

# Evaluation de la Qualité des Applications Web : Etat de l'Art

Ghazwa Malak\* – Nadir Belkhit<sup>1</sup> – Mourad Badri\*\* – Linda Badri\*\*

\* Département d'informatique, Pavillon Adrien-Pouliot, Université Laval, Sainte-Foy (Québec) Canada G1K 7P4, {Ghazwa.Malak, Nadir.Belkhit}@ift.ulaval.ca

\*\* Département de Mathématiques et Informatique, Université du Québec à Trois-Rivières, C.P. 500 Trois-Rivières (Québec) Canada G9A 5H7, {Mourad\_Badri, Linda\_Badri}@uqtr.quebec.ca

## RÉSUMÉ

*Le développement d'applications s'appuyant sur le Web a connu ces dernières années une très forte croissance. Le besoin de mettre au point des principes et méthodes spécifiques au développement de ce type d'applications d'une manière générale, et à l'assurance de leur qualité, notamment au plan de l'ergonomie logicielle, se fait de plus en plus ressentir dans l'industrie du logiciel. En effet, l'évaluation quantitative de la qualité des applications Web (critères et métriques), de leurs interfaces en particulier, est très récente. Dans cet article, nous nous proposons d'établir un état de l'art dans le domaine de l'assurance qualité en matière d'applications Web en évaluant les approches, les critères et les métriques proposés dans la littérature. Les objectifs principaux de ce travail préalable et exploratoire et qui pourront faire l'objet ultérieurement d'un autre article, visent à plus long terme à (1) élaborer un arbre de qualité spécifique aux applications Web, (2) d'identifier les différents aspects quantifiables et (3) de proposer des métriques permettant de les évaluer.*

## ABSTRACT

*Web based applications development has been characterized by a very high growth during last years. The need for elaborating principles and methods which are specific to this kind of applications and to quality insurance, particularly from usability point of view, is more and more pointed out by the software professionals. Indeed, quantitative evaluation of web based applications quality (criteria and metrics) and usability, is quite recent. In this paper, we propose a survey in the field of quality insurance within the framework of web based applications by evaluating proposed approaches, criteria and metrics. The long term goals of this prior and bibliographical survey are (1) to elaborate a quality tree which would be specific to web based applications, (2) to identify the various quantifiable aspects and (3) to propose metrics to evaluate these aspects.*

## MOTS-CLÉS

*Applications Web, évaluation, interface utilisateur, assurance qualité, critères, métriques, arbre de qualité, état de l'art, système d'information Web.*

## KEYWORDS

*Web based applications, evaluation, user interface usability, insurance quality, criteria, metrics, quality tree, survey, Web-based information system.*

---

<sup>1</sup> Ce travail a été effectué durant l'année sabbatique 2001-2002 du professeur Nadir Belkhit à Lyon.

## 1. Introduction

Le développement d'applications liées au Web a connu ces dernières années une très forte croissance. Les applications Web ont été finalement intégrées dans des domaines de plus en plus diversifiés (commerce électronique, éducation, loisir, etc.). Comparées aux applications traditionnelles, elles présentent plusieurs caractéristiques propres tout en étant de taille et de complexité de plus en plus élevées. De plus, force est de constater que ces applications évoluent très rapidement, même si la plupart des sites Web sont pauvrement conçus et présentent des lacunes sur les plans aussi bien de la navigabilité, de la fonctionnalité que de la traçabilité [FLE 98]. Dans la plupart des cas, les sites sont confus, trop lents et ne répondent pas aux besoins des utilisateurs. Des études récentes montrent que 90% des sites ne satisfont pas les objectifs de qualité préalablement établis par les concepteurs [BOL 00a; WAR 99] et ceci pour différentes raisons, comme par exemple et entre autres les suivantes :

- technologie sous-jacente non conçue pour les applications temps réels ;
- tendance des développeurs à utiliser la dernière technologie disponible alors que la plupart des utilisateurs n'ont pas encore l'équipement minimal nécessaire pour la supporter ;
- développement de la plupart des sites Web par des amateurs qui focalisent trop souvent sur le seul aspect visuel des interfaces, ou par des professionnels mais qui manquent de flair artistique. Il en résulte des sites lourds, difficiles à utiliser, et ne répondant pas aux besoins des usagers ;
- manque de standards d'ingénierie associés à la production de sites Web de haute qualité.

La préoccupation vis-à-vis de l'évaluation des sites Web est très récente et demeure encore largement marginale. La recherche de métriques pour l'évaluation quantitative de la qualité des applications Web n'en est qu'à ses premiers balbutiements. Le besoin de mettre au point des principes et méthodes spécifiques au développement de ce type d'applications d'une manière générale, et à l'assurance de leur qualité, notamment au plan de l'ergonomie logicielle, se fait de plus en plus ressentir dans l'industrie du logiciel. L'émergence ces dernières années d'approches basées sur l'application des principes traditionnels et éprouvés du génie logiciel au développement d'applications Web (moyennant une certaine adaptation) va précisément dans ce sens. En effet, des guides de style ainsi que des principes et techniques de conception spécifiques aux applications Web, émergent de plus en plus pour assister les développeurs de sites Web dans leurs tâches, afin de concevoir un produit de qualité. Cependant, très peu de travaux ont été effectués et publiés sur l'évaluation quantitative de la qualité de ce type d'applications.

