	brăilean / buzoian / constănțean / gălățean / tulcean / vrâncean / dobrogean	<i>In the SE region there is no regional identity assumed. Inhabitants of this region use / recognize terms that are identifiable with the 6 NUTS III (which are part of the SE region (NUTS II)). I inserted here the terms, usually recognized within NUTS III regions to identify with these regions</i>
"NATIONALness"	romanesc	românesc	
identity	identitate		
identification	identificare		
European identity	identitate europeana	identitate europeană	
national identity	identitate nationala	identitate națională	
regional identity	identitate regionala	identitate regională	
NUTS 2			
NATION	natiune	națiune	
culture + NUTS2	cultura	cultură	
national culture	cultura nationala	cultură națională	
regional culture	cultura regionala	cultură regională	
culture + NUTS2	braileana / buzoiana / constanteana / galateana / tulceana / vranceana / dobrogeana	brăileană / buzoiană / constănțeană / gălățeană / tulceană / vrânceană / dobrogeană	<i>In the SE region there is no regional identity assumed. Inhabitants of this region use / recognize terms that are identifiable with the 6 NUTS III (which are part of the SE region (NUTS II)). I inserted here the terms, usually recognized within NUTS III regions to identify with these regions</i>
culture + NATION	romaneasca / romanesc	românească / românesc	
country	Romania	România	

nation	romana	română / națiune	
region	regiune		
history + NUTS2			<i>In the SE region there is no regional identity assumed</i>
symbol + NUTS2			<i>In the SE region there is no regional identity assumed</i>
tradition + NUTS2			<i>In the SE region there is no regional identity assumed</i>
landmark + NUTS2			<i>In the SE region there is no regional identity assumed</i>
heritage + NUTS2			<i>In the SE region there is no regional identity assumed</i>

7.6 Keywords in Swedish

KEYWORDS in ENGLISH	KEYWORDS in SWEDISH	COMMENTS
Regional Policy (generic vocabulary)		
regional policy	regionalpolitiken	Translation used by the EU.
cohesion policy	sammanhållningspolitiken	Translation used by the EU.
European solidarity	europeisk solidaritet	Direct translation.
Competitiveness of SMEs	konkurrenskraft hos små och medelstora företag	Translation used by the EU.
Sustainable development	hållbar utveckling	Direct translation.
2020 strategy	2020-strategin	Direct translation with article. Also referred to as <i>Europa 2020, EU 2020, or Europa 2020-strategin</i> .
social exclusion	social utestängning	Translation used by the EU. Also referred to as <i>socialt utanförskap, social exkludering</i> .
investment policy instrument	policyinstrument för investeringar	Functional translation.
public investment	offentlig investering	Direct translation.
national investment	nationell investering; nationell satsning	Direct translation.
national-EU-budget	nationell EU-budget	Direct translation.
EU-investment	EU-investering; EU-satsning	Direct translation.
EU-budget	EU-budget	Direct translation.
financing	finansiering	Direct translation.
Co-financing rate	medfinansieringsgrad	Direct translation.
national program	nationellt program	Direct translation.

PERCEIVE DEL. 5.2: 'DATABASE OF THE TOPICS AND SENTIMENTS TO BE MADE AVAILABLE ON-LINE FOR FURTHER RESEARCH'

regional program	regionalt program	Direct translation.
support	stöd	Direct translation.
promotion	främjande; främjandet	
investment	investering; satsning	Direct translation.
restructuring	omstrukturering	Direct translation.
construction	konstruktion*; konstruerande**; byggnad*; byggande**	* nominative noun: singular, no article ** non- finite verb: present participle tense
building	bygga*; bygger**; byggande***	*non-finite verb: infinite form **finite verb: present tense
infrastructure	infrastruktur	Direct translation.
development	utveckling	Direct translation.
innovation	innovation	Direct translation.
start-up	startup	
sustainability	hållbarhet	Direct translation.
Thematic areas (2014-2020)		
Research and Innovation	forskning och innovation	Translation generated directly from EU Regional Policy, Swedish website.
Information and Communication Technology	informations- och kommunikationsteknik	Translation generated directly from EU Regional Policy, Swedish website.
Competitiveness of SMEs	konkurrenskraft hos små och medelstora företag	Translation generated directly from EU Regional Policy, Swedish website.
Low-Carbon Economy	en ekonomi med låga koldioxidutsläpp	Translation generated directly from EU Regional Policy, Swedish website.
Climate Change Adaptation & Risk Prevention	anpassning till klimatförändringar och riskhantering	Translation generated directly from EU Regional Policy, Swedish website.
Environment Protection and Resource Efficiency	miljöskydd och resurseffektivitet	Direct translation as the one provided at the EU Regional Policy, Swedish website appears insufficient.
Network Infrastructures in Transport and Energy	transport- och energinätverk	Translation generated directly from EU Regional Policy, Swedish website.
Sustainable and Quality Employment	hållbar och kvalitativ sysselsättning	Direct translation as the one provided at the EU Regional Policy, Swedish website appears insufficient.
Social Inclusion	social inkludering	Translation generated directly from EU Regional Policy, Swedish website.
Educational and Vocational Training	utbildning och yrkesutbildning	Translation generated directly from EU Regional Policy, Swedish website.

