

CLEF - IST-2000-31002
Cross-Language Evaluation Forum



Test Collection Report for the CLEF 2002 Campaign: Deliverable 3.2.1

Deliverable Type: Public
Number: D3.2.1
Contractual Date of Delivery: September 2002
Nature: Pre-final

Work Package: WP 3
Task: 3.1 and 3.2

Name of responsible: Michael Kluck (IZ)

Contact info: kluck@bonn.iz-soz.de

Abstract

The topic creation and relevance assessment activity for the CLEF2002 campaign is described. Results are given.

Keyword List

test collections, multilingual comparable corpus, topic creation, relevance assessment

Document Evolution

Version	Date	Status	Notes
1.1	28 October 2002	Draft	Distributed to partners
1.2	04 November 2002	Final	

Table of Contents

1.	Introduction.....	3
2.	Topic Creation	3
2.1	Topic Structure.....	3
2.2	Topic Creation Phases.....	4
3.	Assessment Process and Guidelines	5
3.1	Results.....	6
3.2	Language Specific Statistics	6
4.	Technical Problems.....	7
5.	Conclusion	7
	Appendix	8
1.	CLEF 2002 Topic Creation Rules.....	8
2.	CLEF 2002 Assessment Rules.....	10
2.1	General aspects	10
2.2	Technical and procedural aspects.....	10
3.	English topic set for the main tasks (multilingual, bilingual, monolingual).....	12
4.	Topics for the GIRT task (in English)	20
5.	Topics for the AMARYLLIS task (in English).....	24

1. Introduction

This report covers the topic creation and the relevance assessment phase of creating the test collection for the CLEF 2002 campaign. This work was done within Workpackage 3. The aim of this work package is to create the test data (topic sets and relevance assessments) for the CLEF evaluation campaigns. It is subdivided into two tasks: T3.1 Topic creation; T3.2 Relevance assessment.

2. Topic Creation

2.1 Topic Structure

The participating groups in the evaluation campaign derive their queries in their preferred language from a set of topics which has been prepared by the language groups. The language groups are those consortium partners or associated partners that did the work of topic creation and/or relevance assessment for the collection languages offered in the CLEF 2002 campaign:

Language	Language code	Responsible Partner
Dutch	NL	U Twente
English	EN	NIST
Finnish	FI	U Tampere
French	FR	ELDA
German	DE	IZ
Italian	IT	ISTI-CNR
Spanish	ES	UNED
Swedish	SV	SICS

Table 1: Language Groups

These topics are designed to simulate real user information needs. Each topic consists of three parts: a brief title statement (title); a one-sentence description of the topic (description); a more complex narrative specifying more precisely the range and facets of the topic and giving the definite relevance assessment criteria (narrative). The topics are formatted using SGML tags which also indicate the language in which the topic is presented.

Three distinct sets of topics were created (i) for the mono-/bi-/multilingual tracks, (ii) for the GIRT social science database, (iii) for the Amaryllis multidisciplinary bibliographic database. An example of a topic used for the mono-/bi-/multilingual tasks in the CLEF 2002 Campaign is shown below in its English, French, Finnish and German version (German was the original language):

```
<top>
<num> C124 </num>
<EN-title> Common foreign and security policy (CFSP) </EN-title>
<EN-desc> Find information on establishing a common EU foreign and security policy. </EN-desc>
<EN-narr> Relevant documents will report on meetings, discussions, or political statements regarding the
establishment of common foreign and security strategies (CFSP) for the European Union. </EN-narr>
</top>
```

<top>
<num> C124 </num>
<FR-title> La Politique Étrangère et de Sécurité Commune (PESC) </FR-title>
<FR-desc> Trouver des informations sur la mise en place d'une politique étrangère et de sécurité commune dans l'UE. </FR-desc>
<FR-narr> Les documents pertinents rendront compte des réunions, débats et déclarations concernant la mise en place et les stratégies communes en matière de politique étrangère et de sécurité (PESC) au sein de l'Union Européenne. </FR-narr>
</top>

<top>
<num> C124 </num>
<FI-title> Yhteinen ulko- ja turvallisuuspolitiikka </FI-title>
<FI-desc> Etsi tietoa suunnitelmista EU:n yhteisestä ulko- ja turvallisuuspolitiikasta. </FI-desc>
<FI-narr> Relevantit dokumentit raportoivat kokouksista, keskusteluista tai poliittisista lausunnoista koskien Euroopan Unionin yhteisen ulko- ja turvallisuuspolitiikan perustamista. </FI-narr>
</top>

<top>
<num> C124 </num>
<DE-title> Gemeinsame Außen- und Sicherheitspolitik (GASP) </DE-title>
<DE-desc> Gesucht sind Dokumente zur Etablierung einer gemeinsamen Außen- und Sicherheitspolitik der EU. </DE-desc>
<DE-narr> Relevante Dokumente berichten über Treffen, Diskussionen oder politische Stellungnahmen zur Ausarbeitung einer gemeinsamen Außen- und Sicherheitspolitik (GASP) der Europäischen Union. </DE-narr>
</top>

The English version of all CLEF 2002 topics can be found in the Appendix.

2.2 Topic Creation Phases

The topic creation process consisted of several phases:

- preparation of topic proposals by each language group
- face-to-face discussion of proposed topics and decision on the final topics
- translation of the final topics into all main and additional languages
- cross-checking of topic translations for translation inconsistencies

During the preparation phase, ideas are formulated about possible information needs which could be answered by the corpora in use. This means the topics should be derived from the same time-span (1994 and partly 1995, or for the scientific tasks without time limit) and the same topicality as the corpora (news or scientific documents). After the topic ideas have been generated, they must be checked against the corpora for each language: there should be at least one possible answer and at most 50 possible answers when using the ZPRISE IR system provided by NIST with a carefully formulated search. This is to avoid zero results which can really harm evaluation statistics and also to avoid too many results which could make the task too simple and thus of little significance. Each language group prepared about 10-15 possible topics in their own language and provided a provisional translation into English. They also tested the topics developed by the other language groups on the corpora in their languages.

The topic meeting in which there was a face-to-face discussion between representatives from all language groups enabled each group to understand all topics in depth, to refine them in their original language, and to ensure a good translation into the other languages. After the discussion of the content and a check to ensure searchability of all proposed topics in each collection, the decision on the final set of 50 topics for the main tasks was made unanimously. A master file of the final topics was then produced.

In the third step all language groups translated those topics originally been developed in other languages into their own language. As far as possible the translations have been made from the source language into the target language, otherwise English was used as the pivot language for the translation, additionally other translations were used for cross-reference. The aim in all cases was to produced natural language renderings of the topics for each language rather than literal translations. After this stage, the master file contained a set of all topics in each of the main languages (DE, EN, ES, FR, IT, FI, NL).

At the end of the third stage, all topic sets were cross-checked for translation inconsistencies, unusual language uses and mistakes. The correct tagging was also checked against the CLEF DTD. The topics were then reordered randomly so that participants could not discover the source language of the single topics from the topic list. Translations of the topics into other languages (PT, RU, SV, ZH) were added later and were not cross-checked; Japanese translations were added to the test suite later on and were not used during the campaign. They have, however, been used since by groups for testing purposes..

For the scientific tasks, the topics were produced separately: 25 topics for both AMARYLLIS and GIRT. These topics were checked against the corpora as described above. They were then translated, from French to English for AMARYLLIS, from German to English and Russian for GIRT, and the translations were cross-checked.

In total 100 topics were provided for the different tasks with translations from one up to eleven languages.

3. Assessment Process and Guidelines

The data collections used within the CLEF 2002 evaluation campaign are in Dutch, English, Finnish, French, German, Italian, Spanish, Swedish¹.

The number of documents in large test collections such as CLEF makes it impractical to judge every document for relevance. Instead approximate recall techniques are calculated using pooling techniques. The results submitted by the participating groups are used to form a pool of documents for each topic and

¹ **French dataset:** Le Monde 1994 - Copyright LE MONDE S.A. distribution rights ELDA (Evaluations and Language Resources Distribution Agency).

German dataset: Frankfurter Rundschau, 1994-1995 - Druck- und Verlagshaus Frankfurt am Main, Frankfurt am Main (Germany); Der Spiegel, 1994-1995 - SPIEGEL Verlag, Hamburg (Germany).