L'importance de la qualité de ces applications en général et de leurs interfaces utilisateurs en particulier n'est plus à démontrer. D'une manière générale les concepts définis pour le développement des interfaces de communication humain-

ordinateur pour des applications interactives traditionnelles demeurent tout à fait pertinents et valables pour des applications Web. Cependant, les caractéristiques générales de ces applications nécessitent la prise en compte de considérations additionnelles [PRE 01]. La qualité des interfaces dans ce type d'applications peut avoir un impact important, voire discriminant, sur le succès de l'application même au plan de la convivialité.

Par conséquent, dans le but de concevoir un bon système Web, focalisant aussi bien sur la structure et le contenu que sur la gestion des changements (maintenance), l'une des premières tâches à effectuer serait d'établir une méthode d'évaluation de la qualité des sites Web.

C'est dans ce contexte et dans un objectif d'assurance qualité que s'insèrent nos travaux de recherche. L'objectif de ces derniers est double, à savoir :

- procéder dans une première étape à un état de l'art en matière d'évaluation de la qualité des applications Web en nous intéressant tout particulièrement aux approches, aux critères et aux métriques proposés par les principaux chercheurs oeuvrant dans ce domaine ;
- développer des approches pour les autres facteurs de la qualité, en particulier la maintenabilité. L'objectif principal de cette seconde étape est d'élaborer un arbre de qualité spécifique aux applications Web (intégrant les résultats d'autres travaux), d'identifier les différents aspects quantifiables et de proposer des métriques permettant de les évaluer.

Compte tenu de l'espace imparti à cet article, il nous semblait ambitieux de tenter d'embrasser ces deux aspects dans un espace aussi limité. Par souci de clarté, nous avons donc choisi de ne présenter ici que la synthèse des résultats de nos investigations sur l'étude des principaux travaux de recherche publiés dans le domaine de l'évaluation de la qualité des applications Web. L'étude du facteur de maintenabilité et l'élaboration d'un arbre de qualité (résultant de la synthèse bibliographique présentée ici) seront donc volontairement exclues des limites de cet article.

Pour ce faire, et afin de permettre au lecteur de mieux comprendre ce que nous entendons par « qualité », nous présentons dans la **section 2** qu'est-ce que signifie pour nous la qualité (via ses principales caractéristiques) aussi bien dans le cadre d'un logiciel en général que d'une application Web en particulier. Dans la **section 3**, qui est le cœur de cet article, nous dressons un état de l'art des principaux travaux publiés dans le domaine de la qualité des applications Web. Nous concluons enfin dans la **section 4** par les principaux enseignements que nous avons tirés de ce survol bibliographique approfondi ainsi que par les perspectives envisagées de ce travail.

## 2. La qualité du logiciel, ses caractéristiques et sous-caractéristiques

Lorsque nous parlons de « qualité du logiciel », nous nous référons implicitement à la définition des caractéristiques et des sous-caractéristiques de la qualité telle que désignée par le standard international ISO<sup>2</sup> 9126 [ISO 91] (Tableau 1) dans l'attente de l'adoption définitive du standard similaire pour le Web de l'IEEE<sup>3</sup> (IEEE 2001).

<b>1. Fonctionnalité</b>	- aptitude - exactitude - interopérabilité - sécurité - conformité
<b>2. Fiabilité</b>	- maturité - tolérance aux fautes - possibilité de récupération
<b>3. « Utilisabilité »<sup>4</sup></b>	- compréhensibilité - facilité d'apprentissage - attractivité
<b>4. Rendement</b>	- performance (temps) - ressources
<b>5. Maintenabilité</b>	- facilité d'analyse - facilité de modification - stabilité - testabilité
<b>6. Portabilité</b>	- facilité d'adaptation - facilité d'installation - conformité - interchangeabilité

**Tableau 1.** Caractéristiques et sous-caractéristiques de la qualité

Dans le cas des applications Web, le MMHQ<sup>5</sup> [MMH 99] a proposé une description quelque peu modifiée de ces caractéristiques initialement définies pour le logiciel traditionnel. Cette adaptation en a ainsi facilité l'application à Internet et aux sites Web. Le lecteur intéressé pourra trouver une définition détaillée de chacune de ces caractéristiques et sous-caractéristiques dans [MAL 01].

Ainsi, évaluer la qualité d'une application Web revient à évaluer ces caractéristiques et ces sous-caractéristiques qui peuvent elles-mêmes être subdivisées en critères plus élémentaires pour faciliter leur évaluation.

---

<sup>2</sup> International Standard Organization.

<sup>3</sup> Institute for Electrical and Electrotechnical Engineers.

<sup>4</sup> Traduction du terme anglais "usability".

<sup>5</sup> The Multimedia House of Quality Limited.

Quelques travaux d'évaluation de la qualité des applications Web sont menés par plusieurs équipes de chercheurs un peu partout dans le monde. Le but étant de trouver une méthode d'évaluation quantitative générale permettant de tester et d'améliorer les sites Web existants ou futurs. Dans la section qui suit, nous allons précisément proposer une synthèse des principaux de ces travaux.

