

# Conférence Ilearn forum

## Chaîne éditoriale, accessibilité, mobilité

Date : janvier 2008

Auteurs : Élie Sloïm (Témésis), Stéphane Crozat (UTC)

### Abstract

L'accessibilité est devenue une problématique indispensable de tous les fournisseurs de contenus numériques : comment mettre à disposition une information qui puisse être exploitée pour tous, partout ? Cette problématique inclut les questions du handicap, des zones blanches, des terminaux mobiles, etc.

Afin de répondre à cette problématique il existe deux grandes classes de solution :

- Produire dans un cadre technique très contraignant, sorte de "plus petit dénominateur commun" des modalités de diffusion
- Produire plusieurs fois le contenu en fonction des contraintes des supports cibles

La première solution peut être trop restrictive - et donc frustrante - au regard des supports les plus riches (multimédia, interactivité). La seconde est évidemment une source de travail supplémentaire, et - plus fondamentalement problématique - de duplication de l'information.

Il existe une autre possibilité, qui consiste à définir des structures de contenus, puis à les mettre à disposition et en forme pour des contextes de diffusion différents . C'est la voie choisie par certains CMS avancés ou par les chaînes éditoriales XML.

Celles-ci reposent sur le principe du multi-supports : Un contenu écrit une seule fois (en XML) pourra être transformé plusieurs fois en plusieurs formats de diffusion différents, en prenant en compte les contraintes du contexte cible.

On peut ainsi produire :

- Une version HTML légère pour les faibles débits,
- Une version HTML modeste pour les terminaux mobiles,
- Une version très standard (HTML W3C strict) pour la diffusion Web la plus large
- Une version adaptée pour les lecteurs spécialisés dans le support à un handicap
- Une version PDF pour l'impression
- Une version HTML autement interactive spécifique pour un Intranet (compatible avec un navigateur puissant comme Firefox)
- Une version multimédia en Flash
- Etc.

Cette solution permet de penser la spécificité de chaque support de diffusion afin de profiter au mieux de ses capacités éditoriales. Elle permet de ne pas faire porter à chaque support les contraintes de tous les autres. Elle allie donc puissance de publication des formats "modernes" dans des cadres contrôlés et accessibilité de supports plus "standard" dans des cadres plus larges.

Nous illustrerons ce propos à travers les travaux réalisés dans le cadre de la chaîne éditoriale Scenari/Opale.

# Article : Chaînes éditoriales, accessibilité et mobilité

## ***Les enjeux actuels de la publication de documents numériques***

La publication de documents numériques est passée en quelques années d'une activité réservée à des professionnels avertis à une activité pratiquée par presque tous les utilisateurs de services numériques.

Dans un premier temps, seuls les spécialistes de la création de documents avaient accès à ces activités, notamment à travers l'utilisation de logiciels spécialisés ou de chaînes éditoriales. Avec la diffusion des logiciels de traitement de texte, puis avec l'émergence et le développement du Web puis la mise en place de système de gestion de contenus Web, et enfin l'ouverture de ces CMS aux contributeurs externes, c'est maintenant un ensemble immense de contributeurs qui est amené à publier des documents numériques.

La quantité d'information traitée numériquement augmente mécaniquement. Assez logiquement, des aides technologiques sont progressivement mises en place pour faciliter la recherche et la navigation dans tous ces documents numériques.

Il est toutefois possible que les principales difficultés ne soient pas encore apparues. Le monde de la publication numérique va en effet devoir relever de nouveaux défis, particulièrement sensibles dans le secteur de la formation. Le premier de celui-ci concerne l'accessibilité des contenus Web à tous et leur utilisation en contexte de mobilité, le deuxième concerne l'industrialisation des processus de publication. Le troisième enfin concerne la diffusion de nouvelles pratiques et de nouveaux outils à tous les producteurs d'information, du spécialiste au grand public en passant par les professionnels dans leur pratique quotidienne.

