MyBestQuery – A serious game to collect manual query reformulation

Adrian Chifu, Serge Molina, Josiane Mothe

***Abstract*—This paper presents MyBestQuery, which is a serious game designed to collect query reformulations from players. Query reformulation is a hot topic in information retrieval and covers many aspects. One of them is query reformulation *analysis* which is based on users' session. It can be used to understand user's intent or to measure his satisfaction with regards to the results he obtained when querying the search engine. A*utomatic* query reformulation is another aspect of query reformulation. It automatically expands the initial user's query in order to improve the quality of the retrieved document set. This mechanism relies on document analysis but could also benefit from manually reformulated query analysis. Web search engines collect millions of search sessions and possible query reformulations. As academics, this information is hardly accessible for us. MyBestQuery is designed as a serious game in order to collect various possible reformulation users suggest. The more long-term objective of this work is to analyse the humanly produced query reformulation in order to both analyse manual query reformulation and compare them with the automatically produced reformulations. Preliminary results are reported in this paper.**

***Index Terms*—Information retrieval, Query reformulation, Serious game**

I. Introduction

USERS's queries play an important role in the information retrieval process. A query corresponds to the way a human expresses his information need to the system but also to the way the system matches the user's need to the documents to retrieve.

Understanding or rather not understanding users' queries is a problem search systems face. Some research attempts to discover the semantics of user's query words by using some knowledge resources such as ontologies [1] [2] or by disambiguating query terms [3]. In practice still, the "bag of words" assumption[1] [4] is the most common and does not rely on query understanding.

However, as pointed out by Boldi *et al.* [5], if we could understand query reformulation patterns we may be able to understand user intent and build systems that provide users with efficient assistance. Not only query reformulation can help understanding user expectation, but it can also help measuring user satisfaction. For example, Hassan *et al.* consider the relationship between the user's current query and the next query as implicit signals of query satisfaction [6].

Another important topic related to query reformulation is automatic query reformulation. It aims at improving the search engine effectiveness by automatically building a new query from the user's initial query. Some methods are pre-retrieval methods, while others are post-retrieval such as query relevance feedback [7] [8]. Compaoré *et al.* show that expansion parameters are very important for difficult queries [9]. Considering the reformulated queries, Ermakova *et al.* also show that a linguistic based analysis of queries can help understanding why some automatic query expansion methods work better than others [10].

These examples of query reformulation usage illustrate the importance of harvesting query reformulation examples. Web search engines collect in-house query logs that can be used by company members, but are rarely accessible to academics. Our goal is to collect human based query reformulations at a large scale. However, recruiting users for experiments is a challenge most user studies face. Crowdsourcing is a way to overcome this challenge. Yuen et al. define crowdsourcing as "distributed problem-solving and business production model" [11].

Crowdsourcing has mainly been used in IR to create corpora to be used for evaluation purposes and specifically for relevance assessment [12] or document annotation [13]. In our case, the problem is also to create corpora by humans who provide some annotation-like information. Yuen et al. present a taxonomy of crowdsourcing in which four types of applications are distinguished, namely: voting systems, information sharing systems, games, and creative systems. In our view, games represent an interesting way to collect information which can help reaching our goal of a large scale gathering of annotations if gamification is done. Like in many field, gamification for IR is a hot topic [14].

In this paper, we present MyBestQuery, a serious game we developed that aims at collecting query reformulations and

This paper was submitted on the 29 of May for review. It was supported in part by the French ANR through the CAAS project (Contint Call) under Grant ANR-10-CORD-0001.

A. G. Chifu, Dr. was with the Institut de Recherche en Informatique de Toulouse, UMR5505 CNRS, Université de Toulouse, France. He is now with the LSIS. (e-mail: Adrian.Chifu@lsis.org).

S. Molina , is with the Institut de Recherche en Informatique de Toulouse, UMR5505 CNRS, France (e-mail: serge.molina@irit.fr).

J. Mothe is with the Institut de Recherche en Informatique de Toulouse, UMR5505 CNRS, ESPE, Université de Toulouse, France (e-mail: Josiane.Mothe@irit.fr).

[1] Document and queries are represented as sets of unordered words.

making it available for research purpose. It is available at **http://mbq.irit.fr**. Players challenge the initial query in its capability to retrieve more relevant documents, they get rewards depending on their achievement; padlocks, new levels and leader boards are other means we included in the game.