Italian dataset: La Stampa 1994 - Copyright La Stampa, distribution rights Hypersystems Srl, Torino.

Spanish dataset: Spanish news agency data - Copyright Agencia EFE S.A.

Dutch dataset: NRC Handelsblad - 1994, 1995; Algemeen Dagblad - 1994, 1995 - Copyright PCM Landelijke dagbladen/Het Parool.

English data set: Los Angeles Times, distributed by NIST.

Finnish dataset: Aamulehti - 1994, 1995 - Copyright Aamulehti Oyj

Swedish dataset: TT - Tidningarnas Telegrambyrå, 1994, 1995 - Copyright Tidningarnas Telegrambyrå (TT) SE-105 12 Stockholm, Sweden

GIRT database: InformationsZentrum Sozialwissenschaften, Bonn, Germany.

AMARYLLIS database: INIST, CNRS, France

SDA collection: French, German, Italian Swiss news agency data - Copyright Schweizerische Depeschagentur, Switzerland.

for each language by collecting the highly ranked documents from all submissions. Relevance assessment was distributed over eight different sites and performed in all cases by native speakers. Considerable coordination was needed in order to guarantee consistency of results, over languages and over collections

IZ prepared assessment guidelines; these were based on those used by TREC but adapted to meet the specific needs of CLEF. These guidelines were given to all assessors in all language groups. When uncertainties arose during the assessment process, e-mail discussions were conducted to ensure the homogeneity of relevance judgements.

3.1 Results

Both the number of documents in the collections and the number of runs submitted by the participants increased in CLEF2003. As it is not possible to judge all the top 1000 documents of each submitted run, similarly to CLEF 2000 and CLEF 2001, we were unable to include all runs and all submitted documents in the pool of documents to be judged. The pool depth indicates the maximum number of documents which are taken from each run into the pool of documents which are to be assessed. The pool depth was established at 60 as for CLEF 2001. Therefore we had to select runs to be judged; this is called the restricted pool. Several considerations were taken into account in order to keep the pool at a manageable size: the preferences of the participants, the “originality” of the tasks, or the topics, or the fields used, or the languages/language combinations, the coverage of the participants’ runs and tasks etc. However, all groups are represented in all the target languages they used. At the end of the pooling process, 181 runs of the 285 submitted were included in the restricted pool and judged by the assessors.

language/task	all results		restricted pool		
	number of documents	documents per topic (average)	per cent of all results	number of documents	documents per topic (average)
German	26.049	520	76	19.762	395
English	23.924	478	91	21.874	437
French	26.725	534	65	17.465	349
Italian	23.692	473	75	18.019	360
Spanish	30.119	602	65	19.668	393
Dutch	29.497	589	71	20.957	419
Finnish	9.825	328	100	9.825	328
Swedish	12.841	257	100	12.841	257
GIRT	10.323	412	100	10.323	412
AMARYLLIS	6.950	278	100	6.950	278
sum/average	199.945		79	157.684	

Table 2: Statistics on collections and returned documents per topic

3.2 Language Specific Statistics

After pooling, the delivered result data of the participants have been split into language packages which had to be assessed by the respective language groups. The amount of documents which have been regarded as relevant is shown in the following Table 2.

Task	Language of documents	no. assessed docs	no. relevant docs	% relevant docs
CLEF general (mono-, bi-, multilingual)				
	Dutch	20.957	1.862	8,89
	English	21.874	821	3,75
	Finnish	9.825	502	5,11
	French	17.465	1.383	7,92
	German	19.762	1.938	9,99
	Italian	18.019	1.072	5,95
	Spanish	19.668	2.854	14,51
	Swedish	12.841	1.196	9,31
sum (CLEF general)		140.411		..8,30
GIRT				
	German	10.323	961	9,31
Amaryllis				
	French	6.950	2.018	29,04
sum (all tasks)		157.684		
iCLEF (using CLEF 2001 topics)		only additionally discovered documents	only additionally discovered relevant documents	
	English			
	Finnish	754	35	4,6
	German	1.975	16	0,8

Table 3: Statistics on documents per language

4. Technical Problems

The installation of the Unix based ASSESS software requires considerable familiarity with this operating system. NIST provided remote consulting services on the use of ASSESS. But the installation caused severe problems for ELDA; these were solved by the direct support of IZ who sent a technician to their site.

5. Conclusion

All aims and timelines were met in this work package. Thus, the schedule of the CLEF 2002 campaign was respected.

The work load for the language groups was quite high, but nearly equally distributed (except Finnish and Swedish). The very short time frame for the assessments necessitated very concentrated and hard work. All partners and sub-contractors have done this with additional volunteer support from their staff and institutes.

Appendix

These following rules were circulated to the groups responsible for the topic creation and the relevance assessment work in order to ensure that each group was working in the same way.

1. CLEF 2002 Topic Creation Rules

The topic generation will be carried out by the different CLEF language groups (DE, EN, ES, FI, FR, IT, NL). Topics created should refer mainly to the year 1994 but can also be relevant to news of 1995 (to cover the Finnish data which covers the period 1.11.94 - 31.12.95, and the German data of DER SPIEGEL which covers 1994 and 1995) but can also regard general issues which are not so dependent on a specific year.

Great effort must be made to balance the focus of the topics with respect to general/ international, European (covering real European problems, not only occurring in Europe), national/ regional issues. Each of these topic groups should roughly cover a third of all topics. Additionally other issues such as proper nouns (names, places, etc.), acronyms questions which require factual answers, e.g. "where was John Lennon born?" must also be included; we try to find two factual queries per language group.

Please make sure that you do not repeat topics from earlier campaigns. You can check the topics used for CLEF 2000 and 2001 on the CLEF Web site, in the Workspace for Registered Participants (user name=clef2002; password=wiesp02). It is also useful to look at the former topics to become familiar with the style of wording for the different fields (Title, Description, Narrative). Title contains up to two or three words which gives the main idea of the topic. Description contains a full sentence or question describing the topic in short. Narrative contains a broader description of the topic including examples and perhaps mentions aspects that should not be counted as relevant. The Narrative is also seen as the instruction for the assessor as to how the topic should be interpreted. In case of factual queries, the answers should be provided as an addendum to the Narrative in square brackets and in a separate row which will only be shown to the assessors.

There should be some balance between the document collections with respect to the possible hits in each language. But as in real life not all topics have to have many hits in each language: for instance regional/national topics may not be reported in all countries at length, thus there may be very few or none hits in some collections.

The language groups should test their topic ideas against the data collections of their language. There is no restriction to the number of prospective search results, no upper bound, no lower bound (although for any collection we should avoid a non-hit/one-relevant-document-only hit situation in the pre-tests, because the pooling and the statistics of the results could be negatively affected). But in most cases there should be some hits in each of the main languages and in Dutch and Finnish, and not too many hits in any of them. This means in most of the cases there should be between 5 and 50 hits in the pre-tests.

The pre-tests are carried out using the ZPRISE system (which is provided by NIST and runs under the operating systems Solaris or Unix) on the collection of the respective language. ZPRISE performs a probabilistic search and provides a relevance feedback functionality which allows to mark relevant documents and to recalculate the probabilities after the user feedback. It also makes it possible to derive additional, statistically significant terms automatically. These can be used to expand the original query. Within the first 25 results delivered by the ZPRISE system, there should be 2 to 10 relevant hits, in the first 50 or 100 results, there should be 2 to 25 relevant hits (up to 50 if the topic is a local/regional one).

These numbers are rough indications that there will be a chance for any retrieval system to find relevant documents in the respective collections. If the numbers in the pre-test are too high, there will be no real challenge for systems to find relevant documents.

Each language group should propose about 10 topics for each main language, Finnish, and Dutch (some Japanese topics will also be added: we will get some proposals for topics which are running in the NTCIR campaign). Each group should circulate its topic proposals to the others by 27.02.2002. Before the meeting in Berlin, everyone should do their best to check these proposals against their own document collection in order to know in advance what kind of results can be expected.

From the more than 80 topic proposals 50 topics have to be chosen as final topics. The decision on the final topics will be made in a face-to-face discussion of the language groups at the meeting in Berlin (7.3.-8.3.2002). The main purpose of this intense discussion is to understand the topic content precisely and to give all groups a full understanding of the meaning and of the possible translation into their language. Thus, each group is aware that the respective topic is really feasible and translatable in(to) their language. The participants in this meeting will be the mediators for the assessors in the relevance assessment phase. They should brief them in advance and in case of uncertainty, they must explain the meaning of the topics to try to make sure that all language groups make their decisions in nearly the same way.

