

Archivage à long terme pour l'archive ouverte HAL (Hyper Article en Ligne)

L'archive ouverte pluridisciplinaire HAL (hal.archives-ouvertes.fr/) développée par le Centre pour la communication scientifique directe (CCSD) du CNRS bénéficie désormais d'un archivage à long terme au Centre informatique de l'enseignement supérieur (CINES). Créée en 2000 par le CNRS, la plateforme HAL recueille et diffuse en accès libre plus de 2000 nouveaux articles scientifiques par mois émanant d'établissements d'enseignement et de recherche français ou étrangers ou de laboratoires qu'ils soient publics ou privés. Le contenu de MédiHal¹ l'archive ouverte d'images et de photographies récemment ouverte bénéficie également de cet archivage.

Pourquoi un archivage à long terme ?

L'archivage à long terme va au-delà de la sécurisation physique des données. Il vise le traitement du format dans lequel elles sont représentées. Le but majeur est de s'assurer en permanence que les contenus numériques restent lisibles au cours du temps afin d'éviter le risque majeur d'une perte irréversible des données. Par exemple dans le cas des articles de HAL ou le format PDF a été retenu, il convient de suivre les normes de cette technologie afin de reconverter, le cas échéant, les documents dont la version de ce format est devenue obsolète. Cette reconversion est faite dans le respect du document original et permet à l'utilisateur de lire les documents avec les outils actuels. Par long terme, on entend une période suffisamment longue (30 ans et +) pour qu'il soit effectivement nécessaire de prendre en compte les changements technologiques intervenus et la gestion des nouveaux supports et formats de données.

A propos de HAL

HAL est une archive ouverte, qui s'inscrit dans le mouvement international en faveur du libre accès à l'information scientifique et a pour vocation de faciliter la diffusion des résultats de la recherche. Elle se présente comme une bibliothèque numérique recevant et diffusant les productions intellectuelles de la recherche dans toutes les disciplines. Outil de communication scientifique directe entre les chercheurs, elle permet le dépôt des publications en texte intégral et de leurs notices.

Cette plateforme commune d'archives ouvertes a été choisie par les établissements français de recherche, les universités et les grandes écoles, qui peuvent également extraire et afficher leur propre production, via des portails dédiés comme ceux de HAL-SHS, HAL-INRIA, HAL-INSERM, HAL-UNIV, etc..., Elle fait partie des dix premiers dispositifs d'archives ouvertes les plus visibles dans le monde (placé en 7^{ème} position) selon un récent classement de webometrics auprès des 400 archives ouvertes les plus visibles sur le web.

L'archivage à long terme des documents scientifiques fait partie des missions confiées au CCSD. L'implication du CCSD dans plusieurs projets du Très Grand Equipement ADONIS (<http://www.tge-adonis.fr>) a permis l'accord CINES-TGE Adonis-CCSD, qui rend aujourd'hui possible l'archivage des articles déposés dans HAL depuis 10 ans. Le financement de l'opération est apporté par le TGE Adonis. La mise en production est effective depuis le 3 avril 2010, et 45 000 dépôts sur les 150 000 disponibles dans HAL sont à ce jour déjà traités.

Le démarrage tardif de ce projet a nécessité une première conversion de plusieurs dizaines de milliers de documents dont le format était trop ancien pour être pris en compte par le système du CINES.

L'archivage à long terme au CINES

Le Centre Informatique National de l'Enseignement Supérieur est l'un des principaux opérateurs français d'archivage à long terme de supports numériques issus de la communauté scientifique et technique française assurant ainsi la préservation à long terme de son patrimoine.

La mission d'archivage qui lui a été confiée par le Ministère de l'Enseignement Supérieur et de la Recherche lui permet aujourd'hui de proposer une architecture technique et un soutien dans la gestion de projet des établissements qui souhaitent mettre en place une solution de conservation à long terme de leurs données scientifiques, patrimoniales ou administratives. Cette plateforme d'archivage offre une solution mutualisée permettant ainsi à l'ensemble de la communauté de l'enseignement supérieur et de la recherche, de bénéficier de ce service dès lors que le corpus de documents à archiver est reconnu d'intérêt national.

Le partenariat entre le CCSD et le CINES est d'autant plus naturel que la sauvegarde quotidienne des contenus de HAL est réalisée au CINES depuis novembre 2006, dans le cadre du protocole d'accord inter-établissement sur les archives ouvertes.

Contacts :

CCSD : Daniel Charnay : charnay@in2p3.fr

CINES : Olivier Rouchon : olivier.rouchon@cines.fr

Pour plus d'informations :

Sur HAL : <http://hal.archives-ouvertes.fr/>

Sur le CINES : <http://www.cines.fr/>

Sur les activités du TGE-ADONIS : www.tge-adonis.fr

ⁱ Médihal (medihal.archives-ouvertes.fr), est développée avec le CCSD et le Centre national pour la numérisation de sources visuelles avec le soutien du TGE Adonis