

TER Project M1 IC

Specifications

Table of content

Table of content	1
1. Introduction	2
1. Context	2
2. Background	2
3. Glossary	2
2. Description of the request	3
1. Objectives	3
2. Project product	3
3. Product features	3
3. Constraints	4
1. Deadline constraints	4
2. Material constraints	4
3. Functional constraints	4
4. Other constraints	5
4. Course of the project	6
1. Planning	6
2. Resources	6
5. Organization	6
6. References	6

1. Introduction

1. Context

The computer development project described in this document is part of the TER project ("Travaux Encadrés de Recherche" : supervised research project) of the Master MIASHS Informatique et Cognition.

Francis Jambon and Philippe Mulhem are researchers in computer science at the Laboratoire d'Informatique de Grenoble and have proposed, in collaboration with students of the Master, a system aiming at improving and fluidifying techniques used in WEB research. Indeed, WEB research represents today a major stake for most companies, and a system aiming at making it more effective is very much sought after.

It is now necessary to find a solution to evaluate the functioning of the system. The researchers are therefore collaborating with students of the Master of Computer Science and Cognition to propose a functional solution by June 2022.

2. Background

Francis Jambon and Philippe Mulhem are trying to improve some techniques used in WEB search. To do so, they have developed a sophisticated search engine based on the Terrier engine and a results database as well. This database is able to return relevant results to the user for the query he has sent. At the moment, the relevance of the results is measured using an eye-tracker. Francis Jambon and Philippe Mulhem have conducted a study in which they show that there is an advantage to using the user's gaze as an implicit measure of the relevance of the words proposed in the results [Sungeelee, Jambon, Mulhem, 2020]. In order to improve their search engine, Francis Jambon and Philippe Mulhem aim to have the relevance of the results evaluated by experts in the query domain. The request is to develop an ergonomic graphical interface to allow them to quickly judge the relevance of the results of a query.

An existing application, named "Relevation!" [B. Koopman and G. Zuccon, 2014], dedicated to the same problem, has already been implemented and used in the past, but the Relevation! interface has limitations: it is outdated; the technologies used are not up to date; and it is not directly usable in our case, as the notions of experts and snippet are not present. So it is necessary to create our own software, more robust, scalable, and up to date compared to Relevation!.

3. Glossary

We define here the list of words that will be used along the project:

- Administrator: the person in charge of the management of the overall assessment.
- Annotation: one annotation consists in setting a degree of relevance to one word, snippet or document.
- Assessment: one assessment consists in annotating one snippet/document for a given topic.
- Corpus: set of documents on which the relevance is evaluated.
- Document: a document is a long string, it has an ID. It is independent from the topic.
- Expert: the person who has knowledge about the **topic**. He uses the interface to do the **task**.
- Pool : set of documents on which annotations will be made. It is a subset of the **corpus**.
- Rating campaign : the set of tasks.
- Relevance: the degree of relevance is evaluated on a scale of three levels.
- SERP: search engine result page.
- Snippet: one snippet refers to one document. It is displayed in a frame, gives the title and a short preview of this document.
- Task: set of assessments given to one expert.
- Topic: the query associated with one task that is given to the **expert**. It is a character string. Ex : "health benefits of running".

2. Description of the request

1. Objectives

- Define an evaluation framework: create an interface allowing to manually indicate, when presenting a page of results, what the user finds relevant in order to compare the results with those obtained by the eye tracker.
- Propose an innovative, ergonomic and easy to use solution at the interface level.
- Produce a simple, robust, extensible and documented code in English.
- Aim for the least possible interactions during use. Limit the number of clicks, reduce the annotation time, ...
- Two kinds of users: experts and administrators. The expert

