Toul'AO :
chantier alignement référentiels auteurs et structures

TOUL' AO TOULOUSE Archives Ouvertes

Quel contexte ?

Toul'AO : Groupe de travail « Toulouse Archives Ouvertes »

SICD – réseau des bibliothèques de Toulouse, UFTMiP – Université Fédérale Toulouse Midi-Pyrénées

un « réseau » d'établissements d'enseignement supérieur qui a pour objectifs :

- Promouvoir la diffusion des articles scientifiques en accès ouvert
- Développer la collaboration et le partage d'expériences (enjeux de la science ouverte : publication, édition scientifique ouverte et données de la recherche)
- Organiser des manifestations autour de la science ouverte (Open'Eval en 2016, « Déconfinons la Science ! », Open Access Week en 2021)
- Carnet Hypothèses.org (OpenEdition) https://openarchiv.hypotheses.org

Objectifs du chantier

Groupe de Travail Toul'AO « Référentiels auteurs » lancé en 2019

- Recenser les identifiants auteurs utilisés (champs disciplinaires)
- Intégrer massivement ces identifiants à IdRef
- Contrôler et enrichir les notices IdRef

Méthodologie du chantier

1ère phase

Septembre 2019-juin 2020

- Travail de collecte des identifiants par les différents établissements du réseau Toul'AO, poursuivi jusqu'en début 2021 pour certains établissements

2ème phase

Juillet-août 2020

- Injection des identifiants par l'ABES dans les notices IdRef

3ème phase

Opération qualité des données et enrichissement des notices IdRef

- S'assurer dans un 1er temps que les EC de la communauté toulousaine ont bien tous leur IdRef
- Enrichir les notices selon les recommandations de l'ABES

Le chantier à l'échelle de l'UFTMIP

Bilan de l'opération technique à l'échelle du site toulousain (septembre 2020) :

3960 enseignants-chercheurs concernés par ce projet

Identifiants Chercheurs

- 823 IdHal en entrée => 328 injectés (les autres étaient déjà présents dans les notices)
- 1 721 ORCID en entrée => 1 297 injectés
- 127 univ-droitID en entrée => 3 injectés
- 1 653 Scopus ID en entrée => 1 653 Scopus ID injectés

Identifiants Structures

- 42 ISNI en entrée => 40 injectés
- 116 RNSR en entrée => 9 injectés
- 114 IdHal en entrée => 33 injectés

Qu'est-ce que c'est ?

Depuis 2010 : outil d'ouverture des données d'autorités du Sudoc

- Visibles sur le web
- Réutilisables par tous

Réferentiel à destination des professionnels qui permet :

- D'identifier, discriminer, douter
- Outil de production en temps réel
- Possibilité d'amélioration et rectification par la communauté

Ouvert désormais à d'autres contributeurs que le Sudoc

Le chantier à l'échelle de l'UT3

- 68 structures de recherche en majorité multi-tutelles
- 1839 chercheurs et 1397 doctorants

Le chantier, sur sa phase 1, s’est déroulé du 1er janvier 2020 au 4 mars 2021

Bilan de l’opération au 4/03/2021 :

- 53 laboratoires traités
- 3 257 chercheurs ont vu leur notice IdRef enrichies grâce à ce travail

Les chiffres à l'échelle de l'UT3

	Analyse au 04/03/21	en % du nbre de chercheurs mars 2021
Nombre de laboratoires	53	
Nombre de chercheurs avec un ppn et un autre identifiant	3257	
Nombre d'IdHAL	518	16%
Nombre d'ORCID	2242	69%
Nombre de ResearcherID	865	27%
Nombre de ScopusID	3177	98%
Chercheurs ayant un IdHAL et un ORCID	429	13%
Chercheurs ayant un IdHal et Researcher ID	168	5%
Chercheurs ayant un IdHAL, un ORCID et un ResearcherID	163	5%
Chercheurs ayant un ORCID et un Researcher ID	792	24%

La méthodologie côté UT3

Constitution de la liste des chercheurs

Frein : pas de liste fournie par la DRH pour ne traiter que les chercheurs de l'UT3

Conséquence : besoin de créer d'une méthodologie pour récupérer les noms des chercheurs affiliés aux laboratoires UT3

Utilisation de l'outil « Extracteur Auteurs » créé par Philippe Gambette

Période choisie : 2018-2020

Champs extraits : fullName_s,lastName_s,idHal_i,idHal_s,orcid_id

Téléchargement du 2e tableau intitulé « Tableau des formes auteur sans doublons »

La méthodologie côté UT3

Collecte des identifiants et vérification de l'appartenance des chercheurs à l'UT3

- Ouverture dans un même navigateur d'IdRef, d'ORCID, de Publons, de Scopus et de l'annuaire du laboratoire traité
- Le nom complet de l’auteur est copié dans chaque base pour rechercher les identifiants à collecter et contrôler les données
- Les identifiants sont centralisés dans le tableau issu de l'extracteur

Amélioration des notices IdRef

- Ajout de la note 340 sur la biographie et l'activité et de la note 810 sur la source consulté avec profit

Préparation des séminaires techniques autour de la science ouverte

Création du service Science ouverte en à l'automne 2019 et construction du premier séminaire sur l'identité numérique au printemps 2020

