

Le retour d'expérience du CEA sur la veille technologique autour de XML

Denis Malavieille

denis.malavieille@cea.fr

01 69 08 61 14

Sommaire

- ◆ Notions
- ◆ Usages de XML
- ◆ Fusion des données structurées et des documents
- ◆ Contexte de la veille à DTI/SITI
- ◆ La nature multi-fichiers des documents XML
- ◆ Impacts sur les éditeurs
- ◆ Les types d'éditeurs : textes ; structurés ; wysiwyg
- ◆ Les éditeurs retenus aujourd'hui
- ◆ Présentation de Epic et XMLSpy
- ◆ Présentation de 2 exemples d'édition en ligne

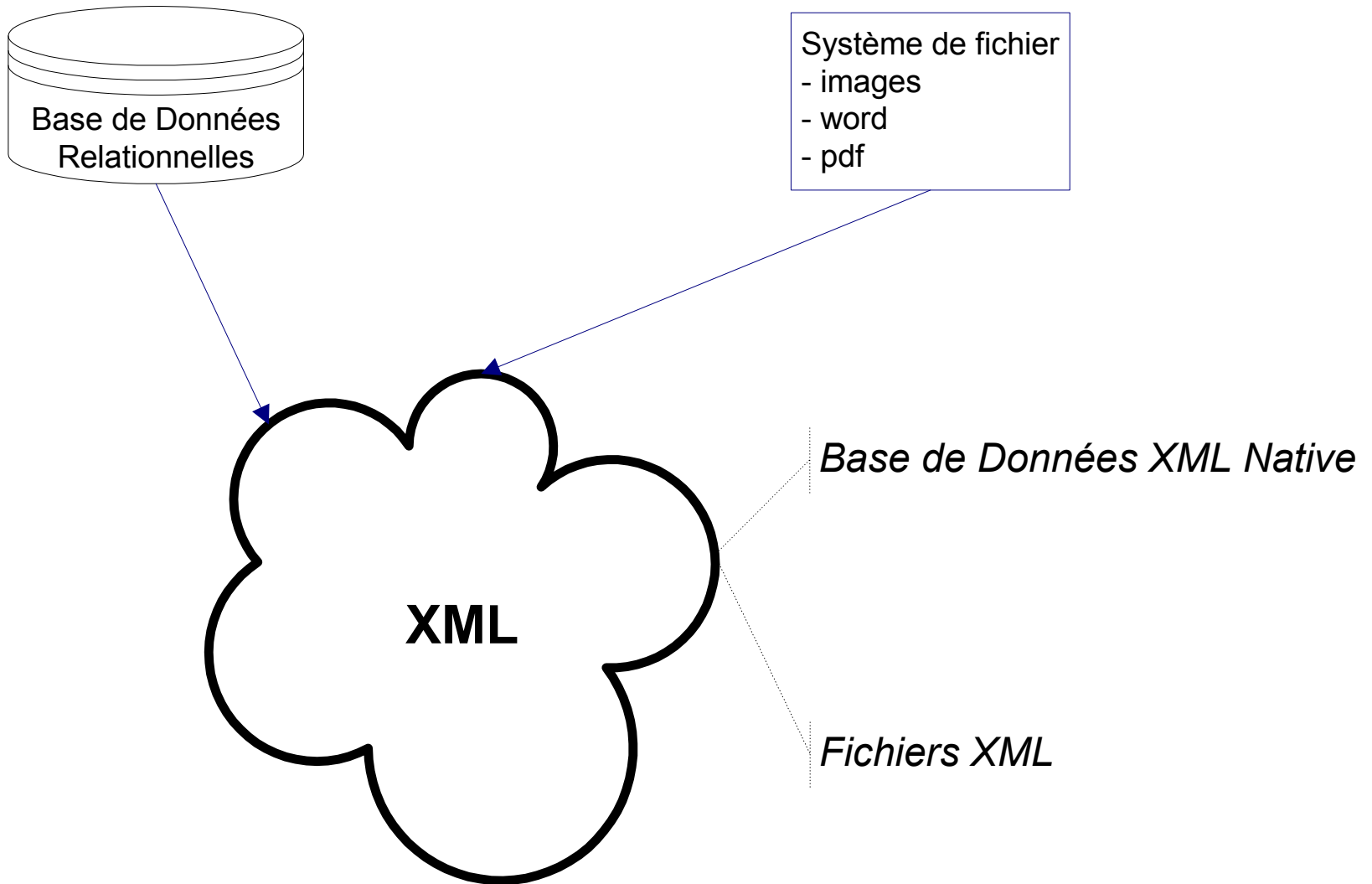
Notions

- ◆ Fichiers textes :
 - ✓ lisibles par l'homme et par la machine
 - ✓ Compatible avec tous les réseaux
 - ✓ Compatible avec tous les systèmes de fichiers
- ◆ Mode de structuration universel : un pari?