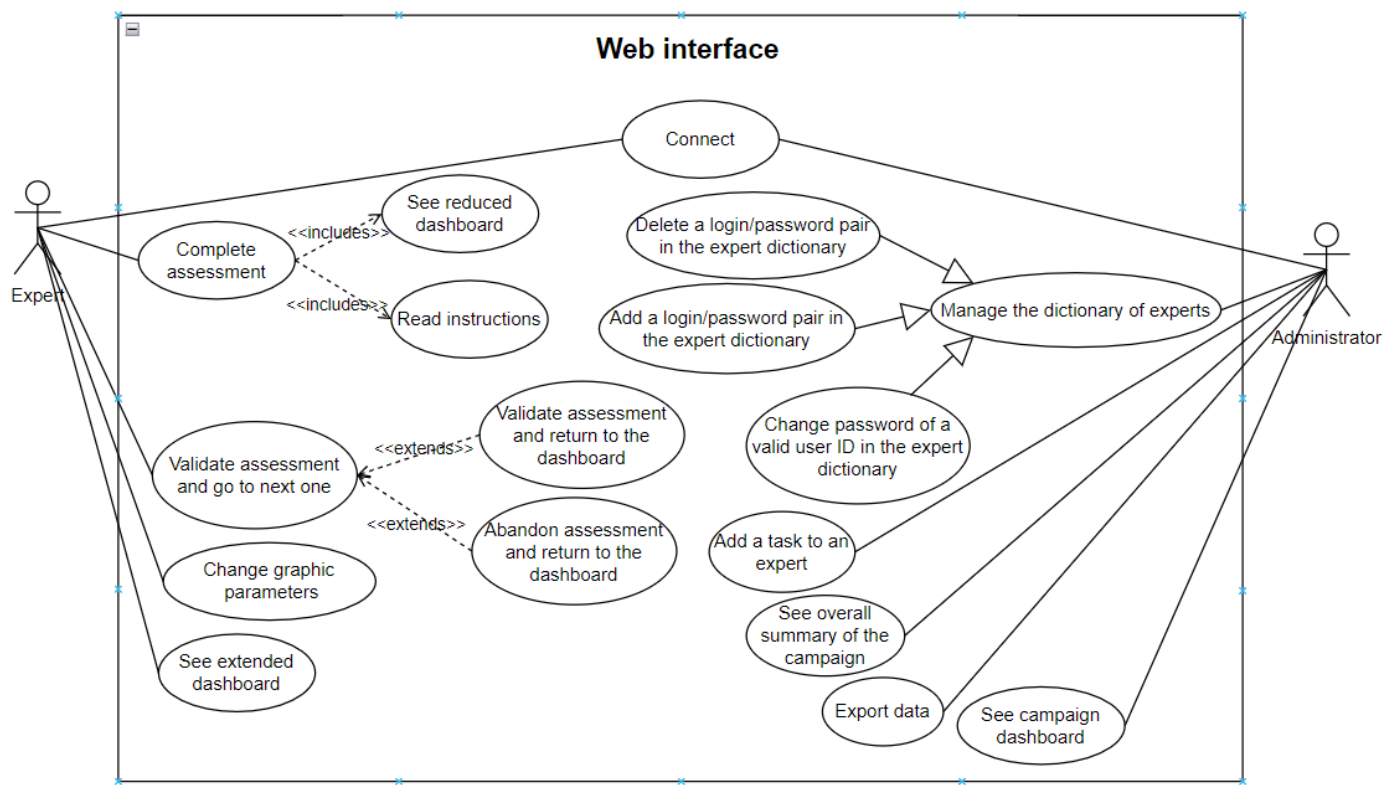
2. Project product

The product is a Web interface where experts in the field of the queries are able to evaluate the relevance of the documents or snippets to the queries, and to assess, when needed, the relevance of the words of the documents.

3. Product features

The interface allows experts to quickly define the words that seem to be the most relevant for a given search. Thus, the product allows to carry out an annotation campaign of the snippets and documents obtained in the result of the query according to their level of relevance. The final objective is to evaluate to what extent the words added to the query, thanks to the eye-tracking system, are helping to refine the query.

- When an expert has finished one assessment, he sends a file containing all his annotations to the remote server.
- The interface is accessible from any remote station.
 - Use of a login/password to identify the experts who connect.
 - Management of concurrent access to the system.
- For the annotation, there are two to four levels of relevance for the words of a snippet/document as well as the document or snippet in its entirety.
- Ergonomic method to quickly assess a word in a snippet/document.
 - Different fonts according to the different levels of relevance depending on the rating campaign.
 - Choice of the color to be made according to the background.
 - Left click on a word : increments the relevance level by 1. Loops on the lowest level when a word has reached the highest relevance level.
 - Right click on a word: directly assigns the highest relevance level.
- The set of results to be annotated by the experts is distributed so that several experts may annotate the same result
 - An expert must annotate the results of several topics.
 - In both cases the assignment is random, an expert does not choose one result to annotate rather than another.
- The data generated by the annotation is accessible through the output files. These files are in XML, json, csv or text format (according to what is used in the information retrieval domain) and can be imported by the administrator to be used on external softwares.
- Dashboards for experts and administrators : reduced version and extended version.



When the expert arrives on the application a connection page is displayed. He can enter his login and password in the two fields. The credential is then verified by the server and the access is either granted or denied.

Once logged, he has access to the main menu. The main menu is composed of a dashboard and a button to start assessing the unfinished snippets/documents. When the expert is done annotating the levels of relevance, he can either click on "validate and go to next assessment", "validate and go back to dashboard" or "abandon assessment and return to dashboard". In general, the user is expected to validate and go to the next assessment.

If there is no more assessment to provide by the expert, he is led back to the dashboard and the "start assessing" button is disabled.

