

ARTICLES, MÉTHODE

## ÉVALUATION DE LA QUALITÉ DES CONTENUS NUMÉRIQUES: REGARD SUR LA MÉTHODE D'AUDIT DE CONTENU

11/05/2017 | ISABELLE SPERANO | UN COMMENTAIRE

*Isabelle Sperano est enseignante-chercheuse dans le Master architecture de l'information à l'ENS de Lyon. Ses activités d'enseignement et de recherche portent sur l'UX design et l'architecture de l'information. Elle s'intéresse particulièrement à l'évaluation, à l'organisation et au design de l'information dans des environnements numériques de grande ampleur. Cet article présente une partie des résultats d'une étude conduite dans le cadre de sa [thèse de doctorat](#) en architecture de l'information à l'école de design de l'Université Laval (Québec) portant sur l'évaluation de la qualité des contenus numériques.*

Depuis une vingtaine d'années, le **Web s'est imposé comme média privilégié** par une variété d'organisations (instances gouvernementales, entreprises, organismes, etc.) pour transmettre une quantité substantielle d'informations à leurs destinataires. La mise en place adéquate de ces imposantes structures informationnelles dépend, entre autres, de **spécialistes de l'information et de la communication de divers champs d'expertise, dont l'architecture d'information (AI)**. Dans la mesure où l'accès à une information particulière constitue l'objectif principal d'une majorité d'activités de recherche sur le Web, ces spécialistes portent normalement une **grande attention au contenu présenté**. Pour ce faire, **différentes méthodes visant à mieux appréhender le contenu ont vu le jour depuis quelques années**. L'audit de contenu est l'une d'elles.

### L'audit de contenu

L'audit de contenu est une **méthode d'évaluation du contenu d'un site Web ou d'un écosystème informationnel plus vaste** (réseaux sociaux, applications, infolettre, etc.). Elle rend possible l'identification, le dénombrement, la description et surtout l'évaluation de chaque élément de contenu et de ses caractéristiques respectives (Jones 2010; Land 2014; Martin et Hanington 2012). Concrètement, **cette méthode consiste à confronter chaque contenu – chaque « page » dans le cas d'un site Web – à une liste de critères d'évaluation sélectionnés par l'auditeur**. Tout au long de son examen diagnostique, l'auditeur note ses observations et consigne ses constatations dans une grille.

Conduit adéquatement et avec rigueur, **l'audit de contenu constituerait un outil indispensable à la révision ou à la conception d'un écosystème informationnel** (Bloomstein 2012; Pernice 2015). Cette méthode d'analyse rétrospective du contenu et de la structure informationnelle non seulement permettrait de **cerner l'étendue et l'ampleur d'un écosystème informationnel, mais contribuerait aussi à engager des discussions et à susciter des débats entre les différents acteurs d'un projet numérique**. Cet examen approfondi rendrait possible l'identification de lacunes à corriger (ex. : redondance du contenu, faible qualité rédactionnelle, manque d'actualisation, faiblesses structurelles) et révélerait corrélativement l'ampleur des problèmes à résoudre (Halvorson et Rach 2012; Land 2014). En outre, l'audit de contenu rendrait possible l'anticipation de certains problèmes (contenus jugés trop sensibles, lesquels ne devraient pas apparaître sur le Web, faiblesses dans la chaîne éditoriale, etc.) (Halvorson et Rach 2012). **De par sa capacité à mettre en lumière différents paramètres du contenu sinon imperceptibles au concepteur**, et ce, dès le début d'un projet d'AI, l'audit de contenu influencerait grandement, selon plusieurs auteurs (Halvorson et Rach 2012; Kissane 2011; Spencer 2010), **la réussite des étapes subséquentes d'une démarche d'AI** (organisation de l'information, conception de la navigation, etc.).

