Project: **The Transformation of Political Mobilisation and Communication in European Public Spheres**

Project acronym: Europub.com
Project website: http://europub.wz-berlin.de

Funded by: 5th Framework Programme of the European Commission
Contract No. HPSE-CT2000-00046
Work package: **WP 4 (Internet)**
WP Coordinator: Ruud Koopmans
Deliverable number: **D 4.3**

Codebook **Codebook for the analysis of political communication on the Internet.
Part 1: Representative sample of websites**

Author: Ruud Koopmans and Ann Zimmermann

Date: 21 October 2002

Europub.com

To get the possibility to examine the Internet as a potentially new arena for political communication on different levels the analysis will take place in two parts with different codebooks. The starting point of analysis in the *first part* is the Internet itself. For structuring the Internet as our object of analysis we will simulate search processes for information about our policy fields in a way that a "common" Internet user would do. In this way we narrow down our research field along the question how people mainly use the medium when they are looking for information. This corresponds to our approach in wp2, where we are not looking directly at "reality" but at the representation of reality by media, which are filtering and selecting the complexity of available information. In wp2, we decided to use widely read newspapers for narrowing down media public sphere and for making it examinable. Thus, the aim of this first empirical part of wp4 is less to see how citizens actually use the Internet for political communication, but to look which actors are "visible" on Internet in regard to specific policy fields. Since prominence is inseparable from the audience it seems to be reasonable to take into account that most of the Internet users use search engines to get information on the Internet.

The expected outcomes of this codebook will allow us:

- to see which political actors are prominent on the Internet in regard to specific political issues
- to compare prominent actors on the Internet to the prominent actors in the traditional mass media in regard to the same political issue at the same time (in consideration of the findings from the wp2 coding).
- to compare claims that are made online with claims that are made offline (in the newspapers) in regard to specific political issues at the same time

The coding of claims on this level is especially interesting because this is our only possibility to make a direct comparison between claims that are made on the Internet and claims that are made in the print media on a specific issue at the same time, because the latest year we will code in wp2 is 2002. The second part of the Internet coding will mainly take place in 2003, so we will not be able to compare forms of claims making offline and online directly.

In the *second part* of analysis we will start from political actors that are important in regard to our policy fields. Here, we will focus on the network of political communication that is facilitated via the Internet. We will analyse the websites of key political actors (e.g. identified from their prominence in the wp2 and wp3 coding) in a detailed way. Beside of formal aspects and the claims that are made on these websites we will focus on forms of interactivity and linkage structures between different actors. Thus, we will be able to analyse the use of the Internet as a communication medium by political actors and the structural and conceptual features of networks of political communication via Internet. Additionally we will analyse the websites of the actors we identified as important on the Internet in the first part of analysis in the same way to see if there are differences in strategies and linkage structures.

The units of analysis

1. The **search process** itself.
2. The returns of the search process, respectively the **websites** found by the search engine.
3. The **texts** that contain information on our search word on the website.
4. The **claims** that are made in these texts.

Except for the search process itself there are several forms of actors defined for each unit. On the website level we distinguish between the actor who serves as a gate to the website with the relevant text and the actor who acts as a medium, that means that he actually offers the relevant text on his website. On the text level we differentiate between the author of the text and the original source of the text. On the claim level the distinction of actors is defined in the same way as in the wp2 codebook which is generally the basis for the claim coding in an adapted form.

Definitions of actors

The abovementioned four new actor forms in this codebook are defined as:

Gate:	the actor who runs the website that provides the access or way to the website that contains the relevant information (text) via an *external* link.
Medium:	the actor who runs the website that actually contains the relevant information
Source:	the actor who originally published the text that contains the relevant information
Author:	the actor who wrote the text.

Example 1:

GOOGLE-SEARCH:	„Monetary politics 2002"
	Yahoo! News Full Coverage-Euro and the European **Monetary** Union **...** Euro and the European **Monetary** Union. Related Full Coverage. **...** (May 6, **2002**). -, Forget those euro 'tests' - and think exchange rate - Scotsman (Apr 22, **2002**). **...** dailynews.yahoo.com/fc/World/European_Monetary_Union/ - 24k - Cached - Similar pages
TADDRESS:	http://news.**bbc**.co.uk/hi/english/business/newsid_2040000/2040432.stm
TTITLE:	Euro soars amid stock gloom
AUTHOR:	no author mentioned
SOURCE:	no source mentioned
MEDIUM:	BBC
GATE:	YAHOO
RESULT:	http://dailynews.**yahoo**.com/fc/World/European_Monetary_Union/

Example 2:

FICTIVE SEARCH:	after following the same RESULT as above
TADDRESS:	http://story.news.yahoo.com/news?tmpl=story&cid=509&509&e=5&u=/ap/20020606/ap_on_bi_ge/europe_interest_rates_8
TTITLE:	ECB Keeps Interest Rates Steady
AUTHOR:	DAVID McHUGH
SOURCE:	Associated Press
MEDIUM:	YAHOO
GATE:	Gate and Medium are the same
RESULT:	http://dailynews.**yahoo**.com/fc/World/European_Monetary_Union/

THE CODING PROCESS

The coding process can be distinguished in three main steps: (1) search, (2) identification of relevant websites and (3) the actual coding. The first step is the search with the search engine. The second step is to identify which of the results the search engine gives you in return to your query are relevant. The relevance of a website is determined by the relevance of the text the website offers about the policy field the search word refers to, and by the country in which the website is located. The third step is the actual coding of the relevant text found on the selected website: the actor who runs the website (medium), if provided the actor who leads via an external link to the website (gate) and finally the claims that are made in the relevant text are to be coded. Thus, apart from the general search information, the starting point of coding is the relevant text that was found on the website.