PERCEIVE DEL. 5.2: 'DATABASE OF THE TOPICS AND SENTIMENTS TO BE MADE AVAILABLE ON-LINE FOR FURTHER RESEARCH'

Efficient Public Administration	bättre offentlig förvaltning	Translation generated directly from EU Regional Policy, Swedish website.
Outermost and Sparsely Populated	Europas yttersta randområden	Translation generated directly from EU Regional Policy, Swedish website.
Additional themes RP		
Urban development	stadsutveckling	Translation generated directly from EU Regional Policy, Swedish website.
Rural development	landsbygdsutveckling	Direct translation.
Northern Ireland: the peace programme	Nordirland: fredsprogrammet	Translation generated directly from EU Regional Policy, Swedish website.
Candidate countries (Enlargement of the EU)	kandidatländer (EU:s utvidgning)	Direct translation.
EU Funds (see Regional Policy web portal)		
European regional development fund (ERDF)	Europeiska regionala utvecklingsfonden (ERUF)	
Cohesion fund (CF)	Sammanhållningsfonden	
European social fund (ESF)	Europeiska socialfonden (ESF)	
European agricultural fund for rural development (EAFRD)	Europeiska jordbruksfonden för landsbygdsutveckling (EJFLU)	
European maritime and fisheries fund (EMFF)	Europeiska havs- och fiskerifonden (EHFF)	
European structural and investment (ESI) funds	Europeiska struktur- och investeringsfonder (ESI)	
5 targets 2020 strategy		
Employment	sysselsättning	Direct translation.
Research and Development	forskning och utveckling	Direct translation.
Climate change and energy sustainability	klimatförändringar och hållbar energi	Functional translation.
Education	utbildning	Direct translation.
Fighting poverty and social exclusion	bekämpa fattigdom och social utestängning	Social exclusion is in Swedish EU language translated as <i>social utestängning</i> ; however in the Swedish context <i>socialt utanförskap</i> and/or <i>social exkludering</i> would be more prevalent.
Cooperation		
interregional cooperation program	interregionalt samarbetsprogram	
territorial cooperation	territoriellt samarbete	
transnational cooperation	transnationellt samarbete	

macro-regional strategies	makroregionala strategier	
Danube region strategy	EU:s strategi för Donauregionen	
Baltic sea region strategy	EU:s strategi för Östersjöregionen; EU:s Östersjöstrategi	
alpine region strategy	EU:s strategi för Alpområdet; EU:s alpina strategi	
Adriatic and Ionian sea strategy	EU:s strategi för området kring Adriatiska havet och Joniska havet	
URBACT	URBACT	
INTERACT (2013 is II)	INTERACT II	
INTERACT (2013 is II)	INTERACT II	
INTERREG (2013 is IV C)	INTERREG IVC	
INTERREG EUROPE	INTERREG EUROPE	
ESPO	ESPO	
European Spatial Planning Observation Network	europiskt observationsnätverk för rumslig planering	
Regions for economic change	regioner för ekonomisk förändring	
REGIOSTARS award	REGIOSTARS-utmärkelse	The use of capitalised letters vary as well as the use of hyphen. The compounds may be closed and written as a single word and some may prefer to keep the English word constituted by open compounds.
EU Identity	EU-identitet	
being European	vara europé	Direct translation; looks rather stale but cannot find any suitable alternative.
Europeannes	europeskhet	Direct translation.
"NUTS2ness"	!	<i>for example for Emilia Romagna would be "emilianità"; see next sheet for thorough explanation on the Swedish case</i>
"NATIONALness"	svenskhet	<i>for example for Italy would be "italianità"</i>
identity	identitet	Direct translation.
identification	identifikation	Direct translation.
european identity	europesk identitet	Direct translation.
national identity	nationell identitet	Direct translation.
regional identity	regional identitet	Direct translation.