**La communauté d'AI se heurte actuellement à un déficit de recherche à propos de l'audit de contenu**. En effet, rares, voire inexistantes, sont les travaux de recherche qui en rendent compte de manière exhaustive. Cette lacune en recherche s'étend d'ailleurs à l'ensemble du champ de l'AI (Hobbs, Fenn, et Resmini 2010). Bien que les approches, les modèles et les concepts proposés à l'heure actuelle en AI soient certainement dignes d'intérêt, ils découlent la plupart du temps de l'externalisation brute et embryonnaire d'une intuition – sans vouloir déprécier cette faculté souvent à l'origine d'importantes découvertes –, d'un pressentiment ou d'un commentaire, sans être suivi d'une démarche de recherche rigoureuse (Madsen 2009). Seulement, ce type de savoir, potentiellement riche, apparaît plutôt éphémère lorsqu'il n'est pas rigoureusement formalisé. Dans ce champ, opinions et savoirs se confondent dans des discussions souvent circulaires au sein de la communauté de pratique, dont le salut dépend en majorité de quelques « maîtres à penser » charismatiques. **Poser un regard descriptif, analytique et critique sur la pratique contribue à la formalisation de l'AI, de ses méthodes et de ses théories, ainsi qu'à munir l'expertise d'une base de connaissances fiable** (Burford 2014). C'est ainsi que Resmini et Instone (2010) avancent qu'il est temps que les questions de recherche en AI portent davantage sur la pratique et examinent certaines prémisses et certains dictats du champ mis

de l'avant par les professionnels. En suivant cette prémisse, ouvrir la voie de la recherche à propos de l'audit de contenu demande l'établissement d'une base de connaissances, d'un état des savoirs actuels relatifs à cette méthode. **Concrètement, un besoin d'inventorier, de formaliser les savoirs actuels, d'établir le socle de connaissances relatif à l'audit de contenu constitue le point d'entrée vers l'étude de cette méthode.**

## Analyse d'un corpus de publications

Toute discipline possède sa collection de savoirs intégrés à sa pratique. Ces connaissances disciplinaires se dévoilent non seulement à travers les objets construits et les artefacts produits, mais aussi à travers les écrits d'experts (Hobbs et al. 2010). Fort d'une validation plutôt informelle effectuée en grande partie par des praticiens, l'audit de contenu se voit justement abondamment décrit et commenté dans une foule d'écrits professionnels. C'est ce qui nous a menée à **examiner un corpus d'écrits d'experts abordant l'audit de contenu** (ouvrages, publications Web, articles).

Le repérage des documents s'étend à trois types de publications :

- Ouvrage (ou chapitre)
- Document Web (article de blogue, page Web, etc.)
- Article de revue ou actes de conférence.

Afin de repérer les ouvrages à intégrer dans le corpus, nous avons effectué des recherches sur les **grands sites de vente en ligne de livres, dans des catalogues de bibliothèques en ligne ainsi que dans des répertoires de livres numériques**. Pour ce qui est des articles de revue scientifiques et des actes de conférence, nous avons entrepris une recherche dans les bases de données d'articles sur les **sites de revues spécialisées**. À partir des moteurs de recherche de ces sites, nous avons cherché les expressions suivantes relatives :

En anglais

- *content audit*
- *content inventory*
- *content assessment*

En français

- audit de contenu
- inventaire du contenu
- audit éditorial

La sélection des publications Web posait un défi supplémentaire : cerner la crédibilité des auteurs. Nous le savons, quiconque (ou presque) est en mesure de publier du contenu sur le Web. Nous craignons qu'une recherche de publications Web à partir de grands moteurs de recherche comme Google ne mène vers des publications de qualité trop variable et incertaine. C'est ce qui nous a menée à adopter une autre voie de sélection. Nous avons fait usage des références citées dans le corpus d'ouvrages comme base de départ pour la recherche de documents Web. **Ainsi, nous avons entrepris d'effectuer la sélection en identifiant les documents Web préalablement cités dans les ouvrages du corpus.**

Cette recherche nous a permis de réaliser un **panorama des publications du corpus ainsi que de leurs auteurs**. Dans un premier temps, nous nous pencherons sur les publications en tant que telles, à savoir les années de publication, leur provenance et leur langue. Nous présenterons ensuite les auteurs de ces publications, leurs expertises et leurs statuts professionnels.