Dans cette intervention, nous traiterons tout d'abord de l'accessibilité des contenus numériques et des avantages de cette approche dans des contextes de formation et de mobilité. Nous montrerons ensuite différentes solutions permettant de produire des contenus numériques. Nous poursuivrons par un focus sur les chaînes éditoriales, et en conclusion, nous verrons les perspectives et les limites des chantiers qui restent à traiter dans le domaine.

## ***Qu'est ce que l'accessibilité ?***

Avant de se pencher plus avant sur les problématiques de production de contenus numériques, il importe d'expliquer en quoi consiste leur accessibilité, à qui et à quoi ce sujet profite, et en quoi les critères d'accessibilité permettent d'améliorer la production de documents numériques.

Pour le W3C, à travers la voix de son directeur Tim Berners Lee, l'accessibilité consiste à « mettre le Web et ses contenus à la disposition de tous les individus, quelque soit leur matériel ou logiciel, leur langue maternelle, leur culture, leurs infrastructures réseau ou leurs aptitudes physiques ou mentales ».

Le principe qui sous-tend l'accessibilité des contenus Web a fait l'objet de directives internationales, produites en 1999, les Web content Accessibility Guidelines, qui définissent la façon dont doit être conçu, structuré et proposé un contenu pour être accessible à tous.

L'accessibilité n'est pas une affaire de sites, mais de contenus. L'accessibilité est bien l'affaire de tous et pas seulement des personnes handicapées.

Pour bien faire comprendre de quoi il s'agit, nous citerons rapidement trois exemples qui montrent comment un document peut être rendu accessible.

- **Exemple 1** : Les images contenant du texte sous forme de représentations graphiques ne peuvent pas être comprises ou décryptées par des machines, des moteurs de recherche, des

lecteurs d'écran. Pour éviter une perte d'information, les producteurs de contenus Web devront s'efforcer d'associer à ces documents graphiques des contenus alternatifs sous forme textuelle, permettant à des machines d'extraire du document le sens ou le contenu textuel des images.

- **Exemple 2** : Pour une personne déficiente visuelle, voire pour certaines personnes déficientes mentales, et finalement d'une certaine manière pour tout usagers des services Web, il peut être essentiel d'accéder et de comprendre la structure d'un document. Pour ceci, il importe d'affecter des éléments d'ordre sémantique à la structure du document. De cette façon, une machine ou un être humain pourra extraire en cas de besoin la structure d'un document.
- **Exemple 3** : Des contenus proposés sous forme d'acronymes ou de sigles ne pourront être compris par tous les usagers que si ces acronymes sont explicités. Ici encore, les sigles et acronymes seront explicités et deviendront dès lors accessibles à des personnes handicapées, ou encore à des machines.

En rendant les contenus accessibles, les créateurs de contenus améliorent leur structure, facilitent leur interprétation par des machines, et notamment les robots de moteurs de recherche. En ce sens, l'accessibilité est un moteur du référencement. En anticipant le traitement des contenus par toutes les machines, l'accessibilité permet de traiter une grande partie des problématiques de la mobilité.

## ***Produire des contenus numériques***

Le domaine de la production de contenus numériques est un secteur industriel. L'objectif est de pouvoir produire des documents structurés et accessibles en grand volume et à grande vitesse. La production doit également être pérenne, c'est à dire qu'elle nécessite d'avoir des garanties sur la capacité de lire et de récupérer les contenus produits non seulement aujourd'hui, mais également dans l'avenir. Le choix des formats ouverts et de l'interopérabilité maximum devient alors décisif. Pour finir, contrairement à ce qu'il se passait il y a quelques années, un contenu Web devra s'afficher dans de multiples configurations, le nombre de facteurs variant dans chaque contexte étant considérable (système d'exploitation, couleurs, écran, résolution, polices, plug-ins, logiciels, navigateurs connexion).