NUTS 2	NUTS2*; riksområde**	* often referred to as NUTS2 ** also referred to as <i>riksområden</i> (plural version or singular <i>riksområde</i>)
NATION	NATION	Direct translation.
culture + NUTS2	kultur + NUTS2	Direct translation.
national culture	nationell kultur	Direct translation.
regional culture	regional kultur	Direct translation.
culture + NUTS2	kultur + NUTS2	Direct translation.
culture + NATION	kultur + NATION	Direct translation.
country	land	Direct translation.
nation	nation	Direct translation.
region	region	Direct translation.
history + NUTS2	historia + NUTS2	Direct translation.
symbol + NUTS2	symbol + NUTS2	Direct translation.
tradition + NUTS2	tradition + NUTS2	Direct translation.
landmark + NUTS2	landmärke + NUTS2; milstolpe + NUTS2	Direct translation.
heritage + NUTS2	arv + NUTS2	Direct translation.

7.7 Keywords in Spanish

KEYWORDS in ENGLISH	KEYWORDS in SPANISH
Regional Policy (generic vocabulary)	Política Regional
regional policy	política regional
cohesion policy	política de cohesión
European solidarity	Solidaridad europea
Competitiveness of SMEs	Competitividad de PYMEs
Sustainable development	desarrollo regional
2020 strategy	Estrategia 2020
social exclusion	exclusión social
investment policy instrument	instrumento política de inversión
public investment	inversión pública
national investment	inversión nacional
national-EU-budget	presupuesto UE nacional
EU-investment	Inversión UE

PERCEIVE DEL. 5.2: 'DATABASE OF THE TOPICS AND SENTIMENTS TO BE MADE AVAILABLE ON-LINE FOR FURTHER RESEARCH'

EU-budget	presupuesto UE
financing	financiación
Co-financing rate	tasa de cofinanciación
national program	programa nacional
regional program	programa regional
support	apoyo
promotion	promoción
investment	inversión
restructuring	reestructuración
construction	construcción
building	construcción
infrastructure	infraestructura
development	desarrollo
innovation	innovación
start-up	start-up
sustainability	sostenibilidad
Thematic areas (2014-2020)	Áreas temáticas (2014-2020)
Research and Innovation	Investigación e Innovación
Information and Communication Technology	Tecnologías de la Información y la Comunicación
Competitiveness of SMEs	Competitividad de las PYME
Low-Carbon Economy	Economía de bajas emisiones de Carbono
Climate Change Adaptation & Risk Prevention	Cambio climático y prevención de riesgos
Environment Protection and Resource Efficiency	Medio Ambiente y Eficiencia de los recursos
Network Infrastructures in Transport and Energy	Redes de Transporte y Energía
Sustainable and Quality Employment	Empleo y mercado laboral
Social Inclusion	Inclusión social
Educational and Vocational Training	Educación y formación
Efficient Public Administration	Mejor Administración Pública
Outermost and Sparsely Populated	Regiones ultraperiféricas y en las regiones con escasa densidad demográfica
Additional themes RP	
Urban development	Desarrollo urbano
Rural development	Desarrollo rural
Northern Ireland: the peace programme	Irlanda del Norte: el programa de paz

Candidate countries (Enlargement of the EU)	Países candidatos (ampliación de la UE)
EU Funds (see Regional Policy web portal)	
European regional development fund (ERDF)	Fondo Europeo de Desarrollo Regional (FEDER)
Cohesion fund (CF)	Fondo de Cohesión
European social fund (ESF)	Fondo Social Europeo (FSE)
European agricultural fund for rural development (EAFRD)	Fondo Europeo Agrícola de desarrollo rural (FEADER)
European maritime and fisheries fund (EMFF)	Fondo Europeo de la Pesca y marítimo (FEMP)
European structural and investment (ESI) funds	Fondos estructurales y de inversión europeos
5 targets 2020 strategy	
Employment	Empleo
Research and Development	Investigación y Desarrollo
Climate change and energy sustainability	Cambio climático y sostenibilidad energética
Education	Educación
Fighting poverty and social exclusion	Lucha contra la pobreza y exclusión social
Cooperation	
interregional cooperation program	Cooperación Territorial Europea
territorial cooperation	Cooperación Territorial
transnational cooperation	Cooperación Transnacional
macro-regional strategies	Estrategias Macrorregionales
Danube region strategy	Estrategia de la UE para la región del Danubio
Baltic sea region strategy	Estrategia de la UE para la región del Mar Báltico
alpine region strategy	Estrategia de la UE para la región Alpina
Adriatic and Ionian sea strategy	Estrategia de la UE para la región adriático-jónica
URBACT	URBACT
INTERACT (2013 is II)	INTERACT (2013 is II)
INTERACT (2013 is II)	INTERACT (2013 is II)
INTERREG (2013 is IV C)	INTERREG (2013 is IV C)
INTERREG EUROPE	INTERREG EUROPE
ESPON	ESPON
European Spatial Planning Observation Network	Red Europea del Observatorio de la Ordenación del Territorio
Regions for economic change	Regiones por el Cambio Económico
REGIOSTARS award	Premios RegioStars

