

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



# Test Collection Report for the CLEF 2003 Campaign: Deliverable 3.2.2

---

*Deliverable Type: REPORT*

*Number: D3.2.2*

*Nature: Public*

*Contractual Date of Delivery: October 2003*

*Actual Date of Delivery: December 2003*

*Status: final*

*Work Package: WP 3*

*Tasks: 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 CLEF2003 campaign is described. Statistics are given.

## Keyword List

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

Document Evolution

Version	Date	Status	Notes
1.1	28 November 2003	Draft	Distributed to partners
2.1	11 December 2003	Final	public

**Table of Contents**

**1. Introduction ..... 3**

**2. Topic Creation ..... 4**

    2.1 Topic Structure ..... 4

    2.2 Topic Creation Phases ..... 6

**3. Assessment Process and Guidelines ..... 7**

    3.1 Results..... 7

    3.2 Language Specific Statistics ..... 8

**4. Technical Problems ..... 9**

**5. Conclusion ..... 9**

**Appendix..... 10**

**1. CLEF 2003 Topic Creation Rules ..... 10**

    1.1 Time Schedule for Topic Creation ..... 12

**2. CLEF 2003 Assessment Rules ..... 13**

    2.1 General aspects ..... 13

    2.2 Technical and procedural aspects ..... 13

**3. Topics for the Main Core Tracks (multilingual, bilingual, monolingual) (in English). 15**

**4. Topics for the GIRT task (in English) ..... 26**

## 1. Introduction

This report covers the topic creation and the relevance assessment phases for the core tracks (monolingual, bilingual and multilingual and domain-specific tasks) in the CLEF 2003 campaign<sup>1</sup>. 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.

All the topic creation and relevance assessment work is done by groups working in their native language. These groups are either members of the CLEF consortium (CM) or associated partners (AP) with sub-contracts, as shown in Table 1. The slight delay in consignment was due to late receipt of reports from some AP.

Language	Language code	Responsible Partner
Dutch	NL	U Twente (AP)
English	EN	NIST (CM)
Finnish	FI	U Tampere (AP)
French	FR	ELDA (CM)
German	DE	IZ (CM)
Italian	IT	ISTI-CNR (CM)
Spanish	ES	UNED (CM)
Swedish	SV	SICS (AP)
Russian	RU	IZ (CM) (work done by native language subcontractor)
GIRT	DE, EN, RU	IZ (CM)

**Table 1: Language Groups**

The main data collections used within the CLEF 2003 evaluation campaign for the core tracks were newspaper and news agency documents in Dutch, English, Finnish, French, German, Italian, Spanish, Swedish for the period 1994-95. In addition a Russian data set was included which only provided news from 1995. The GIRT social science database was used for the domain specific track <sup>2</sup>.

<sup>1</sup> Topics and relevance assessments for additional tracks in CLEF 2003 (multilingual question answering, cross-language image retrieval and cross-language spoken document retrieval), were the responsibility of the groups coordinating these tracks on a voluntary basis. The cross-language interactive track (iCLEF) used CLEF 2002 topics and some additional relevance assessments were performed by WP3 specifically for iCLEF.

<sup>2</sup> **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 and The Herald 1995, SMG Newspapers.

**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.

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

**Russian dataset:** Russika-Izvestia 1995

The work of WP 3 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 native 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.

Two distinct sets of topics were created (i) for the mono-/bi-/multilingual tracks, (ii) for the GIRT social science database. An example of a topic used for the mono-/bi-/multilingual tasks in the CLEF 2003 Campaign is shown below in its English, Finnish, German, Spanish, and Russian version (in this case, German was the original language):

```
<top>
<num> C157 </num>
<EN-title> Wimbledon Lady Winners </EN-title>
<EN-desc> Give the names of any winners of the Wimbledon ladies tennis tournament. </EN-desc>
<EN-narr> Reports on the winners of any of the Wimbledon tennis championships for women are relevant. The names of winners must be explicit. </EN-narr>
</top>
```

```
<top>
<num> C157 </num>
<FI-title> Wimbledonin naisvoittajat </FI-title>
<FI-desc> Etsi Wimbledonin naisten tennisturnauksen voittajien nimiä. </FI-desc>
<FI-narr> Kaikki Wimbledonin naisten tennismestaruuksista kertovat dokumentit ovat relevantteja. Voittajien nimet on mainittava. </FI-narr>
</top>
```

```
<top>
<num> C157 </num>
<DE-title> Siegerinnen von Wimbledon </DE-title>
<DE-desc> Nenne die Namen von Gewinnerinnen des Damentennisturniers von Wimbledon. </DE-desc>
<DE-narr> Relevant sind Berichte über die Gewinnerinnen einer der Tennismeisterschaften für Frauen in Wimbledon. Die Namen der Gewinnerinnen müssen explizit genannt werden. </DE-narr>
</top>
```

```
<top>
<num> C157 </num>
<ES-title> Ganadoras de Wimbledon </ES-title>
<ES-desc> Dar los nombres de las ganadoras del torneo de tenis de Wimbledon femenino. </ES-desc>
<ES-narr> Cualquier documento sobre la vencedora de cualquier campeonato de Wimbledon de tenis femenino es relevante. Los nombres de las ganadoras deben constar explícitamente. </ES-narr>
</top>
```

<top>  
<num> C157 </num>  
<SV-title> Damsegrar i Wimbledon </SV-title>  
<SV-desc> Vilka har segrat i Wimbledon's damtennisturnering? </SV-desc>  
<SV-narr> Dokument med segrare i någon av Wimbledon's tennismästerskap för damer är relevanta. Namn på de segrande damerna måste vara explicit nämnda. </SV-narr>  
</top>

<top>  
<num> C157 </num>  
<RU-title> Победительницы Уимблдона </RU-title>  
<RU-desc> Дать имена победительниц любого из теннисных турниров в Уимблдоне среди женщин. </RU-desc>  
<RU-narr> Релевантны сообщения о победительницах любого из теннисных чемпионатов в Уимблдоне среди женщин. Имена победительниц должны ясно фигурировать в документах. </RU-narr>  
</top>

For the GIRT task topics were created in German and then translated into English and Russian. An example of a GIRT topic in German, English and Russian is shown below:

<top>  
<num> 079 </num>  
<DE-title> Volksentscheide in der Demokratie </DE-title>  
<DE-desc> Finde Dokumente, die über plebiszitäre Elemente in der Demokratie berichten. </DE-desc>  
<DE-narr> Relevante Dokumente thematisieren Möglichkeiten, Vor- und Nachteile der direkten Einflussnahme des Volkes auf die politische Willensbildung in parlamentarischen Systemen. Von besonderem Interesse ist dabei der Vergleich der repräsentativen Demokratie mit dem System der direkten Volksentscheidungen, insbesondere im Hinblick auf Überlegungen, Elemente direkter Demokratie in parlamentarische Systeme zu integrieren. </DE-narr>  
</top>