II. Related Work

A. Crowdsourcing in IR Research Field

The idea of crowdsourcing data using serious games is not new in Information Retrieval. For example, crowdsourcing has been used for relevance assessments [15] and clustering [12]. In IR, relevance assessment consists for human to decide whether a retrieved document is relevant to a query or not. In TREC evaluation campaign, participants' systems retrieve 1000 documents for each query (or topic). This makes relevance assessment a very time consuming task if done by a few people. According to Alonso and Mizzaro, "crowdsourcing is a cheap, quick, and reliable alternative for relevance assessment" [15]. However, some challenges regarding crowdsourcing should not be ignored. The biggest concern remains the constant need for quality control, since there is the risk of contributions from workers that are not qualified enough for the task, even if they might think about themselves that they are. Moreover, the use of gamification in a crowdsourcing application must be done carefully in order to obtain usuable and trustable results. Some of the means used in order to guarantee good results are presented in section IV.

Another very popular type of crowdsourcing application is document annotation. For example, Nowak and Rüger collect image annotations by this mean [13].

B. Crowdsourcing Gamification

Games with a purpose are a way to implement crowdsourcing [16]. Various games exist to allow the annotation of images and videos by players (see http://www.insemtives.eu/games.php for examples). Lafourcade developed the JeuxDeMots game to make human providing semantic relationships between terms [17].

Various features make an application be considered as a game [18], [19]; the main features are developed in section III and we instantiate them for our purpose in section IV.

III. Goals and Means

A. Collecting Query (Re)formulations.

The main goal of the application is to collect information on how humans would (re)formulate a query according to an information need.

More precisely, given an information need and an initial query, we ask the player to provide a query that he think will perform better; that is to that that will make the system retrieving more or better documents according to the information need.

To make the task more concrete and feasible, we first provide the player with a short description of the document collection and a short description of the search engine which is used (basically a Google-like search engine).

B. Users' Information Needs and Initial Queries.

Rather than using our own data, we decided to use reference collections. We choose TREC[2] corpora. TREC provides topics composed of a title part which could simulate a user's query, a description part which provides more information on the user's intent, and a narrative which can be used by assessors to decide on the relevance of a retrieved document. An example of a TREC topic is given Table 1.

The advantage of using TREC reference collections is that we also have access to relevance judgments for each query on the document collection. We can then evaluate the player's queries against the collection and compare the results to the initial query.

In the current version of the game, we use ClueWeb 2012 B document collection (http://lemurproject.org/clueweb12/) and the Robust collection (http://trec.nist.gov), as well as the Indri/Lemur search engine (http://www.lemurproject.org/).

IV. Gamification

A. Main Gamification Feature Description

Various features make an application be considered as a game [18], [19]. Among them, we consider the features described in the following subsections.

There should be a "Sense" to Play. Most of internet users query search engines and have faced the problem of expressing an efficient query. This fact helped us to design a meaningful scenario. It is easy to explain that players will have to try to build a query from an information need and that they will have to try to write the most useful query, the one that will yield as much relevant documents as possible. The action "the player formulates a new query" is clearly related to the outcome "a score that depends on the number of relevant documents retrieved using this query with the given search engine".

Play Implies Interactivity. Interactivity is implemented through various choices the player has in the MyBestQuery game. He can choose among several information needs. He can have a strategy that starts with the information needs he thinks are the easiest and learns from them. He can also decide to reformulate several times the same information need or rather change information need.

Game Exists Within a Frame, Rules, Play, and Culture. Rules are quite simple in this game and are given to the user before he starts playing. The score he obtained depends on the number of relevant documents that are retrieved on the top-10 retrieved documents using his query. The higher this number, the higher the score is. At any time the user can have access to the scores of the other players, making the game more challenging. Indeed, in this game the payers are not playing one against the others.

B. Concrete Gamification of Query Reformulation

Rules. A game consists in reformulating a given query with the objective for the player to help the system retrieving

[2] http://trec.nist.gov provides the details of the TREC evaluation program, tracks and data sets.

relevant documents.

Tutorial. A tutorial for novice players is provided. The tutorial part is also a simple way to explain the rules to players and to give sense to the game. In our case, the tutorial part consists in providing two information needs the new player can choose. The rules are explained and the various choices he has are provided at each step. We also provide the player with some hints on how he can get good scores during the game. During the tutorial, he can try on and he will have a feedback on his try which is different in function of the context (topic, effectiveness of his new query).

