



N° 9 | 2020

Les enjeux et les défis de la fonction d'évaluation en sciences de l'information et de la communication

Utilité, utilisabilité, désirabilité: proposition d'un cadre conceptuel d'évaluation des Systèmes de Recherche d'Information en Ligne (SRIL)

Kouassi Sylvestre Kouakou

Université Cheikh Anta Diop de Dakar

Ecole de Bibliothécaires, Archivistes et Documentalistes

Laboratoire de recherche en sciences de l'information et de la communication.

Djibril Diakhaté

Ecole de Bibliothécaires Archivistes et Documentalistes, Université Cheikh Anta Diop de Dakar

Édition électronique :

URL :

<https://revue-cossi.numerev.com/articles/revue-9/1792-utilite-utilisabilite-desirabilite-proposition-d-un-cadre-conceptuel-d-evaluation-des-systemes-de-recherche-d-information-en-ligne-sril>

DOI : 10.34745/numerev_1561

ISSN : 2495-5906

Date de publication : 29/11/2020

Cette publication est **sous licence CC-BY-NC-ND** (Creative Commons 2.0 - Attribution - Pas d'Utilisation Commerciale - Pas de Modification).

Pour **citer cette publication** : Kouakou, K. S., Diakhaté, D. (2020). Utilité, utilisabilité, désirabilité: proposition d'un cadre conceptuel d'évaluation des Systèmes de Recherche d'Information en Ligne (SRIL) . *Revue COSSI*, (9). https://doi.org/https://doi.org/10.34745/numerev_1561

Évaluer un Système de Recherche d'Information en Ligne (SRIL) revient traditionnellement à interroger son acceptabilité, voire son utilisation. L'acceptabilité d'un SRIL se définit traditionnellement par la perception plus ou moins positive à propos de son utilité et de son utilisabilité. A ces deux dimensions traditionnelles nous ajoutons la désirabilité, c'est-à-dire la nécessité du système et l'envie d'utilisation qu'il suscite. Nous postulons donc pour évaluer un SRIL, il faut s'intéresser à son utilité, à son utilisabilité et à sa désirabilité. L'utilité du système recouvre sa performance et l'avantage relatif qu'il procure. L'utilisabilité se rapporte quant à elle, à l'efficacité du système, à son apprenabilité, à sa mémorabilité et à sa capacité de gestion des erreurs. Enfin, la désirabilité renvoie à l'apparence esthétique du système et à son intégration dans le contexte de l'utilisateur. Nous avançons également que toute relation entre ces trois dimensions est à priori possible et que ces probables relations sont interprétables suivant la rigueur que requiert une évaluation d'un SRIL.

Mots-clefs :

Interactions humain-machine, Bibliothèque numérique, UX design, Expérience utilisateur, Utilisabilité, Évaluation bibliothèque numérique, Ergonomie Web, Système de recherche d'information

Abstract : Evaluating an Online Information Retrieval System (SRIL) traditionally comes down to questioning its acceptability, or even its use. The acceptability of an SRIL has traditionally been defined by the more or less positive perception about its usefulness and usability. To these two traditional dimensions we add desirability, that is to say the necessity of the system and the desire for use that it arouses. We therefore postulate to evaluate an SRIL, we must be interested in its utility, its usability and its desirability. The utility of the system encompasses its performance and the relative advantage it provides. Usability relates to the efficiency of the system, its learnability, its memorability and its ability to handle errors. Finally, desirability refers to the aesthetic appearance of the system and its integration into the context of the user. We also argue that any relationship between these three dimensions is a priori possible and that these probable relationships can be interpreted according to the rigor required by an assessment of an SRIL.

Keywords : human-machine interaction, digital library, UX design, User experience,

usability, evaluation, web ergonomics, information retrieval system.

INTRODUCTION

Évaluer un Système de Recherche d'Information en Ligne (SRIL) c'est-à-dire une *collection en ligne d'objets numérisés (...) créée, rassemblée, gérée (...) et rendue accessible d'une manière cohérente et permanente, aidée par des services nécessaires à la recherche et à l'exploitation de ses ressources* (IFLA, 2010) revient traditionnellement à interroger son acceptabilité, voire son utilisation. L'acceptabilité d'un SRIL se définit généralement par la perception (attitudes, opinions, etc.) plus ou moins positive à propos de son utilité (le but qu'il permet d'atteindre) et de son utilisabilité (facilité d'utilisation). Cependant, nous constatons avec Buchanan & Salako (2009) et Chen & Gross (2017) que la majorité des études d'évaluation des SRIL porte sur leur utilisabilité. Bien que l'évaluation de l'utilisabilité d'un SRIL pourrait améliorer sa prise en main, nous soutenons que sans tenir compte de son utilité, il pourrait s'avérer bien conçu, facile d'utilisation, mais fonctionnellement limité voire inutile. Aussi, on peut interroger la raison d'être d'un dispositif, qui bien qu'à la fois utile et utilisable, n'est pas désirable c'est-à-dire ni attendu ni suscite l'envie d'utilisation. Nous postulons que l'évaluation d'un SRIL revient donc à estimer son utilité, son utilisabilité et sa désirabilité. Que recouvrent exactement ces trois dimensions ? Comment les opérationnaliser ?

Certains auteurs affirment qu'un « bon » SRIL doit obligatoirement remplir ces trois dimensions. Ainsi, il devrait *à la fois servir à quelque chose, être aisé d'accès et donner envie de l'utiliser* (Schmidt & Etches, 2016 :p.18). Même si nous partageons cet avis, nous n'occultons pas le fait qu'un dispositif peut être utile, mais pas facile à prendre en main et par conséquent ne suscitera pas un engouement quant à son utilisation. En outre, il peut être désirable, mais pas facile à pratiquer et par ricochet pas utile à l'utilisateur. Il y a donc un ensemble de scénarios possibles dont l'interprétation permettrait de mieux comprendre l'évaluation.

