GUIDELINES for the PARTICIPANTS in
QA on sense disambiguated data QA-WSD@CLEF 2008
(01/03/2008)

INTRODUCTION

The QA-WSD task will bring semantic and retrieval evaluation together. The participants will be offered the same queries and document collections from the main QA exercise which have been annotated by systems for word sense disambiguation (WSD). The goal of the task is to test whether WSD can be used beneficially for Question Answering.

The exercise scenario is event-targeted QA on a news document collection. In the QA-WSD track only English monolingual and Spanish to English bilingual tasks are offered, i.e. English will be the only target language, and queries will be available on both English and Spanish. The queries will be the same as for the main QA exercise, and the participation will follow the same process, except for the use of the sense-annotated data.

A large number of questions will be topic-related, i.e. clusters of questions which are related to the same topic and possibly contain anaphoric references between one question and the other questions. Contrary to the main QA task, Wikipedia articles are not included, and thus systems need to reply to the questions that have an answer in the news document collection (more details below).

The goal of this task is to evaluate whether word sense information can help in certain queries. For this reason, participants are required to send two runs for each of the monolingual/bilingual tasks where they participate: one which **does not** use sense annotations and another one which **does** use sense annotations. Whenever possible, the only difference between the two runs should be solely the use or not of the sense information. **Participants which send a single run will be discarded from the evaluation**.

The WSD data is based on WordNet version 1.6 and will be supplemented with data from the English and Spanish WordNets in order to test different expansion strategies. Several leading WSD experts will run their systems, and provide those WSD results for the participants to use.

As most of these guidelines are a copy of the QA guidelines, we will only include the sections where there are differences.

ORGANIZING COMMITTEE
- Eneko Agirre (University of the Basque Country, Basque Country)
- Anselmo Peñas (UNED, Spain)
- Richard Sutcliffe (University of Limerick, Ireland)
- German Rigau (University of the Basque Country, Basque Country)
- Xabier Arregi (University of the Basque Country, Basque Country)

IMPORTANT DATES

Release of document collections with automatic sense annotations: **March 1st**

Registration is now open at the website[1] (participants must communicate the tasks in which they are going to participate). Please register before **April 28th**

Question sets for all tasks will be released on: **May 5th**.

Runs must be returned by **May 9th**.

Instructions concerning the submission procedure for the results will be given when the track starts.

Individual results will be released to each participating group from **July 1st**.

DOCUMENT COLLECTIONS

Registered participants can download the corpora from the CLEF website (registration form and end-user agreement must be first filled in). Note that you will need to download the sense annotated versions, which will be available from the 1st of March.

Target document collections are the following Newswire Collections;

TARGET LANGUAGE	COLLECTION	PERIOD
	Los Angeles Times	1994
English (EN)	Glasgow Herald	1995

---

[1] http://nlp.uned.es/clef-qa (Please note that you must also register at the Clef Campaign website http://www.clef-campaign.org/)

QUESTIONS

No changes with respect to main guidelines.

ANSWERS

No changes with respect to main guidelines.

INPUT FORMATS

Check for the format of the sense-annotated document collections

The format of the questions will be the same as those of the main task, but an additional version of the questions will be provided with automatically annotated senses with the following xml format (UTF-8 encoded):

Check for the format of the sense-annotated query collections

OUTPUT FORMAT

The format of the questions will be the same as those of the main task, but the run_id will show clearly which run uses the word sense annotations.

For instance, a group working on the monolingual task should submit two runs 1 and 2, which will be respectively called "XXXX081enen" (system from "XXXX", year "08", run   "1" which does not use sense annotations, and English-English task) and
"XXXX082enens" (system from "XXXX", year "08", run   "2" which does use sense annotations, English-English task, and a last "s" to confirm that this run uses sense information). it is very important that groups use run 1 for the system without sense information, and 2 for the system with sense information.

EVALUATION
The evaluation will be similar to that of the main task, except the fact that we will ignore questions which do not have an answer in the news collection. That is, we will only evaluate the answers of the queries which have at least one correct answer in the news collection.

CONTACT INFORMATION

Eneko Agirre: e.agirre *at* ehu.es

Anselmo Peñas: anselmo *at* lsi.uned.es