

Cahiers **GUT** *enberg*

∞ ÉDITORIAL : MULTILINGUISME ET CODAGE

¶ Jacques ANDRÉ, Michel GOOSSENS, Yannis HARALAMBOUS

Cahiers GUTenberg, n° 20 (1995), p. III-IV.

<http://cahiers.gutenberg.eu.org/fitem?id=CG_1995__20_R3_0>

© Association GUTenberg, 1995, 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 :

Multilinguisme et codage

Jacques ANDRÉ

Michel GOOSSENS

Yannis HARALAMBOUS

CA de l'Association GUTenberg

GUT dans « Association GUTenberg » signifie « Groupement des Utilisateurs (francophones) de T_EX ». Toutefois la dite association ne se donne pas pour seul but de défendre farouchement – les yeux fermés sur le monde extérieur avec de bonnes œillères – le seul logiciel T_EX mais aussi, comme le disent les statuts de cette association, de « favoriser les échanges techniques permettant d'augmenter les possibilités d'impression et de publications scientifiques ». Il se trouve, bien sûr, que nous sommes convaincus que T_EX est une issue pour ces échanges techniques bien souvent malmenés par d'autres produits plus commerciaux, plus « modernes », mais qui restent bien en deça de ce que propose L^AT_EX, du moins dans sa vision de 1995 qui, on ne le répètera jamais assez, n'est plus le L^AT_EX des années 80.

Le dernier numéro des *Cahiers Gutenberg* était consacré à la « diffusion des documents électroniques, de L^AT_EX à WWW, HTML et Acrobat ». Diffusion veut dire diffusion mondiale, et par là multi-linguisme. Nous, francophones, grâce à, ou à cause de, notre désir de voir nos accents préservés sommes un peu sensibilisés à ce problème mais nous ne sommes pas toujours conscients de son ampleur dès lors qu'il s'agit de langues où les caractères dépendent du contexte (comme les ligatures arabes) et à plus forte raison lorsque ces caractères sont des idéogrammes et se comptent alors par milliers. Les *Cahiers Gutenberg* vont donc publier une série de numéros sur ce thème très général de la diffusion électronique de documents à travers le monde dont celui sur WWW aura été le premier. Mais, en fonction de l'importance du sujet, de notre propre connaissance de ce domaine sans oublier des facteurs comme la disponibilité des auteurs, ces numéros paraîtront sans doute dans un ordre un peu hétéroclite. Mais nous essayerons d'être aussi complet que possible, par exemple parler de multi-linguisme, ce n'est pas que parler des glyphes des caractères. C'est parler également de leur codage pour les imprimantes, et du codage entre machines. C'est aussi parler de la façon de les saisir, de les taper sur des claviers appropriés, ou même sur des claviers inappropriés (on a déjà du mal à taper un « é » sur un clavier Ascii, alors comment faire pour un idéogramme chinois ?). Bien plus important : on commence à définir des codages, comme Unicode, permettant de traiter toutes les langues du monde. Mais les formateurs, éditeurs et autres systèmes de manipulation de documents, qu'il s'agisse de T_EX, HTML, BWV ou autres Keushel sont ils prêts à assumer ce multi-linguisme ? La réponse de T_EX est oui (ou plus précisément), c'est bien pour ça que l'on peut sans rougir, défendre une association comme GUTenberg.

Cette série comprendra donc plusieurs numéros. En gros, mais sans nous engager formellement quant aux dates et à la rigueur des numéros, voici ce qui est prévu :

19. *Diffusion des documents électroniques : de L^AT_EX à WWW, HTML et Acrobat*
Numéro paru en janvier 1995 (actes de la journée Gutenberg « de Nanterre » ; 160 pages ; il en reste quelques exemplaires à 75 francs français (s'adresser à l'association).

-
20. *Codage et multi-linguisme* (ce numéro) .
 23. *Omega* ou le traitement multilingue des textes T_EX à 32 bits. À paraître en décembre 1995 ?
 24. *Les ligatures* – numéro consacré aux ligatures avec une certaine emphase sur les ligatures du monde non-latin.
 25. *T_EX et l'arabe* – actes de la journée organisée par l'association à l'Inalco sur ce thème.
 26. *Courier* – un numéro en hommage aux caractères à chasse fixe, style machine à écrire, et sur leurs apports à la typographie mondiale.

Tous ces numéros sont pratiquement prêts. Mais, comme le croyait Zénon, on a beau avancer il y toujours la moitié du chemin qui reste à faire. C'est peut-être pour converger que nous avons accepté de boucler ce numéro en étant conscient qu'il aurait pû être meilleur si on avait encore attendu quelques temps !

Le numéro actuel est donc consacré aux codages et le multi-linguisme ou discute d'initiatives pour améliorer les possibilités de communication par voie électronique dans les différentes langues du monde. La première contribution retrace l'historique des codages de caractères et explique l'évolution du code Acsii vers *Unicode* et la norme universelle ISO/IEC 10646, qui devraient permettre d'échanger sans problèmes des informations écrites dans n'importe quelle langue du monde. La création par Bigelow et Holmes de la police *Unicode Lucida Sans*, à l'origine pour utilisation avec le système informatique Microsoft Windows NT, est le sujet du deuxième article¹. La contribution finale est une description générale du système \TeX , une extension 16-bit de T_EX développée par Yannis Haralambous et John Plaice. La première mondiale de l'annonce du système \TeX eut lieu à Genève le jeudi 16 mars dernier, et une partie du contenu de ce Cahier y fut présenté. Nous tenons à remercier les participants à cette journée \TeX , d'être venu à Genève et d'avoir ainsi montré l'importance qu'ils attachent au développement d'un système qui permettra de mieux traiter les documents multi-lingues. En particulier nous voulons mentionner CSTUG, CyrTUG, NTG et UKTUG, qui ont envoyé des représentants de leurs organisations d'utilisateurs de T_EX respectives à Genève. En plus nous tenons à remercier nos amis britanniques d'UKTUG, ainsi que Malcolm Clark du comité d'organisation de T_EX88 à Exeter, qui ont co-sponsorisé la journée.

¹ On peut ne pas aimer le dessin de *Lucida*, on peut trouver le travail de Bigelow & Holmes trop occidental ; néanmoins, cet article correspond à la première réalisation commerciale d'Unicode et il a semblé utile à certains d'entre nous de le joindre à ce Cahier.