<top>  
<num> 079 </num>  
<EN-title> Plebiscites in Democracy </EN-title>  
<EN-desc> Find documents reporting on plebiscitarian elements in democracy. </EN-desc>  
<EN-narr> Relevant documents discuss possibilities, advantages and disadvantages of direct popular influence on the decision-making process in a parliamentary system. Of special interest are comparisons of representative democracies with systems in which the people participate directly in decision-making, and also considerations on the integration of elements of direct democracy into parliamentary systems. </EN-narr>  
</top>

<top>  
<num> 079 </num>  
<RU-title> Плебисцит в Демократии </RU-title>  
<RU-desc> Найти документы об элементах плебисцита в демократии. </RU-desc>  
<RU-narr> В релевантных документах должны обсуждаться возможности, преимущества и недостатки прямого влияния народа на решающие процессы в парламентарной системе. Особенный интерес представляют сравнения примеров демократических систем, в которых народ непосредственным образом участвует в принятии решений, а также интересные соображения об интеграции элементов прямой демократии в парламентарных системах. </RU-narr>  
</top>

The English version of all CLEF 2003 core track topics can be found in the Appendix. The complete set of topics in all languages is accessible under the password protected area of the CLEF Web site.

## 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 1995, or without time limit for the scientific task) 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: according to this check, there should be more than one and at most 50 possible answers when using the ZPRISE IR system provided by NIST with a carefully formulated search. The aim is to avoid topics with just one relevant result in any language collection which can really harm evaluation statistics and also to avoid topics with too many relevant documents which could make the task too simple and thus of little significance in discriminating between systems.

Each language group prepared about 10-15 possible topics in their own language using the Guidelines provided by the coordinator (see Appendix) and provided a provisional translation into English. They also tested the topics developed by the other language groups on the corpora in their languages.

It is necessary to use the same topic set for the multi-/bi-/monolingual tasks in order to economise on effort involved. It would be far too costly in terms of human resources to prepare and assess separate topic sets for each of these tasks. In the CLEF 2002 campaign, we attempted to select topics that would find relevant documents in each collection. However, this approach has had its drawbacks; it led to the choice of very general topics, that tended to find a large number of documents in many of the collections and thus the tasks were not particularly challenging. It made it somehow easier for systems to find a good number of relevant documents. This meant, for example, that it was difficult to adequately test new solutions to the merging problem in the multilingual task - and the average performance figures tended to level out. For these reasons, we decided to adopt a different approach in CLEF 2003. We increased the size of the topic set (to 60 items) and deliberately selected a variety of topic types that not only included particular features (e.g. proper names, acronyms, cultural idiosyncrasies, terminology, slang) but also ranged from very broad topics, hitting most or all collections, to very narrow topics - hitting just one or two collections. Our only requirement was that, for each collection, there should be at least 40 topics that actually find relevant documents. This information would not be released to the participants in order not to give them information that could be exploited for the multilingual task. Thus, for the multilingual task, the results would be calculated over the entire topic set of 60 items, whereas for the monolingual and bilingual tasks, results could be calculated only for those topics that were actually relevant to that collection. The only exception to this rule has been for the Russian collection. This dataset was added very late - and it was not possible to refer to it during the topic creation stage. Thus, the topics were tested on it subsequently, and a final subset of 37 topics was prepared for the Russian mono- and bilingual tasks (Russian is not included in the multilingual track).

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 60 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 that had originally been developed in other languages into their own language. As far as possible the translations were 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 produce 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 for the multilingual-8 task Dutch, English, Finnish, French, German, Italian, Spanish and Swedish (DE, EN, ES, FR, IT, FI, NL, SV) plus the additional topic languages Russian and Chinese (RU, ZH).

At the end of the third stage, all topic sets were cross-checked for translation inconsistencies, unusual language uses and mistakes. The checking was done at the University of Hildesheim by a group with expertise in translation theory and practice. 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 (RU, ZH) were added later and were not cross-checked.

For the scientific task, the topics were produced separately: 25 topics for GIRT. These topics were checked against the parallel German and English corpora of GIRT4 as described above. They were then translated from German to English and Russian, and the translations were cross-checked.

In total 85 topics have been provided for the different tasks with translations from one up to ten languages.

### **3. Assessment Process and Guidelines**

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 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 (see Appendix); 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 CLEF 2003. For each track, participants must submit runs for every topic consisting of the document ids for 1000 documents, ordered by decreasing relevance to the topic. As it is not possible to judge all the 1000 documents submitted by each group for each topic a pool depth must be decided on. This 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 for CLEF 2003 was established at 60, the same as for CLEF 2002. Tests on past CLEF test collections show that this depth is sufficient to obtain reliable results as long as a sufficiently large number of runs are included. Even with this limited pool depth, the number of documents to be assessed for each language collection per topic was still very large. In order to reduce this number somewhat, we also had to make a selection as to the runs to be judged; this is called the restricted pool. A number of criteria 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. Great care was taken to ensure that runs from all groups were included in all the target languages they used. At the end of the

pooling process, 209 runs of the 415 submitted were included in the restricted pool and judged by the assessors.

language/task	number topics	all results (pool depth = 60 ; all runs included)		restricted pool (pool depth = 60 ; selected runs included)			
		number of documents	docs/topic (average)	% of all results	docs/topic (average)	number of documents	change in workload vs 2002 %
German	60	29.155	486	80	389	23.364	+ 0,3
English	60	34.046	567	70	395	23.674	- 9,6
French	60	31.853	517	64	339	20.358	+ 13,2
Italian	60	39.946	666	62	410	24.601	+ 13,9
Spanish	60	47.660	794	53	424	25.412	+ 8,9
Dutch	60	26.193	437	85	372	22.331	- 11,2
Finnish	60	16.275	271	96	260	15.605	- 25,5
Swedish	60	18.476	308	98	301	18.061	+ 17,1
GIRT	25	17.031	681	100	681	17.031	+ 64,9
Russian	37	15.069	407	100	407	15.069	+100,0
sum/average		275.704		75		205.506	+ 32,6

**Table 2: Statistics on collections and returned documents per topic**

### 3.2 Language Specific Statistics

After pooling, the delivered result data of the participants was split into language packages which had to be assessed by the respective language groups. The amount of documents which have been judged as relevant is shown in the following Table 3.

Very few additional relevance assessment had to be done for the iCLEF task, as the data and assessments of the CLEF 2002 campaign was used.