- ◆ Séparation des éléments de structuration et des données
- ◆ Les fichiers sont plus volumineux qu'un codage spécifique
- ◆ XML tout seul ne définit qu'une syntaxe et pas un vocabulaire
- ◆ Il n'y a pas de présentation graphique donc de mise en page

Les usages d'XML

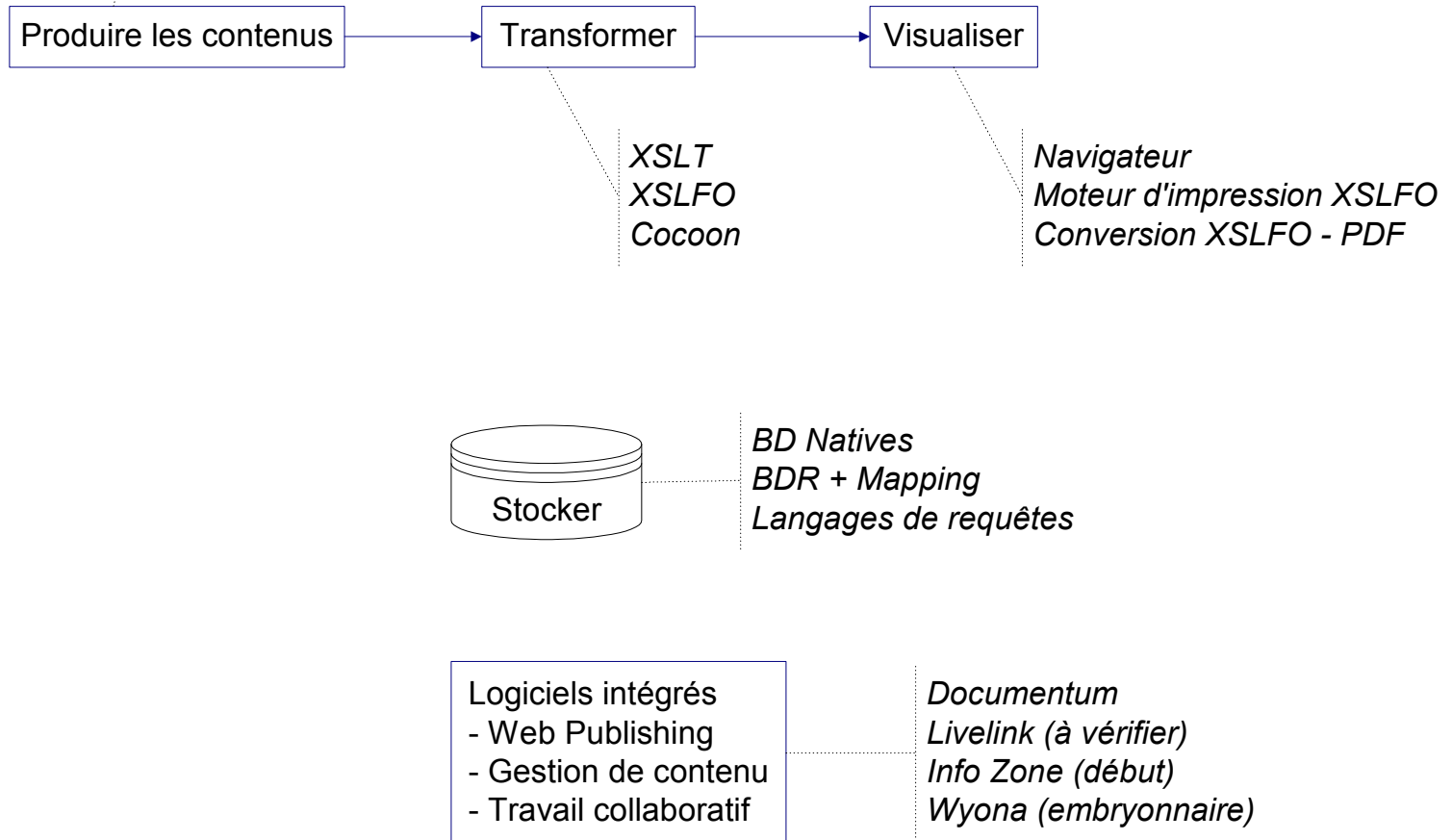
- ◆ Le principal aujourd'hui : échanger des données entre logiciels / intégration d'applications.
 - ✓ Faire communiquer des applications
 - ✓ Faire travailler ensembles des applications
- ◆ Peu utilisé encore : format documentaire
- ◆ Evolution : intégrer les contenus. On manipule techniquement de la même façon des informations très structurées et des informations peu ou non structurées.
 - ✓ Données issues de capteurs
 - ✓ Images
 - ✓ Documents

