

# Table des matières

<b>Introduction</b>	<b>5</b>
<b>I - Principes : Un procédé de création de contenu</b>	<b>7</b>
<b>II - Caractéristiques</b>	<b>9</b>
<b>III - Description technique</b>	<b>11</b>
<b>IV - Enjeux de la chaîne éditoriale</b>	<b>13</b>
<b>V - Bénéfices de l'utilisation de la chaîne éditoriale dans un environnement pédagogique</b>	<b>17</b>
A. Rendre explicites des pratiques pédagogiques.....	17
B. Mutualiser des pratiques pédagogiques.....	18
C. Optimiser la gestion de la production.....	18
D. Faire des économies dans la production documentaire.....	19
<b>VI - Comparaison : exemples et contre-exemples de chaînes éditoriales</b>	<b>21</b>
A. Quelques applications qui ne sont pas des chaînes éditoriales.....	21
B. Exemples de chaînes éditoriales.....	22
<b>VII - Historique : De TeX à XML</b>	<b>25</b>
<b>VIII - Outils de chaînes éditoriales</b>	<b>27</b>
<b>Glossaire</b>	<b>29</b>
<b>Bibliographie</b>	<b>35</b>



# Introduction

Ce document a été élaboré par :

Prénom Nom	Etablissement	UNT
Damien Guillaume	Observatoire d'Astrophysique	UVED
Stéphane Crozat	UTC	UNIT
Laurent Rivet	Rennes 1	UVED
Manuel Majada	UTC	UNIT
Xavier Hennequin	UTC	UNIT

*Tableau 1 Contributeurs*

Prénom Nom	Etablissement	UNT
Franck Rouzé	Lille 1	UNISCIEL
Elodie Perez	INSA Lyon	UNIT
Jean-Marc Dubois	Bordeaux 2	UMVF
Luc Debert	Nancy 2	AUNEGE
Patrice Roturier	Rennes 2	UOH
Romuald Lorthioir	Rennes 1	UVED
Stéphane Poinart	UTC	UNIT

*Tableau 2 Participants à la relecture*



# Principes : Un procédé de création de contenu



## Définition

Une chaîne éditoriale est un procédé technologique et méthodologique consistant à réaliser un modèle de document, à assister dans les tâches de création du contenu et à automatiser la mise en forme [Scenari, la chaîne éditoriale libre]. Son atout premier est de réduire les coûts de production et de maintenance des contenus, et de mieux contrôler leur qualité.

## Séparation des formats de création et de publication

La séparation entre les formats de création et de publication, également appelée séparation fond/forme [Instrumentation numérique des documents : pour une séparation fonds/forme] consiste à utiliser un format informatique pour la production du contenu et un ou plusieurs autres pour sa publication. Dans le cadre d'une chaîne éditoriale numérique, le passage du format de création aux formats de publication est assuré automatiquement par un programme informatique.

L'intérêt de cette séparation est :

- de pouvoir utiliser des formats adaptés en fonction des objectifs opérationnels (créer, archiver, réutiliser, diffuser, etc.),
- de pouvoir disposer de plusieurs formats de publication à partir d'un seul format de création

## Format de création orienté métier

Le format de création d'une chaîne éditoriale s'appuie sur une description métier, c'est à dire qui véhicule une sémantique liée au domaine documentaire à l'intention de l'auteur, et non au traitement informatique à réaliser.

Ainsi l'on n'inscrira pas dans le format de création qu'un paragraphe est en rouge et encadré (ce qui est une indication de mise en forme liée à la publication), mais l'on indiquera que c'est une définition (ce qui est une information liée au contexte d'usage, compréhensible par un être humain connaissant le domaine). Un algorithme informatique permettra alors la transformation du langage métier vers des langages de traitement informatique (par exemple, cet algorithme transformera le paragraphe indiqué comme une définition en rouge).

L'intérêt de cette approche est :

- d'assurer la pérennité de l'information en rendant indépendant le format de création par rapport aux évolutions technologiques (le format devant rester

- humainement interprétable, le fond conservant toute son intégrité)
- de bénéficier d'un plus haut niveau de sémantique (il y a plus d'information dans « définition » que dans « rouge », car on peut appliquer la règle « définition=rouge », mais on ne peut pas déduire « rouge=définition », rouge pouvant renvoyer à une autre intention).

### Édition WYSIWYM

---

La plupart des systèmes d'édition numérique se fondent sur le paradigme du WYSIWYG<sup>⊖</sup>, « What you see is what you get », ou littéralement en français : « Ce que vous voyez est ce que vous obtenez ». Cette approche vise, en résumé, à permettre à l'utilisateur de créer un document tout en composant le rendu final, comme dans sur un support traditionnel non numérique.

 Le WYSIWYM oppose sa démarche à celle du WYSIWYG en partant non pas du résultat graphique, mais de l'information à véhiculer<sup>⊖</sup> [*WYSIWYM (what you see is what you MEAN)*], de sa signification et de l'intention auctoriale<sup>⊖</sup>.



La chaîne éditoriale se fonde, elle, sur le concept du WYSIWYM<sup>⊕</sup>, « What you see is what you mean », c'est à dire : « Ce que vous voyez est ce que vous voulez dire ».

 Le principe est pour un logiciel, de représenter les informations en fonction de leur sens<sup>⊖</sup> [*What you see is what you mean*], de l'information à véhiculer (par opposition à "représenter les informations sous leur forme finale, pour impression, avec mise en forme à l'identique...").



L'édition WYSIWYM est l'instrument technique de la séparation entre le format de création (sémantique et orienté métier) et le format de publication.

# Caractéristiques



## *Modèle*

---

Le modèle sert de guide d'écriture en proposant à l'auteur les éléments de contenu pertinents dans le contexte de sa rédaction. Il assure une qualité standard de structuration, quelques soient les compétences éditoriales des auteurs. Il permet une harmonisation des publications, dans le respect des chartes graphiques et des disparités d'auteurs.

## *Métier*

---

Le modèle de document est contextuel, il s'appuie sur le langage du domaine et non sur un langage technique de publication. Il est donc facilement appropriable par les acteurs. Cela permet également de configurer les interfaces aux justes besoins (plutôt que de disposer de fonctions génériques jamais toutes utilisées), d'archiver les contenus dans un langage non technique et donc d'assurer leur accessibilité dans le temps.