## Panorama des publications

Au terme de la démarche d'élaboration du corpus, nous avons pu sélectionner 200 publications. Le corpus est composé ainsi :

- 105 ouvrages
- 12 chapitres (issus de 3 ouvrages différents)
- 77 documents Web
- 6 articles de revue/actes de conférence

La composition du corpus constitue déjà un résultat révélateur quant à la provenance du discours à propos de l'audit de contenu. En effet, le fait de n'avoir pu relever que six articles de revues ou actes de conférence laisse supposer que le discours relatif à l'audit de contenu n'est que rarement issu de la recherche, comme nous le soupçonnions au départ. Chaque publication est nommée en suivant une nomenclature précise. C'est à ces identifiants que nous ferons référence pour désigner les publications.

Par exemple :

Identifiant	Source complète
ALL11W	Allen, R. (2011). ROT: The Low-Hanging Fruit of Content Analysis. Meet Content. Consulté à <a href="http://meetcontent.com/blog/rot-the-low-hanging-fruit-of-content-analysis/">http://meetcontent.com/blog/rot-the-low-hanging-fruit-of-content-analysis/</a>

Voici les règles de création des identifiants :

- 3 premières lettres du nom de famille du premier auteur (ex. : **ALL11W**).
- 2 derniers chiffres de l'année de la publication (ex. : **ALL11W**).
  - Type de publication (ex. : **ALL11W**)
- O : Ouvrage
- C : Chapitre
- W : Document Web
- A : Article

## Années de publication

Nous avons voulu, dans un premier temps, observer l'évolution chronologique de l'audit de contenu à travers l'examen des années de publication des écrits du corpus. Nous avons pu obtenir cette information pour 173 des 200 publications (86,5 % des publications). En effet, l'année de publication n'était pas indiquée pour 27 documents Web (13,5 % des publications). Les résultats sont présentés dans la Figure 1.