In the end a decision will be taken on the definite set of 50 topics. The discussion has to make sure that there will be a balance in the meaning of the topics in the different languages. The definite set of topics will be prepared as so-called Master Set which, for each topic, will contain the source topic and an English version. This Master Set is the reference set for the next stage which is the preparation of the common topic set in all languages.

As far as possible the translations should be done directly from the source language; otherwise the English version of the topic can be used as a help or as pivot language. In the end, there will be a revision of the translations in all main languages to ensure the consistency and accuracy of the topic sets in each language.

For the scientific data collections, Amaryllis and GIRT, 25 scientific/domain-specific topics will be developed in French (by ELDA/INIST) and German (by IZ). For these, an English and in the case of GIRT also a Russian translation have to be prepared.

At the end of the entire process (creation of Master Set, renderings in all languages, checking and revision), there will be official topic sets in English, German, French, Italian, Spanish, Dutch, Finnish and Russian (for GIRT only) provided by the respective language groups. Additional topic sets will be prepared by volunteer groups for Swedish, Russian, Chinese, Thai, Japanese and Portuguese, and can be used by CLEF participants for multilingual and bilingual runs. Topic sets can also be prepared in other languages by participating groups; runs will be accepted for such topics but will be given the status of "unofficial" runs.

Time Schedule for Topic Creation

17.2. IZ circulates topic creation rules and format

18.2. all groups (DE,EN,FR,IT,ES,NL,FI,JP) start to find topics

27.2. all groups deliver 10 proposals of topics for their own language and distribute them to all other groups including a rough English translation

7.-8.3. meeting on topic creation in Berlin (discussion and decision on the selection of the final topics)

11.3. IZ sends out the list of the final topics and English master set

17.3. partners give back the translation of all topics into their own language to IZ

18.3. IZ sends all topics + translations to Univ. Hildesheim

26.3. checked topics back to IZ

2. CLEF 2002 Assessment Rules

2.1 General aspects

The relevance assessment is binary: relevant - irrelevant, --> no category for undecided, no grading is possible!

The assumption for the assessor is that she or he has to write a report on the given topic and to decide which documents are related to this topic and would be worthwhile to be taken into consideration.

If any part of the document is relevant, then the whole document is regarded as relevant. But be aware of possible restrictions which are mentioned in the narrative, for instance “only mentioning the name is not enough” or “only documents containing figures are relevant”.

The relevance decision has to be taken regardless how other documents are judged, only the present document counts for the relevance decision. If the same documents occurs several times it has to be judged in the same way as at the first occurrence, duplicates are not to be eliminated!

In case of uncertainty consult the whole text of the topic again (title, description, narrative). Especially the narrative should give some clarification such as “if X is mentioned, the documents are also relevant” or “also X is of interest”.

Be strict with your decisions, but “in dubio pro reo”: if a document seems to be somehow related to the topic, count it as relevant.

But do not extend the meaning of the topic. Example: If the topic is on the traffic through the Lötschberg Tunnel, general aspects of the mountain “Lötschberg” like skiing or tourism are not relevant.

Real exclusions are indicated in the narrative: “documents are not relevant, if X” or “documents only giving the name X are irrelevant”.

2.2 Technical and procedural aspects

The Assess software of NIST offers a highlighting functionality for important terms which makes it easier to discover relevant documents as these key phrases are emphasised and you can scroll through the document quite quickly. Enter relevant phrases at the top of the document page and then apply them to

the pooled document set: they will be highlighted whenever they occur in the document texts. But be aware of spelling variants like plurals, umlauts (Ä, Ö, Ü, ß) and combined words or compounds, it is not sure whether they will be processed (try some variants at the beginning).

If there are some sub-regions of languages covered by the documents like American and British English, German in Germany, Switzerland or Austria, French in France or Switzerland, Italian in Italy or Switzerland, you should think of different wordings to find the important terms or key phrases. An example for the German language: in Germany “Abschiebung” equals in Switzerland “Ausschaffung“ which mean in both cases expulsion/deportation/extradition.

Probably you discover such different wordings or additional aspects or facets of the topic after having judged a lot of documents; then you have to rerun the assessment from the beginning to make sure that you did not miss documents containing this wording or these aspects.

If you are not sure about the topic, do a first run discarding all clearly irrelevant documents. In the second round (after rereading the whole topic!) you might be more aware of the topic and you could probably discuss the criteria with your collaborators or the language coordinator who has been involved in the topic creation phase/meeting and might give you additional hints on the common understanding of the topic of all language groups.

Michael Kluck (IZ)

07.06.2002

3. English topic set for the main tasks (multilingual, bilingual, monolingual)

<top>

<num> C091 </num>

<EN-title> AI in Latin America </EN-title>

<EN-desc> Amnesty International reports on human rights in Latin America. </EN-desc>

<EN-narr> Relevant documents should inform readers about Amnesty International reports regarding human rights in Latin America, or on reactions to these reports. </EN-narr>

</top>

<top>

<num> C092 </num>

<EN-title> U.N. sanctions against Iraq </EN-title>

<EN-desc> What measures has Iraq taken to effect the lifting of the U.N. economic embargo and political sanctions imposed after its invasion of Kuwait in 1990? </EN-desc>

<EN-narr> Documents must include ways in which Iraq has attempted to get the sanctions lifted. Mere descriptions of the sanctions or rhetoric against the sanctions are not relevant. Expressions of regret for invading Kuwait by Iraqi officials are relevant. </EN-narr>

</top>

<top>

<num> C093 </num>

<EN-title> Eurofighter </EN-title>

<EN-desc> Find documents which report about the EFA or the "Eurofighter" project. </EN-desc>

<EN-narr> Relevant documents report about the EFA project. The partner countries Germany, the United Kingdom, Italy and Spain formed the "Eurofighter" syndicate. Information on the work share of the groups involved as well as cost estimates also are of interest. </EN-narr>

</top>

<top>

<num> C094 </num>

<EN-title> Return of Solzhenitsyn </EN-title>

<EN-desc> Find documents which report about the return of the Nobel prize winner for literature Solzhenitsyn to Russia. </EN-desc>

<EN-narr> Relevant documents report the reasons and the time of the return of Solzhenitsyn to Russia. They may also talk about the reasons for his emigration to the US. </EN-narr>

</top>

<top>

<num> C095 </num>

<EN-title> Conflict in Palestine </EN-title>

<EN-desc> Find articles dealing with armed conflicts in the Palestinian territories and the involvement of a part of the civil population. </EN-desc>

<EN-narr> Relevant documents will provide information on recent events in Israeli-Palestinian conflicts. Reports on suicidal attacks will also count. </EN-narr>

</top>

<top>

<num> C096 </num>

<EN-title> Should González resign? </EN-title>

<EN-desc> Public requests by political personalities in Spain for

Deliverable 3.2.1: Test Collection Report for the CLEF 2002 Campaign

Felippe González to resign </EN-desc>

<EN-narr> Documents should report on any of the various occasions in which the Spanish president Felipe González was asked to resign by his political opponents. Reasons for these demands must be mentioned. </EN-narr>

</top>

<top>

<num> C097 </num>

<EN-title> Referendum on the independence of Moldova </EN-title>

<EN-desc> Find documents, that report about the referendum on the independence of the Republic of Moldova. </EN-desc>

<EN-narr> Relevant documents report about the referendum in which the majority of the population has voted in favour of the independence of Moldova, i.e. also against a union with Romania. The background on and the consequences of the decision as well as the results of the referendum in different regions, are of interest. </EN-narr>

</top>

<top>

<num> C098 </num>

<EN-title> Films by the Kaurismäkis. </EN-title>

<EN-desc> Search for information about films directed by either of the two brothers Aki and Mika Kaurismäki. </EN-desc>

<EN-narr> Relevant documents give the names of one or more titles of films directed by Aki or Mika Kaurismäki. </EN-narr>

</top>

<top>

<num> C099 </num>

<EN-title> Denying the Holocaust </EN-title>

<EN-desc> Find documents dealing with measures in Germany to limit public actions denying the Shoah. </EN-desc>

