

Crawler thématique

Parties prenantes

- Client : Julien ROUSSEL
- Membres de l'équipe projet (Gestion de projet, MOA, MOE, DEV) : Meriyama BANE, Salah DAHAMNI, Daniele PITROLO
- Validation de la gestion du projet: Georges GROSZ, Stéphane LAMASSÉ

Identification du projet

Nom du projet : Crawler thématique

Nature du projet : nouvelle application

Contexte de la demande

Date de la demande : 26 mars 2016 Motivation du besoin : évaluation, cours de Python, Master 2 MIMO

Objet principal: le projet est consiste en la réalisation d'une application web. Celle-ci demande à l'utilisateur de fournir une URL. Une fois lancée, l'application analyse la page associée à l'URL et elle en identifie les mots clefs. L'application itère l'opération de manière récursive sur les pages dont l'URL est contenue dans la page analysée et effectue ainsi plusieurs niveaux d'analyse. L'application présente enfin les résultats obtenus à l'utilisateur: les pages des niveaux ultérieurs d'analyse ne sont incluses parmi les résultats que s'ils présentent une continuité thématique suffisante avec la première page (le niveau sera déterminé en cours de développement par le biais de tests).

Objectifs

L'objectif principal est de proposer à l'utilisateur une page web permettant de déterminer les thèmes évoqués dans une page web donnée et la continuité thématique au travers des liens présentés.

- documentation client
- documentation utilisateur
- design
- mobile (facultatif)
- export des résultats (facultatif)

Limites

Les parties suivantes de l'application ne seront pas développées dans le cadre du projet, qui réutilisera des éléments déjà existants :

- serveur web
- bibliothèque de crawling
- base de données et interaction avec celle-ci

L'application livrée ne sera pas :

- disponible hors-ligne
- soumise à des tests de charge
- haute disponibilité non requise
- la compatibilité des pages web avec IE ne sera pas testée

Phasage

Le seul jalon prévu avec le client est la livraison du cahier des charges. À cause de la réduite connaissance en la matière de l'équipe de développement, le choix est fait d'engager rapidement le développement pour identifier au plus tôt les difficultés et disposer d'une marge confortable pour ajuster les estimations.