

CAHIERS *GUTenberg*

☞ ÉDITORIAL

☞ Thierry BOUCHE

Cahiers GUTenberg, n° 51 (2008), p. 3-6.

<http://cahiers.gutenberg.eu.org/fitem?id=CG_2008__51_3_0>

© Association GUTenberg, 2008, tous droits réservés.

L'accès aux articles des *Cahiers GUTenberg*

(<http://cahiers.gutenberg.eu.org/>),

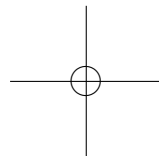
implique l'accord avec les conditions générales

d'utilisation (<http://cahiers.gutenberg.eu.org/legal.html>).

Toute utilisation commerciale ou impression systématique

est constitutive d'une infraction pénale. Toute copie ou impression

de ce fichier doit contenir la présente mention de copyright.



ÉDITORIAL

L^AT_EX ET XML, UNE HISTOIRE INTERMINABLE ?

La journée GUTenberg d'octobre 2008 avait abordé les questions de conversion d'articles ou de métadonnées en L^AT_EX vers du XML, avec un intérêt particulier pour la conversion des expressions mathématiques en MathML. Les actes de cette journée forment le cœur de ce numéro des *Cahiers*.

Voilà plus de dix ans que les normes XML et MathML ont été publiées, plus de dix ans que l'association GUTenberg organise régulièrement des manifestations, des formations sur leur interfaçage avec L^AT_EX, et que les *Cahiers* publient des articles à ce sujet. Si la montée en puissance de XML et de ses avatars est aujourd'hui avérée dans une très grande variété de domaines, il n'en a pas été de même pour MathML qui reste peu déployé, souvent cantonné à des applications internes comme les flux de données des très grands éditeurs, et dans des niches qui nécessitent une représentation structurée des mathématiques en vue de traitements automatiques. La dualité entre MathML de présentation et MathML « sémantique » n'a pas été résolue, les ponts entre les deux versions n'étant toujours que balbutiants. Il s'est ainsi développé deux communautés assez distantes.

— L'une ne jure que par le MathML de contenu et le développement des outils de gestion, voire de démonstration, automatiques que permet ce format, dont l'expansion est limitée par la rareté des ressources ainsi formalisées et l'inextensibilité du standard qui est très loin de couvrir tout ce qui peut être exprimé par des formules mathématiques telles qu'on les écrit ou les réinvente chaque jour dans les laboratoires.

— L'autre cherche à tirer le meilleur parti des formats existants pour la diffusion multimédia et l'échange de textes rédigés par des humains et destinés à leurs semblables. Dans ce domaine, où le MathML de

présentation — éventuellement enrichi par les indications sémantiques que le format permet — règne en maître, de nombreuses avancées ont eu lieu récemment pour sortir enfin les formules mathématiques du statut d’images muettes qui était le leur dans la plupart des formats électroniques. Citons le développement de logiciels de reconnaissance optiques, de lecture à voix haute, de conversion en Braille, de moteurs de recherche qui ne sont plus limités aux seules portions textuelles des documents scientifiques.

Mais l’actualité de MathML pour diffuser des textes scientifiques vient surtout d’une avancée récente : les navigateurs web commencent enfin à pouvoir afficher de façon raisonnable des formules mathématiques lisibles et accessibles (c’est en particulier le cas d’internet explorer muni du plug-in MathPlayer de la société Design Science, et de des navigateurs basés sur le moteur Gecko comme Firefox, modulo l’installation de quelques polices). Il devient envisageable de mettre en ligne des documents scientifiques sous une forme plus dynamique et intégrée que le PDF, plus facile à partager que des sources \TeX ! L’abolition de ces barrières technologiques aura des conséquences considérables pour l’ergonomie des ressources pédagogiques par exemple.

*

* *

Reste à rédiger ces documents, et produire un code MathML de qualité, bien intégré dans une structure HTML propre. Une stratégie est d’utiliser d’emblée un outil dédié à l’édition de XML : ceux qui supportent MathML sont encore peu diffusés, souvent limités à des usages professionnels spécifiques. Signalons néanmoins GELLMU qui a une syntaxe d’entrée très proche de celle de \LaTeX . Mais la stratégie la plus naturelle pour les utilisateurs de \LaTeX est de continuer à rédiger et composer leurs textes avec leur outil de prédilection en profitant du très large éventail de solutions disponibles pour produire des documents de qualité adaptés à toutes les situations, et d’utiliser un convertisseur *ad hoc* pour dériver des versions à diffuser sur le web.

Ce numéro spécial est largement dédié à cette question, peut-être plutôt du point de vue du développeur et de la conversion d'ensembles importants que de l'utilisateur final. Nous espérons publier prochainement des contributions sur tous les aspects qui ne sont qu'effleurés ici.

Le premier article passe en revue les outils utilisables aujourd'hui en se concentrant sur cinq d'entre eux. On verra que, si l'offre dans ce domaine s'est considérablement accrue au cours des dernières années, obtenir l'équivalent fiable en XHTML+MathML d'un manuscrit \LaTeX reste un défi et demandera toujours de peaufiner le fichier source et les paramètres de la conversion en fonction de l'outil utilisé et du but recherché. L'angle d'attaque de cet article pourra surprendre nos lecteurs, car il est celui de chercheurs orientés du côté des mathématiques sémantiques et résolus à sauver le corpus mathématique électronique de l'obscurité absconse des dialectes \TeX , pour le rendre exploitable de façon universelle.

Le deuxième article nous permet d'entrer dans les entrailles de l'un des convertisseurs évalués, Tralics, en décrivant la chaîne de conversion employée pour deux publications.

L'article suivant en donne une troisième basée sur le même Tralics employée pour alimenter la plateforme d'édition électronique CEDRICS développée par la Cellule MathDoc, qui est désormais exploitée également sur notre site `cahiers.gutenberg.org`.

Le dernier article se situe du côté de la sortie : il explique les concepts du formatage de XML avec XSL-FO en des termes adaptés à un lectorat plus fluent en \LaTeX qu'en XSLT.

L'outil LXir, qui avait été l'objet d'une présentation lors de la journée d'octobre, est un peu malmené dans ce numéro : nous espérons publier ultérieurement l'article à son sujet qui n'a pas atteint la rédaction dans les délais extrêmement stricts fixés en vue d'une parution rapide, datée du mois même de ces journées.

© Rédaction des *Cahiers GUTenberg*
Cellule MathDoc
BP 74,
F-38402 Saint-Martin d'Hères Cedex
`cahiers@gutenberg.eu.org`

Cahiers GUTenberg parus sur le même thème

☞ « XML » (actes de GUT'99, Lyon). N° 33-34,
décembre 1999.

☞ « L^AT_EX et XML : coopération pour l'internet »
(actes de GUT 2000, Toulouse). N° 35-36, mai 2000.

☞ Ainsi que des articles substantiels depuis mars 1998
parus dans des numéros monothématiques :
N°s 28-29, 37-38, 39-40.