

Informatique & Bible, asbl - Belgique  
Rue de Maredsous, 11 B5537 Denée - Belgique  
Tél:+32(0)82.69.96.47 Fax:+32(0)82.22.32.69  
cib@cibmaredsous.be



Interface n° e-92 Août 2003

## Comptes-rendus

Le numéro de la revue *Media Development*, 2002/4, sur les Droits de la Communication dans la Société Civile (CRIS) ►

- Un *Dictionnaire féministe de la Bible* ►

- *The One Year Bible* ►

Unicode, écriture du monde?, présenté par ses éditeurs Jacques André et Henri Hudrisier

Vient de paraître: *Unicode, écriture du monde?*, sous la direction de Jacques André et de Henri Hudrisier, numéro spécial de *Document numérique*, vol. 6, n° 3-4, Lavoisier+Hermès, éd., 2002, 364 pages.

Unicode est un codage de tous les caractères du monde. Sont présentées l'histoire et les caractéristiques techniques de ce code, ses relations avec la typographie et les langues (avec un accent sur les langues africaines, arabes, française et japonaise et sur les hiéroglyphes), et ses implications, et appropriations, dans l'information planétaire d'aujourd'hui.

Le monde entier est en train de subir une révolution, l'adoption d'un codage universel de tous les caractères du monde. Ce n'est en fait qu'à peine perceptible pour la majorité des gens et risque d'échapper pour un temps à ceux qui sont déjà exclus de l'Internet (certains pays africains notamment). De quoi s'agit-il? Depuis des siècles, l'homme a enrichi, pour ses besoins liés aux délais et aux distances, ses propres possibilités de transmission d'informations (mode visuel ou sonore, etc.) par des moyens techniques utilisant des codes. Ainsi sont nés le télégraphe de Chappe, le télégraphe électrique et le morse, le télex, etc. L'informatique bien sûr a amplifié ces codages. Mais si, vers 1970, une normalisation très réussie (du moins pour les langues occidentales voire pour la seule langue américaine) s'est imposée avec le très fameux code Ascii, l'anarchie s'est réinstallée dès qu'on a voulu mettre ne serait-ce que des lettres accentuées. Il suffit de considérer la prolifération de codages divers utilisés sur Internet définis dans des normes officielles, comme Latin-1, ou dans des standards privés, comme les codes-pages de Windows incompatibles avec leurs équivalents du Mac, ne permettant plus aucune conversion ni échange vraiment sûr, fiable et international. Parallèlement, des codes ont été définis pour l'arabe, le cyrillique, le chinois ou le japonais, autant de nouveaux étages à cette tour de Babel. Cette situation complexe devenant très difficile à gérer à l'ère d'Internet, deux chantiers mondiaux, formés de scientifiques, d'informaticiens, mais aussi de linguistes, voire de politiciens, etc. ont été entrepris, vers 1990, pour inventorier tous ces caractères et élaborer un système permettant de les coder compte tenu de leurs spécificités (écriture de droite à gauche ou idéographique par exemple). Ces deux chantiers (respectivement celui d'un ensemble de fabricants d'ordinateurs, Unicode, et celui de l'institut de normalisation internationale qu'est Iso) ont fusionné et défini la norme ISO-10646, surensemble du standard Unicode. Et c'est justement cet Unicode qui est donc en passe de devenir un des enjeux majeurs de codage pour toute l'humanité.

Le système métrique a répondu à deux questions: "Quelles unités?" et "Comment les faire fonctionner ensemble?". Unicode a le même soucis pour les écriture, la difficulté étant donc de savoir quels caractères inclure, ce qui impose de répondre de façon univoque à une question complexe: qu'est-ce qu'un caractère? Il nous a alors paru important d'expliquer un certain nombre de concepts sous-jacents aux notions de codage, de caractères, d'écriture et d'œuvrer dans le sens de la compréhension même si parfois le ton des auteurs est un peu polémique, même si chacun a parfois tendance à s'approprier

Unicode avec sa propre culture (pour un typographe, Unicode doit coder tous les caractères qu'il a l'habitude de trouver dans ses casses; un arabisant sera surtout sensible à la bidirectionnalité; un Chinois et un Japonais ne seront peut-être pas d'accord sur l'inventaire des caractères communs à ces deux peuples; un Berbère réclamera le tiffinagh, etc.).