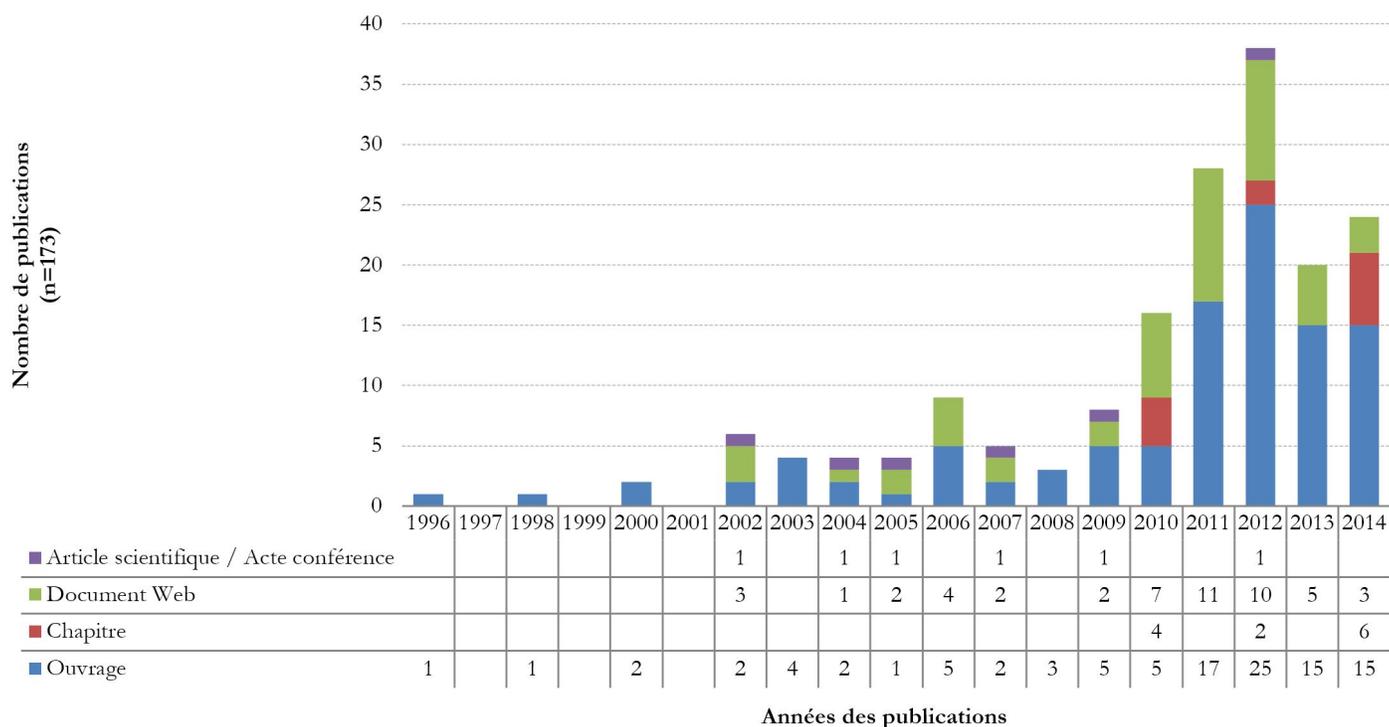


Figure 1 - Nombre de publications en fonction des années de publication

On observe que la première occurrence de l'audit de contenu dans le corpus est un ouvrage publié en 1996. Cette publication [ALL96O] traite du logiciel de conception de CD-ROM Director et du Lingo, le langage de programmation associé à ce logiciel.

L'audit de contenu apparaît comme une méthode relativement récente. En effet, si quelques publications abordent l'audit de contenu avant les années 2000, nous observons que c'est réellement à partir de l'année 2002 que la fréquence des écrits devient plus régulière. **On remarque un accroissement substantiel du nombre de publications entre 2010 et 2012, ce qui laisse supposer une certaine popularité de la méthode à partir de 2010<sup>[1]</sup>.**

Nous avons aussi voulu savoir si les origines de la méthode étaient explicitement abordées dans le corpus. Nous nous sommes toutefois rapidement aperçue que la quasi-totalité des publications du corpus reste muette à ce propos. En effet, seulement deux publications abordent vaguement cet aspect.

EXA12W-1 dit qu'il s'agit d'une méthode en usage depuis une dizaine d'années :

*For at least ten years, Web professionals have been talking about content inventories.*  
[EXA12W-1]

WAC13W abonde dans le même sens :

*People have been talking about content audits and inventories for more than a decade.*  
[WAC13W]

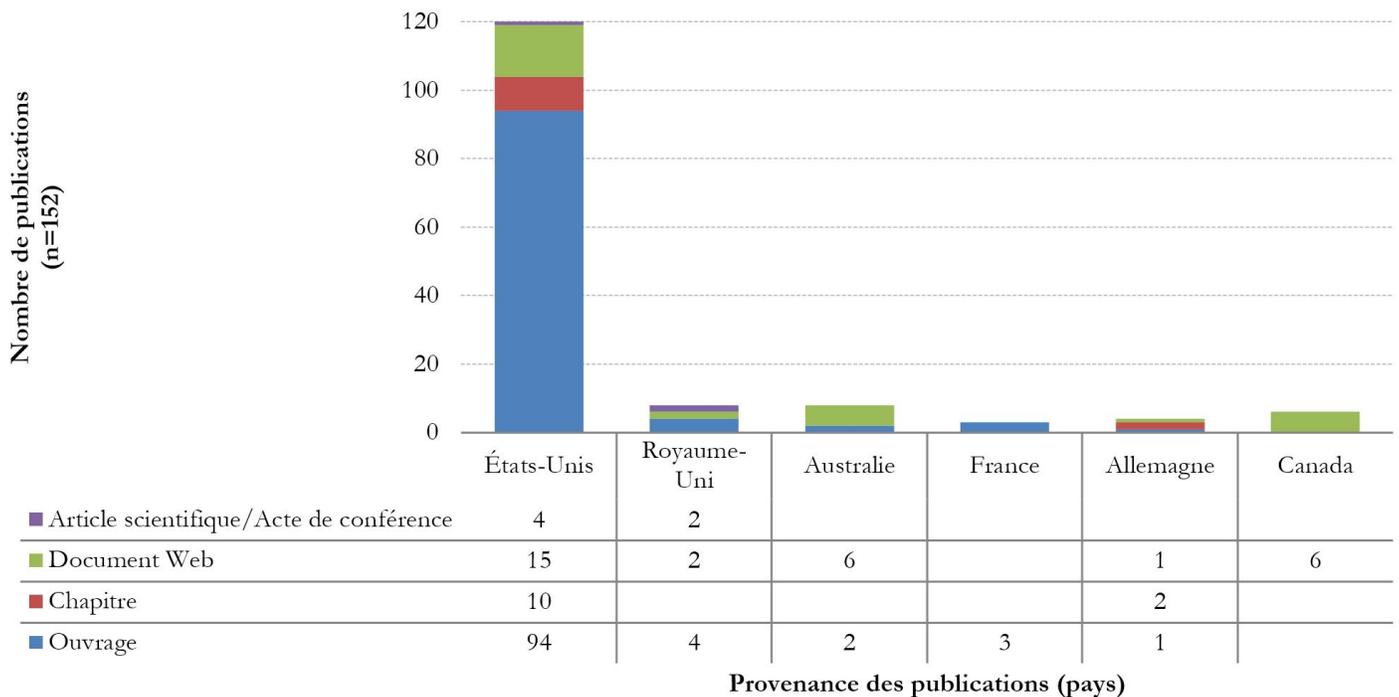
Les dires de ces auteurs semblent s'accorder avec les années de publications des écrits du corpus.