Reward and scores. The player earns points based on the effectiveness of the new formulation of the query he suggests. The score the player gets for a given game (a given query) is from 0 to 3 for each reformulation he tries. The score depends on how much the initial precision (that is to say using the initial query formulation) after the top k retrieved documents (P@k) has been improved. For calculating the score, we process the user query using Indri/lemur system on the document collection and calculate the precision at 10 documents (number of relevant documents in the set of the 10 first retrieved documents). We compare this result to the precision using the initial query (topic title).

Padlocks and levels. When the player gets enough points thanks to the games he played, padlocks are opened, giving access to new queries. Currently the points are easy to obtain so that padlocks can be unlocked when the player starts playing. Depending on the enthusiasm of the players we may make the opening of new padlocks more and more difficult.

Leader boards. A board displays the scores of the best players in order to maintain some challenge in the game. Indeed, since the player does not play against another player, leader board is a mean to keep players active. In future versions, we would like to improve the scenario of the game so that players could play together.

Ensuring the collection of valid data. We tried to avoid as much as possible bias in the collected data. The first mean was by displaying the queries in a random order so that the learning effect is cancelled out. The second mean consisted in carefully choosing the wording in the tutorials and advices given during the game.

TABLE I
EXAMPLE OF TREC TOPIC

Element	Detail
Title	Falkland petroleum exploration
Description	What information is available on petroleum exploration in the South Atlantic near the Falkland Islands?

V. DISPLAYS

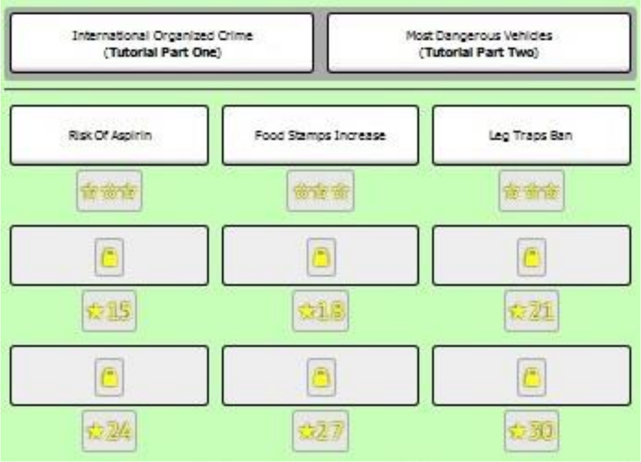
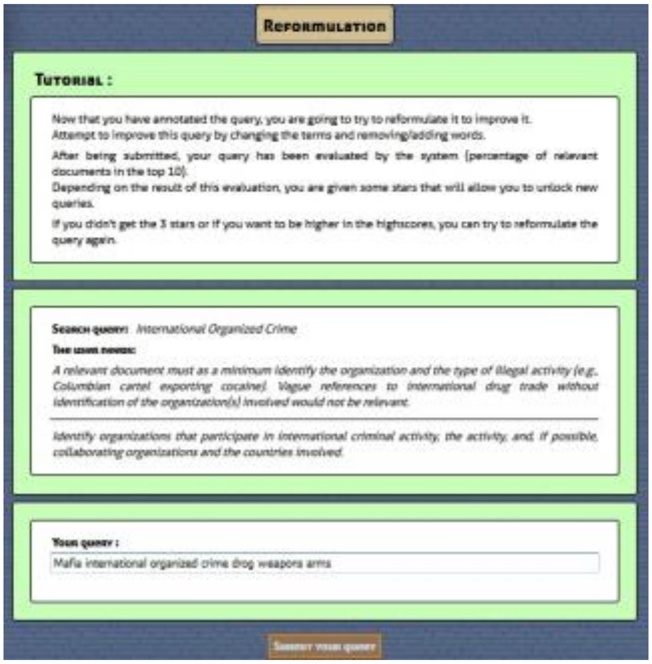


Fig. 1. Main screen of the game: the player has access to various queries for which he can suggest reformulations. Some are not accessible until the player gets a certain score.

Fig. 1 presents a typical display a player will see when connected. He can access the tutorial that explains what the goal of the game is and provides two examples with some suggestions of effective query reformulations. The tutorial part can be accessed without being logged in. As soon as the player makes some successful games, he gets a sufficient score to open lockers that give access to new queries and information needs to play with.



Tutorial :

Now that you have annotated the query, you are going to try to reformulate it to improve it.
Attempt to improve this query by changing the terms and removing/adding words.