<EN-narr> Information is required on measures that have been taken to limit or forbid speeches or publications that clearly deny the Holocaust. Relevant documents can also discuss issues regarding freedom of speech in Germany with respect to expression of opinions that deny the existence of the Holocaust. </EN-narr>

</top>

<top>

<num> C100 </num>

<EN-title> The Ames espionage case </EN-title>

<EN-desc> Find documents that show the impact of the Ames espionage case on U.S.-Russian relations. </EN-desc>

<EN-narr> A relevant document will mention explicitly some effect of the Ames espionage case on U.S.-Russian relations. Simple accounts of the case are not relevant. </EN-narr>

</top>

<top>

<num> C101 </num>

<EN-title> Cyprus and the EU </EN-title>

<EN-desc> Can Cyprus become a member of the EU? </EN-desc>

<EN-narr> Political relations between Turkey and Greece make the integration of Cyprus into the EU problematic. Relevant documents will provide information on the problems and difficulties. </EN-narr>

</top>

<top>

<num> C102 </num>

<EN-title> Victories of Alberto Tomba </EN-title>

<EN-desc> Find documents that report on ski races won by Alberto Tomba. </EN-desc>

Deliverable 3.2.1: Test Collection Report for the CLEF 2002 Campaign

<EN-narr> Only documents that mention competitions that Tomba won are relevant. Pertinent documents regard any type of ski race in any kind of championship. It is not necessary that the document cites explicitly the type of race as long as Tomba is mentioned as the winner. </EN-narr></top>

<top>

<num> C103 </num>

<EN-title> Conflict of Interests in Italy </EN-title>

<EN-desc> Find documents that discuss the problem of the conflict of interests of the Italian prime minister, Silvio Berlusconi. </EN-desc>

<EN-narr> Relevant documents will refer explicitly to the conflict of interests between Berlusconi politician and head of the Italian government and Berlusconi business man. They may also contain information on proposals made or solutions adopted to resolve this conflict. </EN-narr>

</top>

<top>

<num> C104 </num>

<EN-title> Super G Gold medal </EN-title>

<EN-desc> Who won the gold medal in the super G in Lillehammer at the Olympic Winter Games 1994? </EN-desc>

<EN-narr> Find the name of the gold medallist in the super G at the Olympic Winter Games 1994 in Lillehammer, Norway. Only documents after the victory in the super G are relevant. Also of interest are documents describing the consequences for German skiing. </EN-narr>

</top>

<top>

<num> C105 </num>

<EN-title> Bronchial asthma </EN-title>

<EN-desc> Find documents that talk about bronchial asthma and recent medical discoveries concerning this respiratory disease. </EN-desc>

<EN-narr> Documents that discuss bronchial asthma and its possible causes (in particular, environmental conditions such as pollution and smog) are relevant. Also of interest are documents that provide information on new therapies to cure this illness. </EN-narr>

</top>

<top>

<num> C106 </num>

<EN-title> European car industry </EN-title>

<EN-desc> Find documents which report about the situation in the European car industry regarding the fall in sales (sales crisis) and possible countermeasures. </EN-desc>

<EN-narr> Relevant documents report about the sales crisis of the European car industry and the countermeasures of the industry, the car dealers and the European states. Information on the connection between the economic downturn and sales problems of the automobile business is also relevant. </EN-narr>

</top>

<top>

<num> C107 </num>

<EN-title> Genetic Engineering </EN-title>

<EN-desc> How does genetic engineering effect the human food chain? </EN-desc>

<EN-narr> Articles must directly address the introduction of genetic engineering, and its effects on the human food chain. They will discuss both pros and cons. Reports on tobacco bioengineering and human gene engineering are not relevant. </EN-narr>

</top>

<top>

<num> C108 </num>

<EN-title> Southern Yemen Secession </EN-title>

<EN-desc> What are the social and economic consequences of the declaration of secession in Southern Yemen?
</EN-desc>

<EN-narr> All documents discussing the effects of the conflict that followed the declaration of secession in Southern Yemen on the economy and society of that country are relevant. </EN-narr></top>

<top>

<num> C109 </num>

<EN-title> Computer Security </EN-title>

<EN-desc> What is the status of computer security in regard to networked access? </EN-desc>

<EN-narr> Documents must refer to computer security with respect to the use of networks. Reports regarding only the physical security of computer systems are not relevant. </EN-narr>

</top>

<top>

<num> C110 </num>

<EN-title> Iranian Suspects Extradited </EN-title>

<EN-desc> France expelled to Tehran two Iranians suspected to have murdered Kazem Radjavi in Switzerland.
</EN-desc>

<EN-narr> Kazem Radjavi, a political opponent of the Tehran Islamic government was murdered in Switzerland. Two suspected Iranians living in France were not extradited to Switzerland but expelled to Iran. Relevant documents will report on the extradition and can also discuss the effects on the relations between France and Iran. </EN-narr>

</top>

<top>

<num> C111 </num>

<EN-title> Computer Animation </EN-title>

<EN-desc> Find discussions of the impact of computer animation on the film industry. </EN-desc>

<EN-narr> Any information related to the effect of computer animation on the film industry is relevant, for instance the economic viability of computer animated films. Documents describing computer animation for applications other than films are not relevant. </EN-narr>

</top>

<top>

<num> C112 </num>

<EN-title> Pulp Fiction </EN-title>

<EN-desc> What prize did the film Pulp Fiction by Quentin Tarantino win at the Cannes Film Festival. </EN-desc>

<EN-narr> The film Pulp Fiction with John Travolta received an award at the film festival in Cannes. Relevant documents must give the name of the award associated with the film. </EN-narr>

</top>

<top>

<num> C113 </num>

<EN-title> European Cup </EN-title>

<EN-desc> Find documents regarding the preliminary stages and other organisational issues regarding the 1996 European Football Cup. </EN-desc>

<EN-narr> Any information on the preparatory phase of the 1996 European Football Cup is of interest. Documents can discuss the organisation of the groups in the preliminary play-off rounds, can list the teams qualifying for the final phase, etc.. </EN-narr>

</top>

<top>

<num> C114 </num>

<EN-title> Civil War in Afghanistan </EN-title>

<EN-desc> Find all documents describing the political situation in Afghanistan that led to civil war after the deposition of the communist leader Najibullah. </EN-desc>

Deliverable 3.2.1: Test Collection Report for the CLEF 2002 Campaign

<EN-narr> The fall in Afghanistan of the communist government led by Najibullah in April 1992 was followed by a civil war, caused by the political disagreements between the factions of the prime minister Hekmatyar and the president of the Republic Rabbani. Only those documents that make explicit reference to the fall of the Najibullah government as a cause of the subsequent political and civil conflicts are of interest. </EN-narr>

</top>

<top>

<num> C115 </num>

<EN-title> Divorce Statistics </EN-title>

<EN-desc> Provide statistics on divorce rates in different countries. </EN-desc>

<EN-narr> Find information on international trends in divorce rates. Relevant documents will cite figures for divorce in any country in the world. </EN-narr>

</top>

<top>

<num> C116 </num>

<EN-title> The Snowboard </EN-title>

<EN-desc> Find documents that discuss the introduction of a new winter sport: the Snowboard. </EN-desc>

<EN-narr> Relevant documents talk about the introduction, use and risks of the Snowboard and may also provide details on accidents on the ski slopes caused by Snowboard users. Documents that talk in favour of this sport asserting that it is no more dangerous than downhill skiing are also to be considered relevant. </EN-narr>

</top>

<top>

<num> C117 </num>

<EN-title> European Parliament Elections </EN-title>

<EN-desc> How many voters were called to the polls to elect the new European parliament in 1994? </EN-desc>

<EN-narr> Relevant documents report on numbers of citizens in the European Union having the right to vote in the elections of the European parliament in 1994. Relevant documents can give numbers of voters for a single country or region or for the whole of the European Union. Documents that mention percentages of electors who actually voted are not relevant. </EN-narr>

</top>

<top>

<num> C118 </num>

<EN-title> Finland's first EU Commissioner. </EN-title>

<EN-desc> Who was appointed as the first commissioner for Finland in the European Union? </EN-desc>

<EN-narr> Give the name of Finland's first European Union commissioner.

