HAL, atelier Référentiels

Journées CasuHal

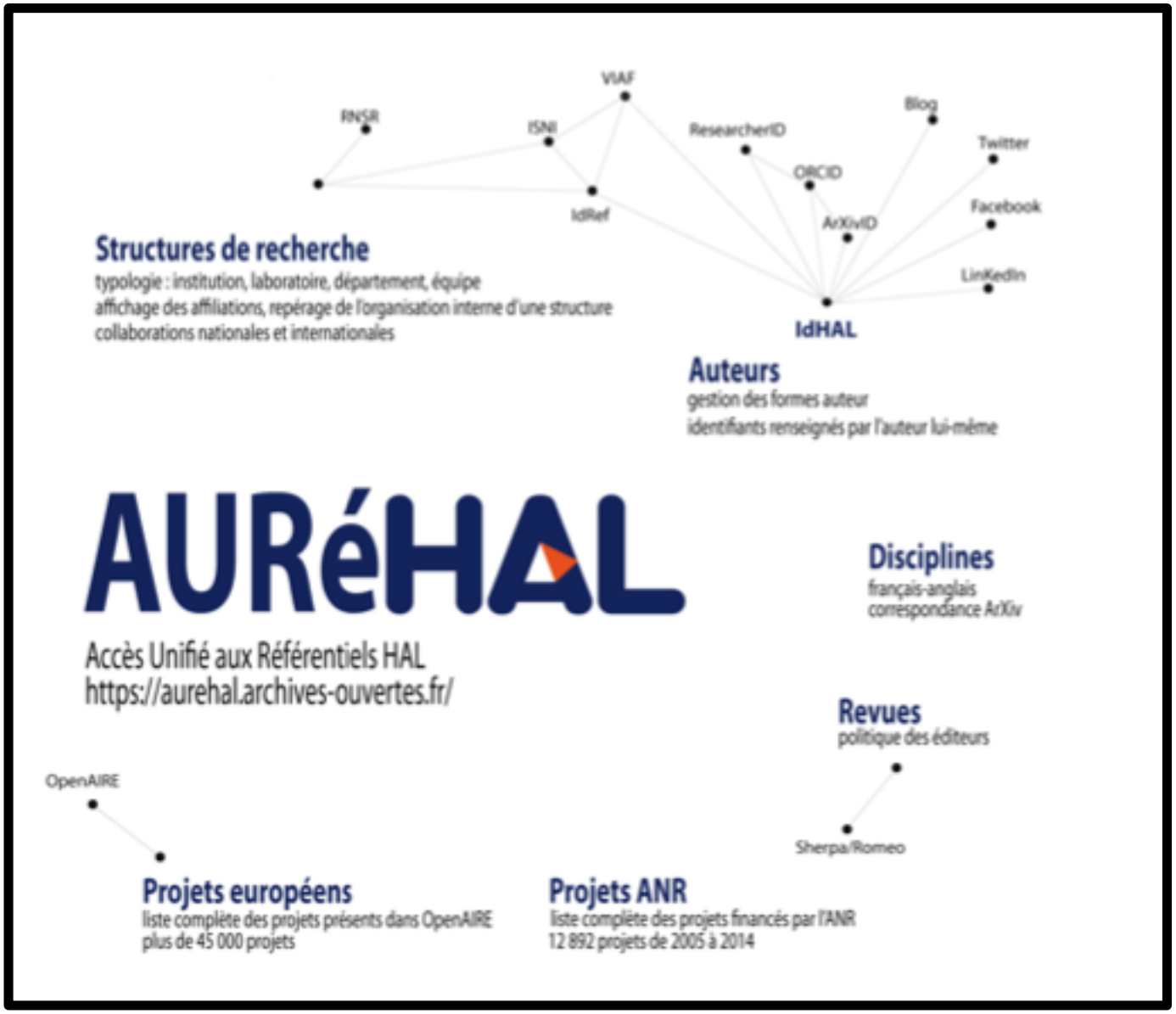
MSH Dijon 31 mai et 1er juin 2018

Bénédicte Kuntziger (CCSD)