After being submitted, your query has been evaluated by the system (percentage of relevant documents in the top 10).
Depending on the result of this evaluation, you are given some stars that will allow you to unlock new queries.

If you didn't get the 3 stars or if you want to be higher in the highscores, you can try to reformulate the query again.

Search query: *International Organized Crime*

The user needs:

A relevant document must as a minimum identify the organization and the type of illegal activity (e.g. Columbian cartel exporting cocaine). Vague references to international drug trade without identification of the organization(s) involved would not be relevant.

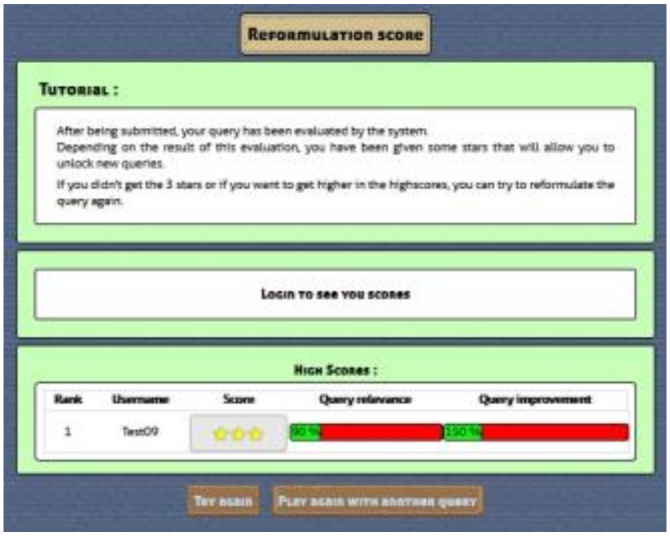
Identify organizations that participate in international criminal activity, the activity, and, if possible, collaborating organizations and the countries involved.

Your query :

Mafia international organized crime drug weapons arms

Fig. 2. Tutorial part of the game regarding query reformulation. Rule, information need, initial query, and a query reformulation suggested by the player.

Fig. 2 shows the tutorial part of the game. Both the query and the information need are displayed (the entire TREC topic



Tutorial :

After being submitted, your query has been evaluated by the system.
Depending on the result of this evaluation, you have been given some stars that will allow you to unlock new queries.
If you didn't get the 3 stars or if you want to get higher in the highscores, you can try to reformulate the query again.

Login to see you scores

High Scores :

Rank	Username	Score	Query relevance	Query improvement
1	Test09	★★★	90 %	190 %

Fig. 3. Tutorial part - scores the user's query obtained.

composed of the title, description, and narrative). The user can then make a try for a new query associated with the information need (see bottom part of the screenshot Fig. 2).

When submitting the query, it is sent to the search engine which retrieved the documents according to its ranking function. The ranked retrieved document list is in turn sent to `trec_eval`[3].

A score is calculated based on the number of relevant documents in the top 10 retrieved. The user has access to the score he obtained with his query (see Fig. 3).

The user can reformulate again the same query or go back to the main screen in which new information needs may be available (depending on the score he obtained).

VI. Results

In this section we present preliminary results on query reformulations as suggested by users.

The game is now online and we have been able to collect the first reformulations from players. We had about twenty different players. The players are mainly students who were advertised by the creators of the game. We use TREC 7 and 8 ad hoc collections. More precisely, in the game, we provide some of the queries in a given order. Twenty-six queries have been played with by the users; seven of them have been reformulated only once. Fig. 4 presents the number of times a query has been played (a player can play the same query several times); thus it indicates the number of reformulations for each query.

For each query formulation, we calculate the effectiveness obtained using Indri/Lemur on the TREC 7 and 8 document collections (ad hoc track with 528,155 documents from newspapers) and employing `trec_eval`. We focused on the precision when the 10 top retrieved documents are considered, also known as P10.

We mention that for several reformulations Indri /Lemur did not retrieve any documents and these reformulations were ignored. An interesting point consists in recurrent reformulations. For instance, the reformulation "international crime mafia" occurred for topic "301 - international organized crime" 10 times from a total of 33 and it also happens to be the best reformulation in terms of P10. We compared the P 10 results of MyBestQuery reformulations with the results obtained by employing automatic retrieval models, such as Query Likelihood (QL), Relevance Model 1 (RM1) [20] and Relevance Model 3 (RM3) [21].

