

RFC 6415 : Web Host Metadata

Stéphane Bortzmeyer

<stephane+blog@bortzmeyer.org>

Première rédaction de cet article le 1 novembre 2011

Date de publication du RFC : Octobre 2011

<https://www.bortzmeyer.org/6415.html>

Ce court RFC spécifie une convention qui permet à un client Web de récupérer sur un serveur une description formelle de métadonnées sur le site servi (par exemple l'auteur du site Web, ou bien la licence du contenu servi). Cela se fait en mettant ces métadonnées au format XRD dans une ressource nommée `/.well-known/host-meta`.

Vous pouvez voir le résultat dans le `</.well-known/host-meta>` de mon site, très inspiré de l'exemple de la section 1.1 du RFC. Il est vraiment minimal et n'inclut que des métadonnées globales (le format XRD permet également de s'en servir pour des métadonnées locales à une ressource, cf. section 4.2).

Notez que le vocabulaire peut être trompeur. Le RFC 3986¹ parle de "*host*" pour le `example.com` de `http://example.com/truc/machin` mais ce terme ne désigne pas du tout une machine. Finalement, le terme marketing flou de « site » est encore le moins mauvais.

Ce nouveau nom de `host-meta` a donc été le premier à être enregistré (cf. section 6.1) dans le registre des ressources bien connues `<https://www.iana.org/assignments/well-known-uris/well-known-uris.xml>` qui avait été créé par le RFC 5785. (Deux ou trois autres sont venus juste après, et le RFC 5785 a été depuis remplacé par le RFC 8615.)

Pour le client, par exemple un navigateur Web, récupérer les métadonnées consiste juste donc à faire une requête GET sur le nom `/.well-known/host-meta` (section 2). Il obtient en échange un fichier XML au format XRD, servi avec le type `application/xrd+xml`.

1. Pour voir le RFC de numéro NNN, <https://www.ietf.org/rfc/rfcNNN.txt>, par exemple <https://www.ietf.org/rfc/rfc3986.txt>

Le contenu du document est détaillé dans la section 3. Ce vocabulaire XRD a été normalisé par OASIS et sert à beaucoup de choses très compliquées <<http://docs.oasis-open.org/xri/xrd/v1.0/xrd-1.0.html>>. Mais l'accent est mis ici sur son utilisation pour fournir des métadonnées. Tout en gardant le modèle de données de XRD, le serveur peut envoyer du XML (la syntaxe habituelle) ou du JSON. Cette dernière syntaxe est normalisée dans l'annexe A, sous le nom de JRD. On peut l'obtenir (n'essayez pas sur mon site, il n'est pas configuré pour cela) en utilisant la négociation de contenu HTTP (champ `Accept: application/json` dans la requête) ou en demandant `/.well-known/host-meta.json` au lieu de simplement `/.well-known/host-meta`.

Qui utilise ou va utiliser ces métadonnées? Pour l'instant, les moteurs de recherche ne semblent pas y accéder mais cela viendra peut-être bientôt (Yahoo et Google ont annoncé travailler sur le sujet). La communauté WebFinger (RFC 7033) est également très impliquée dans ce projet.

Sinon, ce choix d'utiliser une ressource avec un nom spécial (le préfixe `/.well-known` est normalisé dans le RFC 8615) est dû au fait qu'un serveur n'a pas d'URI qui le désigne globalement et on ne peut donc pas utiliser les techniques qui attachent les métadonnées à une ressource (par exemple les en-têtes HTTP comme l'en-tête `Link:` du RFC 8288), sauf à utiliser des trucs contestables comme de considérer que `/`, la racine du site, désigne tout le site.