## Provenance des publications

Nous avons ensuite relevé la provenance (pays) des publications (Figure 2). Cette donnée était disponible pour 152 publications (76 % des publications).

Figure 2 - Nombre de publications en fonction de leur provenance (pays)

On constate que la grande majorité des publications provient des États-Unis. Si cette donnée ne se substitue guère à la provenance des auteurs, elle laisse tout de même penser que le discours actuellement véhiculé est probablement à forte dominance états-unienne. Par ailleurs, les publications sont pour la plupart en langue anglaise (196 publications, soit 98 % des publications), ce qui signifie que le discours présent dans les publications du corpus est quasi uniquement anglophone. Il est à noter que les publications canadiennes et allemandes du corpus sont aussi en langue anglaise. Les autres publications (4 publications, soit 2 % des publications) sont en langue française. Comme notre corpus se



limitait à des publications en langues anglaise et française, nous ne pouvons généraliser ce résultat à l'ensemble des écrits à propos de l'audit de contenu. Toutefois, cet écart important entre les États-Unis et les autres pays mis en lumière par cette étude laisse tout de même penser que la majorité du discours à propos de l'audit de contenu provient de ce pays.

## Portrait des auteurs

Après l'examen des publications en tant que telles, nous nous penchons à présent sur leurs auteurs. Nous avons relevé 197 auteurs différents au sein du corpus. De ces auteurs, 15 ont publié plus d'un des documents du corpus.

Nous avons ensuite fait l'examen du statut professionnel des auteurs. Cette donnée a pu être identifiée pour la quasi-totalité des auteurs, soit pour 188 auteurs sur 197 (95,4 % des auteurs). Cette recension nous a permis d'identifier six statuts distincts :

## Statuts professionnels des auteurs

**Praticien.** Usage de son expertise dans sa pratique professionnelle.

**Auteur.** Auteur d'ouvrages ou de billets de blogues, par exemple<sup>[2]</sup>.

**Conférencier.** Ce statut est octroyé lorsque l'individu précise qu'il réalise des présentations dans le cadre de conférences.

**Formateur.** Ce statut est donné aux individus disant donner des cours, dans un milieu autre que collégial ou universitaire, que ce soit en entreprise ou dans le cadre d'ateliers (ou *workshop* en anglais).

**Enseignant.** Nous attribuons ce statut à ceux qui enseignent dans un collège ou dans une université sans toutefois détenir le titre de professeur.

**Professeur.** Ce statut est réservé aux professeurs d'université.

La Figure 3 présente le nombre d'auteurs se réclamant de chacun de ces statuts professionnels (plusieurs statuts sont souvent cumulés).