## *Maintenance*

---

La maintenance des contenus est assurée dans le temps par l'indépendance technologique des formats de création. Ils sont liés aux évolutions du métier au contraire des formats de publication dépendants des évolutions techniques. La maintenance des publications est facilitée par le principe de génération automatique qui permet d'intervenir systématiquement sur la forme de la totalité du fonds sans retoucher aux contenus. Ainsi, la quantité de travail pour maintenir la publication est indépendante du volume de contenu.

## *Multisupport*

---

Le multisupport permet de n'écrire qu'une seule fois ce qui est publié plusieurs fois. Il permet un gain de temps à la production du contenu, mais surtout il permet de ne maintenir qu'un seul fonds documentaire, sans redondance d'information.

## *Multiusage*

---

Le multiusage consiste à découper le contenu en unités documentaires autonomes et recombinaibles. Il permet de réutiliser ces unités documentaires sans recopie (par référence), afin d'adapter un discours à différents contextes.

## *Multimédia*

---

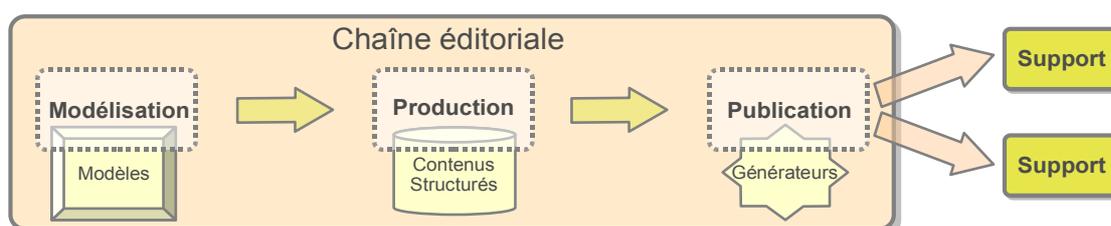
Le multimédia signifie la capacité à intégrer tout type de contenu (du texte à l'audiovisuel) et à structurer aussi bien les contenus textuels, mais aussi spatiaux (par exemple en définissant des zones dans des images) ou temporels (par exemple en définissant des segments dans un flux audio).



# Description technique



## Architecture générale



Graphique 1 Etapes de fonctionnement d'une Chaîne éditoriale

Une chaîne éditoriale se compose classiquement :

D'un ensemble de modèles formels (appelés « schémas documentaires ») permettant de contrôler automatiquement la structure des documents (exemple de technologie : DTD<sup>Ⓔ</sup> ou schema XML<sup>Ⓔ</sup>), les références aux ressources binaires (images, etc.) et les références entre documents (liens entre fichiers) ;

D'un ensemble d'interfaces d'édition<sup>Ⓔ</sup> WYSIWYM<sup>Ⓔ</sup> permettant de créer et modifier des documents dans un format conforme aux modèles, de référencer des ressources binaires depuis ces documents et de créer des liens entre ces documents.

D'un ensemble d'espaces permettant de stocker les fichiers (documents structurés et ressources binaires)

D'un ensemble de moteurs de publications<sup>Ⓔ</sup> permettant de transformer les documents structurés et ressources binaires pour rendre des publications lisibles (HTML<sup>Ⓔ</sup>, PDF, etc.)

## Implémentations

On peut distinguer trois grandes modalités pour implémenter une chaîne éditoriale :

- L'implémentation<sup>Ⓔ</sup> « ad hoc » consiste à mobiliser différentes briques logicielles et à programmer un logiciel spécifique pour un besoin particulier. Par exemple en couplant un éditeur du marché avec un outil de stockage et en réalisant des moteurs de publication sous la forme de feuilles XSL-XSLT<sup>Ⓔ</sup>.
- L'implémentation dédiée consiste à implémenter « en dur » un modèle de chaîne éditoriale pour un besoin assez transversal, tel que la publication de livres ou de modules de cours universitaires.
- L'implémentation générique consiste à réaliser un système paramétrable, indépendant d'un modèle en particulier, et à le configurer en fonction des besoins. Il s'agit alors de ce que l'on peut appeler un Système de Gestion de Chaînes Éditoriale par analogie aux Systèmes de Gestion de Bases de Données.



# Enjeux de la chaîne éditoriale

## IV

### *Polymorphisme et ré-éditorialisation/ré-utilisation*

---

Le polymorphisme consiste en la possibilité technique de disposer d'une source unique (single sourcing en anglais) ☞ [Single Source Publishing] de contenu et de la transformer à volonté selon les supports et mises en formes désirés. Le polymorphisme est un principe technologique qui reste limité dans la pratique : en effet il est rare que l'on souhaite présenter exactement la même information sous deux supports différents pour deux usages différents. Une nouvelle publication implique généralement la sélection du contenu (telle partie en plus, telle partie en moins), sa réorganisation (telle partie avant telle autre), sa remise en contexte (introduction, conclusions, transitions), etc.

L'idée est alors de profiter du découpage logique du contenu formalisé selon un langage XML métier, pour appliquer des césures physiques (découpage de fichiers XML ☞ et utilisation de liens par référence). Il devient alors possible de partager des fragments documentaires entre plusieurs documents, ce qui permet la réutilisation sans recopie.

On appelle ré-éditorialisation (le terme anglais de repurposing étant encore plus adéquat) la remise en contexte de fragments issus d'un fonds documentaire, par leur ré-agencement au sein d'un nouveau document, leur augmentation par une création de contenus spécifiques et leur publication sur un nouveau support et/ou pour un nouveau public..

### *Accessibilité*

---

Le principe d'accessibilité consiste à pouvoir « mettre le Web et ses services à la disposition de tous les individus, quel que soit leur matériel ou logiciel, leur infrastructure réseau, leur langue maternelle, leur culture, leur localisation géographique, ou leurs aptitudes physiques ou mentales » (Tim Berners-Lee, directeur du W3C, cité dans JournalDuNet ☞ [L'accessibilité Web]).