(1) SEARCH

Search engines:

To get an impression of the kind of information the "common" Internet user gets when s/he searches the Internet for information on the policy fields of our project and especially to find out what kind of actors offer this information, we need to simulate the "common" search process. Since most Internet users use search engines when they are looking for information on the Internet (at least if they do not already have specific addresses) it seems to be reasonable to do the same. Therefore the first step for each team is to find out (on the base of secondary information) which are the most often used search engines in their country. The two first most often used search engines will be selected for the search process. Thus, we will conduct a search process with one search word combination two times with the different search engines.

Search word combinations:

There are two search word combinations for each issue field: (1) a general term related to the issue field, (2) a specific term related to the issue field. All search word combinations also include "2002", to get some kind of actuality if possible.

Search process

If the search engine offers more than one possibility to search, then always use the default option. You should not use any forms of advanced search or possibility to specify your search in certain ways (language, content, etc.).

Sample:

The first ten results/websites containing relevant information offered by the search engines in return to each of your entered search word combinations will be included in the analysis. If a returned result contains no relevant information it should be ignored and the next one is to be looked at. The same applies if a website is not available or if you need to register in any way to get to the announced website or text.

Thus, our sample does not automatically correspond to the first ten returned results. Most of the time you will have to look at more than ten results to get ten websites that contain relevant information. However, if you encounter ten or more results in a row that are irrelevant for substantive reasons, you should break off the coding of that search, even if you have not yet reached ten relevant results. 'For substantive reasons' means that the site is rejected because it is not about our issues of research. I.e., this does not include sites that are irrelevant because they are located outside Europe, because they are in a different language, or because the site is not reachable, requires a password, etc. Such types of irrelevant sites do not count among the ten subsequent irrelevant sites that are the criterion for breaking off the coding.

Some search engines offer more than one possibility to click on to get to a website that is indicated as a result on the result list. You should always only click on the first option. E.g., in Google you have always some kind of headline on the top of each result that serves as a link to the actual website. This is the only way you should try to get to the displayed site. In addition Google provides the possibility to click on "cached" and "similar pages". Even if you can't reach the website via the first offered link, you should not use these extra links or any kind of archive search, but go on to the next result.

follow only this link

School Education Survey
... Reports based on the National Public Radio/ Kaiser Family Foundation/ Kennedy **School**
of Government survey on **education** will be aired on NPR beginning Tuesday **...**
Description: Public approves of higher taxes for schools, but wants fairness in funding; support weak for vouchers...
Category: Society > Issues > Education
www.npr.org/programs/specials/poll/ education/education.front.html - 15k - Cached - Similar pages

ignore these links

Search registration

An important step in the search process is to SAVE your search. This will help in future analysis of the search process as well as for quoting in scientific publications. So besides the texts containing the claims, we also produce electronic files displaying the result list of the search engine.
It's just a few clicks.

The *file format* is optional, you can choose between three options according to your equipment or preferences:

a Save or print the search result list as a portable document file (.PDF)
For this option you need to have Adobe installed; you can either go to the normal print menu and choose the virtual printer 'Acrobat distiller', or select the 'Create PDF' button if available.

b Copy the list from the internet browser to a word document (.DOC)
Use the 'edit/select all' command from the menu or press <ctrl><a>, then copy and paste everything in a new word document.

c Save the search list in your internet browser as a website (.HTML)
Use the 'file/save as' command from the menu. In Internet Explorer you have the possibility to choose as 'file type' either 'save complete website' or 'save website as html only'. The second option is sufficient for our purposes and consumes less space on your hard disk, but pictures and logos may be lost. In Netscape you don't have the choice, it's html only.

The *file name* should be structured in a standardized way that will allow us to exchange the files easily between teams:
The first two digits are the letters of your country code (CH, DE, ES, FR, IT, NL or UK).
The third digit is the letter S for Search.
The three following digits are reserved for the SID number (001-999).
The two following positions refer to the results Page (P1-9)
(in a word doc, all pages can be pasted together).

Ex: FRS015P2.pdf = French case, Search ID 015, 2nd page of results list
FRS015.doc   "
FRS015P2.htm   "

Sample scheme:

Data will be collected during the two periods 22 – 30 July 2002 and 18 – 26 November 2002. At both points in time we will collect ten relevant websites from search results of the two most often used search engines in each country for all seven issue fields (14 keywords). It is important that everybody searches with the same keywords on the same day (see tables below).

Sampling scheme July 2002	
Day	**Issue field**
Monday, 22.07.02	Monetary politics
Tuesday, 23.07.02	Agriculture
Wednesday, 24.07.02	Immigration
Thursday, 25.07.02	Troops deployment
Friday, 26.07.02	Retirement and pension schemes
Monday, 29.07.02	Education
Tuesday, 30.07.02	European Integration

Sampling scheme November 2002	
Day	**Issue field**
Monday, 18.11.02	Monetary politics
Tuesday, 19.11.02	Agriculture
Wednesday, 20.11.02	Immigration
Thursday, 21.11.02	Troops deployment
Friday, 22.11.02	Retirement and pension schemes
Monday, 25.11.02	Education
Tuesday, 26.11.02	European Integration

(2) IDENTIFICATION OF RELEVANT WEBSITES

Relevance of websites

The relevance of a website offered as a result by the search engine in return to your search word combination query has to be determined by three criteria:

1) *Language:* The website must be in your own language. This rule is due to the fact that some search keywords are identical in different languages which would lead to biases in regard of Europeanisation/Internationalisation of web communication. If you find directly a text in your own language on a website you should code it, regardless in which text the rest of the website is. If the website is in another language than your own and you have to click on one or more links (that are not in your language) until you get a text in your own language you should ignore the website.