Task	Language of documents	no. assessed docs	no. docs judged as relevant	% relevant docs
<b>CLEF general</b> (mono-, bi-, multilingual)	Dutch	22.331	1.577	7,06
	English	23.674	1.006	4,25
	Finnish	15.605	483	3,10
	French	20.358	946	4,65
	German	23.364	1.825	7,81
	Italian	24.601	809	3,29
	Spanish	25.412	2.368	9,32
	Swedish	18.061	1.080	5,98
	Russian	15.069	151	1,00
	sum (CLEF general)		188.475	
GIRT	German	17.031	3.449	20,25
	sum (all tasks)		205.506	

**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. However, the installation caused severe problems for SICS this time and some difficulties for ELDA.

#### **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 high and as shown has increased over that of 2002, however it was fairly evenly distributed (except Finnish and Russian). The very short time frame for the assessments necessitated very concentrated and hard work. All partners and sub-contractors achieved this with additional volunteer support from their staff and institutes.

## Appendix

### **1. CLEF 2003 Topic Creation Rules**

[These 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.]

The topic generation will be carried out by the different CLEF language groups (DE, EN, ES, FI, FR, IT, NL, SV). Topics created should refer mainly to both years 1994 and/or 1995, but they can also regard general issues which are not that 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 (see list from 2000, 2001 and 2002 at the end). You can check the topics used for CLEF 2000 and 2001 on the CLEF Web site, in the Workspace for Registered Participants (user name=clef2003; password=irbitf03). 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 give 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. Therefore you should think about other countries/language whether your proposed topic might be of any relevance in their context. The topic should hit most of the collections at least with few hits.

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). And it is our experience that some participating groups do not discover any hit for at least one language which they handle, thus they have to exclude this topic from their statistics, although in our pre-test there has been more than one hit in all language. In most cases there should be some hits in each of the 8 languages, 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 8 to 10 topics for each language (some Japanese topics will also be added: we will get some proposals for topics which are running in the NTCIR campaign). Each group should send its topic proposals to me by 09.02.2003, I will circulate the cumulated list by 11.02.2003. 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 (18.02.-19.02.2003). 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 will know whether 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 the same way.

In the end a decision will be taken on the definite set of 50 topics. The discussion has to ensure that there will be a balance in the meaning of the topics in the different languages. The definite set of topics will be prepared in a so-called Master Set, which - for each topic - will contain the topic in its source language 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. This is carried out by the University of Hildesheim with the help of native speaker which are professional translators.

For the scientific data collection GIRT 25 scientific/domain-specific topics will be developed in German (by IZ). For these, an English and a Russian translation have to be prepared by IZ. Then they will be checked by the University of Hildesheim.

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, Swedish and Russian (for GIRT only) provided by the respective language groups. Additional topic sets will be prepared by volunteer groups for Russian, Chinese, Japanese<sup>3</sup> 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.

The format has to be the same as that used for 2002 (see the list for EN at the end), for the proposal topics please number them with the language code and a current number: as DE01, DE02 etc. in the field NUM.

Please start and end the entries within the tags with a blank: this makes it easier to use spell-checkers and tag-checkers!

---

<sup>3</sup> The Japanese set has not been prepared for CLEF 2003.

## 1.1 Time Schedule for Topic Creation

- 30.01.2003 IZ circulates topic creation rules and format
- 31.01.2003 all groups (DE, EN, FR, IT, ES, NL, FI, SV) start to find topics
- 09.02.2003 all groups deliver 8-10 proposals of topics for their own language and send them to IZ including a rough English translation
- 11.02.2003 will circulated the cumulated list of proposed topics
- 18.02.-19.02.2003 meeting on topic creation in Berlin (discussion and decision on the selection of the final topics)
- 20.02.2003 IZ sends out the list of the final topics (first version of the Master Set)
- 23.02.2003 partners give back the translation of all topics into their own language, clean their own topics and deliver a cleaned version of the English translation to IZ
- 26.02.2003 IZ sends all topics + translations to Univ. Hildesheim (second version of the Master Set) (copy to all)
- 14.03.2003 checked topics back to IZ
- 15.03.2003 distribution to all participants of the CLEF 2002 by IZ

## 2. CLEF 2003 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.

However, 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 Y yields” or “documents only giving the name Y 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 document 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, or German in Germany, Switzerland or Austria, or French in France or Switzerland, or 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 again!!) 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 and/or the discussions at the topic creation meeting in Berlin and might give you additional hints on the common understanding of the topic of all language groups.

Michael Kluck (IZ)

20.05.2003

### 3. Topics for the Main Core Tracks (multilingual, bilingual, monolingual) (in English)

[with internal information for assessors, which was released for assessment purposes after the submission of the results (23.05.2003).

For some topics there are already "solutions" (like names) or other information available (such as time-span for result hits) etc]

<top>

<num> C141 </num>

<EN-title> Letter Bomb for Kiesbauer </EN-title>

<EN-desc> Find information on the explosion of a letter bomb in the studio of the TV channel PRO7 presenter Arabella Kiesbauer. </EN-desc>

<EN-narr> A letter bomb from right-wing radicals sent to the black TV personality Arabella Kiesbauer exploded in a studio of the TV channel PRO7 on June 9th, 1995. An assistant was injured. All reports on the explosion and police inquiries after the event are relevant. Other reports on letter bomb attacks are of no interest. </EN-narr>

</top>

time-span: from June 9th 1995

\*\*\*\*

<top>

<num> C142 </num>

<EN-title> Christo wraps German Reichstag </EN-title>

<EN-desc> Find reports on the wrapping of the German Reichstag in Berlin by the artist Christo. </EN-desc>

<EN-narr> The wrapping artist Christo took two weeks in June 1995 to wrap the Berlin German Reichstag in lengths of material. Find reports on this artistic event. Any information on either its preparation or its execution is relevant, including political debates and decisions and technical preparations. </EN-narr>

</top>

time-span: includes discussion already before the event!

\*\*\*\*

<top>

<num> C143 </num>

<EN-title> Women's Conference Beijing </EN-title>

<EN-desc> Controversial positions by a number of delegates meant that the Women's Conference in Beijing risked failure. </EN-desc>

<EN-narr> Relevant documents will discuss any of the many problems or disagreements that arose in connection with the Women's Conference in Beijing. Of particular interest are the positions held by representatives of the Vatican, Muslim communities and the Chinese communist party. </EN-narr>

</top>

time-span: conference in September 1995, discussions already starting before that, but not in 1994.