La chaîne éditoriale permet de repenser totalement la question de l'accessibilité : 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 d'une part produire une information et d'autre part calculer sa ou ses formes accessibles. Cela a comme conséquence qu'il n'y a plus une seule forme « génériquement » accessible, mais 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. Il devient inutile que les contraintes pèsent les unes sur les autres. L'accessibilité est repensée en terme de polymorphisme et non plus en terme de support unique et universel, ce qui rend le défi de l'accessibilité beaucoup moins utopique ☞ [Chaîne éditoriale, accessibilité, mobilité].

## *Archivage et pérennisation*

---

L'archivage pérenne des fonds documentaires numériques repose sur trois piliers : un stockage durable qui assure que l'on pourra lire plus tard ce que l'on stocke aujourd'hui ; des formats malléables qui permettent d'utiliser les documents existants dans les usages présents ; des modalités d'indexation<sup>26</sup> qui assurent que l'on peut accéder à ce qui existe en fonction de ce que l'on cherche. Donc il faut pouvoir stocker, chercher, utiliser.

La chaîne éditoriale répond précisément à ces trois problématiques. Le stockage est durable car les formats de création ne sont pas soumis aux évolutions technologiques. Les contenus XML décrits par des langages métier se caractérisent par une indexation native (les balises<sup>26</sup> de description sont déjà là). Les contenus sont utilisables car fragmentables et transformables selon les besoins.

Par exemple, la modification de la charte graphique d'un organisme entraîne une modification de forme de l'ensemble des documents valides de cet organisme. Si cet organisme utilise une chaîne éditoriale, celle-ci lui permettra de créer les nouveaux supports de publication répondant aux nouvelles exigences de la charte graphique et permettre ainsi la publication de l'ensemble des documents valides sans toucher au fonds documentaire stocké.

## *Réduction des coûts et démarche qualité*

---

L'objectif d'une chaîne éditoriale numérique est d'instrumenter l'industrialisation d'une production documentaire. On retiendra dans le concept d'industrialisation les notions de :

Massification : être capable de produire de grands volumes (plusieurs milliers de pages), malgré la rareté de compétences techniques (comme la capacité à mettre correctement en forme un document selon les canons du domaine).

Économie d'échelle : être capable de réduire les coûts de production et de maintenance

Démarche qualité : être capable d'assurer a priori un niveau de qualité requis (homogénéité, respect de règles éditoriales, graphiques, métiers, accessibilité, etc.)

On peut faire le parallèle de ce point de vue entre une chaîne éditoriale et une chaîne de production de produits manufacturés : l'objectif est de rationaliser pour massifier.

## *Les limites du paradigme de la bureautique*

---

La bureautique<sup>26</sup> a permis la démocratisation de l'usage du numérique pour les pratiques documentaires, en rendant accessibles les outils (traitement de texte, tableurs, base de données, outils de mise en forme, d'édition, logiciel de présentation...) à tous. Mais dans son instrumentation, elle s'est majoritairement limitée à calquer les pratiques antérieures au numérique (comme la machine à écrire), en les améliorant, mais sans les repenser. Les raisons étant essentiellement opérationnelles, cette approche promettant de toucher le plus grand nombre le plus rapidement.

L'approche bureautique n'apporte pas une satisfaction universelle pour la production de document en masse [...]. C'est ce constat qui est un des principaux facteurs de la motivation pour un public exigeant de chercher une alternative à la bureautique dans leurs cas d'usage (Wikipédia, 2008c).

La chaîne éditoriale propose de changer le paradigme fondateur, et plutôt que de copier les pratiques antérieures, elle propose une approche originale en symbiose avec les principes du numérique. Après l'objectif quantitatif atteint par la bureautique, l'enjeu est de remettre en avant des considérations qualitatives qui commencent à faire défaut dans les usages.

### *La séparation des métiers*

---

Traditionnellement la production documentaire fait appel à plusieurs métiers (auteur, rédacteur, correcteur, éditeur, diffuseur, etc.). L'outil informatique, en facilitant certaines tâches (correcteurs orthographiques, outils simplifiés de mise en page, etc.), a tendu à fusionner tous les métiers en un seul « auteur-rédacteur-éditeur ». Mais, au delà de l'aspect technique, ces métiers sous-tendent des compétences qui font en général défaut à l'auteur (savoir écrire n'est pas savoir éditer typiquement), et la conséquence en est une dégradation importante des publications réalisées par les éditeurs « amateurs » que nous sommes (presque) tous.

L'objectif poursuivi par la chaîne éditoriale est de réintroduire ces métiers, en réorganisant une chaîne de production où chaque compétence est mise à profit pour ce qu'elle est.

Une chaîne éditoriale numérique [...] est un outil ou une suite d'outils permettant d'accompagner un processus éditorial depuis l'écriture jusqu'à la publication finale. À l'inverse de la bureautique (qui fusionne l'ensemble des étapes du processus), la chaîne éditoriale les maintient séparées dans l'objectif d'offrir l'environnement le plus adéquat pour chaque type de tâche (Crozat, 2007).

Ainsi la chaîne éditoriale vise d'une part à rompre avec les techniques traditionnelles de production prolongées par la bureautique, et d'autre part à réhabiliter les processus professionnels de production éditoriale, mis à mal par la bureautique, permettant ainsi à chaque acteur du processus de se réappropriier son domaine et ses compétences.



# Bénéfices de l'utilisation de la chaîne éditoriale dans un environnement pédagogique

V

Cette partie aborde les bénéfices de l'usage des principes de la chaîne éditoriale dans le cadre particulier de la production des ressources pédagogiques numériques.

## A. Rendre explicites des pratiques pédagogiques

Pour des enseignants, le choix de l'utilisation d'une CE repose sur la volonté de produire un document porteur de la stratégie pédagogique et des intentionnalités pédagogiques qu'ils auront défini, plus encore s'il s'agit de produire un document qui puisse être utilisé en dehors d'une relation de type face à face.

Cela se traduit :

- en éléments de cohérence : plan, structure, grain pédagogique.
- en éléments d'appui pédagogique qui facilitent l'assimilation : informations sur la nature de l'apprentissage, balises sémantiques de mise en valeur de contenu. Par exemple les balises suivantes : important, remarque, exemple, définition, etc.
- en éléments qui orientent l'action de l'apprenant avec une volonté de provoquer une action / réaction pédagogique : à apprendre par cœur, à lire, à survoler, questions auxquelles il faut répondre, etc. et éléments de gestion du rythme de l'apprentissage : alternance cours exercice, séquençement. indications temporelles, etc..