Dans ce papier, nous évoquons de façon détaillée, mais pas exhaustive les composantes et les critères d'évaluation d'un SRIL en les répartissant suivant deux grandes approches méthodologiques que sont : l'empirisme et l'heuristique. L'évaluation heuristique est réalisée par un « expert » et vise à inspecter une interface suivant des principes et des critères ergonomiques de base reconnus. Ce type d'évaluation présente l'avantage de la détection rapide des problèmes ergonomiques à moindre coût (financier, matériel et humain). Toutefois, il lui est reproché des biais possibles liés à la subjectivité des évaluations : les problèmes identifiés pour le même système pouvant varier considérablement d'un évaluateur à un autre. L'évaluation empirique quant à elle, consiste à soumettre à un échantillon représentatif des usagers réels ou potentiels une série de tests dans le but d'identifier des problèmes d'utilisabilité, de vérifier la conformité du dispositif avec leurs tâches et besoins informationnels, mais plus globalement à observer et à interpréter leurs

comportements, attitudes, opinions vis-à-vis du système. Ces deux approches bien que distinctes restent complémentaires. Nous n'évoquerons pas les conditions de leur mise en œuvre (elles sont déjà bien documentées, cf. Nielsen, 1993 ; Boutin & Martial, 2001), mais montrerons que les trois dimensions de l'évaluation (utilité, utilisabilité et désirabilité) se prêtent à ces deux approches.

Pour les besoins de cette étude, nous avons adopté une méthodologie fondée sur l'examen de la littérature concernant l'analyse ergonomique des systèmes d'information, l'utilisabilité des bibliothèques numériques (critères et méthodes d'évaluation) et l'acceptation des technologies. Le recours à l'examen de la littérature avait deux objectifs : premièrement, faire l'état des connaissances sur le sujet en évaluant et en résumant les concepts clés à l'étude à savoir l'utilité, l'utilisabilité et la désirabilité ; et deuxièmement, identifier les composantes et les critères établis pour chacune des trois dimensions à l'étude afin de concevoir notre cadre d'évaluation et établir des liens interprétatifs.

À partir du corpus collecté (non pas en fonction de la portée temporelle des articles, mais bien de l'impact de leur contribution dans le domaine à travers les indices de citations) et tenant compte des questions de recherche, nous avons subdivisé notre papier en deux parties. Ainsi, dans la première partie, nous décrivons et analysons les composantes, les critères ainsi que les mesures d'évaluation d'un SRIL. Puis, dans la seconde partie, nous établissons, de façon rudimentaire, un cadre interprétatif des liens possibles entre les différentes dimensions de l'évaluation que sont l'utilité, l'utilisabilité et la désirabilité.

COMPOSANTES, CRITERES ET MESURES DE L'UTILITE, DE L'UTILISABILITE ET DE LA DESIRABILITE

Composantes, critères et mesures d'évaluation de l'utilité

Le concept d'utilité renvoie à la capacité offerte par le système à l'utilisateur pour atteindre ses objectifs (Senach, 1990). Davis (1989) dans son Modèle d'Acceptation des Technologies (TAM) le définit comme *le degré auquel une personne croit que l'utilisation d'un système pourrait accroître sa performance au travail*. Nielsen (1993), lui, appréhende l'utilité sous un double angle théorique et pratique. L'utilité théorique renverrait aux objectifs potentiellement atteignables avec le système. L'utilité pratique, elle, aux objectifs effectivement atteints. Il y aurait donc une utilité perçue comme Davis l'évoque et une autre que nous qualifierons de tangible. Nous n'établirons pas de distinction entre ces deux angles de l'utilité. Quelle qu'en soit la conception, nous considérons l'utilité d'un système comme se rapportant à l'adéquation entre son contenu, ses fonctionnalités et les besoins de l'utilisateur. Par conséquent, l'évaluation de l'utilité d'un SRIL consiste à vérifier s'il y a bien adéquation entre les besoins informationnels définis par l'utilisateur à travers sa requête et le contenu du système

perceptible par le biais des résultats obtenus. Aussi, elle recouvre l'apport du système dans la résolution du problème informationnel. À partir de là, nous avançons que l'utilité d'un SRIL repose sur deux composantes : *la performance du système* (Senach, 1990) et *l'avantage relatif* (Rogers, 1995).

La performance d'un SRIL se rapporte à son efficacité, c'est-à-dire à la capacité du système à ramener des réponses pertinentes à la demande de l'utilisateur. Elle se réfère à la comparaison entre les résultats obtenus et les résultats qui doivent être idéalement obtenus. La question qui se pose est de savoir comment appréhender ces résultats idéaux. Il est clair que la performance est tributaire d'une part de la structure informatique et de l'organisation des index du système et d'autre part des connaissances implicites de l'utilisateur (Lefèvre, 2000). Ainsi, le système peut ramener des documents en parfaite réponse à la requête formulée et conformément aux algorithmes donc logiquement pertinents, alors que l'utilisateur peut les trouver moins pertinents par rapport à ses acquis antérieurs. En revanche, des documents considérés par le système comme relativement loin de la requête (classement à l'affichage des résultats), peuvent être trouvés par l'utilisateur comme pertinents.

