

Projet Tutoré – Master IC (parcours CCI) :
Identification automatique du plan des sites web

Responsables : Mathieu Roche, Matthieu Faure

Équipe TAL, LIRMM, Université Montpellier 2
mroche@lirmm.fr
EvalAccess (start-up en cours de création)
mfaure@open-s.com

Contexte du travail

Lors d'un audit d'accessibilité des pages Web, plusieurs questions doivent être posées :

- Dans le site Web, y a-t-il une page "plan du site" ?
- Dans chaque page Web, la page "plan du site" est-elle atteignable à partir d'une fonctionnalité identique ?
- Dans chaque page Web, la fonctionnalité vers la page "plan du site" est-elle située à la même place dans la présentation ?
- Dans chaque page Web, la fonctionnalité vers la page "plan du site" est-elle située à la même place dans le code source ?

Travail à réaliser

Pour répondre à ces questions, des tests doivent être mis en œuvre. Aujourd'hui ces tests sont effectués manuellement. Notre objectif est de maximiser l'automatisation et donc minimiser l'intervention humaine. Les objectifs à atteindre sont les suivants :

1. identifier la page plan du site. Pour cela, on peut déterminer ses caractéristiques, parmi lesquelles on trouvera :
 - le fait qu'elle comporte des liens vers l'ensemble des pages du site (du moins jusqu'à une profondeur donnée).

Projet Tutoré – Master IC (parcours CCI)

- souvent aussi, le fait qu'un lien vers le plan du site est présent en pied de page sur toutes les pages du site. Une fois ces caractéristiques déterminées, on devrait pouvoir identifier l'URL de la page "plan du site".
- 2. déterminer si la fonctionnalité vers la page "plan du site" (typiquement un lien html vers son URL) est toujours "globalement au même endroit" dans le code source. On pourra aussi chercher quelle méthode suivre pour vérifier que la fonctionnalité est toujours au "même endroit visuellement" (des outils comme Perceptual Diff pourront être étudiés : `<http://pdiff.sourceforge.net/>`)
- 3. appliquer la même démarche pour la page aide, et le formulaire de recherche.

Le corpus de travail pourra être constitué des sites labelisés :

(`<http://www.accessiweb.org/fr/Label_Accessibilite/galerie_sites_web_accessible/>`), ainsi que de pages / sites de contre-exemples créés pour l'occasion.

Collaboration LIRMM/EvalAccess

Ce sujet est proposé dans le cadre d'une collaboration entre l'équipe TAL du LIRMM et une start-up innovante en cours de création (incubée à Cap Omega).

Le but de cette start-up (EvalAccess) est l'accessibilité du web, discipline consistant à rendre accessibles les sites web aux personnes handicapées (mots clés : W3C, WCAG, AccessiWeb).

Cette société doit fournir des outils aux webmasters pour les aider dans cette tâche. Le sujet proposé s'insère dans l'ensemble des outils qui seront mis à disposition des webmasters. Les stagiaires seront en prise directe avec des problématiques techniques concrètes, opérationnelles et réelles.

Un stage de fin d'étude pourra faire suite à ce travail de TER.