La chaîne éditoriale apporte l'ensemble de ces éléments au travers de la modélisation documentaire. C'est dans la confrontation entre la pratique de l'enseignant, le choix du modèle pédagogique documentaire et l'exercice de sa rédaction dans ce modèle que s'élaborent la formalisation et la structuration pédagogique du document.

En soi, le document ne sera toujours que l'un des éléments d'un dispositif pédagogique. Il ne peut être autonome et doit donc s'inscrire dans un cadre

pédagogique plus large. Même dans le cadre de l'auto formation c'est l'apprenant qui élabore et se contraint à ce cadre pédagogique. Sans cet effort, en général il abandonne.

## B. Mutualiser des pratiques pédagogiques

Dans la pratique habituelle, le document est un auxiliaire pédagogique auquel l'enseignant attribue les fonctions qui lui sont strictement nécessaires. Il est conçu pour un besoin donné dans un contexte précis : support mnémotechnique, résumé du cours, cours, éléments d'approfondissement, exercices, etc.. Dans ce cadre, le document est indissociable de l'activité de l'enseignant et du contexte pédagogique de celle-ci, cela se traduit par une formalisation légère et un complément oral. Ce document n'est donc pas autonome et est uniquement destiné à l'usage individuel de l'auteur.

Quand on veut passer à un usage mutualisé (entre enseignants) de ce document on est confronté à un besoin de formalisation :

- de la nature de ce document : support mnémotechnique, cours, etc.
- du contexte pédagogique dans lequel il est inclus : formation initiale/formation continue, cours magistraux/travaux pratiques/ travaux dirigés, etc.
- d'une explicitation forte des intentionnalités pédagogiques de l'auteur. : structure, nature et rythme de l'apprentissage.

D'un point de vue pédagogique la CE, au travers du modèle documentaire permet de passer d'un document individuel à un document mutualisé par la formalisation qu'il induit.

Par ailleurs, cette formalisation apporte aussi un apport réflexif sur la propre pratique pédagogique de l'auteur qui peut enrichir le document lui-même.

## C. Optimiser la gestion de la production

### *Création et gestion*

---

Le concept de chaîne éditoriale est porteur d'une notion de rationalisation de la production et d'instrumentation technologique de celle-ci. Elle se situe donc dans un contexte de création mais également de gestion.

En effet, la chaîne éditoriale a pour ambition de produire de la documentation multisupports et multiusages en grande quantité, en facilitant la gestion de cette production, conception et maintenance, et en gardant des exigences hautes en terme de qualité pédagogique et multimedia.

Toutefois, la chaîne éditoriale n'oblige pas à cette approche de rationalisation.

### *Choc des cultures*

---

La culture professionnelle des informaticiens multimedia, même dans le cas d'une extrême rationalisation (le CDrom) reste la culture de l'objet unique, « sur mesure ». Elle rentre en conflit avec la standardisation qu'induit la chaîne éditoriale. Elle n'induit pas les mêmes compétences : certaines tâches sont automatisées entraînant une élévation du niveau de compétence technique des concepteurs, et une baisse du niveau technique de certains opérateurs. Ces modifications remettent en cause les positions individuelles des acteurs de la production documentaire multimedia.

En ce qui concerne la méthodologie de production, c'est à dire l'organisation du travail, c'est l'objectif que la structure entend atteindre qui est déterminant, l'outil ne fait qu'accompagner cette organisation.

## D. Faire des économies dans la production documentaire

Si la chaîne éditoriale a été modélisée en accord avec les besoins, les gains économiques se traduisent avant tout par des gains de productivité :

- **Productivité directe – nombre de documents produits / temps de travail.** En effet, la chaîne éditoriale permet de :
  - diminuer le temps d'intégration technique permettant le multisupport.
  - diminuer le temps de rédaction auteur en guidant celle-ci.
- **Productivité de maintenance – nombre de documents maintenus / temps de maintenance.** Le temps de maintenance sur chaque document est diminué par :
  - l'unicité de la source documentaire par rapport à ses applications multisupport.
  - la réutilisabilité du contenu numérisé : dans différents contextes pédagogiques et/ou dans différents contextes technologiques.
  - Des formats sources choisis pour leur facilité de gestion et pour leur caractère pérenne dans le temps (exemple : le XML<sup>(6)</sup>).
- **Productivité indirecte – standardisation : gain de gestion, gain qualitatif, gain en conception.**



# Comparaison : exemples et contre-exemples de chaînes éditoriales

## VI

Quelques applications qui ne sont pas des chaînes éditoriales 21

Exemples de chaînes éditoriales 22

En fonction du contexte d'utilisation, un même produit peut être considéré comme une chaîne éditoriale ou pas. Les quelques exemples qui suivent permettent de faire cette différenciation.

Le principal critère est la possibilité de transformation automatique de documents, basée sur leur balisage sémantique, les autres caractéristiques d'une chaîne éditoriale en découlant logiquement.

### A. Quelques applications qui ne sont pas des chaînes éditoriales

#### *Traitements de texte WYSIWYG*

Les traitements de texte WYSIWYG<sup>☹</sup> comme Microsoft Word ou OpenOffice Writer, même s'ils reposent sur XML comme support de stockage et permettent différents types d'export, ne peuvent en général pas être considérés comme des chaînes éditoriales. En effet, le sens et l'intention ne sont pas exprimés explicitement par l'auteur lors de la création d'un document. L'auteur profite de l'interface<sup>☹</sup> WYSIWYG pour obtenir l'affichage qu'il souhaite (un affichage unique), et même l'utilisation de styles ne permet pas ensuite d'analyser automatiquement le sens du contenu du document.

#### *Éditeurs HTML*

HTML, une application de SGML<sup>☹</sup>, était conçu à l'origine pour mettre en avant le sens dans les documents. Les différents navigateurs devaient ensuite afficher les documents de diverses façons, en fonction du contexte. L'apparition de CSS<sup>☹</sup> était

aussi un élément sensé permettre la séparation entre le fond et la forme. Ces idées n'ont pas été suivies par la pratique, et la grande majorité des pages web aujourd'hui sont conçues comme on concevrait des documents avec un traitement de texte WYSIWYG. Les auteurs s'attendent à ce que les documents soient affichés comme ils les ont créés, sans aucune transformation automatique ultérieure. L'évolution du web est à tel point allée dans cette direction, malgré les efforts des concepteurs d'HTML, que les pages web sont souvent aujourd'hui créées avec des éditeurs WYSIWYG comme Dreamweaver, et que le format HTML est considéré vis à vis des chaînes éditoriales comme un format de publication au même titre qu'une impression papier.