\*\*\*\*

<top>

<num> C144 </num>

<EN-title> Sierra Leone Rebellion and Diamonds </EN-title>

<EN-desc> What have been the effects of rebellions and other political instability on the Sierra Leone diamond industry? </EN-desc>

<EN-narr> Documents on the rebellion in Sierra Leone must be specific as to effects on the diamond industry to be relevant. </EN-narr>

</top>

<top>

<num> C145 </num>

<EN-title> Japanese Rice Imports </EN-title>

<EN-desc> Find documents discussing reasons for and consequences of the first imported rice in Japan. </EN-desc>

<EN-narr> In 1994, Japan decided to open the national rice market for the first time to other countries. Relevant documents will comment on this question. The discussion can include the names of the countries from which the rice is imported, the types of rice, and the controversy that this decision prompted in Japan. </EN-narr>

</top>

<top>

<num> C146 </num>

<EN-title> Fast Food in Japan </EN-title>

<EN-desc> What North American fast food chains have a large number of franchise restaurants in Japan? </EN-desc>

<EN-narr> Relevant documents must mention the name of most successful US chains of fast food restaurants in Japan, and may contain additional information about the introduction of this kind of food into Japanese society. </EN-narr>

</top>

Answer: MacDonalds

\*\*\*\*

<top>

<num> C147 </num>

<EN-title> Oil Accidents and Birds </EN-title>

<EN-desc> Find documents describing damage or injury to birds caused by accidental oil spills or pollution. </EN-desc>

<EN-narr> All documents which mention birds suffering because of oil accidents are relevant. Accounts of damage caused as a result of bilge discharges or oil dumping are not relevant. </EN-narr>

</top>

<top>

<num> C148 </num>

<EN-title> Damages in Ozone Layer </EN-title>

<EN-desc> What holes in the ozone layer are not an effect of pollution? </EN-desc>

<EN-narr> Not all damage to the ozone layer is caused by pollution. Relevant documents will give information on other causes for holes in the ozone layer. </EN-narr>

</top>

<top>

<num> C149 </num>

<EN-title> Pope's Visit in Sri Lanka </EN-title>

<EN-desc> Find reports on protests or problems caused by previous declarations of the Pope about Buddhism during his visit to Sri Lanka. </EN-desc>

<EN-narr> Pope John Paul found a mixed reception waiting when he visited largely Buddhist Sri Lanka. The Pope's visit to Sri Lanka proved to be a test of relations between the Catholic Church and Buddhist leaders following comments about Buddhism in a recently published book. </EN-narr>

</top>

<top>

<num> C150 </num>

<EN-title> AI against Death Penalty </EN-title>

<EN-desc> Find reports on specific actions by Amnesty International against the death penalty. </EN-desc>



<EN-narr> Amnesty International is dedicated to abolishing the death penalty throughout the world. Relevant documents must describe specific actions of AI against the death penalty. </EN-narr>  
</top>

<top>

<num> C151 </num>

<EN-title> Wonders of Ancient World </EN-title>

<EN-desc> Look for information on the existence and/or the discovery of remains of the seven wonders of the ancient world. </EN-desc>

<EN-narr> Relevant documents will contain references to any of the seven wonders of the ancient world that still exist today. Of particular interest are reports on the recent discovery of remains believed to belong to one of the seven ancient wonders. </EN-narr>

</top>

Only the Great Pyramid at Giza/Gizeh survives, there have been some reports on remains that could belong to the Lighthouse of Alexandria.

\*\*\*\*\*

<top>

<num> C152 </num>

<EN-title> Children's Rights </EN-title>

<EN-desc> Find information on the United Nations Convention on the Rights of the Child. </EN-desc>

<EN-narr> Relevant documents will provide information on the U.N. Convention on children's rights such as the date of its adoption by the General Assembly of the United Nations, the definition of the rights to be respected, the states that ratified it, etc. </EN-narr>

</top>

<top>

<num> C153 </num>

<EN-title> Olympic Games and Peace </EN-title>

<EN-desc> Find documents which report on how the Olympic Games can contribute to world peace. </EN-desc>

<EN-narr> Documents that identify world peace as one of the objectives of the Olympic Games are relevant, including documents in which athletes or members of the International Olympic Committee call for a suspension of hostilities during the games. </EN-narr>

</top>

<top>

<num> C154 </num>

<EN-title> Free Speech on the Internet </EN-title>

<EN-desc> Find documents which discuss censorship and freedom of speech on the Internet. </EN-desc>

<EN-narr> Documents that discuss subjects such as pornography or racism on the Internet without mentioning issues concerning censorship or freedom of speech will not be considered relevant. </EN-narr>

</top>

<top>

<num> C155 </num>

<EN-title> Risks with Mobile Phones </EN-title>

<EN-desc> What risks and dangers are there with an increased use of mobile phones? </EN-desc>

<EN-narr> Relevant documents will describe problems that have to do with mobile phone usage. Both health risks that may have to do with mobile phone usage, such as radiation related risks, or dangers that are incidental to mobile phones, such as traffic accidents caused by inattentive drivers, are relevant. </EN-narr>

</top>

<top>

<num> C156 </num>

<EN-title> Trade Unions in Europe </EN-title>

<EN-desc> What are the differences in the role and importance of trade unions between European countries? </EN-desc>

<EN-narr> Relevant documents must compare the role, status or importance of trade unions between two or more European countries. Pertinent information will include level of organisation, wage negotiation mechanisms, and the general climate of the labour market. </EN-narr>

</top>

<top>

<num> C157 </num>

<EN-title> Wimbledon Lady Winners </EN-title>

<EN-desc> Give the names of any winners of the Wimbledon ladies tennis tournament. </EN-desc>

<EN-narr> Reports on the winners of any of the Wimbledon tennis championships for women are relevant. The names of winners must be explicit. </EN-narr>

</top>

names (year): Steffi Graf (1988, 1989, 1991, 1992, 1993,1995), Conchita Martinez (1994), Martina Navratilova /1990)

\*\*\*\*

<top>

<num> C158 </num>

<EN-title> Soccer Riots in Dublin </EN-title>

<EN-desc> Find reports on the suspension of the international soccer match between Ireland and England in Dublin after riots. </EN-desc>

<EN-narr> After riots by football hooligans, the Ireland-England international soccer match in Dublin was broken off with the Irish in the lead 1:0 in the 27th minute of play. </EN-narr>

</top>

<top>

<num> C159 </num>

<EN-title> North Sea Oil Environment </EN-title>

<EN-desc> What steps are being taken to protect the environment given the increase in oil exploration and production in the North Sea? </EN-desc>

<EN-narr> Relevant documents must identify and discuss environmental issues surrounding the exploration and production of oil in the North Sea. Documents discussing the decommissioning of an oil production site are only relevant if they mention environmental problems. However, documents that simply list companies involved in drilling, or that report on the amount of production and revenues are not of interest. </EN-narr>

</top>

<top>

<num> C160 </num>

<EN-title> Scotch Production Consumption </EN-title>

<EN-desc> Documents will discuss the amount of scotch consumed by Scots relative to the amount of scotch that is exported from Scotland. </EN-desc>

<EN-narr> Documents must report on the consumption of scotch by Scots relative to the amount produced to be relevant. Documents that report on consumption by other nations or only analyze reasons for declining exports are irrelevant. </EN-narr>