2) *Location:* Foreign websites that are in the same language as yours are only relevant if they are located in Europe (e.g. Austrian websites in Germany, French websites in Switzerland) or if they are websites of transnational organisations. Websites of transnational organisations are coded regardless of their geographical location (e.g. UN sites). However, non-european national websites of transnational organisations are excluded from the detailed coding: e.g. the international homepage of Greenpeace (http://www.greenpeace.org/) or Amnesty International (http://www.amnesty.org/) should be coded in detail, while the website of Greenpeace in Canada (http://www.greenpeace.ca) or Amnesty International in USA (http://www.amnestyusa.org/) should be coded only on the reduced level. The inclusion/exclusion criterion is the geographical location of the website, not the nationality of the owners of the site. Therefore you should for example include the US embassy in London or the website of a Spanish portal containing a text from an African NGO, but exclude the British Council in New York or a Canadian website that offers links to a French website.
<u>There is one exception</u>: websites that are located abroad only for strategic reasons (to avoid repression or censorship) and evidently orientated towards a public in Europe, should be included. For example the French pro-nazi site, that is all about France but the owners claim to be based in the US to avoid French law.

Whether a website is national or transnational you should decide by looking at the website itself and not only by looking at the address. Especially in case of generic top level domains like ".int", ".org", ".com" etc. the country of origin is not derivable from the address. If the country in which the website is located is not clear from the information you get on the website, but you have reason to believe that the website/organisation/information may be transnational, you should include it in the full coding. Thus, in case of doubts rather include a website as transnational than exclude it.

2) *Relevant information:* The website must contain relevant information in textual form in regard to the search issue. Therefore, it is necessary to identify that part of the websites that probably caused that the search engine rated the website as relevant. We are looking for information on the issue in textual form that is meaningful in terms of our seven issue fields. A text that contains for example "European" and "Integration" is not relevant if it is about the integration of immigrants in European countries. In this way, a website is only relevant if the

provided text about the search issue is relevant. Whether or not a text contains a claim is not a criterion for coding the website, just as we code articles without claims in the wp2 coding.

Relevance of texts

The texts must be relevant in terms of content and provide more information than the simple mention of the search word (e.g. "Euro"). The provided information should be in the form of claims, facts, explanation or commentaries on the issue. That does automatically exclude all texts that do not offer any kind of information on the issue itself, like announcements that only say that a seminar about monetary politics will take place. Here, the only information you get on the issue is that you can get information on the issue when you take part in a forthcoming seminar. Thus, announcements, programs etc. are only relevant if they include something like an introduction or classification of the issue in terms of content. Book titles, lists of references or curriculum vitas are not relevant.

Generally, the relevance of a text in regard to the search issue should be assessed in the same way as an article is defined as relevant in the claims analysis in wp2. However, the texts do not need to be necessarily articles in a strict sense. They can also be very short/long or unstructured.

Identification of relevant texts

We are searching for texts that contain the central word of the search word combination (marked in bold type on page 17 and 18) and preferably contain claims about the policy field the search word combination refers to. In a first step you should look for the keyword on the part of the website you see on your screen without scrolling. If the keyword can't be found on this part of the website you can search on the whole site with the search function (Strg+f) for the keyword. Sometimes the search function finds the word you are looking for only if the whole website is selected (marked). So if the word is not found you can try it again by selecting the whole website and search then.

If several texts contain the keyword you should choose at first the text that is most prominent on the website in terms of visibility (larger title than the other texts, colours, picture, at the top of the site, etc.). If this text contains no claims you should go on to the next prominent text and so on until you find a text that contains claims in regard to the policy field the search word refers to. If only texts with no claims can be identified on the website, you should select the first one found of these claim-free texts and code it on the text level.

In a list of text titles where each title is a link to the texts itself you should click on the first title that contains the keyword. If the text the link leads to is not relevant you should go back to the title link list and go on to the next text - regardless of potentially relevant links on the site where you found the first text.

If the search word is only present on the website as a link (regardless if it is an internal or external link), then click on this link and search on the site that appears. In general, you should click on relevant links until you find a text that is about the search issue and contains claims. When only the beginning of the text is shown on the website and a link to the whole text is given, you should click on the link to code the whole text (even if the shown part of the text already contains claims)

The texts themselves are primarily treated as the articles in wp2. However, there is no "150 words rule". In the case of extensive online publications (books, papers, etc.) only the summary needs to be read to decide if the text is relevant. If no summary is provided or if the summary does not refer to the search issue you should search the whole document for the issue specific search word in the search word combination (str + f). These specific search words are marked in bold type in the key word list (e.g.: **Monetary** politics 2002). After the keyword was found in the text you should read the first page on which the keyword appears and the page after. If one page can not be easily identified (e.g. within a html document), you should estimate the length of two pages on the basis of your experience.

