

LA MINUTE DU PSIR

Les identifiants pérennes

La Minute du PSIR de cette semaine s'intéresse aux identifiants pérennes. A quoi servent-ils ? Qui les émet ? Qui les conserve ?

Les identifiants pérennes ont été introduits pour pallier les apparitions trop fréquentes de la page (code erreur 404 : page not found) à la place d'une ressource et d'une volonté de mise à disposition des informations sur le long terme des données diffusées sur le web.

Le web est un milieu assez mouvant, les ressources sont associées à des projets qui arrivent à leur terme, migrent sur de nouveaux serveurs (changement d'hébergeur), sont diffusées sur des sites dont les structures de page évoluent.

Les URLS pérennes sont donc une réponse pour identifier de façon unique et sûre nos données accessibles sur le web malgré ces changements.

Le mécanisme utilisé est basé sur un système de réécriture d'URL qui garantit le suivi de la donnée tout au long de ses évolutions et migrations éventuelles. Les systèmes ARK, HANDLE, DOI utilisent des bases de données qui permettent de faire la relation entre l'URL pérenne et l'URL de la ressource à un instant donné. Cette localisation est bien sûr modifiable dans le temps pour garantir la pérennité de l'accès à la ressource.

En règle générale, les URL pérennes ne sont pas destinées à être mémorisées, elles sont généralement difficiles à retenir et ne permettent pas d'avoir d'indication sur les données vers lesquelles elles pointent. Elles permettent néanmoins d'identifier l'« autorité » émettrice.

Pour la MOM nous utilisons des identifiants de type ark et handle et notre Naan (Name Assigning Authority Number) est 76609, et nos URL pérennes auront donc la forme suivante :

<https://ark.mom.fr/ark:/76609/xxx>, xxx étant l'identifiant unique au sein de notre organisation ([Plus d'informations](#)) Ces identifiants sont stockés sur nos bases de données.

Le système Handle.net est équivalent avec un stockage redondé des identifiants sur différents serveurs (CNRI - Corporation for National Research Initiative, Virginia), nous vous proposons la lecture d'un [article en français](#)

Pour terminer, les DOI - Digital Object Identifier (identifiant numérique d'objet) sont principalement utilisés pour identifier des revues, articles scientifiques, rapports. Ils sont gérés par le projet [DataCite](#) En France, c'est le laboratoire INIST (du CNRS) qui délivre les DOI (PID OPIDoR)

La différence majeure entre les DOI et les autres systèmes est qu'il convient de payer une cotisation annuelle pour l'utiliser et qu'il est accompagné de la mise à disposition d'un service complet de gestion des identifiants. Récemment, toutes les ressources déposées sur nakala (service de diffusion proposé par le TGIR Huma-Num) reçoivent également un DOI.

NIVEAU
DIFFICULTÉ



THEMATIQUE
Données

PERSONNES
CONSEILLÉES



MINUTE_PSIR_#15.pdf 215 Ko.

PDF