Relevant documents may also mention the areas of interest for the new commissioner's assignment. </EN-narr>

</top>

<top>

<num> C119 </num>

<EN-title> Destruction of Ukrainian nuclear weapons </EN-title>

<EN-desc> Find documents which report about the agreement between the Ukraine, Russia and the USA on the removal and the destruction of Ukrainian nuclear weapons before June 1996. </EN-desc>

<EN-narr> Relevant documents report about the agreement between the Ukraine, Russia and the USA about the removal and the destruction of Ukrainian nuclear weapons. The number of weapons and details on the negotiations are also of importance. </EN-narr>

</top>

<top>

<num> C120 </num>

<EN-title> Edouard Balladur </EN-title>

<EN-desc> What is the importance for the European Union of the economic policies of Edouard Balladur? </EN-desc>

<EN-narr> Relevant documents will discuss the importance of the financial policies of Edouard Balladur, the French politician, for the economic unity of Europe. </EN-narr></top>

Deliverable 3.2.1: Test Collection Report for the CLEF 2002 Campaign

<top>
<num> C121 </num>
<EN-title> Successes of Ayrton Senna </EN-title>
<EN-desc> Provide a record of the sporting achievements of Brazilian race car driver Ayrton Senna. </EN-desc>
<EN-narr> Documents should report a list of sporting achievements of Formula 1 pilot Ayrton Senna before his death in 1994. </EN-narr>
</top>

<top>
<num> C122 </num>
<EN-title> North American car industry </EN-title>
<EN-desc> Find documents discussing the efforts of the automobile industry in North America to improve their sales in both domestic and international markets. </EN-desc>
<EN-narr> Information is needed on measures taken by the North American automobile industry to improve sales. Relevant documents can discuss increased productivity, price reduction, relations with trade unions, improvements in quality, etc.. </EN-narr>
</top>

<top>
<num> C123 </num>
<EN-title> Marriage Jackson-Presley </EN-title>
<EN-desc> Find documents that report on the presumed marriage of Michael Jackson with Lisa Marie Presley or on their separation. </EN-desc>
<EN-narr> In May 1994, the famous pop star, Michael Jackson, was reported to have married Lisa Marie Presley, the daughter of the king of rock and roll. Relevant documents must either contain some details regarding the wedding, such as where or when it was held, or must discuss the later separation of the couple. </EN-narr>
</top>

<top>
<num> C124 </num>
<EN-title> Common foreign and security policy (CFSP) </EN-title>
<EN-desc> Find information on establishing a common EU foreign and security policy. </EN-desc>
<EN-narr> Relevant documents will report on meetings, discussions, or political statements regarding the establishment of common foreign and security strategies (CFSP) for the European Union. </EN-narr>
</top>

<top>
<num> C125 </num>
<EN-title> European single currency </EN-title>
<EN-desc> What is the schedule predicted for the European single currency? </EN-desc>
<EN-narr> Speculations by politicians and business figures about a calendar for achieving a common currency in the EU are relevant. </EN-narr>
</top>

<top>
<num> C126 </num>
<EN-title> Actions against the fur industry. </EN-title>
<EN-desc> Find information on protests or violent acts against the fur industry. </EN-desc>
<EN-narr> Relevant documents describe measures taken by animal right activists against fur farming and/or fur commerce, e.g. shops selling items in fur. Articles reporting actions taken against people wearing furs are also of importance. </EN-narr>
</top>

<top>
<num> C127 </num>
<EN-title> Escape of Roldán </EN-title>

<EN-desc> Escape and flight of the Spanish Civil Guard ex-director
Luis Roldán </EN-desc>

<EN-narr> Documents should describe the escape of Luis Roldán from Spanish
justice and/or provide information on various incidents occurring during his flight. </EN-narr></top>

<top>

<num> C128 </num>

<EN-title> Sex in Advertisements </EN-title>

<EN-desc> Find articles dealing with eye-catching, sexually-oriented advertisements, and the consequences on
society. </EN-desc>

<EN-narr> Relevant documents will discuss the problem of the excessive use of sexually explicit advertisements.
Claims that this may lead indirectly to sexual attacks and aggression will also be considered of interest. </EN-narr>
</top>

<top>

<num> C129 </num>

<EN-title> 100 billion stars discovered </EN-title>

<EN-desc> What technological breakthroughs permitted the discovery of an exceptionally high number of stars or
galaxies in 1994. </EN-desc>

<EN-narr> Relevant documents will report about the new technology that has permitted the discovery of new
galaxies and stars, not about satellites or celestial bodies of our own solar system. </EN-narr>
</top>

<top>

<num> C130 </num>

<EN-title> Death of Nirvana leader </EN-title>

<EN-desc> How did the lead singer of the American rock and grunge group, Nirvana, die? </EN-desc>

<EN-narr> Kurt Cobain, lead singer of Nirvana, the famous popular music group, died in April 1994. Documents
that report the death of Cobain without mentioning the cause are not relevant. </EN-narr>
</top>

<top>

<num> C131 </num>

<EN-title> Intellectual Property Rights </EN-title>

<EN-desc> Find information on the rights of the developers of intellectual property and efforts by countries to
provide protection for the developers. </EN-desc>

<EN-narr> Relevant documents will discuss means of protecting intellectual property rights anywhere in the world.
Information on patent regulations is to be ignored. </EN-narr>
</top>

<top>

<num> C132 </num>

<EN-title> Area of Kaliningrad </EN-title>

<EN-desc> Find documents discussing the political or economic future of the Kaliningrad Exclave. </EN-desc>

<EN-narr> Only political or economic information on Kaliningrad is of interest. Prospects for future relations with
Scandinavia, the Baltic countries and Russia. Historical or tourist information is not important. </EN-narr>
</top>

<top>

<num> C133 </num>

<EN-title> German Armed Forces Out-of-area </EN-title>

<EN-desc> Find documents which report on political and juridical decisions on out-of-area uses of the Armed
Forces of Germany. </EN-desc>

<EN-narr> Relevant documents either report about political and juridical decisions on out-of-area uses (i.e. combat
missions outside the NATO area) of the German Federal Armed Forces or discuss consequences of this decision.
</EN-narr>
</top>

<top>

<num> C134 </num>

<EN-title> The Clementine space probe </EN-title>

<EN-desc> Find reports on the American space probe Clementine, launched to collect data about the moon and the near-Earth asteroid 1620 Geographos. </EN-desc>

<EN-narr> The main objectives of Clementine, the American space probe launched by the NASA and the Department of Defense, are to test the durability of sensors and micro-electronic systems and to collect information and data about the Moon and the Geographos asteroid. Relevant documents should provide some information about the technical and scientific aspects of the mission. </EN-narr>

</top>

<top>

<num> C135 </num>

<EN-title> EC presidency candidates </EN-title>

<EN-desc> Who was mentioned as a possible candidate to be president of the European Commission? </EN-desc>

<EN-narr> Documents should give the names of people reported as potential candidates for the presidency of the European Commission. </EN-narr>

</top>

<top>

<num> C136 </num>

<EN-title> Leaning Tower of Pisa </EN-title>

<EN-desc> What is the state of health of the Leaning Tower of Pisa? </EN-desc>

<EN-narr> Relevant documents will discuss the state of health of the Leaning Tower of Pisa and in particular will mention its inclination and variations in this. Documents that report efforts to reduce the inclination are also of interest. </EN-narr>

</top>

<top>

<num> C137 </num>

<EN-title> International Beauty Contests </EN-title>

<EN-desc> Find the names of winners of international beauty contests in 1994 or 1995. </EN-desc>

<EN-narr> Relevant documents will cite the names of males or females who won an international beauty or charm contest in either 1994 or 1995. Body-building competitions are not included. </EN-narr>

</top>

<top>

<num> C138 </num>

<EN-title> Foreign words in French </EN-title>

<EN-desc> Find documents dealing with the role of foreign words in French. </EN-desc>

<EN-narr> The Constitutional Council has reviewed and made less strict the Toubon law on the use of the French language. Relevant documents discuss the role and the use of foreign words in the French language. </EN-narr>

</top>

<top>

<num> C139 </num>

<EN-title> EU fishing quotas </EN-title>

<EN-desc> Find information about fishing quotas in the EU. </EN-desc>

<EN-narr> Documents report figures on or discussion of the fishing quotas of the EU member countries. </EN-narr>

</top>

<top>

<num> C140 </num>

<EN-title> Mobile phones </EN-title>

<EN-desc> Prospects for the use of cellular phones. </EN-desc>

<EN-narr> Relevant documents report on the prospects for the use of cellular phones and the development of the mobile phone industry. </EN-narr>