If the only information you get on the search issue is for example a data table (e.g. statistic on cases of BSE in several countries *without* any explanation or commentary) you should treat that table the same way you treat an article without claims in wp2.

(3) CODING

If a website is identified as relevant because it offers a relevant text the actual coding starts.

PART I: Search Variables

Coding of search information

At first the general search information are to be coded. Who? When? Which search engine? Which search word combination was used?

PART II: Result Variables

Coding of results

For all websites containing relevant information that are offered as a result by the search engine in return to the search word combination query we code their position in the list of search results and the country in which they are located. For websites that are irrelevant because they are non-European and non-transnational the coding stops at this point and only relevant websites will be coded in detail. Websites that contain no relevant information should be ignored completely.

PART III: Text Variables

Coding of text information

On this level general information about the text is to be coded. What kind of text is it? Is it offered in more than one language? More detailed information about the text only needs to be coded if the text does not contain claims.

Coding of author

The author of the text will be coded in a similar way as ACT in the wp2 codebook. If no author is specified we only code "no author mentioned" (0).

Coding of source

Source is defined as the actor who originally published a text. E.g. if an actor presents an article from the Guardian on his own website the source is the Guardian. If no source is indicated we only code "no source mentioned" (0).

Coding of medium

The next step is to identify the medium of the website, that means the actor who runs the website that actually contains the relevant text. Often you find links on the website like „home", „about us" or something like that, that lead you directly to information about the medium. Sometimes you can also click on the logo of the medium to get more information.
If nothing like that is offered (e.g. if the relevant site is in Pdf format) you should use the current address/URL displayed in the address field on the top of the screen. To get to the start site of an online offer of an actor you can "shorten" the current address (e.g.
"http://www.bundesbank.de/de/presse/glossar/e.htm" → http://www.bundesbank.de/).
Since there are sometimes actors that offer a platform for other actors on their website who in turn include the online performance of another actor within their online offer, you should cut the address step by step from one slash to the next. The first actor that is indicated by something like a welcome site that communicates that he provides this website as his homepage should be coded. An example is the website of the "Philadelphia Inquirer"
(http://www.philly.com/mld/inquirer/news /nation/3172611.htm) from the second pretest found in search 2. Here, the medium is the "Philadelphia Inquirer" and *not* "Philly.com", because the "Philadelphia Inquirer" runs an own homepage on the website of "Philly.com". The medium is always the actor who actually offers the relevant text, even if another actor works as a platform for the homepage of this actor.
You should shorten the current address to find relevant texts on a website. The starting point of the search for texts is always the websites the search engines indicate as results and only links that are offered directly on these sites should be used.
The medium will be coded in a similar way as ACT1 in the wp2 codebook.

Coding of gate

If the website the search engine indicates as a result does not contain the relevant text itself but leads via an external link to the website of another actor that contains that information, then the website is defined as a gate. In such cases the actor who runs the gate website is coded additionally to the medium. If the result website offered by the search engine already contains the relevant text (directly or via internal links) you only need to code "gate and medium are the same" (0).

PART IV: Claim Variables

Coding of claims

The coding of claims is based on a reduced version of the wp2 codebook. Besides identification and location we only code the variables of subject actors, action forms and aims. Additionally we only code the main claim of the article. As defined in the wp2 codebook, the main claim is the claim around which the information (in as far as relevant to our seven fields) in the article is organized, often indicated by the headline(s). E.g., in an article 'Minister Schily presents new immigration law' which apart from Schily's law also reports reactions to that law by several other actors, Schily's law is the main claim and coded first. If in doubt, the first reported claim is the main claim.
You should only code claims that are directly related to the issue of the search word combination or to other subissues of the policy field the search word combination refers to (defined in the wp2 codebook). You should not code claims related to other policy fields than to the one your search word combination refers to, even if the other policy field is one of our seven policy fields.

Same results within different results or searches

Sometimes you get the same website within one result list twice or more often. If the search result website is exactly the same and there is more than the one relevant text you have already coded offered, you should choose the second (third) most prominent text for coding. If there is no relevant text offered in addition to the one you have already coded you only need to go back to the record where you have coded it and enter the second (third) position of the website within the result list in the variable POSITION2(3). This coding only on the POSITION variable does not count among the ten relevant websites that should be coded. If you get the same search result website within two different searches you should code the website in a new record.

PART I:
SEARCH VARIABLES

Variable **SID** (numeric)
'search identifcation number'

Variable **NAMECOD**
'name of coder'

Variable **SDAY**
'day of search'
(1-31)

Variable **SMONTH**
'month of search'
(1-12)

Variable **SYEAR**
'year of search'

Variable **SNAME**
'name of search engine/portal/catalogue used'

Variable **SLAN**
'search language'

1   'Dutch'
2   'English'
3   'French'
4   'German'
5   'Italian'
6   'Spanish'

Variable **SKEY**
'keywords of search'
Note: Enter the keywords in your own language.