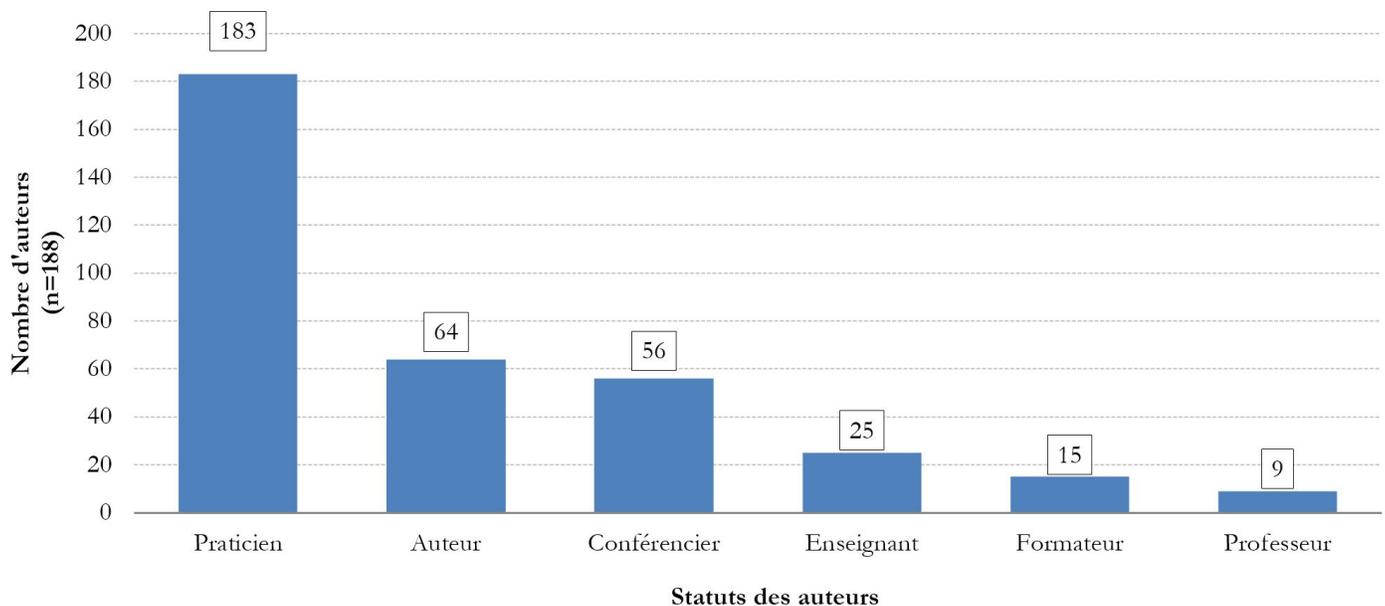


Figure 3 – Statuts professionnels des auteurs figurant dans le corpus

Nous avons vu précédemment que certains indices portaient à croire que **l'audit de contenu était surtout décrit par des praticiens**. Les données illustrées dans la figure ci-haut corroborent effectivement ces avis préliminaires. En effet, nous pouvons dégager de la figure ci-dessus un trait dominant : la quasi-totalité des auteurs du corpus dont nous avons pu identifier le statut se présentent comme des praticiens (183 auteurs, soit 97,3 % des auteurs). De plus, un nombre très restreint évolue dans la sphère universitaire (9 s'identifient comme professeurs, soit 4,8 % des auteurs). Par ailleurs, il paraît essentiel de souligner que plusieurs statuts sont souvent cumulés, **le plus fréquent étant la triade praticien-auteur-conférencier** (26 auteurs du corpus, soit 13,8 % des auteurs).