CCSD Centre pour la Communication Scientifique Directe

La Science partagée

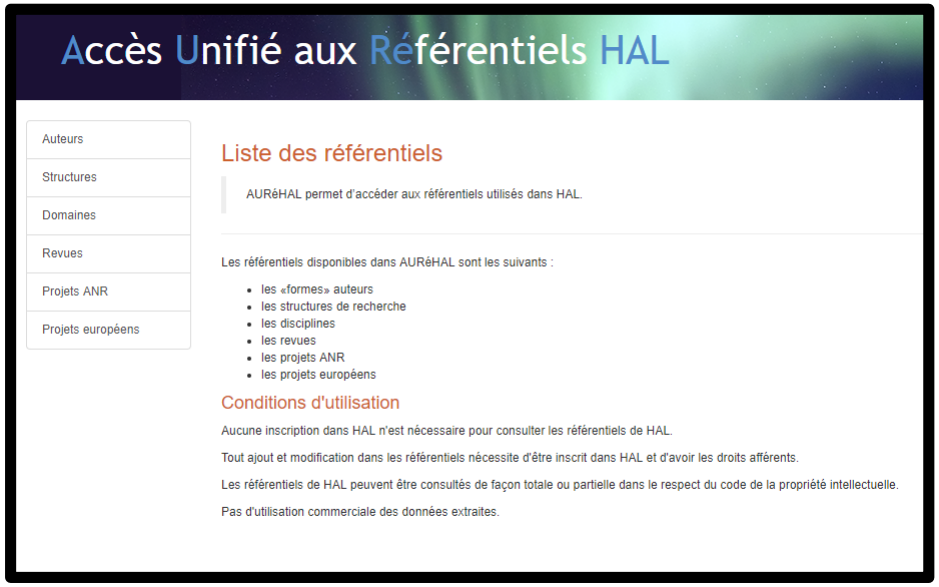
LES RÉFÉRENTIELS DE HAL



AURéHAL : DES RÉFÉRENTIELS PARTAGÉS

Création
Dépôt
Administrateur
Import référentiel

Curation
Auteur (idHAL)
Administrateur
Référent structure



Alignement des référentiels
idRef (auteur, structure)
ORCID, ISNI, VIAF, RNSR
ResearcherID, arXivID

Les services construits sur les données des référentiels

Auteur
Page personnelle

Structures
Portail institutionnel
Collection de laboratoire, d'équipe

Domaines
HAL-SHS
Collection

Revues
Informations via API
Sherpa/Romeo

Projet ANR
Collection

Projet européen
Référencement
OpenAIRE
Collection

- Un des axes de la Feuille de Route du CCSD de 2016
- = Positionner **AURéHAL dans le réseau des référentiels** en cours de construction au niveau national (IdRef, RNSR, …), européen et international

AURÉHAL : ACCÈS UNIFIÉ AUX RÉFÉRENTIELS HAL

https://aurehal.archives-ouvertes.fr

- Auteurs
- Structures de recherche
- Disciplines
- Revues
- Projets ANR
- Projets Européens

Référentiels : caractéristiques communes (sauf disciplines)

- Les référentiels peuvent être consultés, même sans authentification
- L'historique de chaque entrée est visible par toute personne authentifiée : date de création, de mise à jour
- Chaque entrée d'un référentiel permet de lancer une requête dans HAL pour consulter les dépôts liés
- Une mise à jour dans un référentiel génère automatiquement une mise à jour dans les dépôts liés

RÉFÉRENTIELS : PRINCIPES

- Alimentation du référentiel : les nouvelles entrées des référentiels sont créées lors de la saisie au moment du dépôt dans HAL (à l'exception des disciplines)
- (Ou création dans AuréHAL)
- Un administrateur de portail a les droits de création et modification sur tous les référentiels
- Toute personne qui a les droits ‘référent structure’ peut modifier les informations de la structure de recherche dont il est référent (+ structures filles)

RÉFÉRENTIELS : TYPES D'ENTRÉES

- 3 types d'entrées dans le référentiel :
  - Incoming : à valider
  - Valid : formes valides
  - Old : formes 'anciennes'
- Objectifs du travail dans le référentiel
  - Utiliser des formes validées pour exploiter les données
  - => Éliminer les formes incoming par des fusions
- Principes communs :
  - Trouver toutes les formes entrées correspondant à la même entité
  - En compléter et valider une
  - Fusionner les formes non valides avec celle-ci