1   'MONETARY POLITCS: currency politics and interest rate'
   11 **monetary** politics 2002
      DE: **Geldpolitik** 2002
      FR: politique **monétaire** 2002
      NL: **monetair** beleid 2002

   12 **interest rate** decision 2002
      DE: **Leitzins** Entscheidung 2002
      FR: **taux d'intérêt** décision 2002
      NL: **rentebesluit** 2002

2   'AGRICULTURE: subsidies, livestock and dairy quotas, animal disease control'
   21 agriculture **subsidies** 2002
      DE: Landwirtschaft **Subventionen** 2002
      FR: **aides** agriculture 2002
      NL: landbouw **subsidies** 2002

   22 **BSE** cows 2002
      DE: **BSE** Rinder 2002
      FR: **ESB** vaches 2002
      NL: **BSE** koeien 2002

3   'IMMIGRATION: entry and exit'
   31 **immigration** politics 2002
      DE: **Zuwanderungspolitik** 2002
      FR: politique d'**immigration** 2002
      NL: **immigratiebeleid** 2002

   32 **deportation** 2002
      DE: **Abschiebung** 2002
      FR: **reconduite** à la frontière 2002
      NL: **uitzetting** 2002

4   'TROOPS DEPLOYMENT'
   41 **troops** deployment 2002
      DE: **Truppen** Stationierung 2002
      FR: déploiement de **troupes** 2002
      NL: **troepen** stationering 2002

   42 troops **Peacekeeping** 2002
      DE: Truppen **Friedenssicherung** 2002
      FR: troupes **maintien de la paix** 2002
      NL: troepen **vredesmissie** 2002

5   'RETIREMENT AND PENSION SCHEMES'
   51 **pension** politics 2002
       DE: **Rentenpolitik** 2002
       FR: **retraites** politique 2002
       NL: **pensioenbeleid** 2002

   52 pensions **demographic** 2002
       DE: Renten **demographisch** 2002
       FR: retraites **démographique** 2002
       NL: pensioenen **demografisch** 2002

6   'EDUCATION'
   61 **education** politics 2002
       DE: **Bildungspolitik** 2002
       FR: politique d'**éducation** 2002
       NL: **onderwijsbeleid** 2002

   62 education equal **opportunities** politics 2002
       DE: Bildung **Chancengleichheit** Politik 2002
       FR: éducation égalité des **chances** politique 2002
       NL: onderwijs **achterstand** beleid 2002

7   'EUROPEAN INTEGRATION'
   71 EU **reforms** 2002
       DE: EU **Reformen** 2002
       FR: UE **réformes** 2002
       NL: EU **hervormingen** 2002

   72 EU **enlargement** 2002
       DE: EU **Erweiterung** 2002
       FR: **élargissement** UE 2002
       NL: EU **uitbreiding** 2002

Variable **RESULTS**
'number of returned results'

PART II:
RESULT VARIABLES

Variable **RID** (numeric)
'result identification number'

Note: result (search result website) is the website you are directly led to via the result link the search engine offers you.

Variable **POSITION1** (POSITION2, POSITION3)
'position of result within the returned result list'

Note: If you get the same search result website within one result list and if there is no other relevant text on the website than the one you have already coded you should enter the second (third) position of this website within the result list here in POSITION2(3). That does not apply if you get the same result within two different searches. In this case you should code the website in a new record. However, the coding will be reduced by using the drop-down lists.

Variable **RELEVANCE**
'is website relevant in terms of provided information and country where it is located'

1   'relevant'
2   'irrelevant, because non-European[1] and non-transnational'

IF 1 > ON TO **RESULTAD**: IF 2,3 ON TO **WCOUNTRY**

Variable **WCOUNTRY**

END OF CODING FOR IRRELEVANT WEBSITES

---

[1] For this variable, *Europe* includes the EU Member States and candidate countries, other western European countries, the Balkans, and the successor countries of the Soviet Union except the central Asian republics; i. e.: Albania, Andorra, Armenia, Austria, Azerbaijan, Belarus, Belgium, Bosnia & Herzegovina, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Georgia, Germany, Gibraltar, Greece, Greenland, Hungary, Iceland, Ireland, Italy, Latvia, Liechtenstein, Lithuania, Luxembourg, F.Y.R. of Macedonia, Malta, Moldova, Netherlands, Norway, Poland, Portugal, Romania, Russian Federation, Slovakia, Slovenia, Spain, Sweden, Switzerland, Turkey, Ukraine, United Kingdom, Vatican City State, Yugoslavia (Serbia & Montenegro).

Variable **RESULTAD**
'address/URL of the search result website (indicated by the search engine as result)'

Variable **RLAN1** (RLAN2, RLAN3…RLAN7)
'Is the search result website also offered in further languages than your search language?'

0   'no further language'
1   'Dutch'
2   'English'
3   'French'
4   'German'
5   'Italian'
6   'Spanish'
7   'Other regional language(s)' (e.g. Catalan, Welsh, Frisian)
8   'Other EU language(s)'
9   'Other language(s)'

Variable **ROTLAN** (string variable)
'Specification of other languages'

Note: If the website is offered in more than one "other" language you should enter them in the form: Russian, Chinese, Japanese

PART III:
TEXT VARIABLES

Variable **TID**
'text identification number'

Variable **DIRECT**
'is the text offered directly on the search result website'

1   'yes'
2   'no, it is necessary to click on one *internal* link to get to the relevant information'
3   'no, it is necessary to click on more than one *internal* links to get to the relevant information'
4   'no, it is necessary to click on one or more *external* links to get to the relevant information'

Note: The search result website is the website you are directly led to via the link the search engine offers you as a result. "1" should be coded if the whole relevant text is offered on this website directly and you don't have to click on any links at all. "2" should be coded if you need to click on one internal link to get to the relevant text. The same applies if you get the beginning of a text on the search result website and the remaining part of the text on another website via an internal link.

