CLEF-IP 2011

Submission Guidelines for Prior Art and Classification Tasks

Please refer to the Track's Guidelines for a description of the document collection, topic file structures, etc. available on the track's web and download sites:

http://www.ir-facility.org/clef-ip

https://download.ir-facility.org/clef/2011/

The format of the *clefip* submissions is similar to the *trec* format.

A submission consists of a single ASCII text file containing at most 1000 lines per topic, in the standard format used for most t*rec* submissions: five columns separated by whitespace. In addition submissions for the classification tasks must contain a second file („runC") in the format

„subclass patent_document score" respectively „subgroup patent_document score"

Example (parts of) submission files

PAC (Prior Art Candidates search task)

one submission file (extension runP): max 1000 entries per patent

EP-8888888-A2	Q0	EP-0000001	1	1012
EP-8888888-A2	Q0	EP-0000002	2	1011
EP-8888888-A2	Q0	EP-0000002	3	999

….

CLS (Classification task) submission for task1

submission file1 (extension runP): by patent document, max 5 entries per patent

CLS1_EP-9999999-A1	Q0	A20K	1	3010
CLS1_EP-9999999-A1	Q0	B32L	2	3008
CLS1_EP-9999999-A1	Q0	B46A	3	2985

….

submission file2 (extension runC): by subclass

| A20K | CLS1_EP-9999999-A1 | 3010 |

A20K	CLS1_EP-8888888-A1	3008
A20K	CLS1_EP-7777777-A1	2985

CLS2 (Refined Classification task) submission for task2

submission file1 (extension runP): by patent document, max 20 entries per patent

CLS2_EP-9999999-A1_A61K	Q0	A61K9/16	1	3010
CLS2_EP-9999999-A1_A61K	Q0	A61K9/50	2	3008
CLS2_EP-9999999-A1_A61K	Q0	A61K9/38	3	2985

…

submission file2 (extension runC): by subgroup

A61K9/16	CLS1_EP-9999999-A1	3010
A61K9/16	CLS1_EP-8888888-A1	3008
A61K9/16	CLS1_EP-7777777-A1	2985

...

In the trec style submission files

- the first column is the topic number (as specified in the topic file)

- the second column is the query number within that topic. This is to be set to Q0;

- the third column is

    - for the *PAC* task, the document number of the retrieved document, *excluding* the kind code; this is the value of the *country* and *doc-number* attribute in the *patent-document* xml tag, concatenated with a dash;

    - for the classification subtask1, the IPC classification up to *subclass level* to which the topic is assumed to belong to;

    - for the classification subtask 2, the IPC classification up to *subgroup level* to which the topic is assumed to belong to;

- the fourth column is the rank of the document retrieved, respectively the rank of the subclass or subgroup symbol (starting from 1);

- the fifth column shows the score (integer or floating point) that generated the ranking. This score must be in decreasing order.

A run is uniquely identified by the participant id, the method used and the task type.

participantID_method and/or runID_*PAC_*.runP

participantID_method and/or runID_CLS1_.runP

participantID_method and/or runID_CLS1_.runC

participantID_method and/or runID_CLS2_.runP

participantID_method and/or runID_CLS2_.runC

Note that the classification system used by the whole clef-ip collection is IPC-R and later. The results too should be IPC-R codes.

It was agreed at the last workshop that the number of entries per topic/class-code in the result files is to be set to 5 for the first submission fomat. For the refined classification this limit is 20. For PAC up to 1000 results per topic are allowed.

The maximum number of runs per task for a participant is 8. Submissions will be uploaded to an ftp server, access credential are to be communicated at a later time. Together with the run files, and for each of them, we require the participants to upload a concise description of the retrieval methods and document fields used to obtain the respective runs. The naming of the description files should follow the naming of the run files, where the group id, method and task type are part of the file name.

Evaluation Measures
We will mainly use the the *trec_eval* implementations of the precision, recall, F1, MAP and nDcg measures. Further, we plan to try out new measures like, for example, Pres[1] . The relevance judgements to be used during the evaluation will be extracted automatically from the patent documents used to create the topics.

---

1 W. Magdy and G.J.F. Jones. Pres: A Score Metric for Evaluating Recall--Oriented Information Retrieval Applications. In Proceedings of Sigir'10, Geneva, Switzerland.