L'avantage relatif, lui se rapporte à la perception par l'utilisateur que le système est meilleur ou plus performant que les systèmes existants. Cette « performance » est dépendante des besoins particuliers à chaque groupe d'utilisateurs et concerne le gain cognitif et la réduction de l'incertitude. En effet, les utilisateurs dans leur quête d'information identifient un problème spécifique et déterminent le manque qu'ils souhaitent voir le système combler. Un SRIL utile et présentant un avantage est celui qui permet aux utilisateurs de combler leurs lacunes cognitives, de répondre à leurs besoins informationnels et/ou de réduire leurs incertitudes sur un sujet, un problème donné ; et cela dans un laps de temps relativement court, à un coût moindre avec des efforts techniques et cognitifs très réduits.

| Composantes | Indicateurs empiriques | Critères heuristiques |
|------------------------|---|--|
| Performance du système | <ul style="list-style-type: none"> • Taux de rappel. • Taux de précision. | <ul style="list-style-type: none"> • Adéquation réponses /requêtes. • Atteinte des objectifs de recherche. |
| Avantage relatif | <ul style="list-style-type: none"> • Différence entre niveau de connaissance initial et terminal. • Comparaison à l'apport cognitif d'autre SRIL. | <ul style="list-style-type: none"> • Adéquation contenus / besoins informationnels. • Processus cognitifs et méta-cognitifs de résolution de problème. |

Tableau 1. Indicateurs et critères pour l'évaluation de l'utilité d'un SRIL

Le tableau 1 ci-dessus indique que l'évaluation empirique de l'utilité repose d'une part sur la comparaison de l'apport cognitif à des groupes de sujets soumis à des tâches avant et après l'utilisation d'un SRIL ou plusieurs SRIL (Tricot et al., 2003) ; et d'autre part, elle passe par l'évaluation des taux de rappel[1] et de précision[2]. Un SRIL utile est idéalement celui qui présente des taux de précision et de rappel égaux à 1. En clair,

il s'agit d'un système qui retourne à la suite d'une requête tous les documents pertinents (rappel =1), et seulement tous les documents pertinents (précision =1). Rappelons que ces deux mesures ne doivent pas être calculées et interprétées sous un angle binaire, selon lequel un système serait totalement performant ou pas du tout. Ce serait très simpliste dans la mesure où les taux de rappel et de précision sont fortement liés au contenu et à la requête. C'est pourquoi l'évaluation de la performance est caractérisée par des courbes précision-rappel qui indiquent les variations conjointes de ces deux indices complémentaires (Lefèvre, 2000).

Pour l'approche heuristique, nous proposons quatre critères (tableau 1) tirés de la norme ISO 9241-110 (2006), notamment ses principes (1) adéquation à la tâche et (3) conformité aux attentes des utilisateurs. Ces critères se rapportent globalement à la correspondance entre la requête formulée et les réponses fournies par le système, d'une part; et à l'adéquation entre les besoins informationnels et les ressources disponibles dans le SRIL, d'autre part. Par ailleurs, nous voulons souligner que l'évaluation ici prend en compte la mise en œuvre des processus cognitifs et métacognitifs sollicités dans la résolution du problème initial.

Composantes, critères et mesures d'évaluation de l'utilisabilité

L'utilisabilité renvoie à la facilité d'utilisation, c'est-à-dire au degré de facilité associé à l'utilisation d'un système. En d'autres termes, un dispositif utilisable est celui qui offre *des interactions simples et efficaces qui tiennent compte des modes de raisonnement [des utilisateurs], de leur charge de travail et de leurs contraintes [tâche, environnement, poste] et astreintes [fatigue] ... [et qui] garantit un haut niveau de performance humaine et technique* (Brangier et al, 2015). Un système utilisable est donc celui qui est facile à prendre en main, qui guide l'utilisateur au cours de ses interactions et qui anticipe, tolère les erreurs éventuelles de l'utilisateur. Il est traditionnellement associé à l'utilisabilité cinq composantes: l'efficacité, l'apprenabilité, la mémorabilité, la gestion d'erreurs, la satisfaction (Nielsen, 1993, Brangier & Barcenilla, 2003). La satisfaction ne nous est pas apparue comme une composante pertinente de l'utilisabilité. En effet, l'utilisateur peut être insatisfait pour des raisons diverses même si le système lui permet d'accomplir facilement sa tâche. Elle est donc moins liée à la facilité utilisation, qu'à la perception de l'avantage offert par le système c'est-à-dire son utilité. Ainsi, nous avons retenu comme composantes de l'utilisabilité : *l'efficacité, l'apprenabilité, la gestion des erreurs et la mémorabilité*.

L'efficacité est la capacité du système de permettre à l'utilisateur d'accomplir une tâche spécifique en mobilisant un minimum d'efforts (Brangier & Barcenilla, op.cit.). L'efficacité est liée aux efforts à la fois techniques et cognitifs fournis par l'utilisateur dans son interaction avec le système : plus l'effort est faible, plus l'efficacité est élevée.

L'apprenabilité se réfère d'une part à la facilité avec laquelle l'utilisateur peut accomplir des tâches relativement basiques dès sa première interaction avec le système ; et d'autre part, elle recouvre la prise en main aisée du système (Seffah et al, 2006).

La gestion des erreurs désigne quant à elle, les fonctionnalités permettant d'éviter ou de réduire les erreurs, et de facilement les corriger lorsqu'elles surviennent (Bastien & Scapin, 1993). Ainsi, le système devrait aider à reconnaître, à diagnostiquer et à remédier aisément aux erreurs en suggérant une solution appropriée.

La mémorabilité qui est souvent prise en même temps que l'apprenabilité fait référence au degré de facilité pour un utilisateur peu fréquent du système d'accomplir à nouveau une tâche après une période d'inactivité plus ou moins longue. En clair, un utilisateur qui après un long temps d'inutilisation du système à l'impression d'effectuer un nouvel apprentissage, il est très probable que le système présente une mauvaise mémorabilité.