Fusion des données structurées et des documents



Contexte de la veille : Chaîne de production XML pour le documentaire

Editeurs classiques + export import selon 1 schéma (ex : illustrator + svg)
Editeurs XML pour tous les schémas
Aggrégation à partir de sources selon un script ou programme (générateurs cocoon ; JSP ...)

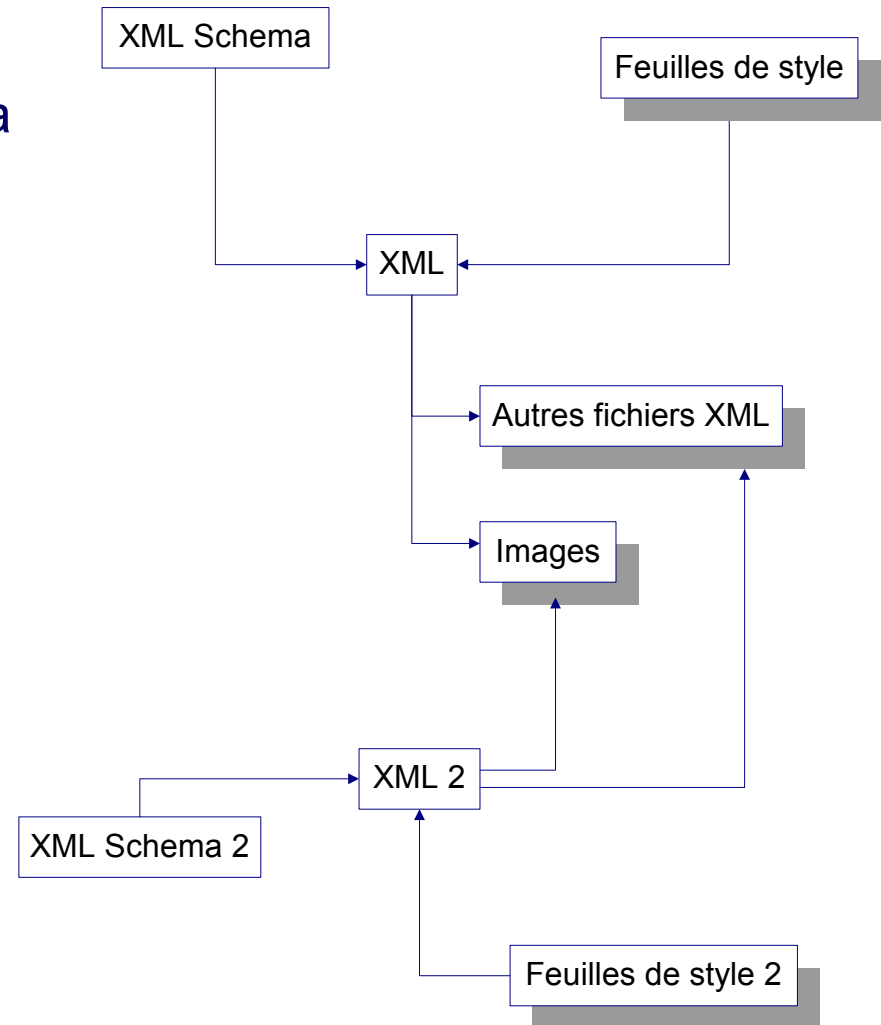


Point de vue qualitatif sur le domaine

- ◆ Traitement : beaucoup d'efforts en java et logiciels libres, réutilisés dans les produits commerciaux. Avance bien.
- ◆ Visualisation : les navigateurs suivent avec beaucoup de retard. XSLFO reste très compliqué et peu utilisé.
- ◆ BD : peu d'offre commerciales ou en logiciels libres. Seules les commerciales seraient (?) efficaces. Un problème de normalisation d'un langage de requête.
- ◆ Logiciels intégrés : XML est de plus en pris en compte. Reste un soucis de stabilisation des normes qui freinent les investissements.
- ◆ Editeurs multi schémas : point bloquant
 - ✓ Peu d'offre
 - ✓ Partie visible et majeure pour l'utilisateur
 - ✓ Complexité inhérente à XML

La nature multi-fichiers des documents XML

- ◆ Maintenir un lien avec le schéma et les transformations
- ◆ Composition de plusieurs contenus de différentes natures
- ◆ Partage de contenus



Impacts sur les éditeurs

- ◆ A priori pas d'éditeur universel : éditer des images, des formules mathématiques, des documents ...
- ◆ Une édition orientée structure et non orientée style.
 - ✓ Le contraire de word
 - ✓ Les feuilles de style dans word sont presque l'équivalent des éléments
- ◆ Nécessité de respecter un schéma : l'éditeur doit contraindre la saisie
 - ✓ Valider le respect du schéma
 - ✓ Aider l'utilisateur à se souvenir des éléments et des règles
 - ✓ Productivité : affranchir l'utilisateur de toutes les saisies fastidieuses automatisables