### 3. Evaluation de la qualité des applications Web : Etat de l'art

#### 3.1 Introduction

Un site Web est un système logiciel fortement interactif. Il interagit au moins avec deux types d'utilisateurs différents [BRA 00] :

- **les utilisateurs finaux** qui sont surtout à la recherche d'information ou de produits à acheter. Ils ont un comportement qui est fortement affecté par leur culture, leur langage, leur niveau de connaissances antérieures dans le domaine et leur expérience sur le Web. Ils interagissent avec les sites Web via une couche de technologie qu'ils ignorent, qui n'est pas sous le contrôle du concepteur du site et qui comporte : les navigateurs, les protocoles, les *plug-ins*, les systèmes d'exploitations, les plate-formes, les dispositifs d'interaction (écrans, multimédia...) et les connexions réseaux ;
- **les développeurs et les professionnels de la maintenance** qui essaient d'offrir un site de qualité en maintenant le caractère opérationnel du système ou en l'améliorant. Ils font face à une technologie qui évolue très rapidement et à un système qui a un cycle de vie très rapide. La maintenance d'un site Web se fait avec un rythme plus important que les autres produits logiciels à cause de la pression du marché et de l'absence d'une barrière de distribution. En plus, la portée de la maintenance devient souvent tellement large qu'une refonte totale du site est parfois nécessaire.

Ainsi, on perçoit actuellement le besoin pressant de ressources à déployer pour assurer la qualité des sites Web. En plus, il faut maintenir ce niveau de qualité tout le long du cycle de vie de ces sites. À noter que ces applications subissent rapidement des changements indispensables en nombres et en variétés.

Les premiers travaux effectués dans ce domaine focalisaient sur des problèmes généraux relatifs à l'évaluation des propriétés de l'hypertexte et de l'hypermédia. Vu l'évolution très rapide d'Internet et l'accroissement important du nombre de sites Web, les travaux plus récents sont plus spécifiques et visent à fournir des tests et des listes de contrôle ("check-list") permettant d'évaluer, d'une façon plus objective, la qualité des sites ou des pages Web. Certains ont même abouti au développement d'outils automatisant, en partie, le processus d'évaluation de l'utilisabilité et de la maintenabilité en vue de contribuer à la résolution de ce problème dont l'intérêt auprès des chercheurs n'a cessé de croître.

Cependant, des efforts importants en terme de recherche demeurent nécessaires, d'une part pour valider et compléter ces travaux et, d'autre part, pour développer des métriques permettant de renseigner au mieux la qualité des applications. Par ailleurs, le développement d'approches basées sur les métriques permettant de piloter des changements dans la conception de ces applications en vue d'assurer leur qualité s'avère nécessaire.

### **3.2 Les premiers travaux**

Depuis que le Web est devenu accessible au public au début des années 90, les études visant à en évaluer la qualité se sont succédées.

A notre connaissance, **Brown** [BRO 90] a été le premier à s'intéresser à la maintenabilité, particulièrement celle des hypertextes qui sont conçus pour survivre à long terme. Il a souligné la nécessité de l'hypertexte pour attiser l'appétit des lecteurs. Il a proposé l'application des principes traditionnels utilisés pour le contrôle de la rédaction et des références afin d'améliorer la qualité de l'hypertexte.

Par la suite, **Botafogo** et ses collaborateurs [BOT 92] ont utilisé la théorie des graphes et autres méthodes du génie logiciel pour identifier les concepts reliés à la cohérence et au couplage de l'hypertexte. En plus, ils ont esquissé quelques métriques pour déterminer la qualité d'une application hypertexte, principalement en ce qui concerne la lisibilité et la maintenabilité, mais aussi l'exactitude, l'intégrité, l'utilisabilité et la testabilité.

**Mehlenbacher** [MEH 93] a évoqué 8 méthodes pour évaluer les systèmes et les documents *Online*. Il en est arrivé à la conclusion qu'un système utilisable doit être un système accessible, maintenable, visuellement consistant, compréhensible, exact et adapté à la nature des tâches de l'utilisateur final.

De plus, **Garzotto** et ses collaborateurs [GAR 95] ont proposé d'autres critères de qualité pour une application hypertexte comme la richesse de l'information, la facilité d'accès, la consistance, l'auto-évidence, la prédictibilité, la lisibilité et la réutilisation.

Des travaux additionnels de **Hatzimanikatis** et de ses collaborateurs [HAT 95] ont produit une analyse très utile des métriques appropriées pour mesurer la lisibilité et la maintenabilité de l'hypertexte en adaptant le génie logiciel au génie de l'hypertexte en général. C'est ainsi que plusieurs métriques logicielles ont été ajustées à l'hypertexte et évaluées.

S'intéressant plus particulièrement au Web et à son étendu, **Bray** a étudié un nombre important de sites pour tenter de fournir « des réponses quantitatives partielles à des questions qualitatives difficiles » [BRA 96]. L'étude portait sur la taille des pages, la richesse des connexions du Web, la visibilité des sites, et les

formats de données. Il a introduit la mesure de la « sincérité » du HTML, en essayant de déterminer l'étendu de son utilisation dans la pratique.