When the administrator arrives on the application, he also has to enter his credentials.

Once logged, the administrator has a similar main menu page as the expert. The "start assessing" button is missing since he has no assessment to do. His dashboard is enhanced with two tabs "experts" and "topics". These tabs allow him to see the experts' advancement and the topics completion. He can see the summary of the campaign on the first line of each table which tells him the number of completed assessments by expert out of total number of assessments by expert and the number of completed topics out of the total number of topics.

3. Constraints

1. Deadline constraints

Part-time from October 25, 2021 to April 29, 2022 (every Monday morning)

Full time from May 2, 2022 to June 24, 2022

Final version of the specifications, recipe book and development plan to be submitted by January 10, 2022

2. Material constraints

- Internet connection
- PC (in principle this will be the only support)

- Access to the remote server to retrieve query results.
- Access to the remote server to send the annotations made by the expert.

3. Functional constraints

- The interface must be :
 - cross-platform (not OS or browser dependent).
 - Remotely accessible (as experts are located around the world).
- Concerning the campaign :
 - Campaigns have IDs. In general, there is only one campaign in progress. The IDs are defined in case another campaign starts while the previous one is not finished.
 - During a campaign, the expert must assess either the snippet and/or its words OR the document and/or its words. Never both at the same time. This will depend on the campaign in progress. The interface must however support a display for all options.
 - The relevance scale may be different for each rating campaign. The relevance scale is set for the whole rating campaign.
 - An expert cannot partially assess a snippet/document. If he leaves it, all changes will be lost.
- Concerning the output file :
 - The data generated through a campaign are all presented according to an abstract model in order to have a single form of output file.
 - The format of the file sent after finishing one assessment is determined by what is used in the information retrieval domain.
- Concerning the input files :
 - The results of a query are all stored in a file containing :
 - the campaign ID
 - the expert ID
 - the topic ID
 - the document/snippet ID
 - and the list of words followed by their annotation
 - An administrator must import a file containing lines of login/password to add experts to the campaign.
- Concerning the assessments:
 - Once it is assessed, an answer cannot be assessed again.
 - All the results of all the topics of a campaign must be assessed by more than 1 expert.
 - The sequence of assessments is static: the order in which the assessments are performed for each expert is defined in advance, not when the expert loads the next assessment.
 - They are distributed among the experts so that the experts work on different topics.
 - The documents/snippets are presented to the expert sorted by topic to make the assessment easier.
- Concerning the users (experts and administrators) :
 - They have unique logins. If a user is an expert and an administrator at the same time, then they have access to two separate accounts.
 - An administrator must be able to add a task to an expert. However, he cannot delete or modify a task.
- Concerning the dashboards:
 - On the homepage of the interface, the extended version of the dashboard is presented. When the instruction and the assessment are displayed, the reduced version of the dashboard is shown. Experts and administrators have different dashboards.
 - The experts reduced dashboard displays :
 - the amount of finished assessments for the current topic / the number of assessments for the current topic
 - The experts extended dashboard displays :
 - one line per topic with the amount of finished assessments for each topic / the number of assessments of the topic.

- The number of completed topics / the total number of topics
- The number of completed assessments / the total number of assessments
- On the admins dashboards there are 2 different views: expert view and topic view. Both are composed of a table in which each line represents a topic or an expert. There can be an option for the administrator to delete a topic (respectively an expert) which removes it from the list of topics of the campaigns (respectively removes him from the experts dictionary). (optional)

4. Other constraints

Use of the English language for every document to make them internationally accessible.

Comply with the European General Data Protection Regulation.

4. Course of the project

1. Planning

- Analysis of the customer's needs
- Define the implicit needs of the user when using the interface
- Writing of the specifications
- Decision on the different possible choices: choice of the atomicity of the record, of the architecture: client-based vs web-based, of an ergonomic method to define the level of relevance, of the technology for the interface
- Drafting of the recipe book and the development plan
- Development
- Writing of the internal documentation
- Writing of the user manual
- Writing of the installation manual
- Writing of the test plan
- Unit tests
- Tests, with real users to compare with the obtained results
- Writing of the project report and summary in French and English
- Realization of the demonstration video

2. Resources

- access with badge to the IMAG building and reserved office
- Git repository shared with the project tutors
- File shared with the project tutors on Google Drive

5. Organization

All parts of this project are done by the two students in collaboration with the research supervisors.

A meeting is held every week so that the research teachers can follow the progress of the work.

A meeting with the referent teacher also takes place regularly in order to make the link with the teaching expectations.

6. References

B. Koopman and G. Zuccon. Relevation!: *An open source system for information retrieval relevance assessment*. In Proceedings of the 37th annual international ACM SIGIR conference on research and development in information retrieval, Gold Coast, Australia, July 2014.