- ◆ Doit respecter en natif plusieurs schémas / normes de base : XMLSchema , XSLT , XSL FO et de préférence MathML, SVG, SMIL
- ◆ Gérer des contenus multi-fichiers

Les types d'éditeurs : textes ; structurés ; Graphiques

- ◆ Textes : éditeurs de texte, de base, avec ou sans aide pour le respect des schémas.
 - ✓ Pour les développeurs
- ◆ Structurés : éditeur d'éléments
 - ✓ Plusieurs niveaux d'aide au respect des schémas
 - ✓ Différentes aides à la productivité
 - ✓ Pour des spécialistes de la production de contenus
 - ✓ Boîtes à outils pour les développeurs (Modes d'édition spécialisés en fonction des schémas les plus importants)
 - ✓ Beaucoup de produits, peu sont crédibles. Des solutions relativement matures.
- ◆ Graphiques
 - ✓ Edition de documents
 - ✓ Travailler en XML comme avec Word
 - ✓ Pour les utilisateurs niveau bureautique
 - ✓ Peu d'offre : une ergonomie qui se cherche

Retour d'expérience : ce qu'on attend d'un éditeur

- ◆ Pour un déploiement de XML :
 - ✓ éditeur Graphique
 - ✓ Faciliter ou masquer le multi fichier
 - ✓ Application web
 - ✓ Cible : utilisateur bureautique
- ◆ Pour les études et la conception de feuilles de styles, de schémas et autres ...
 - ✓ Éditeur structuré
 - ✓ Modes d'édition spécialisés pour les principales normes
 - ✓ Respect des normes et suivi de leur évolution

Les éditeurs utilisés et étudiés

- ◆ EPIC : société Arbortext
- ◆ i4i
- ◆ XML Spy : société Altova
- ◆ XMETAL : société Softquad. Mériterait
 - ✓ Une réévaluation pour surveiller son évolution
 - ✓ Intégration avec Documentum ou des CMS
- ◆ Le dernier : XOPUS ; société Q42
- ◆ Expérimentation : stage de Robert Jimenez
- ◆ Expérimentation apparentée : Refcom
- ◆ Une pléiade d'outils libres : surtout orientés *structuré*. Souvent peu matures, avec un respect moyen des normes.

