

Un projet de site Web pour les étiquettes de langue

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 28 janvier 2007

<https://www.bortzmeyer.org/ltru-site.html>

Les **étiquettes de langue** ("*language tags*") sont des identificateurs servant notamment à indiquer la langue d'un document. Il n'existe pas encore de site Web qui leur soit dédié mais un projet existe.

Ces étiquettes sont normalisées dans le RFC 4646¹. On constate qu'elles sont peu utilisées dans les pages Web, et souvent de travers (combien de pages en anglais utilisent le code `uk` qui est celui de l'ukrainien!). Résultat, les utilisateurs, par exemple les moteurs de recherche s'en servent peu ou pas du tout (voir l'exposé de Google, "*Unicode at Google*" <http://www.macchiato.com/slides/unicode_at_google.ppt>).

De même, on constate que le processus (effectivement pas évident) d'enregistrement de nouvelles sous-étiquettes (notamment pour une variante d'une langue existante) est mal compris de ceux qui connaissent la variante en question, mais ne connaissent pas l'IETF.

Les documents officiels comme le RFC 4646 sont très bien faits mais le besoin d'une communication orientée vers un plus large public se fait sentir. D'où le projet de construction d'un site Web dédié aux étiquettes de langue, actuellement en ligne à l'adresse temporaire <<http://ltru.generic-nic.net/>>.

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc4646.txt>