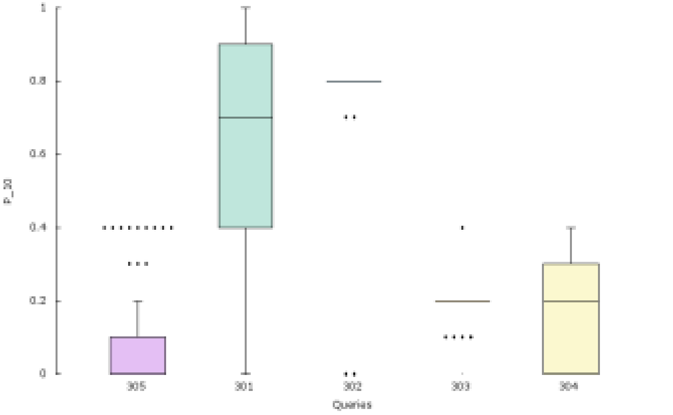


Fig. 5. Boxplots for the 5 queries with most reformulations with respect to P 10 (sorted descending by the number of reformulations made by players).

In Fig. 5 we show the boxplots for the queries with most reformulations, with respect to P10 values. For the topic 305, one can notice that the performance is low and, most of reformulations are below the average point of 0.21. The maximal values are out-liars, gathered at the level of 0.4.

In the case of topics 302 and 303, the values are condensed around the median, with some out-liars mostly towards low performance. The results for topic 301 and 304 have a wider distribution, with topic 301 oriented towards good performance.

TABLE 2
Performance Analysis
for reformulations, in terms of P 10. The three first columns are based on Indri language modelling-based search engine. The first column reports the results when using the initial query (TREC topic title part), the second uses RM1 automatic relevance feedback model, with 100 documents and 100 terms and 10 documents and 10 terms, the third uses RM3 model. Finally the last column uses the human reformulation.

Topic	QL	RM1		RM3				XYZ reformulation			
		100/100	10/10	100/100/0.2	100/100/0.5	100/100/0.8	10/10/0.5	#	min	max	average
301	0.40	0.40	0.30	0.70	0.50	0.40	0.30	33	0.00	1.00	0.606
302	0.80	0.80	0.80	0.80	0.80	0.80	0.80	18	0.00	0.80	0.700
303	0.10	0.10	0.10	0.10	0.10	0.10	0.10	17	0.10	0.40	0.188
304	0.40	0.30	0.30	0.20	0.30	0.40	0.30	11	0.00	0.40	0.155
305	0.10	0.00	0.10	0.00	0.10	0.10	0.10	58	0.00	0.40	0.105
306	0.10	0.70	0.60	0.70	0.80	0.80	0.60	4	0.20	0.40	0.275
307	0.10	0.30	0.70	0.40	0.30	0.60	0.70	1	0.30	0.30	0.300
308	0.20	0.30	0.30	0.20	0.30	0.20	0.30	5	0.00	0.00	0.000
309	0.00	0.00	0.00	0.00	0.00	0.00	0.00	5	0.00	0.00	0.000
310	0.30	0.10	0.20	0.30	0.30	0.20	0.20	3	0.00	0.20	0.133
311	0.90	1.00	1.00	1.00	1.00	1.00	1.00	2	0.00	0.00	0.000
312	0.60	0.80	0.90	0.90	0.80	0.90	0.90	2	0.70	0.90	0.800
313	0.10	0.40	0.20	0.10	0.20	0.20	0.20	2	0.80	1.00	0.900
314	0.00	0.00	0.00	0.00	0.00	0.00	0.00	2	0.10	0.50	0.300
317	0.50	0.50	0.70	0.40	0.70	0.70	0.70	1	0.50	0.50	0.500
419	0.30	0.30	0.10	0.10	0.20	0.10	0.10	4	0.10	0.50	0.350
435	0.30	0.50	0.60	0.10	0.40	0.60	0.60	1	0.70	0.70	0.700
436	0.30	0.20	0.30	0.20	0.20	0.30	0.30	1	0.50	0.50	0.500
432	0.10	0.00	0.00	0.00	0.00	0.00	0.00	6	0.30	0.30	0.300

These results are displayed in Table 2. RM1 and RM3 have parameters that can be tuned, such as the number of

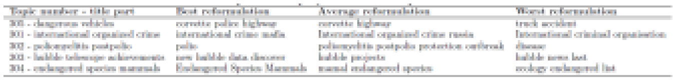
[3] http://trec.nist.gov/trec_eval/ calculates various effectiveness measures considering a specific run and query relevance data.

documents, or the number of terms, considered for reformulation. For RM3 there is an extra parameter, called *Lambda*, which anchors the initial query in the reformulation (interpolation of QL and RM1).