</top>

<top>

<num> C161 </num>

<EN-title> Diets for Celiacs </EN-title>

<EN-desc> Find reports discussing diet problems of celiacs? </EN-desc>

<EN-narr> Relevant documents will discuss nutritional issues for celiacs. Of particular interest are reports on whether celiacs can eat oats. </EN-narr>

</top>

<top>

<num> C162 </num>

<EN-title> EU and Turkish Customs </EN-title>

<EN-desc> Find documents about the problems posed by Greece concerning the abolishment of customs restrictions between the European Union and Turkey. </EN-desc>

<EN-narr> Relevant documents must explicitly mention the Greek position with reference to the proposal for the union of customs regulations between the EU and Turkey. </EN-narr>

</top>

<top>

<num> C163 </num>

<EN-title> Smoking Restrictions </EN-title>

<EN-desc> Find documents about regulations or laws that aim to prohibit or limit smoking in restaurants. </EN-desc>

<EN-narr> All documents which talk about laws or regulations concerning the prohibition or the limiting of smoking in restaurants are relevant, as well as documents describing proposals for or consequences of this kind of law. </EN-narr>

</top>

<top>

<num> C164 </num>

<EN-title> European Drug Sentences </EN-title>

<EN-desc> What sentences are given in Europe for illegal dealing in drugs? </EN-desc>

<EN-narr> Relevant documents give information on sentences passed in European countries for trading in drugs. Information will include the type of sentence, the length of confinement, or alternative penalties. Punishment for drug consumption is not relevant. </EN-narr>

</top>

<top>

<num> C165 </num>

<EN-title> Golden Globes 1994 </EN-title>

<EN-desc> What movie won the Golden Globe for drama in the 1994 season? </EN-desc>

<EN-narr> Golden Globes for the 1994 film season were awarded in Los Angeles in January 1995. Relevant documents will give the name of the picture that won the drama award in Los Angeles. </EN-narr>

</top>

<top>

<num> C166 </num>

<EN-title> French General and Balkan Security Zone </EN-title>

<EN-desc> Who was the French General responsible for the creation of the security zone during the Balkan conflict? </EN-desc>

<EN-narr> Relevant documents will give the name of the French commander of the IFOR forces who created the "security zone" during the Balkan conflict. </EN-narr>

</top>

Answer General Morillon

\*\*\*\*

<top>

<num> C167 </num>

<EN-title> China-Mongolia Relations </EN-title>

<EN-desc> Find information on recent relations and agreements for cooperation between China and Mongolia. </EN-desc>

<EN-narr> Documents that provide information on political and/or economic relations between China and Mongolia in the twentieth century are relevant. </EN-narr>

</top>

<top>

<num> C168 </num>

<EN-title> Assassination of Rabin </EN-title>

<EN-desc> Who shot Yitzhak Rabin and why? </EN-desc>

<EN-narr> Relevant documents must give the name of the murderer of Yitzhak Rabin, the prime minister of Israel, and explain some of the reasons for the assassination. </EN-narr>

</top>

<top>

<num> C169 </num>

<EN-title> Advent of the CD-Burner </EN-title>

<EN-desc> Find documents about the CD-burner and the reactions of the audio industry to it. </EN-desc>

<EN-narr> Documents about the consequences for the audio industry of the use of CD-burners and issues concerning intellectual property rights are especially relevant. </EN-narr>

</top>

It might be named CD-Writer!

\*\*\*\*

<top>

<num> C170 </num>

<EN-title> Official EU Languages </EN-title>

<EN-desc> Find documents about French plans for reducing the number of official languages in the European Union to five languages. </EN-desc>

<EN-narr> France has suggested restricting the core languages in the Union to French, English, German, Italian and Spanish. After much opposition the proposal was dropped. Relevant documents must refer to this proposal. </EN-narr>

</top>

<top>

<num> C171 </num>

<EN-title> Lillehammer Ice Hockey Finals </EN-title>

<EN-desc> What teams played in the ice hockey finals at the 1994 Olympic games in Lillehammer? </EN-desc>

<EN-narr> Relevant documents report which teams played in the final ice hockey game at the 1994 Olympic games in Lillehammer. Documents, which inform as to which teams won first and second place in this event without mentioning the finals specifically are also relevant. </EN-narr>

</top>

<top>

<num> C172 </num>

<EN-title> 1995 Athletics World Records </EN-title>

<EN-desc> What new world records were achieved during the 1995 athletic world championships in Gothenburg? </EN-desc>

<EN-narr> Relevant documents must mention which records were broken and describe which event, what result and give the name of the new world record holder. </EN-narr>

</top>

<top>

<num> C173 </num>

<EN-title> Proof of Top Quarks </EN-title>

<EN-desc> Find reports on the experimental proof of top quarks by US researchers. </EN-desc>

<EN-narr> Scientists of an American research center reported that the existence of the so-called top quark could be proved experimentally. The elementary particles named quarks are the fundamental constituents of matter. Relevant documents must discuss this experimental evidence and not elementary particles in general. </EN-narr>

</top>

<top>

<num> C174 </num>

<EN-title> Bavarian Crucifix Quarrel </EN-title>

<EN-desc> Find reports on the crucifix quarrel in Bavarian schools. </EN-desc>

<EN-narr> In the crucifix quarrel the Federal Constitutional Court of Germany confirmed that individual parents had a right to complain about the hanging of crucifixes in the classrooms of Bavarian schools. </EN-narr>

</top>

time-span: 1995

\*\*\*\*

<top>

<num> C175 </num>

<EN-title> Everglades Environmental Damage </EN-title>

<EN-desc> Find reports on environmental damage in the Everglades, for example caused by the Florida sugar industry. </EN-desc>

<EN-narr> Reports should talk about damage to the Everglades environment. Of particular interest are discussions on problems caused by the sugar industry. References to environmental damage by run-off from farms in general are also relevant. </EN-narr>

</top>

<top>

<num> C176 </num>

<EN-title> Shoemaker-Levy and Jupiter </EN-title>

<EN-desc> Find reports on the break-up of the Shoemaker-Levy comet and its impact on the planet Jupiter. </EN-desc>

<EN-narr> Relevant documents will provide evidence of the break up of the Shoemaker-Levy comet and the collision of parts of the comet with Jupiter. </EN-narr>

</top>

<top>

<num> C177 </num>

<EN-title> Milk Consumption in Europe </EN-title>

<EN-desc> Provide statistics or information concerning milk consumption in European countries. </EN-desc>

<EN-narr> Relevant documents must provide statistics or other information about milk consumption in Europe, or in single European nations. Reports on milk derivatives are not relevant. </EN-narr>

</top>

<top>

<num> C178 </num>

<EN-title> Military Service Denial </EN-title>

<EN-desc> Find cases of people who are willing to risk imprisonment rather than perform military service. </EN-desc>