Par ailleurs, des enquêtes quantitatives et des évaluations spécifiques sont peu à peu apparues par la suite. En particulier, **Lohse** et **Spiller** [LOH 98] ont identifié et mesuré 32 critères qui influencent les ventes par Internet (E-commerce).

Quelque temps plus tard, **Lowe** et ses collaborateurs [LOW 99] ont fait état d'un petit nombre de métriques qui existent pour la mesure de la qualité des applications hypermédia mais qui sont difficiles à interpréter. Il ont dégagé les problèmes d'interprétation des mesures de l'équipe de Botafogo et ont proposé d'autres métriques liées directement à l'utilisabilité : navigabilité, validité des liens et organisation des concepts. Focalisant sur le développement des applications hypermédia, ils se sont intéressés à la taille et à la durée de vie de ces applications.

A trois ans d'intervalle, **Shum** [SHU 96] et **Morville** [MOR 99] ont fait remarquer qu'un nombre considérable de travaux de recherche sur les interfaces homme-machine et sur l'utilisabilité focalisent sur le contenu d'une page individuelle, sans considérer la structure du contenu informationnel du site en entier. On remarque d'ailleurs que les cinq critères retenus pour l'évaluation du contenu d'une page (*Authority, Currency, Coverage, Objectivity, Accuracy*) se retrouvent dans les travaux de plusieurs chercheurs, à savoir [ALE 96], [JAC 96], [EVE 96], [KAP 98], et [LAN 00] pour ne citer que ceux-ci.

Dans le même ordre d'idée, **Kirakowski** et ses collaborateurs [KIR 98] ont travaillé plus particulièrement sur l'utilisabilité des sites Web. Ils ont développé un questionnaire de 60 items basé sur la satisfaction de l'utilisateur, en suivant un modèle de facteurs utilisé avec succès pour l'évaluation des logiciels conventionnels. Ce modèle est basé sur les 9 principes heuristiques de Nielsen et Molich [NIE 90]. Le questionnaire est nommé WAMMI (pour "Website Analysis and MeasureMent Inventory). Ils démontrent que ce type de questionnaire, administré à travers Internet, peut constituer une source d'information utile sur l'utilisabilité des sites Web, mais aussi, il peut fournir des éléments de diagnostic sur la manière d'améliorer un site Web en évaluation. Les critères d'évaluation qui ont été retenus sont les suivants : l'attractivité, la contrôlabilité, l'efficacité, l'aide et la facilité d'apprentissage.

S'intéressant à un aspect plus général, **Ricca** et **Tonella** [RIC 00] ont simplement classé les sites Web en trois grandes catégories : éducationnels, commerciaux et institutionnels afin de les évaluer selon le domaine d'application.

Outre ces travaux, rappelons qu'il existe également le standard ISO 9126 [ISO 91] mentionné précédemment dans ce texte et qui propose des directives de qualité pour le logiciel en général. Dans la même lignée, le nouveau standard de l'organisation IEEE pour les applications Web (IEEE 2001) qui s'adresse au formatage du contenu décrit les bonnes pratiques pour les sites Web à contenu

hautement informationnel. Ce standard propose des recommandations pour les intranets et pour les extranets en ce qui concerne la conception des sites (processus de développement, environnements, langages), ainsi que des serveurs et du contenu (DTD, XML, méta-données, propriétés intellectuelles,...). Pour le moment, il ne prévoit rien sur la sécurité.

Il nous apparaît pertinent de remarquer que dans les travaux cités ci-dessus, chacun des auteurs privilégie certaines caractéristiques de la qualité par rapport à d'autres. Chacun a sa propre vision des critères affectant la qualité d'une application Web, sans pour autant se conformer à un standard donné. Le tableau 2 ci-après présente un résumé des différents points de vue dans un cadre ISO 9126.

Par la suite, un aperçu plus détaillé de quelques recherches, les plus récentes, les plus développées et les plus utiles pour ce travail sera présenté. Certaines de ces recherches sont toujours en cours, dans le but de définir des méthodes et des métriques permettant l'évaluation quantitative de la qualité.

**LEGENDE :** **BW** : Brown; **BO** : Botafogo; **GA** : Garzotto; **HA** : Hatzimanikatis; **BY** : Bray; **LW** : Lowe; **MH** : Mehlenbacher

<b>Caractéristiques</b>	<b>Sous-caractéristiques</b>	<b>a BW</b>	<b>u BO</b>	<b>t GA</b>	<b>e HA</b>	<b>u BY</b>	<b>r LW</b>	<b>s MH</b>
<b>Fonctionnalité</b>	Facilité d'accès			•				•
	Téléchargement							
	Navigabilité						•	
	Liens						•	
	Info.			•				
<b>Fiabilité</b>	...	...	...	...	...	...	...	...
<b>Utilisabilité</b>	Contenu							•
	Contrôle de la rédaction	•						
	Richesse de l'info.			•				
	Lisibilité		•		•			
	Exactitude		•					•
	Intégrité		•					
	Consistance			•				•
	Structure du site							
Taille des pages						•	•	
<b>Rendement</b>	Performance							
<b>Maintenabilité</b>		•	•		•			•
	Testabilité		•					
<b>Portabilité</b>	...	...	...	...	...	...	...	...