OUVERTURE DES DONNÉES DES RÉFÉRENTIELS

- API
  - https://api.archives-ouvertes.fr/docs
  - pour chaque référentiel : liste des métadonnées
- Exemple, dans le référentiel auteur :

lastName_t:frobert AND hasCV_bool:true

OUVERTURE DES DONNÉES DES RÉFÉRENTIELS

- Open data : RDF format

https://data.archives-ouvertes.fr

Le RDF est un modèle de graphe destiné à décrire de façon formelle les ressources Web et leurs métadonnées, de façon à permettre le traitement automatique de ces descriptions

Chaque élement (document, laboratoire, ...) est représenté sous le forme de triplets RDF (**sujet, prédicat, objet**)

Un graphe a été construit pour chaque référentiel d'AuréHAL

RÉFÉRENTIEL DISCIPLINES

- Accessible uniquement en consultation
- Référentiel « maison », aligné sur ArXiv pour les domaines concernés.
- La classification scientifique est faite par les chercheurs et pour les chercheurs. Le but recherché est d'avoir un maximum de 2 niveaux arborescents contenant une trentaine d'item au maximum pour une discipline
- Disciplines spécifiques à un portail (+interdisciplinarité)

RÉFÉRENTIEL PROJETS ANR

- Fichiers csv
- Importés dans le référentiel

RÉFÉRENTIEL PROJETS EUROPÉENS

- Récupérés via OpenAire
- OpenAire moissonne HAL
- Si projet européen, alors identification dans Openaire

RÉFÉRENTIEL REVUES

- Lien via API de Sherpa-Roméo

RÉFÉRENTIEL AUTEUR

- Création d'une forme auteur : champs
  - Prénom * et Nom * + autres prénoms (initiales)
  - Adresse courriel
  - URL site personnel
- Établissement d'appartenance = institution dont dépend le chercheur
- Code couleur lors du dépôt (et bientôt dans le référentiel)
- Statuts : Incoming/ Valid/ Old (couleur au dépôt)
  - Valid : forme auteur par défaut d'un IdHAL (vert)
  - Old : autre forme auteur d'un IdHAL (jaune)
  - Incoming : forme auteur sans IdHAL (rouge)

FORMES AUTEUR

Consultation des formes auteurs

Ce module vous permet de consulter la liste des formes auteurs.

fressengeas | Rechercher

Voir ▾ | Trier ▾ | Afficher ▾

ind	idHAL	Nom	Prénom	Domaine e-mail	ACTIONS
62740	0	Fressengeas	C.		
396830	0	Fressengeas	Claude		
134976	309	Fressengeas	N.	fressengeas.net	
173145	309	Fressengeas	Nicolas	fressengeas.net	
62690	0	Fressengeas	Claude	univ-metz.fr	
528679	309	Fressengeas	Nicolas	fressengeas.net	
1006847	309	Fressengeas	Nicolas	fressengeas.net	
783100	309	Fressengeas	Nicolas	fressengeas.net	
124665	309	Fressengeas	Nicolas	Supelec.fr	
550111	309	Fressengeas	Nicolas	fressengeas.net	

FORMES AUTEUR

Fiche d'une forme auteur

idhal

Identifiants extérieurs

ALIGNEMENTS FORMES AUTEUR

- Via la création de l'IdHal par le chercheur
- Exception : dépôt de thèses par l'ABES (STAR) => alignement avec l'IdRef de l'auteur (forme auteur)
- Alignement avec identifiant Orcid : permet ensuite d'utiliser Orcid pour se connecter à HAL