EU Identity	
being European	europeo
Europeannes	europeos
"NUTS2ness"	extremeña
"NATIONALness"	Española
Europeannes	española
Europeannes	español
european identity	identidad europea
national identity	Identidad española
regional identity	Identidad extremeña
NUTS 2	Extremadura
NATION	España
culture + NUTS2	cultura extremeña
national culture	cultura española
regional culture	cutura extremeña
culture + NUTS2	cultura extremeña
culture + NATION	cultura española
country	España
nation	España
region	Extremadura
history + NUTS2	historia extremeña
symbol + NUTS2	símbolo extremeño
tradition + NUTS2	tradicón extremeña
landmark + NUTS2	Lugares emblemáticos de Extremadura
heritage + NUTS2	patrimonio extremeño

7.8 English Stopword list

a	ask	come	everything
able	asking	comes	everywhere
about	associated	concerning	ex
above	at	consequently	exactly
according	available	consider	example
accordingly	away	considering	except
across	awfully	contain	f
actually	b	containing	far
after	be	contains	few
afterwards	became	corresponding	fifth
again	because	could	first
against	become	course	five
all	becomes	currently	followed
allow	becoming	d	following
allows	been	definitely	follows
almost	before	described	for
alone	beforehand	despite	former
along	behind	did	formerly
already	being	different	forth
also	believe	do	four
although	below	does	from
always	beside	doing	further
am	besides	done	furthermore
among	best	down	g
amongst	better	downwards	get
an	between	during	gets
and	beyond	e	getting
another	both	each	given
any	brief	edu	gives
anybody	but	eg	go
anyhow	by	eight	goes
anyone	c	either	going
anything	came	else	gone
anyway	can	elsewhere	got
anyways	cannot	enough	gotten
anywhere	cant	entirely	greetings
apart	cause	especially	h
appear	causes	et	had
appreciate	certain	etc	happens
appropriate	certainly	even	hardly
are	changes	ever	has
around	clearly	every	have
as	co	everybody	having
aside	com	everyone	he

hello	keeps	near	ourselves
help	kept	nearly	out
hence	know	necessary	outside
her	knows	need	over
here	known	needs	overall
hereafter	l	neither	own
hereby	last	never	p
herein	lately	nevertheless	particular
hereupon	later	new	particularly
hers	latter	next	per
herself	latterly	nine	perhaps
hi	least	no	placed
him	less	nobody	please
himself	lest	non	plus
his	let	none	possible
hither	like	noone	presumably
hopefully	liked	nor	probably
how	likely	normally	provides
howbeit	little	not	q
however	look	nothing	que
i	looking	novel	quite
ie	looks	now	qv
if	ltd	nowhere	r
ignored	m	o	rather
immediate	mainly	obviously	rd
in	many	of	re
inasmuch	may	off	really
inc	maybe	often	reasonably
indeed	me	oh	regarding
indicate	mean	ok	regardless
indicated	meanwhile	okay	regards
indicates	merely	old	relatively
inner	might	on	respectively
insofar	more	once	right
instead	moreover	one	s
into	most	ones	said
inward	mostly	only	same
is	much	onto	saw
it	must	or	say
its	my	other	saying
itself	myself	others	says
j	n	otherwise	second
just	name	ought	secondly
k	namely	our	see
keep	nd	ours	seeing

seem	tends	tried	what
seemed	th	tries	whatever
seeming	than	truly	when
seems	thank	try	whence
seen	thanks	trying	whenever
self	thanx	twice	where
selves	that	two	whereafter
sensible	thats	u	whereas
sent	the	un	whereby
serious	their	under	wherein
seriously	theirs	unfortunately	whereupon
seven	them	unless	wherever
several	themselves	unlikely	whether
shall	then	until	which
she	thence	unto	while
should	there	up	whither
since	thereafter	upon	who
six	thereby	us	whoever
so	therefore	use	whole
some	therein	used	whom
somebody	theres	useful	whose
somehow	thereupon	uses	why
someone	these	using	will
something	they	usually	willing
sometime	think	uucp	wish
sometimes	third	v	with
somewhat	this	value	within
somewhere	thorough	various	without
soon	thoroughly	very	wonder
sorry	those	via	would
specified	though	viz	would
specify	three	vs	x
specifying	through	w	y
still	throughout	want	yes
sub	thru	wants	yet
such	thus	was	you
sup	to	way	your
sure	together	we	yours
t	too	welcome	yourself
take	took	well	yourselves
taken	toward	went	z
tell	towards	were	zero