<EN-narr> Relevant documents talk about people who refuse totally to do military service or a civil equivalent and are willing to risk imprisonment for their beliefs. </EN-narr>

</top>

<top>

<num> C179 </num>

<EN-title> Resignation of NATO Secretary General </EN-title>

<EN-desc> Why was the secretary general of NATO forced to resign in 1995? </EN-desc>

<EN-narr> There was strong political pressure in 1995 for the secretary general of NATO to resign. Relevant documents will mention at least one reason for this. </EN-narr>

</top>

<top>

<num> C180 </num>

<EN-title> Bankruptcy of Barings </EN-title>

<EN-desc> What was the extent of the losses in the Barings bankruptcy case? </EN-desc>

<EN-narr> Relevant documents must quantify in some way the losses caused by the collapse of the oldest bank in Great Britain. </EN-narr>

</top>

<top>

<num> C181 </num>

<EN-title> French Nuclear Tests </EN-title>

<EN-desc> Find reports of international pressure to put an end to French nuclear testing. </EN-desc>

<EN-narr> Greenpeace and many other ecologist groups have denounced the decision by Jaques Chirac after his election to break the international moratorium on nuclear testing. Each test resulted in strong protests throughout the world, especially in Australia and in New Zealand. Both Japan and the USA expressed their regrets about the French attitude. </EN-narr>

</top>

<top>

<num> C182 </num>

<EN-title> 50th Anniversary of Normandy Landings </EN-title>

<EN-desc> Find reports on parachuting by veterans over Sainte-Mère-Église during the celebrations of the fiftieth anniversary of the Normandy landings. </EN-desc>

<EN-narr> About 40 veterans parachuted over Sainte-Mère-Église during the Normandy anniversary celebrations as they had done 50 years earlier on D-day. Any information on plans for this action or the event itself are relevant. </EN-narr>

</EN-narr>

</top>

<top>

<num> C183 </num>

<EN-title> Asian Dinosaur Remains </EN-title>

<EN-desc> In what parts of Asia have dinosaur remains been found? </EN-desc>

<EN-narr> Relevant document will report on findings of any kind of fossilised remains of dinosaurs, including dinosaur eggs, anywhere in Asia. Documents that only mention the sighting of dinosaur footprints are not relevant. </EN-narr>

</EN-narr>

</top>

<top>

<num> C184 </num>

<EN-title> Maternity Leave in Europe </EN-title>

<EN-desc> Find documents that give information on provisions concerning the length of maternity leave in Europe. </EN-desc>

<EN-narr> Documents must report exact figures with respect to the time allowed for maternity leave for working women in Europe. Discussions, laws or proposals of laws concerning the length of maternity leave to which women are entitled in any European country are also relevant as long as numbers are given. Comparisons between regulations in different states are also of interest. </EN-narr>

</top>

<top>

<num> C185 </num>

<EN-title> Dutch Photos of Srebrenica </EN-title>

<EN-desc> What happened to the photographs and films that Dutch soldiers made in Srebrenica which provided evidence of violations of human rights? </EN-desc>

<EN-narr> After the fall of Srebrenica, photos and videos made by Dutch military with evidence of violations of human rights disappeared. Remaining film showed only a few discernible images. Relevant documents will provide any kind of information on this incident and may include details on the content of the photographs and film and discussions on what happened to them. </EN-narr>

</top>

<top>

<num> C186 </num>

<EN-title> Dutch Coalition Government </EN-title>

<EN-desc> Which political parties formed the ruling coalition, the so-called purple cabinet, in the Netherlands, in 1994-1995? </EN-desc>

<EN-narr> In 1994, an improbable combination of parties formed a coalition government in the Netherlands. Relevant documents must mention the names of these three parties. </EN-narr>

</top>

<top>

<num> C187 </num>

<EN-title> Nuclear Transport in Germany </EN-title>

<EN-desc> Find reports on the protests against the transportation of radioactive waste with Castor containers in Germany. </EN-desc>

<EN-narr> Relevant documents report about radioactive waste shipments with Castor containers to Gorleben in Lower Saxony and the protest actions against it. Any document reporting on protests against about these shipments is relevant even if it does not specifically mention the location or the type of container used. </EN-narr>

</top>

<top>

<num> C188 </num>

<EN-title> German Spelling Reform </EN-title>

<EN-desc> Find documents which report on the introduction of a spelling reform for the German language. </EN-desc>

<EN-narr> A spelling reform was approved to become valid from August 1, 1998, in all German speaking countries. The spelling and punctuation rules for the German language have been simplified considerably. The reform nevertheless is controversial. </EN-narr>

</top>

<top>

<num> C189 </num>

<EN-title> Hubble and Black Holes </EN-title>

<EN-desc> What was the role of the Hubble telescope in proving the existence of black holes? </EN-desc>

<EN-narr> Relevant documents must specifically address the evidence provided by the Hubble telescope on black holes. General discussions on Hubble or black holes are not pertinent. </EN-narr>

</top>

<top>

<num> C190 </num>

<EN-title> Child Labor in Asia </EN-title>

<EN-desc> Find documents that discuss child labor in Asia and proposals to eliminate it or to improve working conditions for children. </EN-desc>

<EN-narr> Documents discussing child labor in particular countries in Asia, descriptions of working conditions for children, and proposals of measures to eliminate child labor are all relevant. </EN-narr>

</top>

<top>

<num> C191 </num>

<EN-title> Ebro Delta Farming </EN-title>

<EN-desc> What was the predominant agricultural activity in the Ebro Delta in the early 1990s? </EN-desc>

<EN-narr> Relevant documents must mention the main farming activity in the Ebro delta in the nineties. They can also mention changes foreseen depending on GATT agreements. </EN-narr>

</top>

<top>

<num> C192 </num>

<EN-title> Russian TV Director Murder </EN-title>

<EN-desc> What was the name of the head of a Russian TV company who was shot in the staircase of his home in Moscow? </EN-desc>

<EN-narr> Relevant documents will report on the assassination of a Russian TV director and former journalist in Moscow and must provide the name of the murdered man. </EN-narr>

</top>

name: Vladislav Listjev, Listyev and other spellings possible!

\*\*\*\*

<top>

<num> C193 </num>

<EN-title> EU and Baltic Countries </EN-title>

<EN-desc> Find documents telling about the negotiations between Baltic countries (Estonia, Latvia or Lithuania) and the European Union concerning EU membership. </EN-desc>

<EN-narr> Relevant documents describe the negotiations between the EU and Estonia, Latvia or Lithuania. Reports of agreements for cooperation between single European nations and any of the Baltic countries are not relevant. </EN-narr>

</top>

<top>

<num> C194 </num>

<EN-title> Italian Royal Family </EN-title>

<EN-desc> Find information on the exile from Italy of male members of the House of Savoy. </EN-desc>

<EN-narr> Relevant documents will provide information concerning the terms of the exile of the Italian royal family, the House of Savoy, and/or the conditions imposed. Documents that report discussions and opinions concerning a possible end to the banishment of the heirs and their return to Italy are also to be considered pertinent. </EN-narr>

</top>

The male heirs to the throne (Umberto, Victor Emmanuel, Emmanuel Filiberto) were exiled at the end of the Second World War because of their support to the fascism etc. Conditions for return generally regard swearing loyalty to the Italian republic.