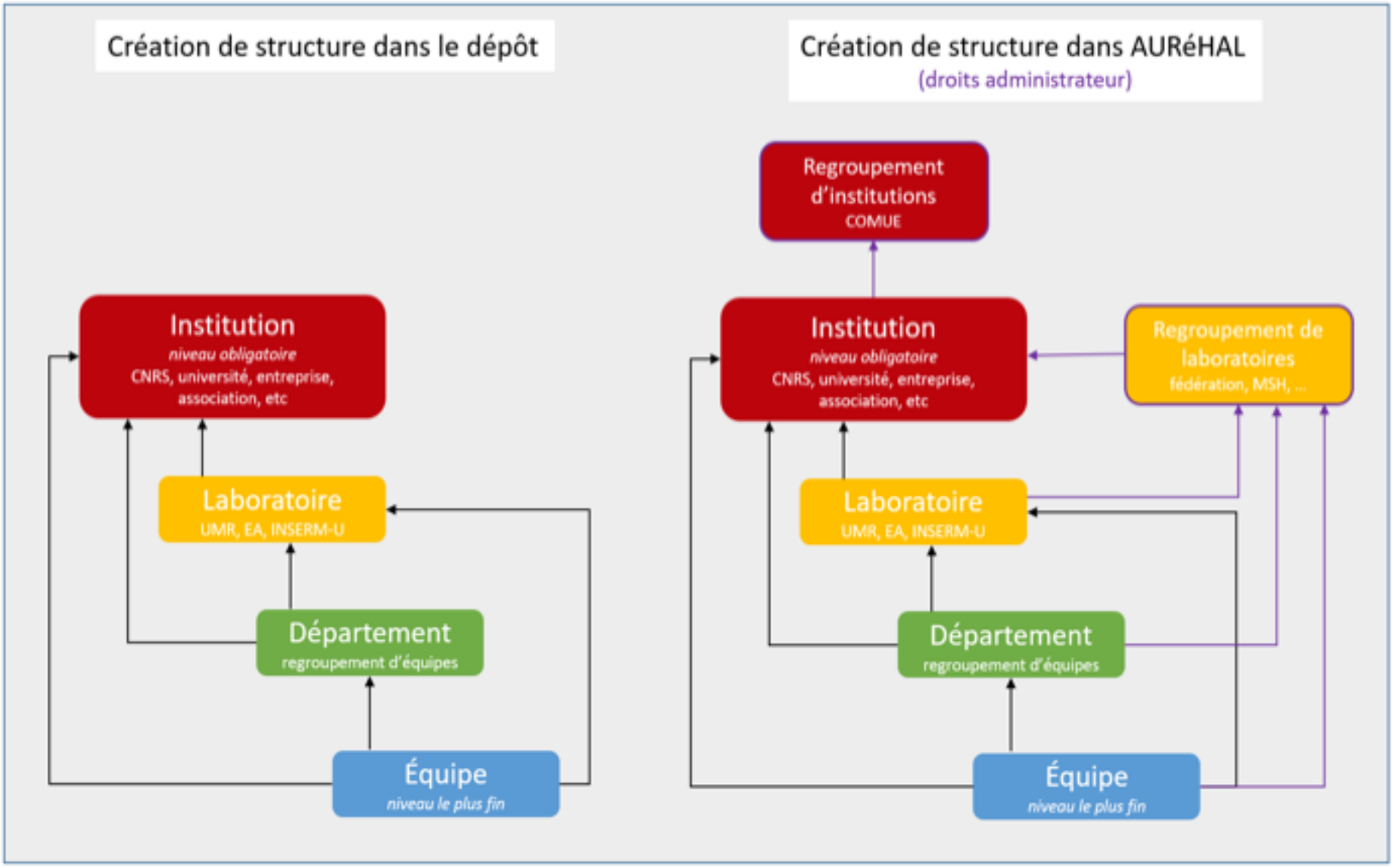
‘FUSION’ DES FORMES AUTEUR

- Démarche complémentaire de celle de l’auteur
- Par l’administrateur de portail
- Remplacer des formes auteur qui identifient le même auteur
- Séparer des publications d’homonymes rassemblées sous une seule forme auteur

RÉFÉRENTIEL DES STRUCTURES DE RECHERCHE

- 6 niveaux au lieu de 4 : 2 niveaux supplémentaires
  - **Regroupement d'institutions** : COMUE par exemple
  - **Institution** : ex : CNRS, Université de Caen,etc....
  - **Regroupement de laboratoires** : fédérations de recherche, instituts de recherche, ou UFR
  - **Laboratoire** : niveau qui correspond à UMR, EA, INSERM-U
  - **Département (ou Regroupement d'équipes de recherche)**
  - **Equipe de recherche** : niveau le plus fin.