</top>

4. Topics for the GIRT task (in English)

<top>

<num> 051 </num>

<EN-title> Self-confidence of girls </EN-title>

<EN-desc> Find documents which report on the loss of self-confidence of young girls during the puberty. </EN-desc>

<EN-narr> The older girls become the more willingly they give up themselves and their demands, the more they lose their vitality and their self-assertion. Relevant documents report on this process of the loss of self-confidence of young girls. </EN-narr>

</top>

<top>

<num> 052 </num>

<EN-title> Youth and violence in films </EN-title>

<EN-desc> Look for documents which discuss the influence of violence in films on the social behaviour of young people. </EN-desc>

<EN-narr> Relevant documents report on connections between experiencing cinematic violence performance and social and biographical experiences of young people themselves. They also deal with the differences in the aggressiveness and propensity to violence and/or sex-specific differences in the reaction to violence oriented cinematic performance. </EN-narr>

</top>

<top>

<num> 053 </num>

<EN-title> Young people and social competence </EN-title>

<EN-desc> Find documents which represent educational materials as well as concepts or methods for the support of social competence of young people. </EN-desc>

<EN-narr> How can social competence or humane competence be promoted with special concepts, methods or educational materials for young people and why is this of such a big importance? </EN-narr>

</top>

<top>

<num> 054 </num>

<EN-title> Topics of the right wing extremists and the centre </EN-title>

<EN-desc> Find documents which report on common topics discussed by the right wing extremists and the political centre. </EN-desc>

<EN-narr> Which are the current topics of the extreme right wing groups? Are there any correspondence between the topics of centre and the right wing extremists? The extreme right and the political centre particularly find propagandistic support in socially relevant debates about German identity, multiculturalism and immigration for the dissemination of racism and nationalism. </EN-narr>

</top>

<top>

<num> 055 </num>

<EN-title> Relation of genders in further education </EN-title>

<EN-desc> Find documents which report on the relationship of genders in further education. </EN-desc>

<EN-narr> Relevant documents report on the existing relationship between the sexes in the area of further education and show its possible causes and reasons. They also may report on concepts which reduce the discrimination of women in this field. </EN-narr>

</top>

<top>

<num> 056 </num>

<EN-title> Interdependence between industrialised and developing countries </EN-title>

<EN-desc> Find documents on the context of consumer wants in highly developed industrial countries and its production possibilities in the developing countries. </EN-desc>

<EN-narr> Relevant documents shall discuss the western consumption and the southern production as well as the resulting economic interdependence between industrialised and developing countries. Only documents, which deal with this topic in the context of the north-south conflict, are relevant. </EN-narr>

</top>

<top>

<num> 057 </num>

<EN-title> Learning disabled pupils and discrimination </EN-title>

<EN-desc> Find documents which discuss learning situations of learning disabled pupils and show the discrimination resulting from the handicaps in a society of global knowledge. </EN-desc>

<EN-narr> The attendance at a special school for the individual support of learning is not caused by lacking intelligence of the majority of the pupils. It is rather a result of deficits in the socialization process of these pupils, which is insufficient with regard to the competences necessary for mastering life within our modern and highly educated society. This means, the conditions and structures, under which these children and young people grow up, do not promote studying sufficiently. </EN-narr>

</top>

<top>

<num> 058 </num>

<EN-title> Learning behaviour of pupils and change in values </EN-title>

<EN-desc> Find documents which report on problems in the learning behaviour of pupils which result from the societal change in values. </EN-desc>

<EN-narr> Which attitudes to social values and to societal models change and which problems arise from that with pupils. Documents can discuss concepts of the reduction or early recognition of these problems as well. Does the change in social values cause obstacles to the social integration of the pupils? </EN-narr>

</top>

<top>

<num> 059 </num>

<EN-title> Environmental awareness in Germany </EN-title>

<EN-desc> Find documents reporting on the development of the environmental awareness and the attitude to the environment in Germany. </EN-desc>

<EN-narr> Relevant documents report on the development of the environmental awareness in Germany, on the attitude towards the environment or the readiness for an environmentally aware behaviour. Documents generally speaking about environment without picking the environmental awareness out as a central theme aren't relevant </EN-narr>

</top>

<top>

<num> 060 </num>

<EN-title> Selection function of the school </EN-title>

<EN-desc> Find documents which report on the selection and assignment function of compulsory schools. </EN-desc>

<EN-narr> How do schools perform their selection function and which negative effects are resulting? Relevant documents can also discuss the predetermination of the social position which the pupils will reach later in his life. Discussions on the legitimacy of the school by perception of these functions are also important. </EN-narr>

</top>

<top>

<num> 061 </num>

<EN-title> Importance of vocational education </EN-title>

<EN-desc> Find documents reporting on the importance of vocational education in the present time. </EN-desc>

Deliverable 3.2.1: Test Collection Report for the CLEF 2002 Campaign

<EN-narr> Can one nowadays still talk about a career within a certain profession or rather of a work career, during which different professions are practised. Which consequences arise from this social development and does the vocational education still fulfil these tasks? </EN-narr></top>

<top>

<num> 062 </num>

<EN-title> Constraint by freedom of choice </EN-title>

<EN-desc> Find documents which discuss problems of the freedom of choice or curse and blessing of the existing choices. </EN-desc>

<EN-narr> In today's societies, people are offered a vast amount of individual possibilities for the arrangement of their own. One can decide which profession one wants to choose, control when and whether one would like to have a baby and even which sex one prefers. This freedom of choices or decisions leads many people into a coercive situation of choice. Concepts to avoid such a decision can be described as well. Documents which deal with elections in the political meaning aren't relevant. </EN-narr>

</top>

<top>

<num> 063 </num>

<EN-title> Collective processes of identity formation of Muslims </EN-title>

<EN-desc> Finds documents, in which the collective processes of identity formation of Muslims are discussed. </EN-desc>

<EN-narr> How do the respective doctrines of salvation of different Muslim groups (orthodoxy, modernism, fundamentalism, sufism) have an effect on the identity formation with regard to their public behaviour? </EN-narr>

</top>

<top>

<num> 064 </num>

<EN-title> Cultural conflicts </EN-title>

<EN-desc> Finds documents reporting on causes of problems and conflicts between members of different religions and/or cultures. </EN-desc>

<EN-narr> Relevant documents describe the reasons for conflicts between different religious groups. It is all about the mutual lack of understanding with regard to actions and opinions which are based on basically different conceptions of the world and then can lead to conflict situations: for example completely divergent opinions on the death in connection with suicidal attacks. </EN-narr>

</top>

<top>

<num> 065 </num>

<EN-title> Multicultural conflicts in school </EN-title>

<EN-desc> Find documents reporting on multicultural conflicts in school and their causes. </EN-desc>

<EN-narr> Relevant documents discuss multicultural conflicts in schools. What are the reasons for problems at schools with pupils of different cultural origin. Why do these problems arise and how could one prevent them? </EN-narr>

</top>

<top>

<num> 066 </num>

<EN-title> Psychosomatic complaints of nursing staff </EN-title>

<EN-desc> Find documents which report about significant psychosomatic complaints of outdoor nursing staff and rescue service employees. </EN-desc>

<EN-narr> Do significant psychosomatic troubles appear with rescue staff? Why and which stress is indicated? How can these troubles be prevented? </EN-narr>

</top>

<top>

<num> 067 </num>

<EN-title> Drug abuse </EN-title>

Deliverable 3.2.1: Test Collection Report for the CLEF 2002 Campaign

<EN-desc> Drug abuse with so-called designer drugs, Ecstasy etc. </EN-desc>

<EN-narr> Relevant documents report on the abuse of so-called designer drugs such as Ecstasy. General considerations on the drug abuse or reports on other drugs aren't relevant. </EN-narr></top>

<top>

<num> 068 </num>

<EN-title> Job satisfaction </EN-title>

<EN-desc> Find documents which discuss the different aspects of job satisfaction and of identification with the job or the profession. </EN-desc>

<EN-narr> Relevant documents discuss the satisfaction at work and the identification of the employees with their activity field. Which consequences does a lacking work satisfaction have and what does it arise from? </EN-narr>

</top>

<top>

<num> 069 </num>

<EN-title> Aggressiveness of car drivers </EN-title>

<EN-desc> Find documents reporting on the aggressiveness at driving a car and its causes. </EN-desc>