\*\*\*\*\*

<top>

<num> C195 </num>

<EN-title> Strikes by Italian Flight Assistants </EN-title>

<EN-desc> Find reports on disruptions or problems caused by strikes by flight assistants in Italy. </EN-desc>

<EN-narr> Relevant documents will give information on disruptions that are a direct result of strikes by Italian flight assistants. Pertinent information can include details on delays and cancellations of flights and difficulties for passengers anywhere in Europe or beyond due to effects of trade union action by flight assistants in Italy. </EN-narr>

</top>

<top>

<num> C196 </num>

<EN-title> Merger of Japanese Banks </EN-title>

<EN-desc> Find reports on the merger of the Japanese banks Mitsubishi and Bank of Tokyo into the largest bank in the world. </EN-desc>

<EN-narr> The two Japanese credit institutions Mitsubishi Bank and Bank of Tokyo agreed to merge. The new bank is the largest in the world. </EN-narr>

</top>



<top>  
<num> C197 </num>  
<EN-title> Dayton Peace Treaty </EN-title>  
<EN-desc> Find reports on the Dayton agreement for the maintenance of peace in Bosnia-Herzegovina. </EN-desc>  
<EN-narr> In accordance with the Dayton agreement, NATO undertook responsibility for the military maintenance of peace in Bosnia-Herzegovina. Relevant documents report about the negotiations, the contents and the conclusion of the agreement. </EN-narr>  
</top>  
time-span: 1995  
\*\*\*\*

<top>  
<num> C198 </num>  
<EN-title> Honorary Oscar for Italian Directors </EN-title>  
<EN-desc> Find information on Italian film directors that have been awarded an Honorary Oscar for their lifetime work. </EN-desc>  
<EN-narr> Relevant documents will provide information on which Italian film directors have received an Honorary Oscar at the Academy Awards ceremony in Los Angeles for their lifetime work in film making. </EN-narr>  
</top>  
it is sufficient that the documents contain the name of the director and specific reference to the Honorary Oscar (and not just for a single film). The answer should be either Federico Fellini or Michelangelo Antonioni or both.  
\*\*\*\*\*

<top>  
<num> C199 </num>  
<EN-title> Ebola Epidemic in Zaire </EN-title>  
<EN-desc> Find reports on preventive measures taken after the outbreak of the Ebola epidemic in Zaire. </EN-desc>  
<EN-narr> In May 1995, in the Zairian city of Kikwit, there was an outbreak of one of the most dreaded epidemics: the Ebola epidemic. Relevant documents report on measures taken to prevent the spreading of this disease. </EN-narr>  
</top>  
time-span: 1995  
\*\*\*\*

<top>  
<num> C200 </num>  
<EN-title> Flooding in Holland and Germany </EN-title>  
<EN-desc> Find statistics on flood disasters in Holland and Germany in 1995. </EN-desc>  
<EN-narr> Relevant documents will quantify the effects of the damage caused by flooding that took place in Germany and the Netherlands in 1995 in terms of numbers of people and animals evacuated and/or of economic losses. </EN-narr>  
</top>

## 4. Topics for the GIRT task (in English)

<top>

<num> 076 </num>

<EN-title> Political Culture in Democracy </EN-title>

<EN-desc> Find documents analysing or describing types of political culture in democratic systems. </EN-desc>

<EN-narr> Relevant documents provide information on theoretical aspects of political culture and/or give concrete examples of political culture within party activities, parliamentary work, or political life in general. </EN-narr>

</top>

<top>

<num> 077 </num>

<EN-title> Political Participation of Women in the GDR </EN-title>

<EN-desc> Find documents reporting on the participation of women in the political life of the German Democratic Republic. </EN-desc>

<EN-narr> Relevant documents analyse the role and position of women within the political system of the GDR. Documents should provide information on the opportunities of women for participation in the political life of the state. Documents reporting on female employment or the participation of women in political life after the political changes (in 1989/90) in Eastern Germany are irrelevant. </EN-narr>

</top>

<top>

<num> 078 </num>

<EN-title> Education in Turkey </EN-title>

<EN-desc> Find information about the educational system in Turkey. </EN-desc>

<EN-narr> Relevant documents report on educational systems in Turkey with respect to preschool and school programmes and professional training. Of particular interest are reports describing the Turkish system or comparing it with other systems, or analysing problems and discussing political reforms. Documents reporting the integration of Turkish emigrants into the national educational system of a migration country are irrelevant. </EN-narr>

</top>

<top>

<num> 079 </num>

<EN-title> Plebiscites in Democracy </EN-title>

<EN-desc> Find documents reporting on plebiscitarian elements in democracy. </EN-desc>

<EN-narr> Relevant documents discuss possibilities, advantages and disadvantages of direct popular influence on the decision-making process in a parliamentary system. Of special interest are comparisons of representative democracies with systems in which the people participate directly in decision-making, and also considerations on the integration of elements of direct democracy into parliamentary systems. </EN-narr>

</top>

<top>

<num> 080 </num>

<EN-title> Industrial and Economic Development of India </EN-title>

<EN-desc> Find documents reporting on the industrial and economic development of India. </EN-desc>

<EN-narr> Relevant documents discuss the economic progress of India and the importance of the industrialization of this country on both the national and international scale. Reports on problems regarding the economic development of this developing country are also of interest. </EN-narr>

</top>

<top>

<num> 081 </num>

<EN-title> Vocational Training Dropout </EN-title>

<EN-desc> Find documents on the dropout from TVE training in companies. </EN-desc>

<EN-narr> Relevant documents analyse reasons and facts for the dropout from Technical and Vocational Education training in companies and report on concepts to avoid the dropout from vocational training. Questions of vocational rehabilitation are not relevant. </EN-narr>

</top>

<top>

<num> 082 </num>

<EN-title> Professional Training of Immigrants </EN-title>

<EN-desc> Find documents on the professional integration of immigrants through vocational training in Western Europe. </EN-desc>

<EN-narr> Relevant documents discuss the difficulties of immigrants in Western European countries when they are integrated into the professional system of the host country. Studies treating problems of education at school and of professional training are of specific interest. Political discussions about foreigners in the context of the national labour market situation are irrelevant. </EN-narr>