RÉFÉRENTIEL STRUCTURES



RÉFÉRENTIEL STRUCTURES

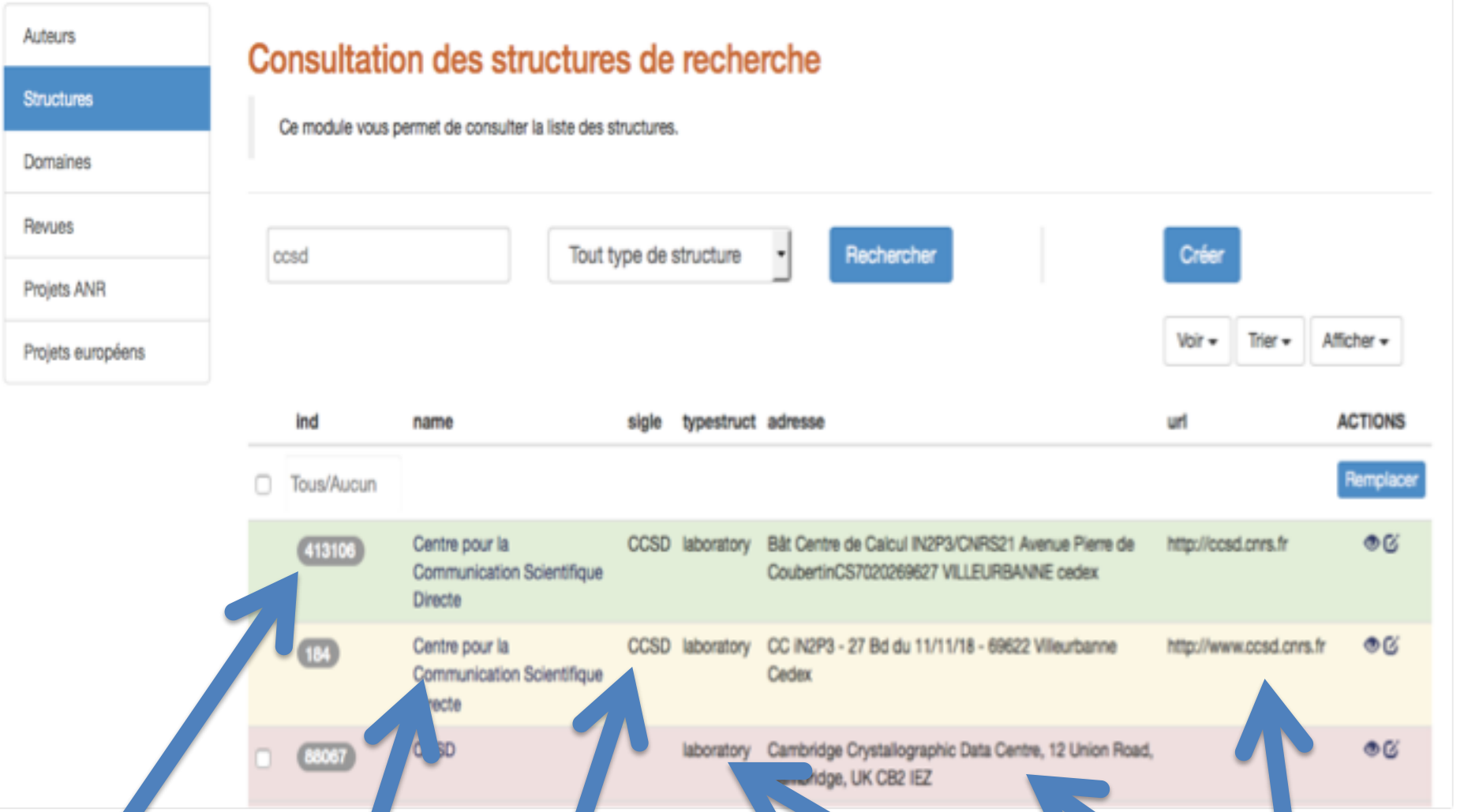
- Ces nouvelles structures ne peuvent pas être créées lors du dépôt
- Elles sont créées dans AuréHal par des admin
- Puis utilisées lors du dépôt (sélection liste auto-complétion
- Et affichées dans le dépôt (détail des affiliations)
- **Seul le niveau Institution est obligatoire**