<EN-narr> Relevant documents report on the aggressiveness at driving a car. They also can pick out the possible connection between gender and aggressiveness at driving a car. </EN-narr>

</top>

<top>

<num> 070 </num>

<EN-title> Work and mothers </EN-title>

<EN-desc> Find documents which discuss the influence of the mothers' work within the first months after the delivery on the emotional relationship between mother and child. </EN-desc>

<EN-narr> What are the long-term consequences does mothers' work have on the baby's life in early months, on his emotions and on the way his forms his future interpersonal relations? </EN-narr>

</top>

<top>

<num> 071 </num>

<EN-title> East German youth and right-wing extremism </EN-title>

<EN-desc> Find documents which analyse reasons for the stronger slope of East German teenagers to the right-wing extremism. </EN-desc>

<EN-narr> Relevant documents discuss the stronger slope of East German teenagers to the right-wing extremism. Why are those more strongly endangered and which avoidance concepts exist? </EN-narr>

</top>

<top>

<num> 072 </num>

<EN-title> Welfare aid and labour market politics </EN-title>

<EN-desc> Look for documents which report on considerations for the integration of the welfare aid into labour market politics. </EN-desc>

<EN-narr> Relevant documents describe considerations and discussions on the inclusion of the welfare aid in labour market politics. Documents aren't relevant, if they merely discuss in general on active labour market politics, without concretely naming the inclusion of the welfare aid, or they generally talk about the change of the welfare aid. </EN-narr>

</top>

<top>

<num> 073 </num>

<EN-title> Security politics and Swiss neutrality </EN-title>

<EN-desc> Look for documents on the neutrality of Switzerland as the base of security politics. </EN-desc>

<EN-narr> Relevant documents discuss the security policy of Switzerland under the aspect of its neutrality. Documents discussing the security politics without an explicit reference to the neutrality of Switzerland aren't relevant. </EN-narr></top>

<top>
<num> 074 </num>
<EN-title> Slovenian minority in Austria </EN-title>
<EN-desc> Find documents which deal with the Slovenian minority in Austria and their rights. </EN-desc>
<EN-narr> Relevant documents report on the Slovenian minority in Austria, their problems and the discussion about their integration and their rights. Documents generally reporting on minorities in Austria or on other ethnic groups aren't relevant. </EN-narr>
</top>

<top>
<num> 075 </num>
<EN-title> Life without work </EN-title>
<EN-desc> Find documents which report on the social situation of unemployed citizens. </EN-desc>
<EN-narr> Relevant documents offer insights into the current life situation and the perspective of the persons affected by unemployment. Documents of interest are also, which give attention to special social problem constellations or labour market problem groups. </EN-narr>
</top>

5. Topics for the AMARYLLIS task (in English)

<top>
<num>001</num>
<EN-title>The impact of diesel engine on environment</EN-title>
<EN-desc>Air pollution by the exhaust of gas from diesel engines and methods of controlling air pollution. Pollutant emissions (NOX, SO2, CO, CO2, unburned product, ...) and air pollution control
<EN-narr>
<c>Concentration and toxicity of pollutant</c>
<c>Pollutant formation mechanism</c>
<c>Pollution prevention and reduction</c>
<c>Motor fuel selection</c>
<c>Combustion control</c>
<c>Exhaust gas treatment</c>
<c>Legislation and regulation</c>
</EN-narr>
</top>

<top>
<num>002</num>
<EN-title>The treatment of drinkable water and the purification of waste water through membranous technologies</EN-title>
<EN-desc>A relevant document will treat one or several membranous engineering used in the treatment, whether they are combined with other processes or not. It will describe one or several fundamental aspects of these technologies and their applications in laboratories.</EN-desc>
<EN-narr>
<c>Drinking water</c>
<c>Waste water</c>
<c>Industrial waste water</c>
<c>Urban waste water</c>
<c>Domestic waste water</c>
<c>Microfiltration</c>
<c>Nanofiltration</c>
<c>Ultrafiltration</c>
<c>Reverse osmosis</c>
<c>Dialysis</c>

<c>Electrodialysis</c>
<c>Pervaporation</c>
<c>Membrane</c>
</EN-narr>
</top>

<top>
<num>003</num>
<EN-title>The mechanism of fluids: The flow</EN-title>
<EN-desc>A relevant document will treat fluid flow caused by natural convection in a cavity, including practical applications</EN-desc>
<EN-narr>
<c>Natural convection</c>
<c>Flow</c>
</EN-narr>
</top>

<top>
<num>004</num>
<EN-title>Continuous training of teachers</EN-title>
<EN-desc>Information about further education of teachers including training methods, formation site and pedagogical materials</EN-desc>
<EN-narr>
<c>Formation continue</c>
<c>Continuous Training</c>
<c>Adult education</c>
<c>Permanent training</c>
<c>Life-Long Training</c>
<c>Staff Training</c>
<c>Further-Training</c>
</EN-narr>
</top>

<top>
<num>005</num>
<EN-title>Underground storage of natural gas</EN-title>
<EN-desc>Construction and equipment of storage tanks for natural gas: technical, economic and environmental aspects of underground storage of natural gas and LNG</EN-desc>
<EN-narr>
<c>Gas holders</c>
<c>Storage tanks</c>
<c>Geological structure</c>
<c>Geotechnics</c>
</EN-narr>
</top>

<top>
<num>006</num>
<EN-title>Fish breeding</EN-title>
<EN-desc>Physiological and ecological aspects of fish breeding and breeder behaviour. Relevant documents will concern gonad's development, sexual maturation, reproductive cycle, fecundity, influence of environmental factors on its breeding and its sexual behaviour</EN-desc>
<EN-narr>
<c>Reproductive cycle</c>
<c>Fish</c>

Deliverable 3.2.1: Test Collection Report for the CLEF 2002 Campaign

<c>Breeding</c>
<c>Fecundity</c>
<c>Breeder behavior</c>
<c>Sexual behavior</c>
<c>Courtship behavior</c>
</EN-narr>
</top>

<top>
<num>007</num>
<EN-title>Newborn and child diarrhoea</EN-title>
<EN-desc>Documents about epidemiology, diagnostic, complications, treatment and prevention infantile and neonatal diarrhoeas regardless of its causes. Documents about diseases in which diarrhoea is a usual manifestation like gastroenteritis, enteritis are considered as relevant </EN-desc>
<EN-narr>
<c>Diarrhea</c>
<c>Pediatrics</c>
<c>Dehydration</c>
<c>Rehydration</c>
<c>Bacteriology</c>
<c>Parasitology</c>
<c>Virology</c>
</EN-narr>
</top>

<top>
<num>008</num>
<EN-title>Pain during cancer evolution (solid tumours and malignant homeopathies)</EN-title>
<EN-desc>Relevant documents will include therapeutic aspects and will treat the pain as an epiphenomenon appearing during a malignant affection and during its therapy as well</EN-desc>
<EN-narr>
<c>Cancerology </c>
<c>Analgesia</c>
<c>Palliative treatment</c>
<c>Terminal stage</c>
<c>Quality of life</c>
</EN-narr>
</top>

<top>
<num>009</num>
<EN-title>Resistance to anticancer agents</EN-title>
<EN-desc>Documents about tolerating medicines which are used in chemotherapy of solid tumours or malignant homeopathies and also in chemo-resistance mechanisms</EN-desc>
<EN-narr>
<c>Cancerology</c>
<c>Chemotherapy</c>
<c>Multiple resistance</c>
<c>Cross resistance</c>
<c>Chemoresistance</c>
</EN-narr>
</top>

<top>
<num>010</num>
<EN-title>Psychology and social identity of athletes</EN-title>

Deliverable 3.2.1: Test Collection Report for the CLEF 2002 Campaign

<EN-desc>To be relevant, a document will include psychological factors and the role of sport in the socialization and the construction of the social identity of athletes. Sport intake in social rehabilitation of physically or mentally handicapped persons is also relevant. Documents about sport medicine, sport equipment and leisure sport are not considered as relevant</EN-desc>

<EN-narr>
<c>Sportsman</c>
<c>Sportswoman</c>
<c>Psychology</c>
<c>Sociology</c>
</EN-narr>
</top>