Plusieurs réponses ont été proposées pour résoudre cette problématique de la production de contenus numériques à grande échelle.

Le *traitement de texte* est bien évidemment une solution fort intéressante pour produire des contenus. Elle a fait ses preuves sur un certain nombre de points, mais elle a tout de même deux défauts majeurs. Premièrement, seul un petit nombre d'utilisateurs de traitement de texte savent élaborer une structure documentaire correcte. Très peu d'utilisateurs savent réellement séparer le contenu de la présentation en utilisant les styles. Ensuite, le traitement de texte intègre la navigation en mode hypertexte.

La *production de pages Web en mode WYSIWYG*<sup>1</sup> est une version Web de nos logiciels de traitement de texte. Malheureusement, au moins dans leurs premières versions il s'agissait surtout de produire des pages individuelles, ce qui posait de nombreux problèmes de cohérences entre pages et de maintenance. De plus la plupart de ces logiciels ne permettaient pas de séparer de manière véritablement complète le contenu et la présentation.

Les *Content Management System (CMS)*, nés du développement des sites dynamiques, présentent de nombreux avantages. Ils permettent d'extraire le travail technique de la publication électronique pour ne laisser le travail d'écriture et uniquement le travail d'écriture aux contributeurs. Plus besoin de connaître les arcanes du Web pour publier sur le Web. Ils offrent ainsi un premier niveau de

---

1 What You See Is What You Get

séparation entre le contenu et la présentation. En revanche, il est encore extrêmement difficile de leur faire produire des contenus intégralement structurés et accessibles, en particulier car la structure de rédaction reste fortement liée à une structure de présentation préalable : on écrit pour un support, ici et maintenant.

Les *chaînes éditoriales* partent du principe assez simple qu'un document est avant tout une structure sémantique. Il s'agit dès lors de délimiter avec précision la structure du document, pour pouvoir plaquer *ensuite* sur cette structure une couche de présentation. Ce mode de travail possède plusieurs avantages. Tout d'abord, le document est structuré nativement. Ensuite, la présentation est totalement séparée du contenu, ce qui permet de proposer plusieurs vues de publication pour un même contenu. Le rôle du rédacteur n'est pas de présenter, mais bien de rédiger des contenus. Le rôle du système informatique devient d'interpréter cette rédaction pour la présenter sur plusieurs supports, pour plusieurs contextes : ici et maintenant, mais aussi là-bas et plus tard.

## **Focus sur les chaînes éditoriales**

Les chaînes éditoriales sont nées en même temps que les traitements de texte à la fin des années 70, avec les technologies TeX/LaTeX et GML qui deviendra la norme SGML en 1986. En 1998 le standard XML actualise la technologie en la rendant beaucoup plus accessible, ce qui permet l'extension des chaînes éditoriales, là où elles étaient confinées à des contextes de pointes (documentation technique des industries de l'armement ou de l'aéronautique, publication scientifique par exemple).

La chaîne éditoriale repose fondamentalement sur deux principes :

- Il est possible grâce à l'ordinateur de séparer le contenu de sa présentation, puisqu'un algorithme peut appliquer une mise en forme *a posteriori* à un contenu.
- Il est possible d'abstraire et de décrire la structure d'un document, afin de la représenter formellement (interprétation par une machine) et non plus en terme de mise en forme (interprétation par l'humain seulement).

Par exemple je peux écrire : « XML est fondé sur un langage de *balise* ». Le lecteur aura interprété que le mot balise était important dans ce texte.

Mais – disposant d'un ordinateur – je peux écrire à la place dans la machine : « XML est fondé sur un langage de `<important>balise</important>` »; et par ailleurs utiliser un algorithme qui dit « Si un mot est entre `<important>` et `</important>`, alors l'afficher en italique ». Le résultat graphique sera la même pour le lecteur final, mais par contre l'information stockée en machine sera plus riche dans ce second cas. En effet je sais que le mot « balise » est important, donc je peux décider de le mettre en italique pour une publication papier noir et blanc, mais également en gras et en rouge pour une publication Web en couleur, de le donner en priorité à un moteur d'indexation, etc. Il suffit d'appliquer des algorithmes différents à la *même* information.