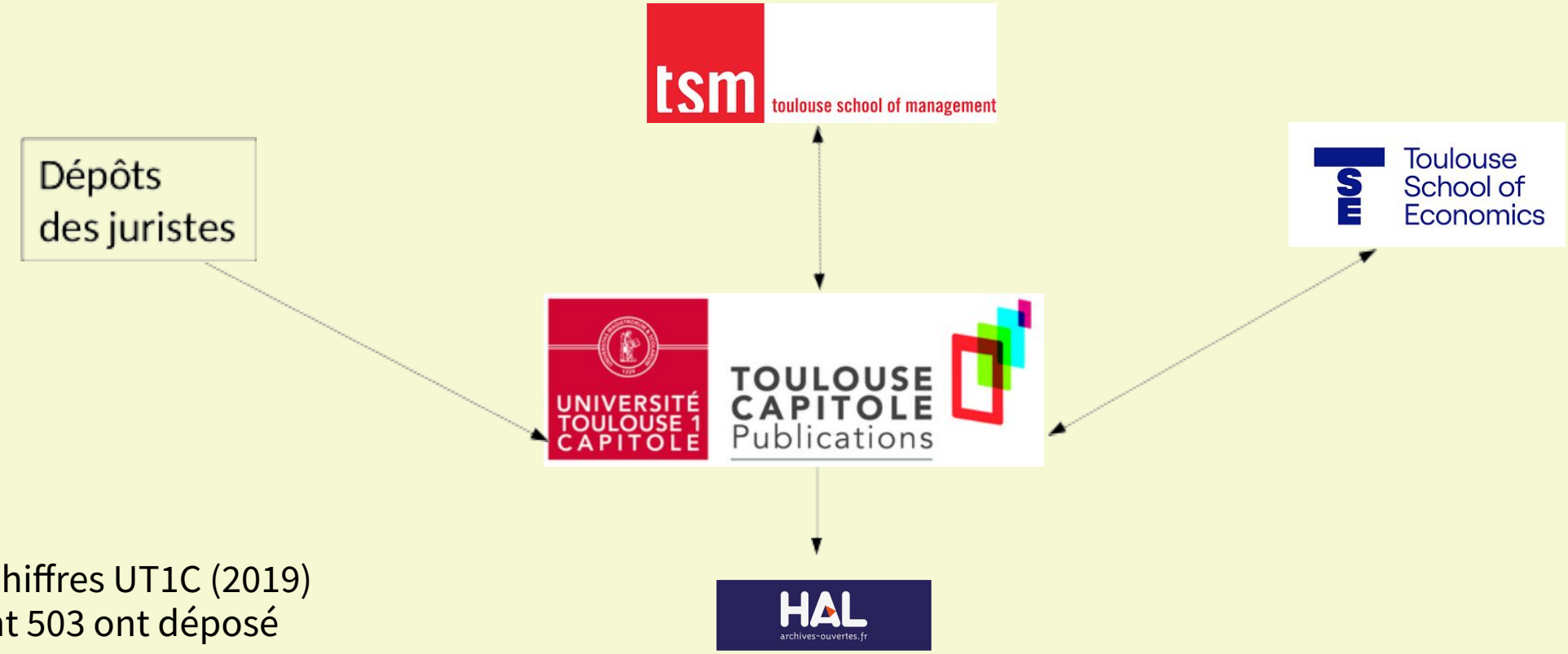
La collecte des données à permis :

- un état des lieux des identifiants pour la communauté
- un dialogue sur des bases concrètes avec les chercheurs

Actuellement deux séminaires techniques autour de l'identité numérique sont proposés :

- Déposer dans HAL : astuces pour faciliter le dépôt (dont création de l'IdHAL)
- Créer son profil ORCID (dont remonter ses publications depuis HAL)

Contexte UT1



Quelques chiffres UT1C (2019)
560 EC dont 503 ont déposé dans l'AO
12 unités de recherche

Contexte à l'UT1C

2019 : lancement du chantier d'alignement de Toulouse Capitole publications avec IdRef

Objectifs :

- En finir avec les doublons dans l'index auteur de l'archive institutionnelle
- Améliorer l'identification des enseignants-chercheurs de l'université qui ont déposé

⊠ 2 chantiers complémentaires à mener en parallèle

Acteurs des chantiers : ABES, informaticien UT1C, correspondante autorité UT1C, service des publications de l'UT1 (3 pers. sur l'enrichissement des identifiants)

Le chantier identifiants UT1C en chiffres

	Analyse au 31/12/2020	en % du nbre de chercheurs UT1	en % du nbre de chercheurs UT3
Nombre de laboratoires	12		
Nombre de chercheurs UT1C avec un IdRef	560	100%	
Nombre d'IdHAL	23	4,1%	16%
Nombre d'ORCID	130	23%	69%
Nombre de ResearcherID	17	3%	27%
Nombre de ScopusID	0	0%	98%
Nombre de Univ-droit	127	23%	0%
Nombre de Repec	147	26%	0%

Aligner pour quels résultats ?

Depuis janvier 2021 : mise en production

- Création d'IdRef et amélioration des données d'autorité au fil de l'eau
- Au-delà de l'objectif : création systématique d'un IdRef ou mise à jour pour tout nouvel EC à l'UT1C qu'il dépose ou pas dans l'AO

  - Moins de doublons dans l'index auteur pour nos enseignants-chercheurs

  - Meilleure identification (exemples Pierre Dubois)

Conclusion

Ce chantier a permis :

- Meilleure identification des enseignants-chercheurs à l'échelle du site toulousain
- Amélioration du taux de couverture d'IdRef à l'échelle du site
- Quantification et identification des différents identifiants
- Meilleure connaissance des usages des identifiants par les enseignants-chercheurs
- Harmonisation des pratiques à l'échelle du site concernant IdRef
- Dynamique de travail autour des référentiels (Orcid)
- Poursuite de la collaboration Toul'AO/ABES

TOUL' AO TOULOUSE Archives Ouvertes

Des questions ?

Alicia Benson-Rumiz
alicia.benson-rumiz@univ-tlse3.fr

Marie-Françoise Bremond
marie-francoise.bremond@ut-capitole.fr

Carole Vaccari
carole.vaccari@ut-capitole.fr