In Table 2 these parameter setups are denoted "#documents/#terms" for RM1 and "#documents/#terms/lambda" for RM3, respectively. In the case of MyBestQuery reformulations we show the number of reformulations together with minimum, maximum and average P10 performance, per considered topic. The best values for each topic are in bold. One can notice that topic 309 remains difficult for both approaches (automatic and human reformulations) with P10 of 0 throughout its corresponding table line. However, for topic 311 the automatic methods are clearly better.

For the topics with several query reformulations, we also give examples of good, average and bad MyBestQuery reformulations (see Table 3). The title part of the topic is also mentioned as initial query reference. One can notice that term misspelling affects performance ("mamal" instead of "mammal"), as well as giving the inappropriate synonym for a particular term ("disease" for "polio"). On the other hand, good human comprehension, followed by a thoughtful term choice could enhance performance.

TABLE 3
REFORMULATION SAMPLE FROM PLAYERS FOR THE 5 QUERIES WITH THE LARGER NUMBERS OF REFORMULATIONS.



Topic number - title part	Best reformulation	Average reformulation	Worst reformulation
300 - dangerous vehicles	corvette police highway	corvette highway	truck accident
301 - international organized crime	international crime mafia	International organized crime russia	International criminal organization
302 - poliomyelitis postpolio	polio	poliomyelitis postpolio protection outbreak	disease
303 - hubble telescope achievements	new hubble data discover	hubble projects	hubble news last
304 - endangered species mammals	Endangered Species Mammals	mamal endangered species	ecology endangered list

Query reformulation performance also has robustness issues, meaning that some queries may be harmed, even though on average the effectiveness is improved. To analyse this robustness, we computed the Robustness Index (RI) [22], for the best reformulations collected by MyBestQuery against the best automatic retrieval (QL, RM1 and RM3), taken on a per topic basis. The RI has the value of 0.365, with 13 topics out of 19 improved by MyBestQuery reformulations.

VII. CONCLUSION

In this paper we presented the first version of the MyBestQuery serious game; it aims at collecting query reformulations from users. The query reformulation so far is based on the information need description. We collected the first reformulations and have shown some features from these preliminary results. These results have to be supplemented by analysing the future data we will collect thank to MyBestQuery.

In future work, we would like to study how much considering the first retrieved documents can help a user to reformulate the query that provides more relevant documents at high ranks in the retrieved document list. We will also deeply analyse the terms the users used in the formulation of the query in order to try to understand their behaviour and to extract some information that could be useful for automatic query reformulation.

REFERENCES

[1] R. Trillo, J. Gracia, M. Espinoza, and E. Mena. Discovering the semantics of user keywords. Journal of Universal Computer Science, JUCS, 13(12):1908-1935, 2007.

[2] Hernandez, N., Mothe, J. An approach to evaluate existing ontologies for indexing a document corpus. In *Artificial Intelligence: Methodology, Systems, and Applications*, pp. 11-21. Springer Berlin Heidelberg. 2004.

[3] Chifu, A. G., Hristea, F., Mothe, J., Popescu, M. (2015). Word sense discrimination in information retrieval: A spectral clustering-based approach. *Information Processing & Management*, *51*(2), 16-31.

[4] K. Sparck Jones. A statistical interpretation of term speci_city and its application in retrieval. Journal of documentation, 28(1):11-21, 1972.

[5] P. Boldi, F. Bonchi, C. Castillo, and S. Vigna. Query reformulation mining: models, patterns, and applications. Information retrieval, 14(3):257-289, 2011.

[6] A. Hassan, X. Shi, N. Craswell, and B. Ramsey. Beyond clicks: Query reformulation as a predictor of search satisfaction. In Proceedings of the 22nd ACM international conference on Conference on Information & Knowledge Management, CIKM, pages 2019-2028. ACM, 2013.

[7] C. Carpineto and G. Romano. A survey of automatic query expansion in information retrieval. ACM Computing Surveys (CSUR), 44(1):1, 2012.

