

ATHAR

Indexation

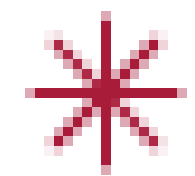
Numérisation

Diffusion enrichie

Un corpus en histoire de l'art

#1 Le projet

Mettre à disposition des outils spécifiques de navigation et d'exploitation scientifique d'un corpus



Faire converger des sources documentaires différentes concernant un édifice de l'art arabe

Développé par InVisu et Persée, le projet Athar [de l'arabe "traces" ou "empreintes" et par extension monuments ou antiquités] met à disposition, en open access, l'intégralité des travaux du Comité de conservation des monuments de l'art arabe avec une indexation fine des toponymes

Perspectives scientifiques

Historiographie des arts de l'Islam

Histoire de la restauration, de la patrimonialisation et de la transformation des monuments en Égypte

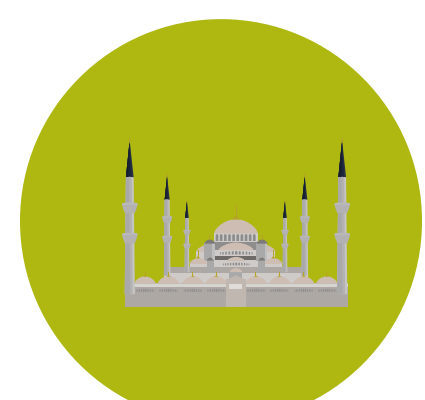
Histoire sociale et culturelle par la compréhension de la gestion du patrimoine architectural arabe

Histoire des relations entre les habitants du Caire, le Comité de conservation et le bâti monumental

#2 Le corpus



Le Bulletin du Comité de conservation des monuments de l'art arabe



Publication contenant la totalité des procès-verbaux et rapports du Comité de conservation des monuments de l'art arabe, instance créée pour inventorier, décrire et restaurer les édifices islamiques puis coptes du Caire



1882 - 1953

Publication annuelle

8 000 pages

800 photos et illustrations

600 monuments décrits



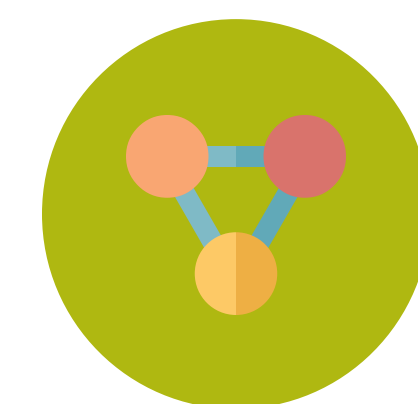
Des collections éparées et incomplètes

Des documents fragiles imprimés sur du papier acide

Des reliures de mauvaise qualité

De multiples variantes des toponymes arabes issues de la translittération en caractères latins

#3 L'indexation



Le Cairo Gazetteer

Un référentiel multilingue en arabe, en français et en anglais pour les toponymes

sous la forme d'une ontologie qui recense les variantes orthographiques de ces toponymes

qui permet d'identifier, de décrire et de localiser les 600 édifices classés du Caire

et d'aligner les entités décrites avec Geonames, IdRef, data.bnf.fr, les Subject Headings de la Bibliothèque du Congrès et DBpedia



L'encodage TEI

Un module spécifique développé par Persée et implémenté dans la plateforme de gestion de la production

qui permet d'encoder les alignements vers IdRef pour les auteurs des articles

et les alignements vers le Cairo Gazetteer pour les monuments décrits dans les articles (balisage au niveau des titres, des intertitres et des illustrations)

#4 Les formats



Numérisation

Respect du format original des documents

Numérisation de toutes les pages

400 dpi couleurs ou niveaux de gris



Structuration

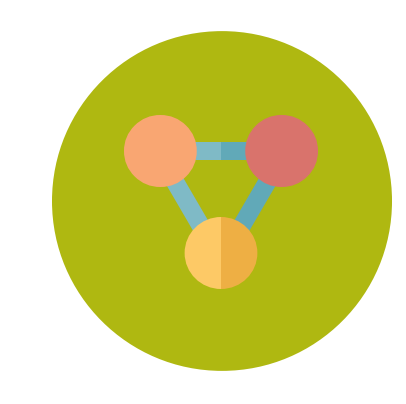
Saisie des notices des édifices avec Gingo