## Champs d'expertise des auteurs



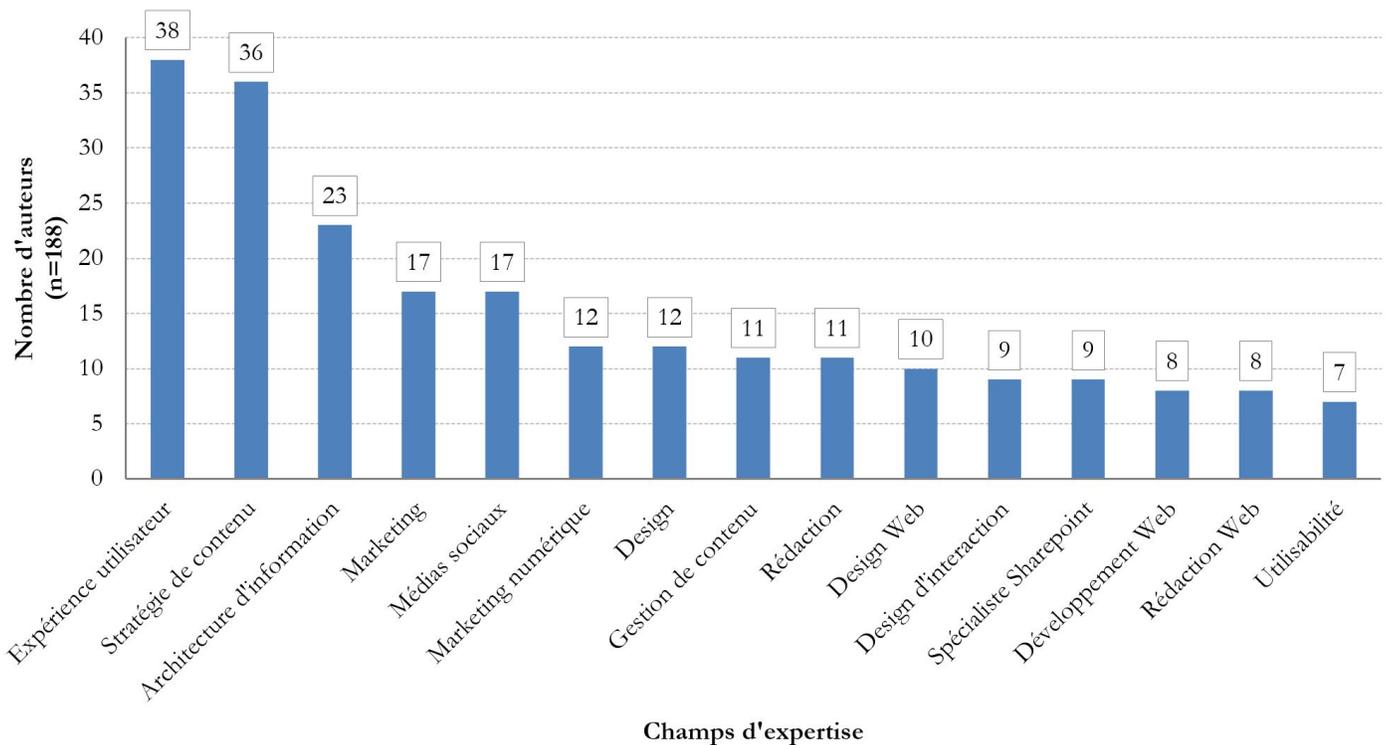


Figure 5 – Champs d'expertise les plus populaires en fonction des auteurs

On remarque que l'expérience utilisateur (38 auteurs, soit 20,2 % des auteurs), la stratégie de contenu (36 auteurs, soit 19,1 % des auteurs) et l'architecture d'information (23 auteurs, soit 12,2 % des auteurs) sont les champs d'expertise les plus revendiqués. Cet examen plus attentif des champs d'expertise les plus revendiqués révèle que **la rédaction** (rédaction, rédaction Web) et **le design** (design, design Web, design d'interaction) se positionnent aussi de façon centrale.

On remarque aussi la forte présence d'une expertise relative au **logiciel SharePoint** (spécialiste SharePoint)<sup>[6]</sup>. La forte présence de cette expertise s'explique principalement par le fait que les trois publications traitant de ce logiciel sont rédigées par de nombreux auteurs qui s'identifient tous comme spécialistes SharePoint.

Par ailleurs, il n'est pas rare que les auteurs précisent leur zone d'intervention dans leur expertise (ex. : Rédaction **Web**, design **Web**, marketing **numérique**), mais que d'autres demeurent plus généraux (ex. : Rédaction, design, marketing).

## Conclusion

Cette étude a permis de dessiner un **portrait des publications relatives à l'audit de contenu**. Les résultats présentés montrent que l'audit de contenu serait **une méthode relativement récente**. Nous avons aussi pu constater que le discours véhiculé à propos de l'audit de contenu semble être à **forte dominance états-unienne et est quasi exclusivement anglophone**. Enfin, nous avons pu observer

que l'audit de contenu est peu issu de la recherche et est surtout décrit par des praticiens, lesquels se réclament de nombre d'expertises. L'expérience utilisateur, la stratégie de contenu et l'architecture d'information sont les trois plus fréquentes.