Epic

- ◆ Le plus avancé pour le wysiwyg et la productivité de l'utilisateur final
- ◆ Editeur d'équations et de tables
- ◆ Le plus avancé pour l'impression de document : mais technologie propriétaire et aucune prise en compte de XSLFO
- ◆ Intégration avec Documentum et d'autres CMS
- ◆ Version Citrix
- ◆ En retard sur les normes : vient de SGML
- ◆ Coût élevé pour un déploiement massif
- ◆ Conception ancienne :
 - ✓ vient d'UNIX,
 - ✓ mal intégré à Windows
 - ✓ une ergonomie qui date.
- ◆ Temps d'apprentissage long pour le paramétrage wysiwyg et l'impression.
- ◆ Spécialisé documentaire
 - ✓ approche structurée limitée
 - ✓ Aucunement une boîte à outil pour le développeur

XML Spy

- ◆ Un bon compromis :
 - ✓ Edition structurée
 - ✓ Edition Wysiwyg
 - ✓ Boîte à outil développeur
- ◆ Le plus à jour et pointu dans le respect des normes
 - ✓ XSLT et XSL FO
 - ✓ Editeur spécialisé pour XML Schéma
- ◆ Ergonomie intuitive : temps d'apprentissage court
- ◆ Une approche intéressante pour le Wysiwyg : mais encore peu mature et trop de bugs.
- ◆ Une approche originale du structuré : déroutante et parfois contre productive
- ◆ Plutôt économique
- ◆ Un module d'édition wysiwyg via un navigateur : n'a pu être testé. Les évolutions sont à suivre.
- ◆ Un environnement de développement pour créer des éditeurs spécialisés.

Outils d'édition en ligne

◆ Stage de Robert Jimenez

- ✓ Environnement Application Web J2EE : pas d'autres équivalents identifiés
- ✓ Mi chemin graphique et structuré : pour des raisons de temps et de technologie
- ✓ Outil autonome : un composant interopérables avec d'autres applications et processus J2EE

◆ XOPUS

- ✓ Javascript : dépendant du navigateur
- ✓ Orienté post-processeur : on génère des pages statiques avec beaucoup de javascript (à vérifier)
- ✓ Une ergonomie très séduisante
- ✓ A étudier de près

Prototype : édition par formulaires

The screenshot shows a web browser window titled "Bienvenue dans l'éditeur XML en ligne... - Microsoft Internet Explorer". The page header features the CEA logo and the text "Direction des Technologies de l'Information". The main content area is a form with the following sections:

- document**
 - documentid: 157494
 - titre: Exemple de document
- auteur**
 - nom: Jimenez
 - prenom: Robert
 - email: robert.jimenez@cea.fr
- contenu**
 - para: Le CEA est un acteur clef de la recherche, du développement et de l'innovation en matière d'énergie, de défense, de technologies de l'information, ainsi que de santé. Depuis sa création en 1945, il relève des défis scientifiques majeurs dans un ensemble de domaines : programmes électronucléaires, dissuasion nucléaire, micro et nanotechnologies, astrophysique, imagerie médicale, toxicologie ou biotechnologies..
 - para: Sa capacité à développer et à innover est le résultat d'une culture croisée ingénieurs-chercheurs et trouve sa source dans une recherche fondamentale de haut niveau, indispensable à l'émergence de nouveaux concepts.

On the left side of the form, there is a vertical menu with the following options: "Actions...", "Insérer en dessus... prenom", "Insérer en dessous... prenom", "email", and "Copier".

Conclusion

- ◆ Outil de travail retenu : XML Spy
 - ✓ Polyvalence
 - ✓ Gain de temps
 - ✓ Suivi des normes
- ◆ EPIC Editor: continuer la maintenance annuelle et le suivi du produit
- ◆ Maintien d'une veille
 - ✓ sur les produits graphiques en priorité, avec une préférence pour les solutions WEB.
 - ✓ Sur les BD libres (essais d'une commerciale?)
 - ✓ Sur les normes
- ◆ Expérimentations propres :
 - ✓ Opérationnaliser XSLFO
 - ✓ Sites web dynamique en XML avec édition en ligne