**Tableau 2.** Critères de qualité selon les auteurs



### 3.3 Les travaux détaillés

Les méthodes d'évaluation de la qualité se divisent principalement en deux catégories : **qualitatives** (qui sont de nature informelle) et **quantitatives** (qui sont de nature plutôt formelle).

En général, les méthodes d'évaluation qualitatives sont basés sur une liste de caractéristiques à analyser pour préciser les avantages et les désavantages d'une application et les comptabiliser en pourcentage. Cette approche est évidemment attractive, mais seulement dans le cas d'un problème suffisamment simple. Elle n'est pas très concluante dans le cas de problèmes complexes.

Par contre l'évaluation quantitative permet de mieux structurer le processus d'évaluation tout en étant relativement simple et précise. Elle fournit des indicateurs quantitatifs globaux qui sont utilisés pour trouver et justifier une décision optimale [DUJ 96].

En fait, les 2 méthodes sont complémentaires et les travaux exposés ci-dessous rejoignent l'un et l'autre de ses modes d'évaluation.

#### 3.3.1 Travaux de Jakob Nielsen

Jakob Nielsen suggère de ne pas inventer de règles spécifiques à votre site mais plutôt de suivre la « loi de Jakob sur l'expérience des utilisateurs » : « les utilisateurs passent la plus grande partie de leur temps sur d'autres sites que le vôtre, et c'est là qu'ils construisent leur expérience et apprennent à se faire une idée du mode de fonctionnement du web... ».

Comme nous avons pu le constater dans ce travail rétrospectif de synthèse bibliographique que l'espace ne nous permet pas de développer ici (voir [MAL 01] pour le détail), les travaux de Nielsen [NIE 90] [NIE 00] portent seulement sur une caractéristique de la qualité qui est l'utilisabilité. Ils couvrent tous les aspects de l'utilisabilité mais sous forme de directives et de conseils. Nielsen a proposé très peu de métriques, sans proposer de méthode quantitative permettant de mesurer objectivement l'utilisabilité des sites Web. Il s'agit essentiellement d'un travail empirique basé sur l'expérience du chercheur. Il reste que les autres caractéristiques de la qualité sont tout aussi importantes pour l'assurance de la qualité des applications Web, et il nous apparaîtrait particulièrement pertinent de développer une méthode quantitative d'évaluation de cette qualité.

#### 3.3.2 Travaux de Cornelia Boldyreff

Boldyreff et ses collaborateurs étudient l'évolution des sites Web. Ils proposent de développer un modèle de mesures unifiées combinant les différentes métriques existantes. En plus, ils pensent étendre ces mesures afin de couvrir les aspects de l'utilisabilité aussi bien que ceux de l'évolution. S'inspirant des recherches déjà effectuées dans ce domaine, l'équipe de l'université Durham (UK) a reconnu d'abord l'existence des deux principaux problèmes suivants :

1. l'absence de standards et de directives pour supporter le développement de documents Web maintenables;
2. le fait que le développement des documents Web ne repose sur aucune approche d'ingénierie.

Dans ses travaux [BOL 00b], l'équipe a identifié huit dimensions selon lesquelles les sites Web et les pages Web peuvent être classés, à savoir : *la taille, le domaine, l'objectif, la technologie, la fonctionnalité, l'âge, le taux de changement, et la stratégie d'évolution.*

L'étude fait la distinction entre un site Web, une page Web et un composant Web : « un site peut comprendre un certain nombre de pages, et une page peut être composée d'un ou de plusieurs composants...Nous croyons (les auteurs) que la présentation apparente à l'utilisateur est un facteur important du succès ou de l'échec d'un site Web. Ainsi, il est vital de comprendre l'apparence d'une page Web et non pas de la réduire simplement à ses composants ». Chacune des huit dimensions citées ci-dessus, a été analysée en tenant compte des facteurs de niveau inférieur qui lui sont associés et qui sont mesurables ou qui peuvent être déterminés par analyse ou par inspection. Le lecteur intéressé pourra trouver les résultats détaillés de cette analyse dans [MAL 01].

Dans les publications de cette équipe disponibles sur Internet, nous avons trouvé l'objectif et les approches qui sont adoptées pour réaliser le projet. En revanche, la méthodologie elle-même n'est pas décrite. Quant aux articles que nous avons pu retrouver, ils ne signalent aucune métrique ni existante, ni proposée par l'équipe elle-même. En outre, nous n'avons retrouvé aucune référence à un standard de l'ISO ou de l'IEEE.

### 3.3.3 Travaux de *Luis Olsina*

L'équipe de l'École d'ingénierie à UNLPam en Argentine travaille depuis quelques années sur l'évaluation et la comparaison de la qualité des sites Web en phase opérationnelle. Olsina soutient qu'il existe un besoin important de développer des méthodes d'évaluation flexibles, bien définies et basées sur les standards d'ingénierie, mais également des modèles et des outils pour supporter le processus d'évaluation des exigences de qualité si complexes dans le cas du Web. En plus, « Pour comprendre, évaluer et améliorer la qualité des systèmes basés sur le Web, on doit utiliser de plus en plus les méthodes, les modèles et les techniques du génie logiciel » [OLS 98] [OLS 99].