La réutilisation sémantique automatique de documents créés de façon traditionnelle avec un éditeur comme Dreamweaver étant impossible, on ne peut pas considérer cet éditeur comme une chaîne éditoriale.

### *LaTeX "standard"*

---

Lorsqu'on ne se limite pas à un modèle de document avec LaTeX<sup>☞</sup>, l'édition de documents s'oriente vers l'affichage (en général l'impression papier) : LaTeX permet de contrôler précisément l'affichage, jusqu'aux positions de coupures dans les mots. La syntaxe des équations ne code pas le sens mathématique, mais la façon dont les équations doivent s'afficher. Il est possible de définir quelques environnements et commandes dans un document, mais ces fonctionnalités sont souvent utilisées comme des raccourcis ou une façon pratique d'uniformiser l'affichage de certaines parties du texte plutôt qu'un moyen de structurer et donner du sens au document.

## **B. Exemples de chaînes éditoriales**

### *Exemple minimaliste : ensemble d'applications avec un éditeur de texte, un validateur XML et un processeur XSLT*

---

Une chaîne éditoriale ne possède pas nécessairement des fonctionnalités graphiques avancées et un environnement d'édition convivial. A la limite, on peut considérer qu'un simple éditeur de texte couplé à un outil de validation XML et un processeur XSLT<sup>☞</sup> puisse être une chaîne éditoriale : un schéma XML dédié à un métier peut être créé avec, il est possible de créer des documents valides par rapport à ce modèle et de réaliser divers exports avec des feuilles de style XSLT. Dans ce cas, l'édition se révèle souvent pénible et lente pour la plupart des utilisateurs.

### *Certains environnements d'édition WYSIWYM basés sur LaTeX*

---

Si l'utilisation sans contrainte de LaTeX<sup>☞</sup> ne permet pas de profiter des avantages d'une chaîne éditoriale, LaTeX et les outils associés comme le compilateur n'en ont pas moins la capacité technique d'une chaîne éditoriale. L'utilisation d'un modèle de document, par exemple sous la forme d'un fichier de style définissant un balisage sémantique, associée au compilateur pour la validation syntaxique, correspond bien à une chaîne éditoriale. Les documents générés peuvent ainsi être transformés automatiquement dans des langages XML, et aboutir à différentes formes de publication. Un exemple de chaîne éditoriale basée sur LaTeX est Polytex.

### *Jaxe*

---

Jaxe est un éditeur XML qui valide les documents en temps réel et permet d'effectuer des transformations XSLT sur les documents. De part ses fonctionnalités graphiques, il pourrait parfois être utilisé comme un simple éditeur WYSIWYG<sup>☞</sup>,

mais son approche résolument WYSIWYM de l'édition pousse à l'utilisation sous forme de chaîne éditoriale. Quand on l'associe à une configuration particulière (schéma XML orienté contenu, fichier de configuration Jaxe et feuilles de style XSLT), il présente toutes les caractéristiques d'une chaîne éditoriale dédiée à un métier.

Jaxe est un outil libre et open source.

### *Scenari et OpaleSup*

---

Scenari est une suite logicielle libre de conception de chaînes éditoriales numériques.

Opalesup est une chaîne éditoriale, libre et open source, développé avec le système SCENARI pour la production de ressources pédagogiques numériques principalement dans l'enseignement supérieur.

Opalesup permet l'édition WYSIWYM de documents XML, le stockage des fichiers informatiques correspondants et la publication multisupport : site Web pour diffusion Internet, respectant SCORM pour une diffusion dans des plate formes de formation, document odt imprimable, et exports vers différents schémas pivot : UNIT, UVED, inter UNT.

### *ChainEdit*

---

Cette plateforme, développée pour une intégration dans les espaces numériques de travail, est une chaîne éditoriale permettant une édition WYSIWYM de documents, stockés au format XML, complétés par différents médias (images, sons, animations, vidéos, ...) intégrés dans la ressource produite. Elle permet la génération dans plusieurs formats cibles (Html, ODF, Pdf, Schéma pivot inter-UNT, ..) en fonction des paramétrages effectués dans l'application. Elle permet également l'import de projets qui peuvent être transformés lors de cette opération (importer un module au format inter-UNT vers un module au format UVED par exemple).

Elle nécessite l'utilisation d'un navigateur internet standard (IE/FireFox) et est multi-plateforme. Bien que destinée à être installée sur un serveur pour une gestion centralisée et permettre la saisie à plusieurs d'un document, elle peut être installée de façon autonome sur un poste client. C'est un outil libre, gratuit et OpenSource.



# Historique : De TeX à XML

## VII

### *Les grandes étapes*

---

Le concept de chaîne éditoriale trouve son origine dans l'édition et dans la presse : il consiste à organiser les tâches de production et de publication, en séparant les métiers intervenant dans le processus. A l'origine une chaîne éditoriale est donc un processus avant d'être un outillage, qui trouve son origine dans un besoin d'industrialisation.

Historiquement, les deux technologies ayant permis l'implémentation de chaînes éditoriales numériques sont LaTeX (créé en 1982 sur la base du langage d'édition TeX, lui-même créé en 1978) et SGML (norme ISO depuis 1986, issue des travaux d'IBM initiés en 1979 avec GML). XML (standard W3C depuis 1998) est aujourd'hui la technologie de référence pour la réalisation de chaînes éditoriales. C'est sa maturité qui a permis de sortir ce procédé des domaines auxquels il était confiné pour en élargir les usages.