</top>

<top>

<num> 083 </num>

<EN-title> Media and War </EN-title>

<EN-desc> Find documents on the commentatorship of the press and other media from war regions. </EN-desc>

<EN-narr> Relevant documents discuss the role of war reports and the role and the function of media in war times. </EN-narr>

</top>

</top>

<top>

<num> 084 </num>

<EN-title> New Media in Education </EN-title>

<EN-desc> Find documents reporting on benefits and risks of using new technology such as computers or the Internet in schools. </EN-desc>

<EN-narr> Relevant documents discuss the use of computers or the Internet in schools, the emergence of virtual classrooms, the use of learning software or PC-based simulation of teaching environments. Of particular interest are critical discussions of the new media. </EN-narr>

</top>

<top>

<num> 085 </num>

<EN-title> European Education Systems </EN-title>

<EN-desc> Find documents that report on comparative analyses between the education systems of different countries in Europe. </EN-desc>

<EN-narr> Relevant documents give comparisons of school and vocational education systems between any two European countries. Comparisons can be on a descriptive or quantitative basis and will discuss advantages and disadvantages of different systems. Otherwise, they can be made in the context of political discussions regarding the reform of educational systems. </EN-narr>

</top>

<top>

<num> 086 </num>

<EN-title> Political Symbols in Eastern Europe </EN-title>

<EN-desc> Find documents discussing the change in political symbols and rituals in the former socialist countries of Eastern Europe. </EN-desc>

<EN-narr> Relevant documents report studies on the destruction of monuments, icons or statues of everyday socialist culture in the former countries of the Warsaw Pact. Documents reporting on the change of political symbolism in these countries as well the importance of symbols for political thinking after the decline of socialism are also of interest. </EN-narr>

</top>

<top>

<num> 087 </num>

<EN-title> Integration of the Disabled in Schools </EN-title>

<EN-desc> Find documents reporting on the integration of pupils with a physical or emotional disability into the regular school system. </EN-desc>

<EN-narr> Relevant documents discuss problems of integrative education in comparison with the system of special schools for disabled pupils and/or report advantages and disadvantages of incorporating disabled pupils into the regular school system. </EN-narr>

</top>

<top>

<num> 088 </num>

<EN-title> Sports in Nazi Germany </EN-title>

<EN-desc> Find documents about the role of sports in the German Third Reich. </EN-desc>

<EN-narr> Relevant documents treat the functional aspects of sports in Nazi Germany between 1933 and 1945 in a political context. Of interest are both documents reporting on international sports competitions or popular sports for the masses in Nazi Germany. </EN-narr>

</top>

<top>

<num> 089 </num>

<EN-title> Professional Ethics in Journalism </EN-title>

<EN-desc> Find documents reporting on the profession of journalists with respect to moral issues and values in modern society. </EN-desc>

<EN-narr> Journalism is frequently criticized because it has degenerated to sensational reporting in the daily fight for circulation and audience ratings. Relevant documents will discuss social ethics with regard to journalism. </EN-narr>

</top>

<top>

<num> 090 </num>

<EN-title> Newspaper Reading </EN-title>

<EN-desc> Find documents discussing the importance of reading newspapers in the age of competitive media like television, radio and the Internet. </EN-desc>

<EN-narr> Relevant documents describe studies of newspapers and magazine reading. Documents comparing the printed media with its competitors are also of interest. </EN-narr>

</top>

<top>

<num> 091 </num>

<EN-title> Quality of Educational Systems </EN-title>

<EN-desc> Find documents discussing the quality control of educational systems in different countries. </EN-desc>

<EN-narr> Relevant documents report on models, political aims or theoretical approaches to guarantee the quality of the results of educational systems. For example, measures for process or finance controlling, evaluation models or systems for network formation between institutions involved in education will be of interest. Relevant documents must regard qualitative aspects of education. </EN-narr>

</top>

<top>

<num> 092 </num>

<EN-title> Party Members </EN-title>

<EN-desc> Find documents reporting on the membership of political parties. </EN-desc>

<EN-narr> Relevant documents report on the work of the members of political parties and types of active and passive membership. Reports on the assumption of representative tasks and the possibilities of participation in political activities within a party are also of interest. </EN-narr>

</top>

<top>

<num> 093 </num>

<EN-title> Burnout Syndrome </EN-title>

<EN-desc> Find documents reporting on the burnout syndrome. </EN-desc>

<EN-narr> Relevant documents describe or analyse the exhaustion syndrome within different professional groups or sections of the population. </EN-narr>

</top>

<top>

<num> 094 </num>

<EN-title> Homosexuality and Coming-Out </EN-title>

<EN-desc> Find documents reporting on the coming-out of homosexuals. </EN-desc>

<EN-narr> Relevant documents describe the process of admitting publicly to homosexuality and the effect of such admissions on the lives of people involved. </EN-narr>

</top>

<top>

<num> 095 </num>

<EN-title> Sunday Work </EN-title>

<EN-desc> Find documents discussing the problem of working on Sunday. </EN-desc>

<EN-narr> Relevant documents discuss economic and political issues with regard to the necessity for work on Sunday and arguments against a widespread introduction of Sunday working. </EN-narr>

</top>

<top>

<num> 096 </num>

<EN-title> Costs of Vocational Education </EN-title>

<EN-desc> Find documents reporting on the costs and benefits of vocational education. </EN-desc>

<EN-narr> Relevant documents analyse the costs, benefits and success of vocational education and further education in terms of quality and quantity. </EN-narr>

</top>

<top>

<num> 097 </num>

<EN-title> Religious Sects </EN-title>

<EN-desc> Find documents reporting on the existence of religious sects. </EN-desc>

<EN-narr> Relevant documents discuss the existence, the spreading, the distinguishing features and the views of religious sects. </EN-narr>

</top>

<top>

<num> 098 </num>

<EN-title> Canadian Foreign Policy </EN-title>

<EN-desc> Find documents discussing the foreign policy of Canada. </EN-desc>

<EN-narr> Relevant documents describe or analyse the foreign policy of Canada. Of particular interest are documents that discuss the Canadian position regarding the USA, but also its world politics. </EN-narr>

</top>

<top>

<num> 099 </num>

<EN-title> Sports and Youth Work </EN-title>

<EN-desc> Find documents reporting on successes but also on problems arising from the involvement by youth workers of young people in sporting activities. </EN-desc>

<EN-narr> Relevant documents provide information about adolescents and children in the context of youth work and the function of sports. Of particular interest are reports on social work with young people and their inclusion in regional sporting activities as a success concept of the social work. </EN-narr>

</top>

<top>

<num> 100 </num>

<EN-title> Pub Culture </EN-title>

<EN-desc> Find documents reporting on the habits of visitors to pubs. </EN-desc>

<EN-narr> Relevant documents analyse the habits of pub visitors, and in general pub culture and environment. </EN-narr>

</top>