C'est à dire que l'on devient capable d'*adapter* la présentation d'une information à un contexte de diffusion. L'extension de cet exemple simple ne portant que sur un seul mot à l'ensemble des structures du document (parties, titres, blocs de texte, images, glossaire, etc.) permet de reconfigurer globalement un document pour un usage.

Les bénéfices des chaînes éditoriales sont nombreux et ne seront pas tous repris ici. Mentionnons par exemple le polymorphisme et la réditorialisation (on écrit une seule fois, mais l'on exploite plusieurs fois) et d'une façon générale les gains de productivité liée à la production documentaire (passage du stade artisanal au stade industriel).

Mais du point de vue de l'accessibilité – au sens global préalablement défini – la chaîne éditoriale permet de repenser totalement la question : Tandis que l'approche « classique » consiste à produire une information de telle façon que sa forme soit la plus accessible possible, l'on va pouvoir dans le cas des chaînes éditoriales, en vertu de la séparation entre contenu et présentation, d'une part

produire une information et d'autre part penser sa ou ses formes accessibles.

Cela implique en partie qu'il n'y a une seule forme « génériquement » accessible, mais bien plusieurs formes pour plusieurs accessibilités : Une forme pour tel handicap, une autre pour des contraintes de débit, une troisième pour des contraintes d'écran, etc. Mais inutile que les contraintes des unes pèsent sur les autres. L'accessibilité et la mobilité peuvent alors être repensées en terme de multisupports – plusieurs supports pour plusieurs usages et plusieurs contraintes – et non plus en terme de support universel (utopique à obtenir).

Cela permet ensuite de penser l'accessibilité dans le temps : les contenus produits avec une chaîne éditoriale pourront respecter les normes d'accessibilité de demain ! Tandis qu'il n'est pas possible d'intégrer à une information écrite dans un langage technique (comme HTML) des contraintes qui n'existent pas encore, il sera possible d'appliquer plus tard aux contenus XML structurés sémantiquement des algorithmes qui intégreront ces contraintes lorsqu'elles existeront, *sans reprendre les contenus*.

Nous illustrerons ces concepts avec la chaîne éditoriale Opale, réalisée avec Scenari<sup>2</sup>.

## **Perspectives**

La publication numérique n'en est sans doute qu'à des débuts. Dans le domaine de la formation, les contenus devront être accessibles, interopérables, publiés très rapidement, modifiables indépendamment sur la forme et sur le fond, et bien entendu, rendus et maintenus accessibles. Autant dire que le choix de formats, de techniques, de processus et de méthodes compatibles avec ces exigences va s'avérer fondamental.

Le concept de chaînes éditoriales intègrent le *processus* de façon intrinsèque. Les autres outils peuvent s'adapter pour répondre aux mêmes exigences, en réintégrant les concepts de modèles documentaires et de processus de production :

- Ainsi, un CMS peut proposer des modèles documentaires et permettre aux utilisateurs d'assembler un ensemble de briques structurées. C'est par exemple le cas du CMS Ezpublish, qui pose par ailleurs d'autres problèmes du point de vue de la conservation des données.
- Des logiciels de traitement de texte ou des logiciels de production de pages Web peuvent également à travers des fonctions avancées permettre la production de contenus structurés. Ces solutions restent encore toutefois complexes à mettre en oeuvre, et, plus fondamentalement, elles remettent en cause la philosophie même du traitement de texte, qui se construit sur le paradigme de la page blanche, a priori peu compatible avec l'écriture structurée d'une part (on ne part plus de rien) et le Web d'autre part (le concept de page propre à l'imprimé disparaît).