Métadonnées encodées en Dublin Core, MarcXML et MODS

OCR et structure du document en TEI

Container METS

Jeux de données en RDF



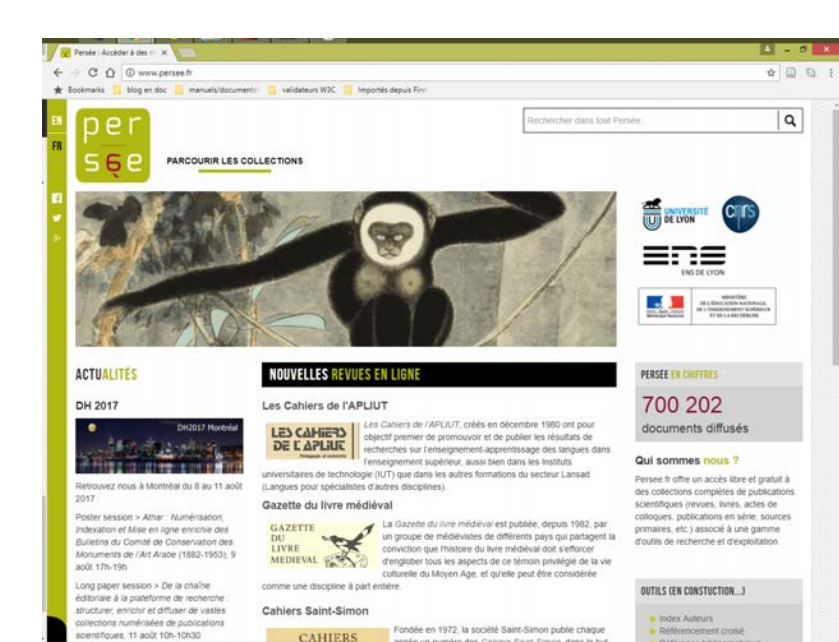
Indexation enrichie

Données du Cairo Gazetteer en Skos

Exploitation des entités nommées dans le corpus en TEI

#5 La diffusion

Juillet 2016. le portail Persée



<http://www.persee.fr/collection/ccmaa>

Diffusion de la totalité de la collection en open access au même niveau que les autres collections du portail

Feuilletage par volume (classement par date d'exercice) et par article ou accès via le moteur de recherche du portail

Au niveau du volume, listes des auteurs et des figures

Au niveau de l'article et des illustrations, lien vers la notice du monument dans le Cairo Gazetteer

Index Auteurs pour la collection

Recherche par facettes (langue, type de document, date de publication, auteur) et affichage par mots-clés

Février 2017. Le triplestore de Persée



<http://data.persee.fr>

Diffusion des données sous la forme d'un graphe RDF enrichi des alignements vers IdRef et le Cairo Gazetteer

Accès via un SPARQL Endpoint et via Sparklis

Téléchargement du jeu de données complet au format RDF

Décembre 2017. La Perséide



<http://athar.persee.fr>

Un site web dédié au corpus avec une éditorialisation propre

Les mêmes fonctionnalités de navigation et de recherche que sur le portail Persée

Des pages composites pour chaque auteur et monument qui agrègent l'ensemble des données relatives à ces autorités qu'elles soient produites par Persée ou récupérées via les alignements

L'ajout, à terme, de nouveaux corpus portés par le même partenaire

#6 Les partenaires



Valorisation du patrimoine scientifique

Unité mixte de services de l'Université de Lyon, de l'ENS de Lyon et du CNRS depuis 2013

14 années d'expérience et de savoir-faire en numérisation, description et diffusion de publications scientifiques

Diffusion de collections patrimoniales complètes : revues, livres, actes de congrès

700 000 documents en texte intégral et en open access

Pour les chercheurs en SHS

<http://www.persee.fr>



L'information visuelle et textuelle en histoire de l'art : nouveaux terrains, corpus, outils

Unité de service et de recherche de l'INHA et du CNRS depuis 2008

Veille, accompagnement, expérimentation et conseil en matière d'acquisition, de traitement et de mise à disposition de l'information

Ancrage thématique dans le domaine orientaliste

Recherches sur l'histoire de l'architecture et du patrimoine en Méditerranée aux XIXe et XXe siècles

<http://invisu.inha.fr>