[8] Liana Ermakova, Josiane Mothe, Elena Nikitina. *Proximity Relevance Model for Query Expansion (regular paper). ACM Symposium on Applied Computing (SAC 2016), Pisa, Italy*, ACM, 2016

[9] Compaoré, J., Déjean, S., Gueye, A. M., Mothe, J., Randriamparany, J. Mining information retrieval results: Significant IR parameters. *Advances in Information Mining and Management. 2011*.

[10] Ermakova, L., Mothe, J., Ovchinnikova, I. (2014). Query expansion in information retrieval: What can we learn from a deep analysis of queries?. In *International Conference on Computational Linguistics-Dialogue 2014* (Vol. 20, No. 13, pp. pp-162).

[11] M.-C. Yuen, I. King, and K.-S. Leung. A survey of crowdsourcing systems. In Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom), 2011 IEEE Third International Conference on, pages 766-773. IEEE, 2011.

[12] C. Eickhoff, C. G. Harris, A. P. de Vries, and P. Srinivasan. Quality through flow and immersion: gamifying crowdsourced relevance assessments. In Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval, pages 871-880. ACM, 2012.

[13] S. Nowak and S. Rüger. How reliable are annotations via crowdsourcing: a study about inter-annotator agreement for multi-label image annotation. In Proceedings of the international conference on Multimedia information retrieval, pages 557-566. ACM, 2010.

[14] G. Kazai Lumi, F. Hopfgartner, U. Kruschwitz, and M. Meder. Ecir 2015 workshop on gamification for information retrieval (gamifir'15). In ACM SIGIR Forum, volume 49, pages 41-49. ACM, 2015.

[15] Alonso, O., & Mizzaro, S. Using crowdsourcing for TREC relevance assessment. Information Processing & Management, 48(6), 1053-1066. 2012

[16] Von Ahn, L. Games with a purpose. Computer, 39(6), 92-94. 2006

[17] Lafourcade, M. Making people play for Lexical Acquisition with the JeuxDeMots prototype. In SNLP'07: 7th international symposium on natural language processing (p. 7). 2007.

[18] K. Salen and E. Zimmerman. Rules of play: Game design fundamentals. MIT press, 2004.

[19] J. Schell. The Art of Game Design: A book of lenses. CRC Press, 2014.

[20] V. Lavrenko and W. B. Croft. Relevance based language models. In Proceedings of the 24th international ACM SIGIR conference on Research and development in information retrieval, pages 120-127. ACM Press, Sept. 2001.

[21] N. Abdul-Jaleel, J. Allan, W. B. Croft, F. Diaz, L. S. Larkey, X. Li, M. D. Smucker, and C. Wade. UMass at TREC 2004: Novelty and HARD. In E. M. Voorhees and L. P. Buckland, editors, TREC, volume Special Publication 500-261. National Institute of Standards and Technology (NIST), 2004.

[22] T. Sakai, T. Manabe, and M. Koyama. Flexible pseudo-relevance feedback via selective sampling. ACM Transactions on Asian Language Information Processing, 4(2):111-135, June 2005.

Adrian Chifu. is Associate Professor and he is 28 years old. In 2009 he obtained a Bachelor Degree in Mathematics and Computer Science from "Valahia" University, Târgoviște – Romania. He has a Master Degree in Databases and Web Technologies, degree obtained in 2011 and issued by Universitatea din București, Romania. He received the title of PhD in Computer Science from the Université de Toulouse in 2015. Then he was post-doctorate at Aix-Marseille Université. His research focuses in information retrieval, term discrimination and query difficulty prediction.

Serge Molina is a 21 years old Ingeneering Student at the Upssitech engineering school, from the Université de Toulouse. He completed a Bachelor degree in computer science at the Blagnac's Institut Universitaire de Technologie.

Josiane Mothe is Professor in computer science at the ESPE (teacher training school) Université de Toulouse since 2002. She is a specialist in information retrieval, data mining and Big Data. Since 2012, she has been leading the Information System team of the French IRIT-CNRS lab. She is the leader of the FabSpace 2.0 H2020 project and participated to 3 European projects (WS-Talk, eStage, and IRAIA 2000-02). She was the scientific responsible for UPS-IRIT in the federation FREMIT (collaboration between IRIT and IMT, one topic being Big Data) until 2013. From 2004 to 2014, she was the editor in chief for Europe and Africa of the international Information Retrieval Journal, (Springer). She was co-general chair of the CLEF 2015 conference. Eleven PhD students were supervised to successful completion by her.