Pour ce qui concerne le comment évaluer ces différentes composantes de l'utilisabilité, le tableau 2 ci-dessous montre que pour l'évaluation empirique, les indicateurs sont relatifs au temps d'accomplissement des tâches, à l'affordance, ainsi qu'au taux, à la nature et à la gravité des erreurs.

| Composantes | Indicateurs empiriques | Critères heuristiques |
|---------------------|---|--|
| Efficienc | <ul style="list-style-type: none"> • le nombre de fonctions utilisées pour réaliser une tâche. • le nombre de modes opératoires exécutés pour accomplir une tâche donnée. • les ressources cognitives nécessaires pour atteindre les objectifs de recherche. • le temps mis par l'utilisateur pour réaliser la tâche. | <ul style="list-style-type: none"> • Fonctionnalités de raccourcis et de navigation. • Fonctionnalités de recherche: <ul style="list-style-type: none"> ◦ Sélection de bases de données. ◦ Opérateurs de recherche. ◦ Formulation de requêtes. ◦ Index, langage d'interrogation. • Affichage/Affinement et exploitation des résultats. |
| Apprenabilité | <ul style="list-style-type: none"> • la capacité du système à être découvert naturellement (affordance). • la nature des difficultés rencontrées par les utilisateurs. • la disponibilité et accessibilité de la documentation d'utilisation. • la facilité d'usage de l'aide informatisée. | <ul style="list-style-type: none"> • Guidance du nouvel utilisateur. • Feedback. • Contrôlabilité. • Aide à la recherche et à l'exploitation des résultats. |
| Mémorabilité | <ul style="list-style-type: none"> • le nombre de fois où l'usager a recours à une aide interne ou externe. • le nombre d'erreurs commises. • le temps d'exécution d'une tâche préalablement accomplie. | <ul style="list-style-type: none"> • Sauvegarde des requêtes et résultats de recherche. • Mémorisation. |
| Gestion des erreurs | <ul style="list-style-type: none"> • le nombre d'erreurs commises par l'utilisateur et le nombre de récupérations. • Évaluer la compréhension des messages de correction • les temps nécessaires à la correction des erreurs. | <ul style="list-style-type: none"> • Prévention des erreurs. • Exactitude et clarté des messages d'erreur. • Récupération d'erreur. |

Tableau 2. Indicateurs et critères pour l'évaluation de l'utilisabilité d'un SRIL

L'indicateur temps d'accomplissement d'une tâche ne peut être considéré qu'en fonction de la nature des informations recherchées et du contexte d'utilisation. En effet, pour un SRIL, les contraintes temporelles sont différentes lorsqu'il s'agit de rechercher des informations pratiques sur les horaires de la bibliothèque ou de trouver des références bibliographiques dans le catalogue. Aussi, le temps d'exécution peut être influencé par la vitesse de la connexion de l'utilisateur et la densité du trafic sur le système au moment de la réalisation de la tâche. Le temps est également utile pour évaluer la complexité de l'interaction avec le système en termes de nombre de commandes, d'opérations ou de procédures requises. Nous proposons en outre de mesurer le nombre de clics et de raccourcis nécessaires pour accéder à une information, un formulaire, ou à une ressource.

L'approche empirique concerne également l'affordance du SRIL c'est-à-dire sa capacité à proposer des fonctionnalités tellement évidentes que l'utilisateur au cours de son interaction (même pour sa première) identifie correctement ce qu'il doit faire (Brangier & Barcenilla, 2003). En outre, le système s'il propose des aides, celles-ci sont faciles à appréhender et le processus d'apprentissage est aisé, ne nécessitant pas « trop » d'efforts cognitifs et techniques.

Pour l'évaluation heuristique, les critères proposés sont largement empruntés à Chowdhury (2004), Silvis, Bothma et De Beer (2019). Ces critères reposent sur l'inspection des options de navigation et de raccourcis, des fonctionnalités de recherche et de celles d'aide fournies par le système. Les fonctionnalités de recherche recouvrent trois étapes fondamentales de la recherche d'information : la sélection des bases de données, la formulation des requêtes et la manipulation des résultats. Les différentes possibilités et options de recherche et d'affinement des résultats proposées ont un impact significatif sur l'utilisabilité du système (Chowdhury, op.cit.). En outre, au cours de sa recherche d'informations, le système devrait proposer différentes aides à l'utilisateur. Ces aides se traduisent par la mise à disposition d'outils du type guide ou tutoriel à la prévention et à la correction d'erreurs éventuelles.

Composantes, critères et mesures d'évaluation de la désirabilité

La désirabilité correspond aux affects personnels que l'utilisateur peut ressentir à l'égard d'un système. Pour établir un lien entre un système et un utilisateur, il faudrait que ce dernier soit en ait besoin soit veuille l'utiliser ou idéalement les deux (Schmidt et Etches, 2016). Un système, bien qu'utile et utilisable, peut être largement sous-utilisé s'il n'est pas plaisant à utiliser. En effet, les utilisateurs sont des êtres à la fois rationnels et émotifs. Ils recherchent des systèmes qui en plus des aspects fonctionnels et pratiques procurent du plaisir et des émotions à l'utilisation. À partir de là, un système désirable est donc un système pour lequel l'utilisateur a une certaine attirance, un plaisir à son utilisation et qui est conforme en termes fonctionnels à ses attentes. En partant des travaux de Norman (2012), Desmet, Hekkert et Hillen (2004) et Hassenzahl & Tractinsky (2006), nous avançons que la désirabilité est perçue à deux niveaux d'émotions. Premièrement, il s'agit de l'émotion antécédente à l'utilisation du système. Elle est éprouvée à partir d'un jugement à priori fondé sur l'apparence du système. Ce qui induit l'existence d'une relation « directe » entre la manière de concevoir le système, c'est-à-dire le design, l'esthétique et les émotions qu'il peut susciter. Deuxièmement, l'émotion est la résultante de l'interaction avec le système. Elle apparaît à postériori et se fonde sur l'expérience d'utilisation. Elle implique une relation entre les caractéristiques de l'utilisateur et celles du système. Deux composantes se dégagent donc pour mesurer la désirabilité d'un système : *l'apparence esthétique et l'adéquation contexte-utilisateur.*