Cet ouvrage réunit donc des articles abordant l'Unicode selon diverses cultures:

- Trois articles entourent ce numéro en faisant le point sur l'Unicode:
  - Jacques André, en se basant sur l'évolution historique de ces normes de codage des caractères, explique les concepts souvent mal compris de codage, de caractère, de norme et standard, etc.;
  - Patrick Andries présente globalement Unicode et ses principales techniques;
  - Ken Whistler, directeur technique d'Unicode, en relate l'histoire et la philosophie.
- Deux articles (respectivement d'Olivier Randier et Yannis Haralambous) donnent l'avis de typographes ayant chacun leur propre questionnement par rapport à l'Unicode.
- Plusieurs langues ou systèmes d'écriture font l'objet d'articles spécifiques:
  - Rachid Zghibi montre comment l'Unicode résout le traitement des écritures arabes;
  - Serge Rosmoduc s'intéresse, plutôt dans le cadre d'ISO/CEI 10646 que dans celui d'Unicode, aux écritures anciennes et notamment aux hiéroglyphes (égyptiens);
  - Andrei Popescu-Belis, partant d'une expérience sur le terrain, fait part de problèmes de multilinguisme en Afrique auxquels Unicode pourrait remédier;
  - Pour Nadine Lucas enfin, Unicode est une aubaine pour le Japon qui tient à son écriture et à l'ouverture au monde.
- Deux articles annoncent quelques possibilités et problèmes de l'utilisation d'Unicode dans des applications sur Internet:
  - Emmanuel Giguet et Nadine Lucas montrent les avantages et problèmes de la recherche d'informations en Unicode sur Internet;
  - Michel Bottin présente le projet d'un site consacré aux écritures du monde.
- Deux synthèses font enfin le point:
  - Sylvie Baste reprend Unicode et son appropriation par la communauté francophone: pour elle il n'y a pas d'informatisation des écritures sans approche culturelle;
  - Henri Hudrisier cherche à montrer l'importance de ces normes comme paradigmes (sinon mécanismes ou composantes) fondamentaux d'une nouvelle techno-culture de l'informatique rendant enfin possible la montée en charge d'échanges planétaires multilingues et multi-écritures harmonieux. Il insiste sur la nécessité d'une prise en compte non seulement technique mais linguistique et surtout culturelle de cette nouvelle technologie.

En rassemblant ces textes les auteurs font partager aux lecteurs leur enthousiasme vis à vis d'Unicode et de ISO/CEI 10646 qui s'imposent désormais comme héritiers obligatoires de ce vieil Ascii désormais insuffisant pour affronter une informatique devenue planétaire. Ils espèrent que la nécessité d'adopter définitivement ce nouveau format de codage deviendra évidente pour chacun: non seulement pour les constructeurs informatiques et les typographes qui en sont persuadés depuis longtemps, mais pour les développeurs de logiciels et de sites Web, pour les professionnels du commerce et de l'échange de données international, pour les conservateurs de musées ou de bibliothèques, pour les enseignants de langues et à terme pour toute personne destinée à traiter des données écrites.

C'est en maintenant actifs des échanges fructueux entre utilisateurs spécialistes, experts d'Unicode ou d'ISO/CEI 10646 et les industriels constructeurs que nous parviendrons, nous utilisateurs potentiels, à maintenir la pression pour qu'Unicode soit le produit parfait que l'on attend et que nous éviterons une main mise toujours possible sur Unicode qui équivaldrait à la confiscation industrielle d'une des pièces maîtresses des fondements de nos civilisations: l'écriture.