Variable **TADDRESS**
'address/URL of the website which contains the relevant text'

Variable **MAINTOP**
'is the issue main topic of the text'

1   'yes'
2   'no'

Variable **TDAY**
'day of text release on the website'
(1-31)

Variable **TMONTH**
'month of text release on the website'
(1-12)

Variable **TYEAR**
'year of text release on the website'
(two digits, e.g. 85=1985; 99=1999; 0=2000, 1=2001)

NOTE: If there are no information about the day, month or/and year, simply let the cells blank.

Variable **TEXTYPE**
'type of text'

1	'newspaper/magazines article'
2	'article by a press agency'
3	'article by a television station'
4	'article by a radio station'
5	'information source/data base'
6	'descriptions / information on competences, responsibilities'
7	'news-style article by a political actor on an issue (e.g. party, NGO report about a protest event, etc.)'
8	'statement on position/aims by a political actor (e.g. party programme, NGO self description etc.)'
9	'conference papers'
10	'research report'
11	'seminar paper (Hausarbeit)'
12	'presentation of survey/poll results'
13	'online publication of a book or paper'
14	'press release'
15	'interview'
16	'print version of a public speech'
17	'contribution in a discussion forum'
18	'announcement of / call for a demonstration, protest event, etc. '
19	'petitioning/signature collection'
20	'information about a seminar (university)'
21	'presentation / advertisement for publications (without the original text)'
22	'information about a meeting, conference (announcement, program)'
23	'table, chart without text'
24	'publication of a political decree/decision'
25	'publication of a legal act'
26	'unspecified text, article'
27	'parliamentary and governmental report/record'

Variable **CONTOP**
'in which temporal context is the issue discussed/presented/commented'

1 'current context'
2 'general context' (not time specific)
3 'historical context'

Note: The decisive factor is the publishing date of the article. An article written in 2000 has a current context if it comments on an issue that was current at this time. If the issue is discussed in several temporal contexts you should code the most recent one (current → genera l→ historical)

Variable **THEAD** (string variable)
'headline of text'

Note: only the main (largest print) headline. If the text has no headline enter 0.

Variable **TLAN1** (TLAN2, TLAN3…TLAN7)
'Is the text also offered in further languages than your search language?'

0 'no further language
1 'Dutch'
2 'English'
3 'French'
4 'German'
5 'Italian'
6 'Spanish'
7 'Other regional language(s)' (e.g. Catalan, Welsh, Frisian)
8 'Other EU language(s) '
9 'Other language(s) '

Variable **TOTLAN** (String Variable)
'Specification of other Languages'

Note: If the text is offered in more than one "other" language you should enter them in the form: Russian, Chinese, Japanese

Filter-variable **TCLAIM**
'Does text contain claims?'

    0   'no'
    1   'yes'

IF 0 > ON TO **TTOPSCOP**: IF 1 > ON TO **AUT**

Variable **TTOPSCOP**
'scope of text topic'

1   'supranational: United Nations'
2   'other supranational'
3   'European Union'
4   'other European supranational'
5   'multilateral'
6   'bilateral'
7   'national'
8   'regional'
9   'local'
99 'unclassifiable'

Variable **TCOUNTRY**

AUTHOR

Variable **AUTS**
'summary author of the text'

0	'no author indicated / author and medium are the same'
10	'whole polities'[2]
20	'politicians' (if unspecified and unclear whether referring to government, parliament or parties)
30	'former states(wo)men'
40	'government/executive'[3]
50	'legislative'[4]
60	'judiciary'[5]
70	'police and internal security agencies'[6]
80	'military'[7]
90	'central banks'
100	'social security executive organizations' (incl. state pension funds)
110	'other state executive agencies'[8]
120	'political parties'[9]
130	'unions and employees'[10]
140	'employers organizations and firms' (incl. private pension funds; excl. agriculture)
150	'churches and religious organizations and groups'
160	'general media / journalists'
161	'issue specific media / journalists'
170	'farmers and agricultural organizations'

---

[2] Whole countries/polities: e.g., 'the EU', 'Britain', if used not to refer to the government or any other specific institution, but to the polity as a whole; note that 'the Europeans', 'the British', etc. are coded in 900: general public.

[3] Governments and government representatives (spokespersons, ministers, royalty etc.) irrespective of territorial scope. The EU-Commission and Council of Ministers, the UN General Secretary and Security Council are coded as governments. Other examples: mayor, Landesregierung, ministry of education.

[4] Legislatives and parliaments (all chambers), including individual members thereof, including parliamentary fractions of political parties. Examples: Bundestag, House of Lords, local councils, parliamentary fraction of the SPD, Labor MPs. Intergovernmental organizations which draw up international treaties on the basis of unanimous consent of the signatories are coded among executive/government The European Parliament and the General Assembly of the UN are, however, coded as legislatives, because they have (limited, but still) the power to make binding decisions on the basis of majority decisions.

[5] E.g., European Court of Justice, openbaar ministerie (public prosecutor), individual judges, juries.

[6] E.g., police, marechaussee, Bundesgrenzschutz, secret service, Verfassungsschutz, Europol. Note: the Police Union is coded as a union.

[7] E.g., Bundeswehr, NATO.

[8] E.g., ILO, WHO, Einwohnermeldeamt, Schulaufsichtsbehörde.

