

Nous nous plaçons ici dans un contexte de personnalisation d'une page Web. Cette personnalisation pourra se faire sur les différentes propriétés physiques (comme la taille, ou la couleur) ou le type des éléments de la page (comme menu ou bloc de texte). Dans ce sens nous voulons donc acquérir des informations d'une part sur la page en question que l'utilisateur chercherait à modifier et d'autre part sur les personnalisations que l'utilisateur souhaite réaliser sur les pages.

Le cœur du sujet réside dans l'étude et la réalisation des parties autour du module de calcul et de traitement principal. Ces parties permettent de faire le lien entre quatre éléments : l'utilisateur, ses souhaits, les pages Web dans un navigateur et le module de traitement principal.

Ce projet comprend donc trois axes majeurs plus ou moins dépendants :

- La conception et le développement des bases logicielles (extension pour navigateur et application locale en C/C++). Cette partie est plutôt technologique mais nécessaire pour mettre en place l'infrastructure.
- la conception d'un module de détection d'objets types dans les pages peu ou mal balisées. La question de recherche relève en partie de la recherche de motifs dans un arbre DOM afin de requalifier des noeuds de l'arbre et de leur associer des types qui nous intéressent pour une analyse ultérieure et pour la personnalisation. Cette étape pourrait aider à faire de la migration semi-automatique vers HTML5 même si ce n'est pas le but principal. **Elle se concrétisera dans la définition d'un méta-modèle de pages HTML restreint aux aspects concernant le projet et dans la construction d'un modèle de page HTML, dans lequel les motifs d'intérêt apparaîtront comme des fragments de modèles.**
- L'acquisition, la représentation et le stockage des souhaits de personnalisation de l'utilisateur. Dans cette partie on cherchera à définir un protocole et une interface d'acquisition des souhaits. La question de recherche consistera à étudier diverses approches dont une approche d'acquisition et d'apprentissage à partir d'exemples. Dans cette partie, on utilisera les modèles de pages HTML et les transformations de ces pages effectuées par l'utilisateur pour lequel on cherche à apprendre les souhaits. **On mettra en correspondance les fragments de modèles source (dans la page initiale) et cible (dans la page transformée) puis on appliquera une méthode d'apprentissage sur ces correspondances pour en tirer des règles de transformation d'ordre général exprimant les souhaits de l'utilisateur.**

Technologies utilisées : HTML/CSS, Javascript, (PHP), C/C++.