En effet, les travaux de recherche de cette équipe sont parmi les plus avancés dans ce domaine et sont basés sur les standards IEEE et ISO/IEC en ce qui concerne les métriques et les directives de qualité du logiciel. L'équipe a développé le *Website QEM (Quality Evaluation Method)* [OLS 98] [OLS 99], comme une approche quantitative, afin d'évaluer les caractéristiques de qualité tout au long des différentes phases du cycle de vie d'un système d'information Web (SIW).

### 3.3.4 Travaux de Chignell et Keevil

Les travaux de Chignell et Keevil sont plutôt informels. Ils s'intéressent seulement à la caractéristique « utilisabilité » des sites Web et en particulier à l'exactitude, la précision et l'objectivité du style quand il s'agit d'écrire pour le Web.

S'appuyant sur les travaux de Jakob Nielsen et sur les travaux de l'équipe de Ravden relatifs aux interfaces humain-ordinateur, les auteurs ont converti les recommandations et les directives éprouvées sur l'utilisabilité des sites Web, en une liste de vérification (« check-list ») de 203 questions qui sont plus orientées utilisateurs que développeurs. Cette liste est utilisée pour mesurer l'index d'utilisabilité d'un site et quantifier la qualité du document Web. Ainsi, ils ont procédé aux regroupement des critères selon :

- 1- la facilité à retrouver l'information
- 2- la compréhensibilité de l'information
- 3- la capacité à répondre aux besoins de l'utilisateur
- 4- la capacité à présenter une information exacte
- 5- la présentation de l'information

Toujours selon Chignell et Keevil, l'inconvénient du questionnaire est qu'il peut être biaisé par des évaluations subjectives. Dans le but de diminuer cette subjectivité, les auteurs ont choisi des questions auxquelles il faut répondre par « Oui » ou « Non » seulement, ce qui facilite aussi l'automatisation du test. Ils proposent, en plus, dans leur étude une mise à jour hebdomadaire des sites Web et l'utilisation des techniques interactives [KEE 98].

La liste de Chignell et Keevil est trop axée sur la langue écrite, sur le format, la présentation et sur le style, avec des questions très détaillées (Keevil est un écrivain technique). La précision des questions pourra contribuer à diminuer la subjectivité des réponses. Pour notre part, nous pensons que cette liste pourrait être intéressante à utiliser dans le cadre d'une méthodologie d'évaluation plus large incluant les autres caractéristiques de la qualité.

### 3.3.5 Travaux de Guay et Poulin

Guay et Poulin présentent un prototype expérimental d'évaluation quantitative de systèmes d'information Web (SIW). Ils s'inspirent essentiellement de nombreux travaux dans les domaines des systèmes d'information et des systèmes documentaires, ainsi que de leur propre expertise personnelle. Leurs travaux se limitent à la dimension « utilisabilité » des sites Web.

Les auteurs proposent un ensemble de critères et de métriques [GUA 01] en puisant, entre autres, dans les travaux de Chignell et Keevil, dans les normes ISO 9126 et IEEE 2001 (alors à l'état de document de travail).

Pour ces auteurs, le SIW est un ensemble de logiciels, de données, d'information et du *design* graphique, sans oublier qu'il s'agit d'une « expérience différente pour chaque visiteur ». Ils évoquent la possibilité d'évaluer la qualité d'un SIW selon 4 perspectives : *utilisation, maintenance, développement, technologie*.

#### **4. Première tentative de fusion de critères de qualité**

Les travaux effectués par les différentes équipes constituent un point de départ intéressant. Ils proposent des directives, des critères et des métriques à l'intérieur d'une méthodologie propre à chacune d'elles. A la lumière de cet examen de l'état de l'art, il serait maintenant intéressant de faire un regroupement des critères proposés dans les différentes études. L'objectif est le développement d'approches pour les autres facteurs de la qualité, spécifiquement, la maintenabilité et l'élaboration d'une base pour un *arbre de qualité spécifique aux applications Web*, en identifiant les différents aspects quantifiables et en proposant un prototype de test pour les évaluer.

Ainsi, après un essai de rassemblement, de regroupement et d'extension des critères soulevés par les auteurs, nous avons pu trouver plus de 290 critères et sous-critères différents, à partir desquels nous avons, dans un premier temps, développé une *check-list* [MAL 01]. À la plupart des questions de cette liste les réponses possibles sont « *Oui* », « *Non* » ou « *ne s'applique pas* ». Un premier embryon d'*arbre de qualité pour l'évaluation des applications Web* est proposé dans l'Annexe 1.

Dans une prochaine étape, nous nous proposons de raffiner cet arbre de qualité et de développer un prototype de test en Visual Basic supportant notre *check-list* et permettant de tester une ou plusieurs caractéristiques de la qualité ainsi que la qualité totale d'un site Web ainsi que de démontrer la faisabilité d'une telle approche.

Pour mesurer l'efficacité de ce prototype, nous envisageons enfin de l'expérimenter pour tester des sites reconnus pour leur bonne ou mauvaise qualité.