L'apparence esthétique se rapporte aux aspects visuels du système présentés de manière attrayante afin de susciter des affects positifs (amusement, plaisir, joie, etc.) chez l'utilisateur. Elle est également perçue sous l'angle de l'agencement interface-contenu. Les interfaces de recherche et de présentation des résultats sont présentées

de manière à inciter l'utilisateur soit à compléter les formulaires de recherche, soit à cliquer sur les liens hypertextuels ou hypermédias à la présentation des résultats. Ainsi, un SRIL présentant une apparence esthétique acceptable est celui qui d'une part, incite à son utilisation de par son attrait visuel et d'autre part, établit chez le chercheur d'information, une attitude affective (au travers des interfaces conviviales et agréables), au moment où ce dernier mobilise ces capacités cognitives et techniques pour trouver l'information (Tricot et al., 2003).

L'adéquation contexte-utilisateur est issue du construit « compatibilité » de Rogers (1995) et se traduit par le degré auquel le système est considéré comme étant conforme aux valeurs existantes, aux expériences antérieures ainsi qu'aux besoins des utilisateurs potentiels. En clair, le SRIL doit parfaitement s'imbriquer dans les pratiques courantes de l'utilisateur. L'adéquation passe également par la perception que le système et son contenu sont importants, attendus ou souhaités par l'utilisateur dans ses activités.

| Composantes | Indicateurs empiriques | Critères heuristiques |
|---------------------------------|--|---|
| Apparence esthétique | <ul style="list-style-type: none"> • les remarques positives et/ou commentaires négatifs sur l'aspect esthétique du système. • le degré d'appréciation de l'agréabilité du système. | <ul style="list-style-type: none"> • Fonctionnalités générales de l'interface <ul style="list-style-type: none"> ◦ Type d'interface et langue ◦ Options de navigation ◦ Couleurs, typographie ◦ Mise en page et les graphiques. |
| Adéquation contexte-utilisateur | <ul style="list-style-type: none"> • le nombre de fois qu'un utilisateur exprime de la frustration et/ou de la satisfaction. • le degré de satisfaction du besoin informationnel par le système. | <ul style="list-style-type: none"> • Adaptabilité du système • Compatibilité fonctionnalités système /caractéristiques socio-techniques des utilisateurs. |

Tableau 3. Indicateurs et critères pour l'évaluation de la désirabilité d'un SRIL

L'évaluation empirique, ici, se fonde en général sur l'affect, le sentiment des utilisateurs par rapport au design du système, et l'enjouement ressenti à son utilisation. Il s'appuie également sur le degré de satisfaction de l'utilisateur par rapport à son besoin informationnel initial. Ces indicateurs sont mesurables à partir d'une enquête de satisfaction par questionnaire, par entrevue ou tout simplement à travers l'observation des attitudes et gestes de l'utilisateur au cours de son interaction avec le système.

L'évaluation heuristique quant à elle concerne la capacité du dispositif à prendre en compte les caractéristiques socio-techniques des utilisateurs (perceptions, habitudes, attentes, etc.) et à s'intégrer dans leur activité réelle. Pour ce faire, les critères fonctionnalités de l'interface de Chowdhury (2004), l'aspect esthétique de Buchanan & Salako (2009) et la compatibilité de Bastien & Scapi (op.cit.) semblent pertinents.

INTERPRETATION DES RELATIONS ENTRE LES TROIS DIMENSIONS DE L'ÉVALUATION D'UN SRIL

Dans cette partie, nous étudions les relations entre utilité, utilisabilité et désirabilité pour en proposer un modèle conceptuel. Nous partons du fait que l'évaluateur est libre de ne choisir que quelques critères ou indicateurs de chacune des composantes avec pour but de déterminer le niveau de satisfaction de chaque dimension. Les modèles antérieurs ont considéré que l'acceptabilité d'un système était essentiellement fonction de l'utilité et de l'utilisabilité. Ils n'ont donc examiné que quelques relations standards entre l'utilité, l'utilisabilité et l'acceptabilité. Contrairement à ces modèles, celui (figure 1) que nous proposons intègre la désirabilité (l'un des piliers du design thinking selon Schmidt & Etches, 2016) et considère l'acceptabilité comme une étape de l'évaluation de l'utilisation qui s'inscrit dans le continuum de l'adoption de la technologie (Kouakou, 2015). L'originalité de notre modèle repose également sur le fait qu'il associe à chacune des dimensions, différentes composantes, pour lesquelles il propose différents critères d'évaluation.