Les chaînes éditoriales existent donc depuis le début des années 80, mais sont restées confinées, jusqu'à la fin des années 90 à usage de niche (documentation technique stratégique dans l'aviation avec SGML, publication scientifique avec LaTeX, etc.). On peut citer Arbortext Epic (1982), Adobe Framemaker (1986), ou Grif (1983). Les premières démocratisations de l'approche se manifestent sur la base d'applications LaTeX avec LyX [LyX - The Document Processor, LyX], commencé en 1995 (sous le nom de Lyrix) ou PolyTeX [PolyTeX : un environnement pour l'édition structurée de photocopiés électroniques multisupports], également débuté en 1995. En 1999, l'Université de Technologie de Compiègne réalise Scenari, premier environnement de conception de chaînes éditoriales XML [Le procédé SCENARI : Une chaîne éditoriale pour la production de supports numériques de formation] proposant un langage déclaratif de modélisation et de publication de haut niveau. D'autres outils émergent ensuite, tels que ChainEdit, Jaxe, La Poule ou l'Oeuf, Publimap, etc. s'appuyant notamment sur des standards tels que DocBook ou DITA et des méthodes comme Information Mapping.

Bruno Bachimont peut être considéré comme un des pères fondateurs de la démocratisation de la chaîne éditoriale numérique. Il en est à l'origine en réalisant l'un des premiers prototypes en 1995 (PolyTeX) et en introduisant à cette occasion pour la première fois le terme de « chaîne éditoriale » dans le sens d'environnement de production de documents numériques. Il en est également le théoricien le plus avancé, avec son ouvrage fondateur sur l'ingénierie documentaire qui pose le concept de calculabilité [Ingénierie des connaissances et des contenus] si particulièrement exploité dans la chaîne éditoriale.

### *Fondement théorique*

---

La raison computationnelle est un concept posé par Bachimont pour décrire le principe de calculabilité propre au numérique : tout contenu numérique est le produit d'une manipulation par l'ordinateur, c'est à dire d'un calcul (notons que le raisonnement s'inscrit dans la continuité de « la raison graphique » de Jack Goody,

qui montrait en quoi le passage de l'oral à l'écrit était constitutif de modes de pensée et pas seulement de modes de transmission).

L'ordinateur est utilisé pour sa fonction de numérisation<sup>②</sup>, c'est à dire sa capacité à enregistrer de l'information sous forme de nombres. A chaque information du monde « physique » (lettre d'un mot, pixel d'une image, etc.) est associé un nombre dans la machine, et à la demande ce nombre peut être restitué sous forme d'information lisible par l'homme (sur un écran, une imprimante, etc.).

L'ordinateur calcule un nombre à partir d'une information signifiante pour l'homme et calcule une information signifiante à partir d'un nombre. Si l'information entrée est exactement l'information restituée, alors le calcul est uniquement utilisé comme vecteur pour passer du monde sémiotique au monde numérique, et l'ordinateur n'est qu'une photocopieuse améliorée.

Mais l'ordinateur peut également être utilisé pour sa fonction d'informatisation, c'est à dire sa capacité à transformer les informations. La séquence n'est alors plus seulement :

« information1 » => numérisation => « information1 », mais

« information1 » => numérisation => transformation => « information2 ».

Ce que l'on obtient en sortie de processus est une transformation par le calcul de ce que l'on avait en entrée.

Par analogie une base de données relationnelle n'est pas juste un enregistrement de données que l'on peut lire séquentiellement, c'est également une représentation numérique qui promet des capacités de manipulation puissante (synthèses, recherches, etc.). L'enjeu est le même pour le document, au delà de l'enregistrement, la manipulation.

Les chaînes éditoriales s'inscrivent dans ce paradigme<sup>③</sup> de la transformation de l'information, au delà de sa numérisation.

# Outils de chaînes éditoriales

## VIII

Arbortex [<http://www.ptc.com/products/arbortext>]  
ChainEdit [<http://www.chainedit.fr>]  
FrameMaker [<http://www.adobe.com/fr/products/frame maker>]  
Grif [<http://fr.wikipedia.org/wiki/Grif>]  
Jaxe [<http://jaxe.sourceforge.net>]  
La Poule ou l'Oeuf [<http://www.lescomplexes.com>]  
Lyx [<http://www.lyx.org>].  
PolyTeX [<http://www.lmac.utc.fr/polytex>]  
Publimap [<http://www.takoma.fr>]  
Scenari [<http://scenari-platform.org>]



# Glossaire

## Algorithme

Ensemble d'opération permettant d'effectuer une tâche précise.

Succession de tests, décisions et actions dans le but de décrire le comportement d'une entité (objet, programme, personne).

Souvent appliqué au domaine informatique, l'algorithme n'en est pas moins utilisable dans d'autres domaines. Exemple : "Si la boulangère n'a pas livré le pain, aller acheter une baguette à la boulangerie d'en face et ensuite faire la vaisselle" est un algorithme décrivant le comportement que doit adopter la personne à qui est destiné le mot. Un algorithme n'est pas forcément infallible, il se peut que certains cas ne soit pas prévus. Dans notre exemple, que doit faire la personne si la "boulangerie d'en face" est fermée?

Dans le domaine informatique, l'algorithme est indépendant du langage de programmation, il peut s'énoncer en langage clair, il existe également une symbolique couramment admise pour la représentation graphique de l'algorithme (losange pour un test, rectangle pour une action).

Tout logiciel peut être décrit par un algorithme (très complexe). En pratique on utilise uniquement l'algorithme pour décrire une partie sensible du logiciel, souvent le coeur du système ou d'une fonctionnalité.

Les algorithmes les plus classiques sont les algorithmes de tri qui permettent d'ordonner des éléments dans un ordre croissant ou décroissant. Chaque algorithme de tri réalise la même opération mais de manière différente, chacun possède ses avantages et ses inconvénients (en termes de complexité de mise en oeuvre, de mémoire utilisée, et de rapidité du classement).

La liste des algorithmes classiques est infinie est en constante évolution, citons encore quelques exemples :

- les algorithmes de moteur de représentation en 3 dimensions utilisés notamment dans les jeux vidéo,
- les algorithmes génétiques,
- les algorithmes de résolutions d'équations,
- les algorithmes de cryptage,
- les algorithmes d'indexation et de recherche utilisés dans les bases de données,
- les algorithmes d'intelligence artificielle

(dico du net, <http://dicodunet.com>)

## Auctorial

De l'auteur.

## Balise, balises sémantique

Unités sémantiques délimitant chacune un ensemble à l'intérieur d'un fichier texte. (wikipedia)

## Bureautique - Outil de bureautique

Traitement de texte, tableurs, base de données, outil de présentation, logiciel de mise en page d'édition... (Suite OpenOffice.org, Suite Bureautique Microsoft...)