[9] This category should be used only for parties as parties, e.g., party chairman, party congress, "die SPD", "a Labor party spokesman", as well as for sub-organizations of parties (e.g., Junge Sozialisten). Note that the same person may be coded differently according to the way in which her or his position is described: e.g., Bundeskanzler Schröder is coded as government, Mitglied des Bundestages Schröder is coded as legislative, SPD-Parteivorsitzender Schröder is coded as political party.

[10] Includes the general categories "workers" and "employees".

180 educational professionals and organizations' (incl. schools, universities in their educational capacity)
190 'economists and financial experts'
200 'other scientific and research professionals and institutions' (e.g., experts, research institutes, universities in their research capacity)
210 'students, pupils, and their parents'
220 'other professional organizations and groups'[11]
230 'consumer organizations and groups'
240 'migrant organizations and groups'[12]
250 'pro- and anti-European campaign organizations and groups'
260 'solidarity and human rights organizations'[13]
270 'welfare organizations'[14]
280 'racist and extreme right organizations and groups'[15]
290 'peace movement organizations and groups'
300 'organizations and groups of the elderly'
310 'women's organizations and groups'
320 'environmental organizations and groups'
330 'terrorist groups'
340 'rebel forces/guerrilla' (e.g., UCK, Northern Alliance)
350 'other civil society organizations and groups'[16]
360 'whole economies'
370 'Private individual'
380 'ISP, portal, search engine'
900 'the general public' (e.g., 'citizens', 'the citizenry', 'die Öffentlichkeit', 'the electorate', 'the Germans', 'the population', 'taxpayers'; only if explicitly mentioned!)
999 'unknown'

Variable **AUTNAM**
'name of author'

Note: not only the name of individuals but of all actors (institutions, groups, etc.)

---

[11] E.g., Deutscher Ärtztekammer, Berufsverband Deutscher Psychologen, Deutscher Sportbund, doctors, football players, writers, sollicitors, musicians. Note: unions are always coded as unions, non-union organizations of police and judges are coded under their respective institution.

[12] Organizations or groups of immigrants, asylum seekers, ethnic minorities of migrant origin. Includes migrant organizations that would also fit another category. E.g., Islamic or Jewish organizations in Europe are coded as migrant organizations, not as religious groups, the European Association of Turkish academics is coded as a migrant organization, not as a professional organization, etc.

[13] This includes only private organizations such as Pro Asyl, Anti-Racist Alliance, Arbeiterwohlfahrt, Amnesty International, Terre des Hommes, medecins sans frontières etc.

[14] E.g. Red Cross, Arbeiterwohlfahrt, not state welfare agencies (these are coded as state executive agencies).

[15] Includes vague descriptions such as "skinheads" or "right-wing extremists". Radical right parties should be coded as political party, unless the party label is merely window dressing and does not indicate significant involvement in the electoral process (e.g. the Freiheitliche Arbeiterpartei in Germany).

[16] Including not already mentioned social categories such as 'youth', 'the unemployed', 'children', etc.

Variable **AUTTYP**
'Type of author'

1 'unorganised collective or anonymous representatives thereof' (e.g., 'farmers', 'a farmer')
2 'named representative(s) of an unorganised collective' (e.g., 'farmer X')
3 'organization or institution' (e.g., the 'National Union of Farmers – NUF')
4 'anonymous spokesperson(s) for organization/institution' (e.g., 'critics within the NUF leadership')
5 'named spokesperson(s) for organization/institution' (e.g., 'X, the President of the NUF')
6 'private individual'
99 'not identifiable'

Variable **AUTSCOP**
'Scope of author'

1 'supranational: United Nations'
2 'other supranational'
3 'European Union'
4 'other European supranational'
5 'multilateral'
6 'bilateral'
7 'national'
8 'regional'
9 'local'
99 'unclassifiable'

Variable **AUTCOUN**
'Country of author'

Variable **AUTPAR1**
'Party affiliation of author'
Categories, see data entry file.

Variable **SOUS**
'summary source of the text'
same categories as AUTS except of "0 no source mentioned / source and medium are the same"

Variable **SOUSNAM**
'name of source'

Note: not only the name of individuals but of all actors (institutions, groups, etc.)

Variable **SOUTYP**
'Type of source'
same categories as AUTTYP

Variable **SOUSCOP**
'Scope of source'
same categories as AUTSCOP

Variable **SOUCOUN**
'Country of source'

Variable **SOUPAR1**
'Party affiliation of source'
Categories, see data entry file.

MEDIUM

Variable **MEDINFO**
'source of information on medium who offers/runs the website'

1   'provided on the website containing the text or via an internal link' (e.g. by clicking on the
    logo, "home", "about us", "who we are", etc.)[17]
2   'own research'[18]
3   'no information on medium available'[19]

Variable **MEDADD**
'address/URL of the homepage of the medium'

Variable **MEDS**
'summary of medium who offers/runs the website'
same categories as AUTS except of there is no "0   no author mentioned"

Variable **MEDTYP1**
'Type of medium'
same categories as AUTTYPE

Variable **MEDNAME** (string variable)
'name of medium who runs the website'

Note: not only the name of individuals but of all actors (institutions, groups, etc.)

---

[17] The offered information must contain more than the name of the medium. It should allow to code at least one of
the following Medium variables. Thus, you should not code "1" if the only information you get on the website that
contains the relevant text is the name or logo of the medium.