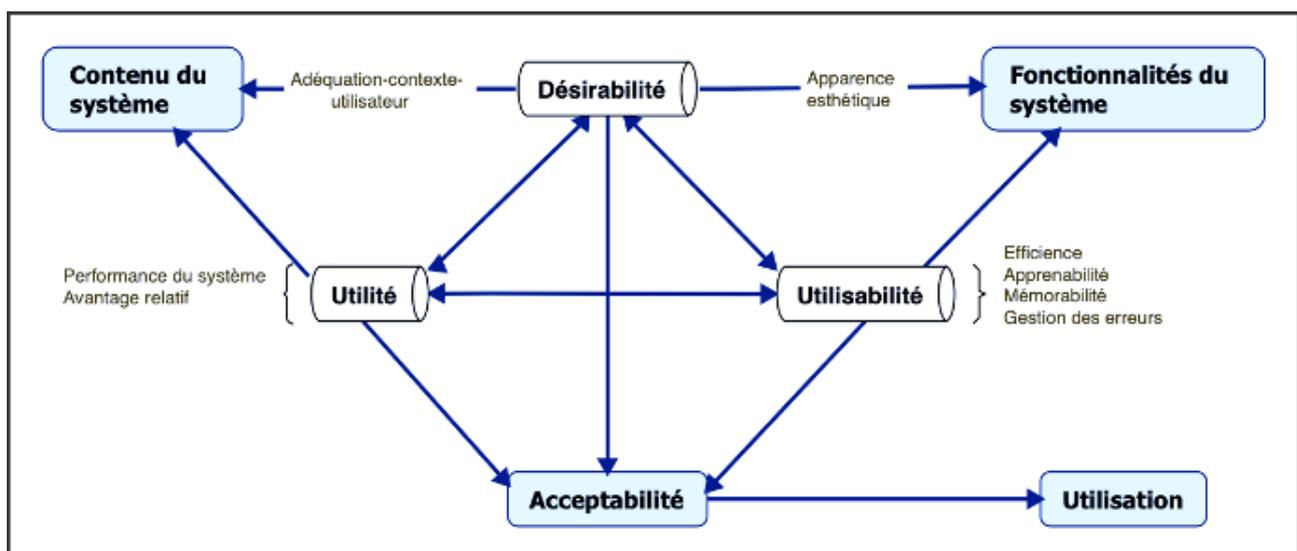


Figure 1. Modèle conceptuel des relations formellement possibles entre l'utilité, l'utilisabilité et la désirabilité

Relations entre les composantes des trois dimensions

À l'instar de plusieurs études (Xu & Du, 2018 ; Tsakonas & Papatheodorou, 2008 ; Brangier & Barcenilla, 2003) nous postulons que les différentes composantes de notre modèle sont liées les unes des autres. Pour l'utilisabilité, des recherches (Xu, op.cit. ; Buchanan & Salakko, 2009) ont montré qu'il existe une corrélation entre l'efficience, l'apprenabilité et la gestion des erreurs. Généralement, ces relations peuvent se

formuler comme suit :

- Un système facile à apprendre et qui apporte une aide à la prévention et à la correction des erreurs est nécessairement efficient.
- Un système efficient présente nécessairement une bonne gestion des erreurs.
- Un système offrant une aide à la correction des erreurs n'est pas nécessairement mémorisable.
- Un système mémorisable n'est pas forcément efficient.
- Un système efficient et facile à apprendre n'est pas nécessairement mémorisable.

En ce qui concerne l'utilité, il est à noter que seule la performance du système est prise en compte par les études (Xu & Du, 2019) comme étant significativement associée à la qualité du système. Dans notre modèle, nous admettons qu'un système performant ne présente pas systématiquement un avantage relatif au besoin informationnel de l'utilisateur. En effet, un système peut retourner en réponse à une requête une grande partie des ressources contenues dans la base de données donc avoir un bon taux de rappel et de précision sauf que ces réponses peuvent ne pas satisfaire le besoin informationnel de l'utilisateur. Dans ce cas, les résultats doivent être interprétés en tenant compte des paramètres tels que le contenu global de la base de données et la qualité de l'indexation.

Enfin pour la désirabilité, rappelons que jusqu'ici aucune étude ne l'a intégrée comme dimension de l'évaluation des SRIL. Contrairement aux deux autres dimensions, il n'existe pas de corrélation entre les deux composantes. En effet, un système esthétiquement bien élaboré ne signifie pas qu'il s'intègre parfaitement dans le contexte de l'utilisateur. En clair, un système présentant un aspect visuel attrayant n'est pas nécessairement celui attendu par l'utilisateur. Il n'y a donc pas de relation de causalité entre les deux composantes. L'évaluateur doit donc les appréhender séparément, mais de façon complémentaire en gardant à l'idée le caractère subjectif de la composante aspect esthétique.

Utilité - utilisabilité - désirabilité : le tryptyque de l'acceptation d'un SRIL

Il pourrait apparaître que les trois dimensions d'évaluation (utilité - utilisabilité - désirabilité) sont relativement simples et se résument pour chacune d'entre elles à l'addition de leurs composantes respectives. En tenant compte de façon sommaire de ces composantes et de leurs critères associés, nous pouvons être tentés d'affirmer qu'il serait aisé d'évaluer l'acceptabilité d'un SRIL. Cependant, le fait de satisfaire les critères ne garantit en rien l'acceptabilité. Il convient de noter que malgré les relations de proximité qui s'établissent entre elles, l'utilité n'est pas l'utilisabilité, pas plus que l'utilisabilité n'est la désirabilité. Ainsi, un SRIL utile et utilisable n'est pas nécessairement désirable ou encore un SRIL désirable n'est pas forcément utile ni même utilisable. Aussi, la facile utilisation d'un système ne garantit pas de son utilité. En effet, un SRIL dont la prise en main est aisée, mais ne permet pas aux utilisateurs de combler leur besoin informationnel parce que ne ramenant pas de réponses précises

aux requêtes sera difficilement accepté et par ricochet moins utilisé.