## Computational, raison computationnelle

Les supports numériques de connaissance font l'objet d'une ingénierie systématique permettant d'articuler finement conception des supports et production du sens. Le numérique, en proposant un système technique homogène, engendre des conséquences cognitives que l'on estampille du label de « raison computationnelle », à l'instar de la raison graphique [Goody, 1979] qui qualifie la rationalité propre à l'écriture.

## CSS

Le langage informatique CSS (Cascading Style Sheets : feuilles de style en cascade) sert à décrire la présentation des documents HTML et XML. Les standards définissant CSS sont publiés par le World Wide Web Consortium (W3C). Introduit au milieu des années 1990, CSS devient couramment utilisé dans la conception de sites web et bien pris en charge par les navigateurs web dans les années 2000.

(Wikipédia)

## DTD

La Document Type Definition (DTD), ou Définition de Type de Document, est un document permettant de décrire un modèle de document SGML ou XML.

Une DTD indique les noms des éléments pouvant apparaître et leur contenu, c'est-à-dire les sous-éléments et les attributs. En dehors des attributs, le contenu est spécifié en indiquant le nom, l'ordre et le nombre d'occurrences autorisées des sous-éléments. L'ensemble constitue la définition des hiérarchies valides d'éléments et de texte. En revanche, les DTD ne permettent pas de poser des contraintes sur la valeur du texte comme «le contenu de l'élément X est un entier en décimal», ou encore «dans l'élément Y, toutes les séquences de blancs sont équivalentes à un seul espace».

Définir ce qui est valide est aussi le rôle des « schémas » comme Schéma XML, Relax NG et Schematron mais ceux-ci sont préférentiellement exprimés en syntaxe XML alors que les DTD ont une syntaxe spécifique.

Une DTD peut être divisée en 2 parties: le sous-ensemble « interne », placé dans le document SGML ou XML lui-même, est une modification, une extension ou un paramétrage s'appliquant au sous-ensemble « externe ». Le sous-ensemble interne est optionnel.

Une DTD peut aussi définir des entités générales. Celles-ci ont l'un des rôles suivants :

- une référence à un fragment de document externe, typiquement un autre fichier.
- une abréviation pour un fragment de texte répétitif. Pour cette utilisation, la définition est plutôt dans le sous-ensemble interne.
- un synonyme de caractère permettant des références par nom plutôt que par un code numérique.

Note : Les entités générales ne sont pas définissables dans la syntaxe XML Schema du W3C mais cela peut être contourné.

(wikipédia)

## Environnement de conception

Interface permettant le développement efficace d'un type de projet (d'une chaîne

éditoriale par exemple).

### Grain pédagogique

Le grain pédagogique est le plus petit élément indivisible d'un parcours de formation. Le grain désigne une notion à appréhender, une entité possédant une unité de sens dont le contenu, la présentation et le traitement peuvent être séparés (texte, image, lien, animation, vidéo, simulation, question, etc). Articulés avec d'autres grains pédagogiques, ils contribuent à la construction d'un parcours individualisé de formation.

L'intérêt de granulariser la formation est donc de pouvoir répondre à tous les profils d'apprenants, en permettant de composer des parcours de formation individualisés originaux. En représentant les notions (concepts) comme des grains de connaissance, cela permet d'indexer les documents les concernant en laissant une grande souplesse de rédaction, de présentation et de représentation et donne la possibilité de réutiliser des documents existants.

(Wiki Université Paris Descartes Paris 5)

### HTML

L'Hypertext Markup Language, généralement abrégé HTML, est le format de données conçu pour représenter les pages web. Il permet notamment d'implanter de l'hypertexte dans le contenu des pages et repose sur un langage de balisage, d'où son nom. HTML permet aussi de structurer sémantiquement et de mettre en forme le contenu des pages, d'inclure des ressources multimédias dont des images, des formulaires de saisie, et des éléments programmables tels que des applets. Il permet de créer des documents interopérables avec des équipements très variés de manière conforme aux exigences de l'accessibilité du web. Il est souvent utilisé conjointement avec des langages de programmation (JavaScript) et des formats de présentation (feuilles de style en cascade). HTML est initialement dérivé du Standard Generalized Markup Language (SGML).

(Wikipédia)

### Implémentation - Implémenter

Implémenter : mettre en œuvre une analyse informatique (programmer) ou un programme informatique (paramétrer).

Implémentation : action d'implémenter; son résultat.

<http://dictionnaire.sensagent.com/impl%C3%A9menter/fr-fr/>

### Indexation

L'indexation, en sciences de l'information, consiste à identifier, pour un document, plusieurs éléments significatifs (tels que le titre de l'ouvrage, le nom de l'auteur, la date de publication et les sujets traités) afin de pouvoir retrouver ce document dans une base de données.

Plus généralement, l'indexation consiste à décrire une ressource par une liste structurée ou non de mots-clés en vue de la représenter, de l'identifier, de la localiser et de faciliter son accès ultérieur via une requête basée sur ces mots-clés. La liste des mots-clés est appelée "index de la ressource".

Indexer une ressource permet de la retrouver ultérieurement, et ce, afin :

- de la consulter
- ou de la réutiliser dans un contexte différent.

(Wiki Paris Descartes)

### Interface, Interface Homme-Machine

Ensemble des programmes gérant l'utilisateur et les rapports qu'ils peut entretenir avec sa machine et les logiciels qu'elle abrite.

<http://www.linux-france.org/prj/jargonf/>

## ISO

Organisation internationale de normalisation.

## Langage métier

Langage dédié généralement orienté vers une application professionnelle précise.

## LaTeX

LaTeX est un système logiciel de composition de documents créé par Leslie Lamport, ou plus exactement : une collection de macro-commandes destinées à faciliter l'utilisation du « processeur de texte » TeX. Le nom est l'abréviation de Lamport TeX.

Du fait de sa relative simplicité, il est devenu la méthode privilégiée d'écriture de documents scientifiques employant TeX ; la version actuelle est LaTeX2 $\epsilon$  (LaTeX 2 epsilon). Il est particulièrement utilisé dans les domaines techniques et scientifiques pour la production de documents de taille moyenne ou importante (thèse ou livre). Néanmoins, il peut aussi être employé pour générer des documents de types variés (par exemple des lettres ou des transparents).