#### **5. Conclusion et perspectives**

A l'évidence, cet examen bibliographique approfondi nous a permis de constater que les travaux effectués par les différentes équipes sont très intéressants. Ceux de Nielsen couvrent plusieurs aspects de l'utilisabilité, mais ne proposent que 3 métriques. L'équipe de Boldyreff expose le problème sans détailler la méthode ou les métriques proposées ou utilisées. La méthodologie d'Olsina est une base intéressante qu'il faut élargir pour couvrir le plus de caractéristiques possibles et enrichir l'arbre de qualité. Plusieurs métriques sont proposées par Olsina, et ses travaux sont parmi les plus avancés dans ce domaine. La *check-list* de Chignell et Keevil donne l'exemple d'un questionnaire détaillé utile pour diminuer la subjectivité de l'évaluation. Le prototype de Guay et Poulin est un modèle pratique

de type de test qu'il est possible d'implanter et d'utiliser. Isaak évoque les bonnes pratiques à suivre pour la conception du Web et appelle les développeurs à se conformer au standard IEEE 2001. Offut (dont nous n'avons pas présenté les travaux ici faute d'espace) [OFF 02] retient les caractéristiques de qualité qu'il juge importantes actuellement. Il soulève le problème de l'évolution très rapide d'Internet et de la technologie associée au Web et présente un résumé de l'état actuel de la recherche dans ce domaine. L'outil d'évaluation de la qualité d'Ivory et Hearst (travaux également non présentés ici pour les mêmes raisons) [IVO 02] compile et calcule plus de 150 mesures pour un site Web. Les outils de tests étudiés par Brajnik peuvent être utilisés pour automatiser la collection et l'évaluation de quelques critères de qualité relatifs aux sites Web.

Nous remarquons qu'un nombre important de travaux de recherches se font actuellement sur la qualité des applications Web. Le besoin d'un standard d'ingénierie pour ce type d'application se fait de plus en plus sentir à travers toutes ces études. Les tendances convergent vers l'adoption des caractéristiques de la qualité proposées par ISO 9126 moyennant une certaine adaptation. Nous percevons aussi l'écart qui se creuse entre les logiciels traditionnels et ceux du Web et qui est dû au dynamisme de ces derniers. Les changements rapides dans les applications Web rendent plus critique et plus important le problème de la maintenabilité des sites Web. Les exigences du marché, la compétition, et les transactions monétaires via Internet font surgir l'épineux problème de la fiabilité et de la sécurité de ces applications.

Ainsi, après avoir procédé à un état de l'art dans le domaine d'évaluation des applications Web, et après avoir constaté l'ampleur du travail qui reste à réaliser, nous nous proposons d'apporter une modeste contribution à l'avancement des travaux dans ce domaine en nous inspirant de toutes ces études antérieures pour proposer un *arbre de qualité* dont un tout premier jet fait l'objet de l'Annexe 1. La section précédente fait mention des prochaines étapes permettant d'atteindre un tel objectif en guise de perspectives envisagées de ce travail.

## Remerciements

Nous aimerions remercier le Conseil de Recherches en Sciences Naturelles et Génie (CRSNG) du Canada pour le financement partiel de ce travail ainsi que les trois évaluateurs de cet article pour leurs commentaires et suggestions ayant permis d'en améliorer la version soumise initialement.

## 6. Bibliographie

[ALE 96] Alexander J., & Tate M.A. (1996). "Checklist for an informational Web Page".

<http://www2.widener.edu/Wolfgram-Memorial-Library/webevaluation/inform.htm>

[BOL 00a] Boldyreff C., Warren P., Gaskell C., & Marshall A. (2000). "WEB-SEM Project: Web Site Evaluation Metrics", Janvier 2000.