Par ailleurs, bien que l'utilité soit une dimension nécessaire pour l'acceptabilité d'un SRIL, elle n'est pas suffisante si l'individu n'arrive pas à l'utiliser ou encore doit fournir beaucoup d'efforts cognitifs et techniques pour l'utiliser ou tout simplement s'il n'apprécie pas l'aspect visuel de l'interface. Dans cette perspective, l'acceptabilité d'un SRIL dépendra de la relation entre les fonctionnalités du système et leur facilité d'utilisation. Cependant, évaluer l'acceptabilité ne peut tenir qu'à ce seul critère de facile prise en main des fonctionnalités. Le SRIL peut proposer les fonctions de recherche et de présentation des résultats, ainsi que l'indexation requise, qui par ailleurs seraient facilement utilisables, pourtant, il peut ne pas être accepté par l'utilisateur. Le système doit au préalable satisfaire les besoins informationnels, mais également doit être bon, nécessaire, profitable, indispensable, voire salubre, c'est-à-dire désirable. Ainsi, l'utilisation de certains systèmes peut se faire de façon « naturelle » parce que s'intégrant dans le contexte de l'utilisateur (son cadre de vie) ou se situant dans la continuité des systèmes existants et que par ailleurs leur utilité « va de soi » (à titre d'exemple, l'OPAC de la bibliothèque d'une université pour un étudiant, celui de la médiathèque d'une commune pour un résident, celui du centre de documentation d'une institution pour un employé, etc.). En revanche, quand il s'agit d'un système dit émergent pour lequel l'utilisateur n'y a pas régulièrement recours, même s'il intègre le contexte de travail, son acceptabilité se fera progressivement au fur et à mesure que l'utilité du système se construira.

Au total, notre modèle dispose que l'évaluation d'un SRIL ne repose pas sur une relation binaire entre les dimensions, mais bien sur un ensemble de relations discrètes et continues entre elles et en leur sein.

« Quand » s'applique notre modèle ?

Notre modèle est applicable à différentes étapes de la « vie » du système : en partant de la conception (à priori) au déploiement (à posteriori) en passant par l'expérimentation (en fin de conception). L'évaluation en phase de conception vise à faire vérifier par les concepteurs que le système permettra aux utilisateurs d'atteindre sans grande difficulté le but pour lequel il est construit. À ce stade, l'évaluation permet de corriger rapidement et à moindre coût les éventuels problèmes qui pourraient être onéreux techniquement et financièrement, voire impossibles à corriger à la fin de la conception (Tricot et al., 2003). L'évaluation en fin de conception, c'est-à-dire au moment de l'expérimentation du système s'intéresse à sa version bêta. Durant cette étape, l'évaluation se fait avec des utilisateurs-testeurs qui n'ont aucun lien avec l'équipe de conception. Ces derniers seront conduits d'une part à exécuter des tâches que les utilisateurs destinataires du système seraient amenés à réaliser ; et d'autre part à évaluer la performance du système. À ce stade, il s'agit également de vérifier le temps d'apprentissage du système et les émotions ressenties à la vue et à l'utilisation du système. Les résultats permettent d'apporter des corrections avant la mise en service du système. Enfin, après le déploiement, il est question d'évaluer le produit fini et en cours d'utilisation. À cette étape, il est question de vérifier en situation réelle

comment le système répond aux besoins des utilisateurs, comment ils l'appréhendent et comment le système s'intègre dans leur contexte. Les résultats issus de cette évaluation permettent d'apporter des améliorations au système. Toutefois, quand les problèmes soulevés sont très importants, une telle évaluation peut nécessiter une refonte plus ou moins complète du système.

CONCLUSION

Dans cet article, nous avons postulé que pour évaluer l'utilisation ou pas d'un SRIL, il faut s'intéresser à ses conditions d'acceptabilité c'est-à-dire à son utilité, à son utilisabilité et à sa désirabilité. L'utilité du système recouvre sa performance et l'avantage relatif qu'il procure. L'utilisabilité se rapporte quant à elle, à l'efficacité du système, à son apprenabilité, à sa mémorabilité et à sa capacité de gestion des erreurs. Enfin, la désirabilité renvoie à l'apparence esthétique du système et à son intégration dans le contexte de l'utilisateur. Nous avançons également que toute relation entre ces trois dimensions est a priori possible et que ces probables relations sont interprétables suivant la rigueur que requiert une évaluation d'un SRIL. Il nous apparaît également que l'évaluation suivant une approche itérative au cours des différentes phases de développement du système partant du processus de conception au déploiement en passant par l'expérimentation permet de prévenir ou de corriger une part importante des éventuels défauts du SRIL.

Finalement, considérant la nouveauté de l'association de la dimension désirabilité aux dimensions traditionnelles de l'évaluation d'un SRIL que sont l'utilité et l'utilisabilité, nous admettons que notre proposition de cadre conceptuel et interprétatif est rudimentaire. Il mérite d'être affiné, en prenant en compte les inférences statistiques, mais aussi en intégrant des dimensions culturelles et sociales.

BIBLIOGRAPHIE

Bastien, J.M.C. & Scapin, D. (1993). *Ergonomic Criteria for the Évaluation of Human Computer interfaces*. Institut National de Recherche en Informatique et en Automatique. Repéré à <https://hal.inria.fr/inria-00070012/document>

Boutin, M. & Martial, O. (2001). *Évaluation de l'utilisabilité d'un site web: tests d'utilisabilité versus évaluation heuristique*. Communication présentée au 36e Congrès annuel de la SELF et au 32e Congrès annuel de l'ACE. Montréal. Repéré à : https://www.crim.ca/Publications/2001/documents/plein_texte/ERG_BouMal_0101-08.pdf

Brangier, E., Desmarais, M., Alexandra, N. et Tep, S. P. (2015). Évolution de l'inspection heuristique : vers une intégration des critères d'accessibilité, de praticité, d'émotion et de persuasion dans l'évaluation ergonomique. *Journal d'Interaction Personne-Système*, AFIHM, 2015, 4 (1), pp.69-84.