(Wikipédia)

## Moteur de publications

Outil de transformation d'un contenu en un document publié sous un format destiné à la lecture / à la consultation.

## Numérisation

La numérisation est le procédé permettant la construction d'une représentation discrète (Le terme discret vient des mathématiques.) d'un objet du monde réel.

Dans son sens le plus répandu, la numérisation est la conversion d'un signal (vidéo, image, audio, caractère d'imprimerie, impulsion, etc.) en une suite de nombres permettant de représenter cet objet en informatique ou en électronique numérique[1]. On utilise parfois le terme franglais digitalisation (digit signifiant chiffre en anglais).

(Wikipédia)

## Paradigme

Un paradigme est une représentation du monde, une manière de voir les choses, un modèle cohérent de vision du monde qui repose sur une base définie (matrice disciplinaire, modèle théorique ou courant de pensée). C'est en quelque sorte un rail de la pensée dont les lois ne doivent pas être confondues avec un autre paradigme.

En informatique : un paradigme de programmation est un style fondamental de programmation informatique qui traite de la manière dont les solutions aux problèmes doivent être formulées dans un langage de programmation (à comparer à la méthodologie, qui est une manière de résoudre des problèmes spécifiques de génie logiciel).

(wikipedia)

## SGML

Standard Generalized Markup Language (langage normalisé de balisage généralisé -

SGML) est un langage de description à balises, de norme ISO (ISO 8879:1986).

### WYSIWYG

What you see is what you get, signifiant littéralement en français « ce que vous voyez est ce que vous obtenez » ou de façon plus concise « tel affichage, tel résultat ». L'acronyme est couramment utilisé en informatique pour désigner les interfaces utilisateur graphiques permettant de composer visuellement le résultat voulu, typiquement pour un logiciel de mise en page, un traitement de texte ou d'image. L'acronyme désigne, de fait, une interface « intuitive » : l'utilisateur voit directement à l'écran à quoi ressemblera le résultat final (imprimé).

(wikipedia)

### WYSIWYM

WYSIWYM, acronyme utilisé en informatique qui vient de l'anglais « What You See Is What You Mean », signifiant « ce que vous voyez est ce que vous voulez dire », par opposition au WYSIWYG : What You See Is What You Get.

Le principe est pour un logiciel, de représenter les informations en fonction de leur sens, de l'information à véhiculer (par opposition à « représenter les informations sous leur forme finale, pour impression, avec mise en forme à l'identique... »). Un des objectifs des interfaces WYSIWYM est de permettre une meilleure séparation du fond et de la forme lors de la création de documents.

(wikipedia)

### XML

Extensible Markup Language (en), « langage de balisage extensible ») est un langage informatique de balisage (sémantique) générique. Le World Wide Web Consortium (W3C), promoteur de standards favorisant l'échange d'informations sur Internet, recommande la syntaxe XML pour exprimer des langages de balisages spécifiques. De nombreux langages respectent la syntaxe XML : XHTML, SVG, XSLT, etc.

Son objectif initial est de faciliter l'échange automatisé de contenus entre systèmes d'informations hétérogènes (interopérabilité). XML est une simplification de SGML dont il retient les principes essentiels comme :

la structure d'un document XML est définissable et validable par un schéma, un document XML est entièrement transformable dans un autre document XML.

Cette syntaxe est reconnaissable par son usage des chevrons (< >).

(Wikipédia)

### XSLT, eXtensible Style Language Transformation

eXtensible Style Language Transformation

L'objectif principal est la transformation d'un document XML vers un autre, ou un dialecte XML (XHTML, XSL-FO, HTML, etc.). Cependant, le langage XSLT permet aussi les transformations vers tout autre type de document, au format texte ou dans un format binaire (bien que ceci ne soit pas nativement prévu par la recommandation XSLT).

XSLT s'appuie sur XPath (une autre partie de la recommandation XSL) pour désigner une partie d'un arbre XML. XSLT est lui-même un dialecte XML. Un programme XSLT est donc, avant tout, un document XML :

L'une de ses principales particularités est d'être centré sur les données. Un XSLT doit s'appuyer sur un XML, c'est un couple obligatoire, et on ne peut créer en XSLT que des boucles parcourant des données sélectionnées par XPath .

(Wikipédia)



# Bibliographie

***[Ingénierie des connaissances et des contenus]***

BRUNO BACHIMONT, Lavoisier, 2007

***[Instrumentation numérique des documents : pour une séparation fonds/forme]***

BRUNO BACHIMONT, STÉPHANE CROZAT, Revue I3 vol 4, num 1, pp:95-104, 2004.

***[Le procédé SCENARI : Une chaîne éditoriale pour la production de supports numériques de formation]***

BRUNO BACHIMONT, ISABELLE CAILLEAU, STÉPHANE CROZAT, MANUEL MAJADA, SYLVAIN SPINELLI, TICE'02, 2002

***[PolyTex : un environnement pour l'édition structurée de polycopiés électroniques multisupports]***

BRUNO BACHIMONT, JEAN CHARLET, EuroTex'98, 1998

***[Scenari, la chaîne éditoriale libre]***

STÉPHANE CROZAT, Eyrolles, 2007

***[Chaîne éditoriale, accessibilité, mobilité]***

<http://blog.temesis.com/post/2008/02/05/288-ilearnforum-chaines-editoriales>, ÉLIE SLOÏM, STÉPHANE CROZAT, iLearning Forum, 2008

***[L'accessibilité Web]***

<http://www.journaldunet.com/solutions/dossiers/pratique/accessibilite/1.shtml>, ANTOINE CROCHET-DAMAIS, JDN Solutions, 2008

***[LyX – The Document Processor, LyX]***

<http://www.lyx.org>

***[Single Source Publishing]***

[http://en.wikipedia.org/wiki/Single\\_source\\_publishing](http://en.wikipedia.org/wiki/Single_source_publishing), Wikipédia

***[What you see is what you mean]***

<http://fr.wikipedia.org/wiki/WYSIWYM>, Wikipédia

***[WYSIWYM (what you see is what you MEAN)]***

<http://scenari-platform.org/trac/scenari/wiki/WYSIWYM>, SYLVAIN SPINELLI