Sur le plan des formats, ce sont les langages de description sémantique qui sont et qui seront de plus en plus la référence. Rien ne dit qu'une innovation technologique ne remettra pas en cause XML en tant que formalisme dans les années à venir. Mais dans son principe XML assure la pérennité (nous avons l'assurance de pouvoir interpréter des contenus XML indépendamment des évolutions technologiques futures, puisque le format est interprétable par l'humain et pas seulement pas la machine) ainsi que la calculabilité (puisque le format est interprétable, nous sommes également certains de pouvoir le transformer en autre chose si cet autre chose émerge). Ceci représente un changement de paradigme important pour la constitution des fonds documentaires, puisque ceux-ci s'inscrivent alors dans une perspective de très long terme (plusieurs dizaines d'années), rendant le contenu indépendant des évolutions des formats techniques de publication (que celles-ci proviennent d'évolution technique des possibilités informatiques, de nouveaux terminaux de restitution ou de nouvelles exigences d'accessibilité).

---

<sup>2</sup> <http://scenari-platform.org>

Du point de vue de l'accessibilité, les chantiers à venir concerneront au moins autant l'accessibilité des contenus produits (interfaces de diffusion) que l'accessibilité des interfaces de production (Authoring Tools Accessibility Guidelines<sup>3</sup>). Seule les solutions qui sauront combiner efficacement ces deux aspects pourront prétendre à devenir à terme des solutions industrielles de publication numérique réellement accessibles, dans un contexte où la production contributive se généralise et atténue les frontières entre auteurs et lecteurs.

L'accessibilité totale reste néanmoins une utopie, car il existe des limites, inhérentes aux compétences des contributeurs. Le fait d'ouvrir le monde de la publication numérique au grand public a un revers, car il n'est pas possible que toute personne amenée à publier des contenus sur le Web ait une compétence et une culture de l'accessibilité. Dans une logique éditoriale classique, l'auteur professionnel peut être formé à cette culture et le contexte éditorial peut lui imposer les contraintes requises. Dans une logique contributive « ouverte » où chacun est appelé à écrire, cela n'est plus possible. Dès lors, l'outil devient essentiel, pour aider, corriger, accompagner le créateur de contenus numériques à produire des contenus accessibles.

L'émergence d'outils faciles d'usage pour la publication Web (Wiki, CMS, blogues, etc.) a permis une réelle généralisation de la pratique de publication électronique. Maintenant que cet objectif qualitatif est atteint, ou en passe de l'être, un des enjeux majeurs à l'avenir est selon nous de redonner à l'éditorial ses lettres de noblesses, dans ce qu'il apporte de *qualitatif* au document. L'éditeur - dont la fonction disparaît avec le contributif ou l'auto-production – n'est pas seulement celui qui sait publier, il est aussi celui qui sait comment *bien* publier. Il est celui qui assure, par la mise en place de règles éditoriales, que le lecteur saura accéder au contenu publié.

Paradoxalement le risque de la généralisation d'outils de production « pour tous » est que les productions soient de moins en moins lisibles, au sens documentaire. Tout le monde pourra publier ce qu'il a à dire, mais peu pourront en réalité y accéder, au sens de trouver la bonne information, (celle dont j'ai besoin), et d'être capable de la comprendre, de la remettre dans son contexte (de la critiquer).

Des outils reprenant à leur compte les principes de publication automatisée – à l'instar de ceux instrumentés dans les chaînes éditoriales – pourront mobiliser des algorithmes rigoureux, intégrant les savoir-faire des métiers éditoriaux et permettant ainsi la combinaison du quantitatif et du qualitatif, de l'accessible à tous en écriture, mais aussi en lecture !

---

3, voir « Directives 2.0 sur l'accessibilité des outils d'éditions », ébauche candidate W3C, [http://www.culture.gouv.fr/cric/accessibilite/ATAG2\\_0-fr.html](http://www.culture.gouv.fr/cric/accessibilite/ATAG2_0-fr.html)