<top>
<num>011</num>
<EN-title>Athletes diseases</EN-title>
<EN-desc>All documents about accidents, acute or chronic incidents during sport practice</EN-desc>
<EN-narr>
<c>Pathology</c>
<c>Fracture</c>
<c>Sport medicine</c>
<c>Sportsman</c>
<c>Sportswoman</c>
</EN-narr>
</top>

<top>
<num>012</num>
<EN-title>The control of noise that comes from road transport, rail transport and from air traffic</EN-title>
<EN-desc>Relevant documents will concern all types of transport and noise and its health effects, comfort and techniques for estimating, controlling and neutralizing nuisances of a noise</EN-desc>
<EN-narr>
<c>Noise</c>
<c>Acoustics</c>
<c>Noise control</c>
<c>Acoustic monitoring</c>
<c>Acoustic nuisance</c>
<c>Comfort</c>
<c>Disturbance </c>
<c>Quality of Life</c>
<c>Air transportation</c>
<c>Transportation infrastructure</c>
<c>Airport</c>
<c>Jet engine</c>
<c>Road transportation</c>
<c>Motor car</c>
<c>Rail transportation</c>
<c>Urban transportation</c>
<c>Metrology</c>
<c>Standards</c>
<c>Laws and legislation</c>
</EN-narr>
</top>

<top>
<num>013</num>
<EN-title>Linkage mechanism</EN-title>

Deliverable 3.2.1: Test Collection Report for the CLEF 2002 Campaign

<EN-desc>All documents about motion transmission with Mechanical hinge. Other types of transmission like cams, gears and hydraulic control are also relevant</EN-desc>

<EN-narr>

<c>Linkage mechanism</c>

<c>Robotics</c>

<c>Manipulators</c>

<c>Mechanical hinge</c>

<c>Machine design</c>

<c>Lever</c>

<c>Crank</c>

<c>Slider crank mechanism</c>

<c>Bar mechanism</c>

<c>Engineering </c>

</EN-narr>

</top>

<top>

<num>014</num>

<EN-title>Smoking: Epidemiology and health effects</EN-title>

<EN-desc>Documents dealing whether with epidemiology (prevalence, evolution) of any kind of tobacco consumption (active or passive smoking, smoked or quid), or dealing with health consequences (including protectives) which are scientifically established or expressed by populations through surveys (knowledge or public opinion). Documents about nicotine prevention (education, publicity, ...) are also considered as relevant.</EN-desc>

<EN-narr>

<c>Tobacco</c>

<c>Tobacco smoking</c>

<c>Nicotine</c>

<c>Epidemiology</c>

<c>Mortality</c>

<c>Morbidity</c>

<c>Diseases</c>

<c>Health</c>

<c>Risk</c>

</EN-narr>

</top>

<top>

<num>015</num>

<EN-title>Sociology: Violence at school</EN-title>

<EN-desc>Documents about causes, manifestations and consequences of violence in scholastic environment, and solutions to control them</EN-desc>

<EN-narr>

<c>Violence</c>

<c>Aggression</c>

<c>Aggressiveness</c>

<c>Menace</c>

<c>Insult</c>

<c>School</c>

<c>Educational system</c>

<c>Teacher-Pupil Relation</c>

</EN-narr>

</top>

<top>

<num>016</num>

<EN-title>French language teaching </EN-title>

<EN-desc>Documents concerning education / learning of French language (mother tongue or foreign Language) all over the world</EN-desc>

<EN-narr>

<c>Language Teaching</c>

<c>French language</c>

<c>Teaching method</c>

<c>Learning Psychology</c>

<c>Education</c>

</EN-narr>

</top>

<top>

<num>017</num>

<EN-title>Diabetes' therapy</EN-title>

<EN-desc>All documents about therapy of diabetes complications and its prevention are considered as relevant</EN-desc>

<EN-narr>

<c>Diabetes</c>

<c>Therapy</c>

<c>Insulin</c>

<c>Hypoglycemic</c>

<c>Glycemia</c>

<c>Complication</c>

<c>Prevention</c>

</EN-narr>

</top>

<top>

<num>018</num>

<EN-title>Menopause and osteoarticular system</EN-title>

<EN-desc>We also will consider the articles talking about menopause substitutional hormone replacement therapy as relevant, even if the osteoporosis is not explicitly mentioned</EN-desc>

<EN-narr>

<c>Menopause</c>

<c>Ostoarticular system</c>

<c>Bone density</c>

<c>Osteoporosis</c>

<c>Substitutional hormone replacement therapy</c>

<c>Prevention</c>

</EN-narr>

</top>

<top>

<num>019</num>

<EN-title>The polar regions</EN-title>

<EN-desc>Only polar regions have been considered as relevant: This means, all those including arctic and Antarctic regions but excluding the sub-Antarctic islands such as Kerguelen, Crozet...except from Greenland and Lapland...</EN-desc>

<EN-narr>

<c>Arctic region</c>

<c>Antarctic region</c>

<c>Spitzberg</c>

<c>Svalbard</c>

<c>Greenland</c>

<c>Lapland</c>

<c>Polar regions</c>

<c>North pole</c>

Deliverable 3.2.1: Test Collection Report for the CLEF 2002 Campaign

<c>South pole</c>
</EN-narr>
</top>

<top>
<num>020</num>
<EN-title>Biotechnologies: transgenic animals and plants</EN-title>
<EN-desc>Genetic transformation, genetic manipulation and genetic engineering are relevant subjects: genetically modified bacteria, animal and plant have been excluded.</EN-desc>
<EN-narr>
<c>Transgenic plant</c>
<c>Transgenic animal</c>
<c>Genetic transformation</c>
<c>Genetic engineering</c>
<c>Genetically modified organism</c>
<c>Transfection</c>
</EN-narr>
</top>

<top>
<num>021</num>
<EN-title>The biological indicators of pollution</EN-title>
<EN-desc>Pollution indicators were including biological accumulation of pollutants, bio-magnification, air or water quality indicators, were considered in a broad sense </EN-desc>
<EN-narr>
<c>Pollution indicator</c>
<c>Bioaccumulation</c>
<c>Biomagnification</c>
<c>Biosensor</c>
<c>Bioevaluation</c>
<c>Pollutant</c>
<c>Contaminant</c>
<c>Xenobiotics</c>
<c>Water quality</c>
<c>Air quality</c>
</EN-narr>
</top>

<top>
<num>022</num>
<EN-title>Aspects of public health and of AIDS in Africa</EN-title>
<EN-desc>Documents describing symptoms, diagnostic, exploration techniques, evolution, preventive and curative treatments in relation with the case itself and not with the population will not be relevant. Documents about seropositivity that do not mention HIV are also excluded.</EN-desc>
<EN-narr>
<c>Epidemiology</c>
<c>Mortality</c>
<c>Morbidity</c>
<c>Prevalence</c>
<c>Incidence</c>
<c>Prevention</c>
<c>Health policy</c>
</EN-narr>
</top>

<top>
<num>023</num>

<EN-title>Ethology: The social behaviour of insects</EN-title>

<EN-desc>Documents will describe the social behaviour, the social organisation and the social behaviour development of insects. The communication aspect between insect individuals will not be taken into account, unless there are specific cases of social interaction</EN-desc>

<EN-narr>

<c>Social behavior</c>

<c>Interindividual interaction</c>

<c>Social structure</c>

<c>Animal society</c>

<c>Social insect</c>

</EN-narr>

</top>

<top>

<num>024</num>

<EN-title>Gas turbines and steam turbines and turbo-machines for thermal power plants</EN-title>

<EN-desc>All documents about conception, functionality, manufacturing, exploitation and maintenance of turbines for electric power plants except hydroelectric power plants</EN-desc>

<EN-narr>

<c>Gas turbine</c>

<c>Steam turbine</c>

<c>Turbine generator</c>

<c>Turbomachine</c>

<c>Thermal power plant</c>

<c>Engineering conception</c>

<c>Manufacturing technology</c>

<c>Materials</c>

<c>Performance characteristics</c>

<c>Performance</c>

<c>Pollution control</c>

</EN-narr>

</top>

<top>

<num>025</num>

<EN-title>Prehistoric archaeology: European Neolithic</EN-title>

<EN-desc>Documents containing the notion of Neolithic and the notion of European geographical environment, i.e. the contemporary names of the countries from Europe as a geographical reality will be considered as valid

</EN-desc><EN-narr>

<c>Neolithic</c>

<c>Neolithization</c>

<c>Europe</c>

<c>France</c>

</EN-narr>

</top>