<http://www.dur.ac.uk/cornelia.boldyreff/WEB-SEM/websem.htm>

- [BOL 00b] Boldyreff C., Gaskell C., Marshall A., & Warren P. (2000). "Establishing a Measurement Programme for the World Wide Web", Août 2000.  
<http://www.dur.ac.uk/cornelia.boldyreff/WEB-SEM/ias.html>
- [BOT 92] Botafogo R.A., Rivlin E., & Schneiderman B. (1992). "Structural Analysis of Hypertexts : Identifying Hierarchies and Useful Metrics", *ACM Transactions on Information Systems*, 10(2) : 142-180.
- [BRA 00] Brajnik G. (2000). "Automatic web usability evaluation : what needs to be done ? ", Octobre.  
<http://www.tri.sbc.com/hfweb/brajnik/hfweb-brajnik.html>
- [BRA 96] Bray T. (1996). "Measuring the Web", *Proceedings of the Fifth International World Wide Web Conference*, Paris, France.  
[http://www5conf.inria.fr/fich\\_html/papers/P9/Overview.html](http://www5conf.inria.fr/fich_html/papers/P9/Overview.html)
- [BRO 90] Brown P.J. (1990). "Assessing the Quality of Hypertext Documents", *European Conference on Hypertext 1990*, INRIA, France, Cambridge University Press.
- [DUJ 96] Dujmovic J.J. (1996). "A Method for Evaluation and Selection of Complex Hardware and Software Systems". *The 22<sup>nd</sup> International Conference for the Resource Management and Performance Evaluation of Enterprise Computing Systems. CMG 96 Proceedings*, Vol. 1, pp. 368-378.
- [EVE 96] Everhart N. (1996). "Web Page Evaluation Worksheet".  
<http://www.duke.edu/~de1/evaluate.html>
- [FLE 98] Fleming J. (1998). "WEB navigation: designing the user experience", O'Reilly & Associates.
- [GAR 95] Garzotto F., Mainetti L., & Paulini P. (1995). "Hypermedia Design Analysis and Evaluation Issues", *Communications of the ACM*, 38(8).
- [GUA 00] Guay B., Poulin L. (2000). « Evaluation d'un Système d'Information Web (SIW) », manuscrit non publié.
- [HAT 95] Hatzimanikatis A.E., Tsalidis C.T., & Christodoulakis D. (1995). "Measuring the Readability and Maintainability of Hyper documents", *Software Maintenance: Research and Practice*, 7 : 77-90.
- [IEE 92] IEEE Std 1061 (1992). IEEE Standard for a Software Quality Metrics Methodology.
- [ISA 00] Isaac J. (2000). "Internet Best Practices", 2000. <http://www.JimIsaak.com>
- [ISO 91] ISO/IEC 9126 (1991). Information technology – Software Product Evaluation – Quality characteristics and guidelines for their use. International Standard. <http://www.cse.dcu.ie/essiscope/sm2/9126ref.html>
- [IVO 02] Ivory M., & Hearst M. (2002). "Towards Quality Checkers For Web Site Designs". <http://webtango.berkeley.edu/papers/ieeeeic02/ieeeeic.pdf>
- [KAP 98] Kapoun J. (1998). "Five criteria for evaluating Web pages".  
<http://servercc.oakton.edu/~wittman/find/eval.htm>
- [KEE 98] Keevil B. (1998). "Measuring the Usability of Your Web Site".  
<http://www3.sympatico.ca/bkeevil/sigdoc98>
- [KIR 98] Kirakowski J., & Cierlik B. (1998). "Measuring the usability of Website", *HFES Annual Conference*, Chicago.  
<http://www.ucc.ie/hfrg/questionnaires/wammi/research.html>

- [LAN 00] Landsberger J. (2000). "Evaluating website content".  
<http://www.iss.stthomas.edu/webtruth>
- [LOH 98] Lohse G., & Spiller P. (1998). "Electronic Shopping", *Communications of the ACM*, 41(7) : 81-86, Juillet.
- [LOW 99] Lowe D., & Hall W. (1999). "*Hypermedia & the Web: An Engineering Approach*", John Wiley & Sons.
- [MAL 01] Malak G., Belkhiter N., Badri M., & Badri L. (2001). Évaluation de la qualité des applications Web. Rapport de recherche, Département d'informatique, Université Laval, Québec, Canada, Octobre, 48 pages.
- [MMH 99] MMHQ (1999). "Views on Quality : System constraints – is it good enough ? ", Août.  
<http://www.mmhq.co.uk/mmhq/my-quality/measure-system-constraints.shtml>
- [MEH 93] Mehlenbacher B. (1993). "Software Usability : Choosing Appropriate Methods for Evaluating Online Systems and Documentation", *SIGDOC93 : The 11<sup>th</sup> Annual International Conference Proceedings, New York, NY : The Association for Computing Machinery (ACM), Special Interest Group on Documentation*, pp. 209-222.
- [MOR 99] Morville P. (1999). "Information, Architecture and Usability", *Webreview*, Mars.  
[http://www.webreview.com/1999/03\\_12/strategists/03\\_12\\_99\\_3.shtml](http://www.webreview.com/1999/03_12/strategists/03_12_99_3.shtml)
- [NIE 90] Nielsen J., & Molich R. (1990). "Heuristic Evaluation of User Interfaces", *ACM SIGCHI Bulletin*, pp. 249-256.
- [NIE 00] Nielsen J. (2000). « Conception des sites Web : L'art de la simplicité », Campus Press.
- [OFF 02] Offut J. (2002). "Quality Attributes of Web Software Applications". *IEEE SOFTWARE*, March/April, pp. 25-32.
- [OLS 98] Olsina L., Godoy D., Lafuente G., & Rossi G. (1998). "Quality Characteristics and Attributes for Academic Web Sites", *Web Engineering Workshop*, Toronto, Canada.
- [OLS 99] Olsina L. (1999). "Web-site Quality Evaluation Method : A Case Study on Museums". *Proceedings of ICSE 99 – 2<sup>nd</sup> Workshop on Software Engineering over the Internet*, 1999.
- [PRE 01] Pressman R.S. (2001). "Software Engineering: A Practitioner's Approach", Fifth Edition, Mc Graw Hill.
- [RIC 00] Ricca F., & Tonella, P. (2000). "Visualization of Web Site History", *Proceedings of 2<sup>nd</sup> International Workshop on Web Site Evolution WSE'2000*, Zurich, Switzerland, pp. 30-33.
- [SHU 96] Shum S. (1996). "The missing Link: Hypermedia Usability Research & the Web", *Interfaces, British HCI Group Magazine*.
- [WAR 99] Warren P., Boldyreff C., & Munro M. (1999). "The Evolution of Websites", *Proceedings of International Workshop on Program Comprehension, IWPC99*, IEEE Computer Press, pp. 178-185.

**Annexe 1 : Embryon d'arbre de qualité pour l'évaluation d'applications Web**

