

# ShARe/CLEF 2015 eHealth Task 2 Guidelines

Participants will be provided with a collection of documents (health-focused web pages) to be searched, example queries they can use to develop their search systems, and test queries they will have to use to prepare runs for submission. The evaluation will be conducted using these latter test queries. Teams are allowed to use any outside resources in their algorithms. This year's task encourages participants to submit, along with automatic runs, also manual runs, i.e. runs that involve human intervention for example to reformulate the original query or provide relevance feedback.

## *Timeline:*

- Run submission deadline: 15th of April (Hawaii time UTC-10)
- Task result release: 1st of June
- First working notes submission deadline: 7th of June
- Final working notes submission deadline: 15th of June
- CLEFeHealth2015 one-day lab session: 7-11 Sep 2015 in Toulouse, France

## **Run Submission Guidelines:**

### *Runs description*

**Task 2a (monolingual information retrieval):** Participants can submit **up to ten ranked runs for the English (EN) queries** in Task 2a. The runs should include the top 1,000 documents returned for each in the test set. Manual runs should be clearly marked (see below).

**Task 2b (multilingual information retrieval):** Participants can also submit **up to ten ranked runs for each of the other language queries (Arabic (AR), Czech (CZ), French (FR), German (DE), Spanish (ES), Italian (IT), Portuguese (PO), Farsi (FA))**. For example, a participant team may submit up to 10 runs for the English queries, up to 10 runs for the German queries and up to 10 runs for the Czech queries. The runs should include the top 1,000 documents returned for each in the test set. Manual runs should be clearly marked (see below).

For each subtask, the run 1 should be a baseline run, where only the titles of the query can be used, and no external resource (including other corpora, ontology, etc.) or manual intervention can be used. The other runs must be ordered by priority (see below) and can use any resource, manual intervention, or combination thereof. The first run will provide to the participant team and the community with an understanding of the impact the strategies used in the remaining runs have on the retrieval effectiveness.

Submitted runs should use the following naming convention:  
<TeamName>\_<QueryLanguage>\_Run<RunNumber>.<FileFormat>

For example: QUT\_EN\_Run1.dat , QUT\_CZ\_Run5.dat

Please use as TeamName the name you have given us at registration time.

Submitted runs have to follow the TREC format; a tool to validate the format of the runs ('format-script-clefeHealth-task2.zip') is available in Physionet.

### *How to Choose the Priority of the Runs*

Participants can submit up to 10 runs; while organisers will attempt to sample as many runs as possible to build the relevance assessment pools, the actual runs that will be sampled may be less than ten per participants. Runs will then be sampled according to priority; the priority of a run is expressed by the

number that is assigned to the run by the participant, for example run 2 has a higher priority (and thus higher likelihood of inclusion in the assessment pool) than run 8. Run 1 (the baseline) has the highest priority; run 10 the lowest. It is therefore important that participants prioritise the runs as they see more convenient.

### **Runs Submission**

**Runs should be submitted through EasyChair:**

<https://easychair.org/conferences/?conf=clefehealth2015resul>

1. Submit separately to each task by selecting “New Submission”. You will submit all runs for one task at the same time. After you have created a new submission, you can update it, but no updates of runs are accepted after the deadline has passed.
2. List all your team members as “Authors”. “Address for Correspondence” and “Corresponding author” refer to your team leader. Note: you can acknowledge people not listed as authors separately in the working notes (to be submitted by June 7 (see below)) – we wish this process to be very similar to defining the list of authors in scientific papers.
3. Please provide the task and your team name as “Title” (e.g., “Task 2a: Team QUT” or “Task 2a using extra annotations: Team QUT”) and a short description (max 100 words) of your team as “Abstract”. See the category list below the abstract field for the task names. If you submit to multiple tasks, please copy and paste the same description to all your submissions and use the same team name in all submissions.
4. Choose a “category” and one or more “Groups” to describe your submission. We allow up to 10 runs for Task 2.
5. Please provide 3-10 “Keywords” that describe your the different runs in the submission, including methods (e.g., Language Models, BM25, Divergence from Randomness, etc.) and resources (e.g., Unified Medical Language System, expert annotation, etc.). You will provide a narrative description later in the process.
6. As “Paper” please submit a zip file including the runs for this task. Submitted runs should use the following naming convention: <TeamName>\_<QueryLanguage>\_Run<RunNumber>.<FileFormat>  
For example: QUT\_EN\_Run1.dat , QUT\_CZ\_Run5.dat  
Run ID 1 should refer to the mandatory baseline run (mandatory run).
7. As the mandatory attachment file, please provide a txt file with a description of the submission. Please structure this file by using your run-file names above. For each run, provide a max 200 word summary of the processing pipeline (i.e., methods and resources). Be sure to describe differences between the runs in the submission.
8. Before June 7, 2015, please submit your working notes. Formatting and submission guidelines will be available on this website closer to the submission date.
9. Post submission, relevance assessment will be conducted on the test queries to generate the complete result set. The organizers will provide the evaluation results via the Easy Chair for CLEFeHealth2015 runs. This includes your ranking with respect to other teams as well as the value(s) of the official evaluation measure(s).

### **Evaluation Metrics:**

Evaluation will focus on P@5, P@10, NDCG@5, NDCG@10, but other suitable IR evaluation measures will also be computed for the submitted runs.

Evaluation metrics can be computed with the trec\_eval evaluation tool, which is available from [http://trec.nist.gov/trec\\_eval/](http://trec.nist.gov/trec_eval/).

***Working notes:***

Participating groups in Task 2 must submit a report (working notes) describing their Task 2 experiments.

Details on preparing working notes & link to the working notes submission system will be made available soon.