Brangier, E. & Barcenilla, J. (2003). Concevoir un produit facile à utiliser : Adapter les technologies à l'homme. Paris : Éditions d'Organisation. Repéré à https://www.acifr.org/concevoir_un_produit_extraits.pdf

Buchanan, S. & Salako A. Evaluating the usability and usefulness of a digital library. *Library Review*, 58(9),638-651. [doi.10.1108/00242530910997928](https://doi.org/10.1108/00242530910997928)

Chen, T. & Gross, M. (2017, July). Usability modeling of academic search user interface. In A. Marcus and W. Wang (Eds.): International Conference of Design, User Experience, and Usability, 2017, Part III, LNCS 10290, 16-30, 2017. Springer, Cham. [DOI: 10.1007/978-3-319-58640-3_2](https://doi.org/10.1007/978-3-319-58640-3_2)

Chowdhury, G., G. (2004) Access and usability issues of scholarly electronic publications, in Gorman, G.E. and Rowland, F. (Eds): *International Yearbook of Library and Information Management, 2004/2005* (Scholarly Publishing in an Electronic Era, 77-98). London: Facet Publishing.

Davis, F. D. (1989). Perceived usefulness, Perceived ease of use, and User acceptance of Information technology. *MIS Quarterly*, 13(3), 319-339.

Desmet, P. M. A, Hekkert, P. et Hillen, M. G. (2004). *Values and emotion*. In Proceedings of the Fifth European Academy of Design Conference. Barcelona, Spain.

Hassenzahl, M. & Tractinsky, N. (2006). User experience - a research agenda. *Behavior & Information Technology*, 25, 91-97.

ISO 9241-110 (2006). Ergonomics of human-system interaction - part 110: dialogue principles. Repéré à www.iso.org/obp/ui/#iso:std:iso:9241:-110:ed-1:v1:fr.

ISO 2789 (2013) Information and documentation-International library statistics. Repéré à <https://www.iso.org/obp/ui/#iso:std:iso:2789:ed-5:v1:en>

Kouakou, K. S. (2015). Adoption des réseaux sociaux numériques par les bibliothécaires des universités ivoiriennes. *Les Cahiers du numérique*, 11(2), 167-202. doi.org/10.3166/lcn.11.2.167-202

Lefèvre, P. (2000). *La recherche d'informations: Du texte intégral au thésaurus*. Paris: Hermès Science publications.

Nielsen, J. (1993), *Usability Engineering*. Boston MA: Academic Press.

Norman, D. A. (2012) *Design émotionnel, Pourquoi aimons-nous ou détestons-nous les objets qui nous entourent ?* Bruxelles : De Boeck.

Schmidt, A. & Etches, A. (2016). Utile, utilisable, désirable : Redessiner les bibliothèques pour leurs utilisateurs. (traduit par N. Clot). Villeurbanne : Presses de l'ENSSIB, cop. 2016. - 1 vol. (258 p.) Repéré

à <https://www.enssib.fr/bibliotheque-numerique/utile-utilisable-desirable.pdf>

Seffah, A., Donyaee, M., Kline, R. B. et Padda, H.K. (2006). Usability measurement and metrics: A consolidated model. *Software Quality Journal*, 14, 159-178. doi.org/10.1007/s11219-006-7600-8

Senach, B. (1990). Evaluation ergonomique des interfaces homme-machine : une revue de la littérature. [Rapport de recherche]. Institut National de Recherche en Informatique et en Automatique. Repéré à <https://hal.inria.fr/inria-00075378/document>

Silvis I. M., Bothma T.J.D. et De Beer K.J.W. (2019). Evaluating the usability of the information architecture of academic library websites. *Library Hi Tech*, 37(3), 566-590. doi.org/10.1108/LHT-07-2017-0151

Tricot, A., Plégat-Soutjis, F., Camps, J. F., Amiel, A., Lutz, G. et Morcillo, A. (2003, avril). *Utilité, utilisabilité, acceptabilité: interpréter les relations entre trois dimensions de l'évaluation des EIAH*. In colloque EIAH, Strasbourg (pp. 391 - 402). Repéré à <https://edutice.archives-ouvertes.fr/edutice-00000154/document>

Tsakonas, G. & Papatheodorou, C. (2008). Exploring usefulness and usability in the evaluation of open access digital libraries. *Information Processing & Management*, 44(3), 1234-1250. Repéré à <https://www.sciencedirect.com/science/article/pii/S0306457307001318>

Xu, F. & Du, T. J. (2018). Factors influencing users' satisfaction and loyalty on digital libraries in Chinese universities. *Computers in Human Behaviour*, 83(6), 64-72. doi.org/10.1016/j.chb.2018.01.029

Xu, F. & Du, T. J. (2019). Examining differences and similarities between graduate and undergraduate students' user satisfaction with digital libraries. *The Journal of Academic Librarianship*, 45, 1-8. doi.org/10.1016/j.acalib.2019.102072

[1] Il désigne le nombre de documents pertinents ramenés à l'issue d'une requête en rapport avec le nombre de documents pertinents disponibles dans le système. Un rappel de 1 signifie que tous les documents pertinents apparaissent dans le résultat. Un rappel nul indique que le résultat ne contient aucun document pertinent.

[2] Elle se réfère à la capacité du système à ramener que des résultats pertinents. La précision permet donc d'évaluer la proportion de documents pertinents parmi l'ensemble de ceux trouvés par le système. Une précision de 1 indique une absence totale de documents inutiles. Une précision égale à 0 signifie qu'il n'y a aucun document pertinent dans le résultat.