Cette recherche a fourni un ancrage factuel à propos de la méthode plutôt qu'une série de commentaires anecdotiques, comme c'est trop souvent le cas en AI et dans les expertises connexes comme l'UX (*user experience* ou expérience utilisateur en français) et la stratégie de contenu (Hobbs et al. 2010). Elle constitue **un premier pas vers une formalisation de l'audit de contenu et des autres méthodes mises de l'avant en architecture d'information.**

## Bibliographie

Bloomstein, M. (2012). *Content strategy at work real-world stories to strengthen every interactive project*. Waltham, MA: Morgan Kaufmann.

Burford, S. (2014). The Interplay of the Information Disciplines and Information Architecture. Dans A. Resmini (Éd.), *Reframing Information Architecture* (p. 47-59). Springer International Publishing.

Halvorson, K. et Rach, M. (2012). *Content Strategy for the Web* (2e édition.). Berkeley, CA: New Riders Press.

Hobbs, J., Fenn, T. et Resmini, A. (2010). Maturing a Practice. *Journal of Information Architecture*, 2(1), 37-54.

Jones, C. (2010). *Clout: The Art and Science of Influential Web Content*. Berkeley, CA: New Riders.

Kissane, E. (2011). *The Elements of Content Strategy*. New York: A Book Apart.

Land, P. L. (2014). *Content Audits and Inventories: A Handbook*. Laguna Hills, CA: XML Press.

Madsen, D. (2009). Shall We Dance? *Journal of Information Architecture*, 1(1), 1-5.

Martin, B. et Hanington, B. (2012). *Universal Methods of Design: 100 Ways to Research Complex Problems, Develop Innovative Ideas, and Design Effective Solutions*. Beverly, MA: Rockport.

Microsoft. (2016). Microsoft Sharepoint. Consulté le 17 mars 2016, à l'adresse <https://products.office.com/fr-ca/sharepoint/collaboration>.

Pernice, K. (2015). Content Migration Alone Is Not An Effective Content Strategy. *Alertbox*. Consulté à l'adresse <http://www.nngroup.com/articles/intranet-content-strategy/>.

Resmini, A. et Instone, K. (2010). Research and Practice in IA. *ASIS&T Bulletin*. Consulté à l'adresse [https://www.asis.org/Bulletin/Aug-10/AugSep10\\_Resmini\\_Instone.html](https://www.asis.org/Bulletin/Aug-10/AugSep10_Resmini_Instone.html).

Spencer, D. (2010). *A Practical Guide to Information Architecture*. Penarth, Royaume-Uni: Five Simple Steps.

## Notes

[1] Les 12 chapitres proviennent de 3 ouvrages différents.

[2] Évidemment, tous les individus figurant dans le corpus sont l'auteur d'au moins une publication (figurant dans le corpus). Toutefois, nous avons uniquement catégorisé un individu comme tel lorsque celui-ci s'identifiait lui-même ainsi.

[3] Dans cette figure, les expertises revendiquées par plus de six auteurs sont de couleur orangée, les expertises revendiquées par deux à six auteurs sont de couleur bleue et les expertises revendiquées une seule fois apparaissent en gris. De plus, la taille des caractères est proportionnelle au nombre d'occurrences des expertises.

[4] Nous avons systématiquement traduit les noms d'expertise et avons tenté de les identifier le plus fidèlement possible, sans effectuer de regroupement ou de catégorisation, afin de refléter le plus justement possible la réalité observée.

[5] Nous avons remarqué que, à travers la recherche sur les auteurs, les expertises dont chacun se réclame varient parfois d'un site à l'autre et d'un ouvrage à l'autre. L'information à propos de l'expertise demeure tout à fait pertinente, mais ce phénomène de désignation mouvant et multiforme demande de faire preuve de prudence et invite à interpréter ces résultats de façon globale, dans leur ensemble, sans leur attribuer un poids démesuré.

[6] SharePoint est un système de gestion de contenu (SGC ou *content management system*) qui permet aux organisations de créer des sites Web et de mettre en place des Intranets (Microsoft 2016).