[18] Own research does not mean that you should try to get the Information from other sources than the website. You
should only code "own research" when the only way to get to the homepage or some page that offers information
about the actor is to "shorten" the current address (see page 9 in this codebook).

[19] If no information on the medium is available on its website but you are able to code some of the following
Medium variables on the basis of your own knowledge you should code them.

Variable **MEDSCOP**
'Scope of medium'
same categories as AUTSCOP

Variable **MEDCOUN**
'Country of medium'

Variable **MEDPAR1**
'Party affiliation of medium'
Categories, see data entry file.

GATE

Variable **GATS**
'summary of gate'.
same codes as for AUTS except of "0 gate and medium are the same"

Variable **GATTYP**
'Type of gate'
same categories as AUTTYPE

Variable **GATNAME** (string variable)
'name of gate '

Note: not only the name of individuals but of all actors (institutions, groups, etc.)

Variable: **GATSCOP**
'Scope of medium'
same categories as AUTSCOP

Variable **GATCOUN**
'Country of gate'

Variable **GATPAR1**
'Party affiliation of gate'
Categories, see data entry file.

PART IV:
CLAIM VARIABLES

IDENTIFICATION AND LOCATION

Variable **TIDC** (dropdown list)
'number of the text where the claim is reported'

Variable **CID**
'identification number of the claim'

Variable **CTITLE** (string variable)
'description of claim'

Variable **CDAY**
'day of claim'
(1-31)

Variable **CMONTH**
'month of claim'
(1-12)

Variable **CYEAR**
'year of claim'
(two digits, e.g. 85=1985; 99=1999; 0=2000, 1=2001)

Variable **COUNTRY**
'Country where the claim was made'

Variable **REGION** (only to be coded for your own country; otherwise enter 0)
'region in which claim was made'

Variable **CITY**
'city in which claim was made'

Note: only if country is one of our seven and the city is among the twenty largest cities inclusive of the national capital, plus Brussels, Luxemburg, and New York in the case of claims with an EU- or UN-dimension

CLAIMANTS: SUBJECT ACTORS

Variable **ACT1S** (ACT2S, ACT3S)
'summary first actor'.
same categories as AUTS except of "0   none"

Variable **ACTTYP1** (ACTTYP2, ACTTYPE3)
'Type of first actor'.

1   'unorganized collective or anonymous representatives thereof' (e.g., 'farmers', 'a farmer')
2   'named representative(s) of an unorganized collective' (e.g., 'farmer X')
3   'organization or institution' (e.g., the 'National Union of Farmers – NUF')
4   'anonymous spokesperson(s) for organization/institution' (e.g., 'critics within the NUF leadership')
5   'named spokesperson(s) for organization/institution' (e.g., 'X, the President of the NUF')

Variable **ACTNAME1** (ACTNAME2, ACTNAME3)
'name of actor'

Note: not only the name of spokesperson but also the name of institutions/organisations

Variable: **ACTSCOP1** (ACTSCOP2, ACTSCOP3)
'Scope of first actor'.
See wp2 codebook for clarification and examples

1   'supranational: United Nations'
2   'other supranational'
3   'European Union'
4   'other European supranational'
5   'multilateral'
6   'bilateral'
7   'national'
8   'regional'
9   'local'
99   'unclassifiable'

Variable **ACTCOUN1** (ACTCOUN2, ACTCOUN3)
'Country of first actor'

Variable **ACTPAR1** (ACTPAR2, ACTPAR3)
'Party affiliation of first actor'

Variable **ACT1** (ACT2, ACT3)
'first actor' (three digits)
same categories as in wp2

Variable **AUTONOM1** (AUTONOM2, AUTONOM3)
'online presence of claimant'

1   autonomous
2   dependent
99 unclassifiable

*Note: These two different forms in which political claims can become visible on the Net are
defined as: (1) **autonomous online presence of claimants**: the claimant is the same as the actor
who runs the website (medium) and as the actor who wrote the text containing the claim (author),
and (2) **dependent online presence of claimants**: the claimant depends on other actors to make
his claim publicly visible. In this case medium and/or author are different actors than the
claimant and/or an external gate in addition to the search engine exists.*

ACTION FORMS

Variable **FORM**
'form of action'
same categories as in wp2 codebook

Variable **ONOFF**
'online or offline form of action'

1   'offline'
2   'online'
3   'mixed'

NOTE: Variable ONOFF is directly related to FORM and defines if the action takes/took place online or offline (E.g. if the claimant published an letter online the code is FORM = 45 (public) letter and ONOFF = 1 online)

AIMS

Variable **ISSUE1S** (two-digit codes) (ISSUE2S, ISSUE3S)
'summary of ISSUE1'
same categories as in wp2

Variable **ISSUE1** (ISSUE2, ISSUE3) (string variable)
'first issue'

Variable **ISSCOP1** (ISSCOP2, ISSCOP3)
'scope of first issue'
For clarification and examples see wp2 codebook

1   'supranational: United Nations'
2   'other supranational'
3   'European Union'
4   'other European supranational'
5   'multilateral'
6   'bilateral'
7   'national'
8   'regional'
9   'local'
99 'unclassifiable'

Variable **ISCOUN1** (ISCOUN2, ISCOUN3)
'country to which issue refers'

Variable **ISPOS1** (ISPOS2, ISPOS3)
'Relation of issue position (aim) towards European integration'

-1 'negative'
 0   'neutral/ambivalent'
 1   'positive'