

RFC 4647 : Matching of language tags

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 11 septembre 2006

Date de publication du RFC : Septembre 2006

<https://www.bortzmeyer.org/4647.html>

Le RFC 4646¹ qui spécifie les nouveaux "*language tags*", les étiquettes pour indiquer la langue utilisée, étant désormais disponible, ce RFC complète le travail en définissant (partiellement) les algorithmes de correspondance entre une étiquette existante et une étiquette demandée, par exemple lors de la négociation de contenu entre un navigateur Web et un serveur.

Notre RFC définit donc une nouvelle syntaxe, le "*language range*" (intervalle de langues) qui exprime des préférences et que l'algorithme doit mettre en correspondance avec les étiquettes disponibles (par exemple, les valeurs de l'attribut `xml:lang` si le serveur sert des documents XML).

Le RFC définit ensuite deux classes d'algorithmes possibles, le "*filtering*" (filtrage), qui va trouver O, 1 ou N documents parmi ceux disponibles et le "*lookup*" (recherche) qui va trouver un document et un seul. Le filtrage est par exemple utilisé lorsque le processeur veut extraire d'une bibliothèque de documents XML tous les documents qui répondent au même critère (être en Breton, par exemple). La recherche est utilisée lorsque le processeur doit trouver un document et un seul, par exemple pour envoyer un message d'erreur à l'utilisateur.

L'intervalle ressemble à une étiquette mais certains champs peuvent contenir un astérisque qui signifie l'acceptation de n'importe quelle valeur. Ainsi, `*-CH` pourra désigner toutes les langues utilisées en Suisse.

À noter que les détails des deux algorithmes, de filtrage et de recherche, sont délibérément non complets : c'est au protocole qui utilise ce RFC de tout spécifier. Même chose pour le cas où le client souhaite indiquer une liste de langues possibles, comme ce que permet le protocole HTTP avec son champ `Accept-Language`. Notre RFC ne donne pas de syntaxe précise pour de telles listes, ce sera aux futurs RFC utilisateurs de le faire.

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