



Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : <http://oatao.univ-toulouse.fr/>
Eprints ID : 16925

The contribution was presented at CORIA 2016 :
<https://www.irit.fr/sdnri2016/coria.php>

To cite this version : Chifu, Adrian and Molina, Serge and Mothe, Josiane
MyBestQuery : un jeu sérieux pour apprendre des utilisateurs. (2016) In:
Conference francophone en Recherche d'Information et Applications (CORIA
2016), 9 March 2016 - 11 March 2016 (Toulouse, France).

Any correspondence concerning this service should be sent to the repository
administrator: staff-oatao@listes-diff.inp-toulouse.fr

MyBestQuery : un jeu sérieux pour apprendre des utilisateurs

Adrian Chifu¹, Serge Molina¹, Josiane Mothe¹

1. Institut de Recherche en Informatique de Toulouse, Université de Toulouse
118 route de Narbonne, 31062 Toulouse cedex, France, prenom.nom@irit.fr

RESUME. MyBestQuery est un jeu sérieux qui collecte des éléments sur les requêtes soumises à un moteur de recherche: (i) la prédiction de la difficulté de la requête par le joueur (ii) des raisons possibles expliquant cette difficulté (iii) des propositions de reformulation.

ABSTRACT. MyBestQuery is a serious game designed to collect items from queries submitted to a search engine: (i) the query difficulty prediction (ii) the possible reasons for this difficulty (iii) other query formulations.

MOTS-CLES : Jeu sérieux ; Crowdsourcing, Etude utilisateur, Moteur de recherche d'information, Annotation des requêtes, Aide aux utilisateurs.

KEYWORDS: Serious game, Crowdsourcing, User study, Search engine, Query annotation

En recherche d'information, une requête est difficile si les moteurs échouent à répondre de façon satisfaisante. Plusieurs prédicteurs automatiques de difficulté ont été proposés dans la littérature ; ils s'appuient essentiellement sur la façon dont les moteurs fonctionnent (IDF moyen des termes de la requête, écart type entre les scores des documents (Shtok, 2012)). Ils ne sont que peu corrélés avec la difficulté observée (Mothe, 2005), (Shtok, 2012). Nous étudions la perception des utilisateurs sur des requêtes. Nous demandons ainsi à l'utilisateur de prédire la difficulté d'une requête, de répondre à un ensemble de questions sur les raisons de sa prédiction et de proposer une reformulation de requête qu'il pense meilleure. Les informations collectées jusqu'ici ont été analysées et des premiers résultats sont disponibles (Chifu, 2016). D'autres analyses pourront peut-être permettre d'adapter les apprentissages dans le domaine de l'éducation aux médias et à la recherche d'information efficace ou de mettre au point de nouveaux prédicteurs automatiques. Nous avons développé un jeu sérieux pour collecter l'expérience utilisateur ; nous utilisons des collections de références de TREC (adhoc et web track trec.nist.gov).

2. MyBestQuery : un jeu sérieux

MyBestQuery (mbq.irit.fr) possède toutes les caractéristiques d'un jeu sérieux (scénarisation, règles du jeu, récompenses, tableau des gagnants, ...). L'objectif

principal de l'application est de collecter des informations sur la perception des utilisateurs de la difficulté des requêtes soumises à un moteur de recherche d'information (typiquement sur le web) mais également sur des suggestions de reformulation. Les figures 1 & 2 présentent un extrait de l'écran permettant de juger la difficulté d'une requête et d'expliquer son choix. La figure 3 illustre un exemple de moyen mis en œuvre pour rendre captif le joueur : les niveaux se débloquent au fur et à mesure que le joueur gagne.



Figure 1 : Prédiction de la difficulté de la requête

	Reasons related to entire query					
	(-2) Strongly disagree	(-1) Disagree	(0)	(+1) Agree	(+2) Strongly agree	I don't know / Not applicable
The number of query words is too high	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The query is concrete/explicit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The query is broad/vague	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure 2 : Annotation de la requête – raisons de la difficulté (extraits)



Figure 3 : Les cadenas se débloquent au fur et à mesure de l'avancée dans le jeu

Références

- (Hauff, 2010) C. Hauff, D. Kelly, and L. Azzopardi. A comparison of user and system query performance predictions. *International CIKM*, pages 979-988, 2010.
- (Shtok, 2012) A. Shtok, O. Kurland, D. Carmel, F. Raiber, and G. Markovits. Predicting query performance by query-drift estimation. *ACM TOIS*, 30(2):11, 2012.
- (Mothe, 2005) J. Mothe and L. Tanguy. Linguistic features to predict query difficulty. *Workshop SIGIR*, pages 7–10, 2005.
- (Chifu, 2016) A. Chifu, S. Mizzaro, J. Mothe, Human-Based Query Difficulty Prediction: Is There Any Hope? (submitted).