RÉFÉRENTIEL STRUCTURES

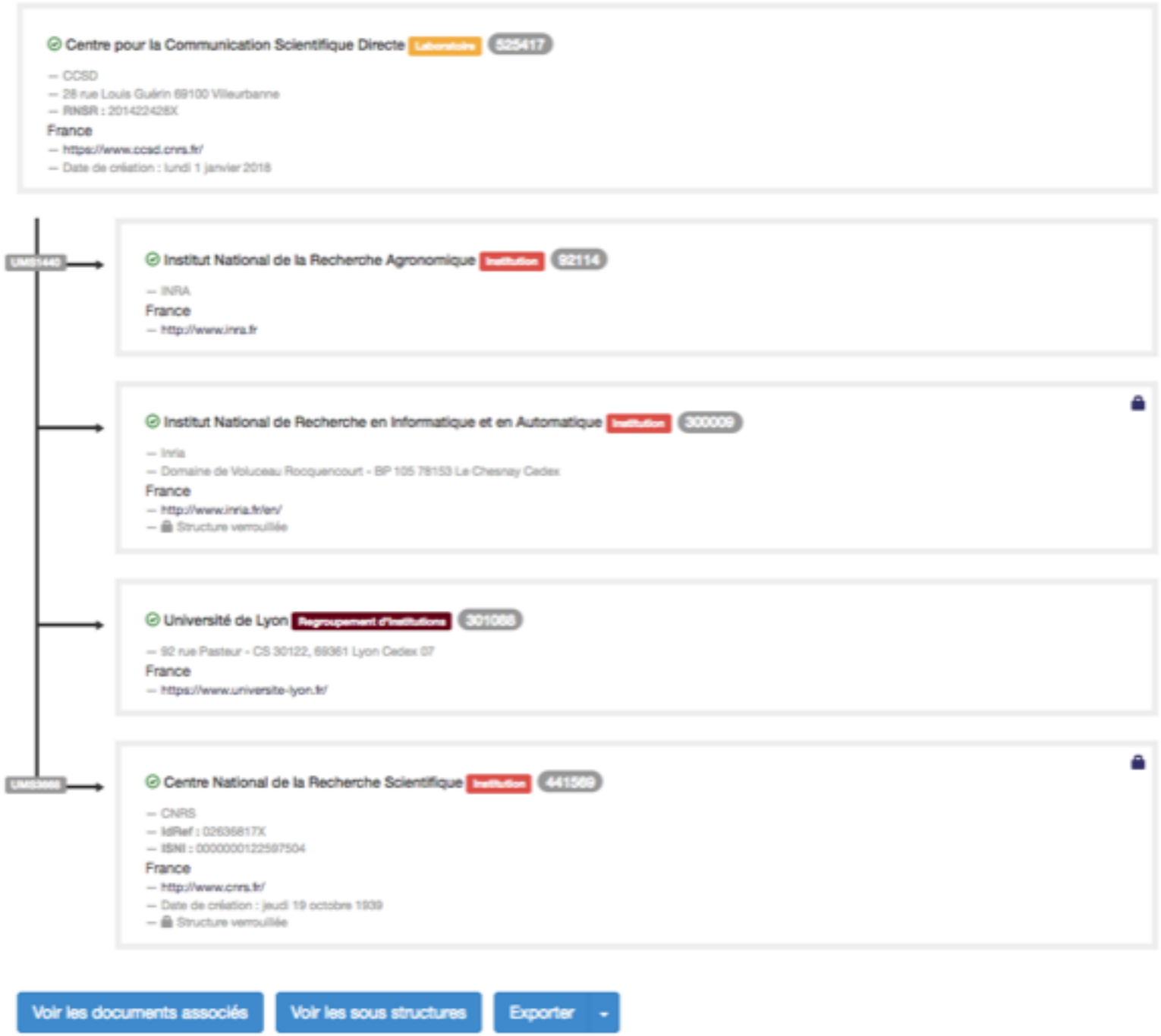
- Création d'une structure : Champs
  - Nom *
  - Sigle
  - Adresse
  - Pays
  - Type structure * (équipe, département, laboratoire, institution)
  - URL
  - **Date de création**
  - **Date de fermeture/fusion**
  - **Identifiants externes : RNSR (**Répertoire National des Structures de Recherche) **/ IDREF** (SUDOC) **/ ISNI (**International Standard Name Identifier)

- Validité = Statut + code couleur dans AURéHAL et lors du dépôt:
  - Valide (vert)/ N'existe plus (jaune)/ A valider (rouge)

Accès Unifié aux Référentiels HAL



Fiche d'une structure



QUE FAIRE SELON SON RÔLE (PRIVILÈGE) D'UTILISATEUR

- Gestionnaire de collections : pas d'action possible

=> action principale sur sa collection

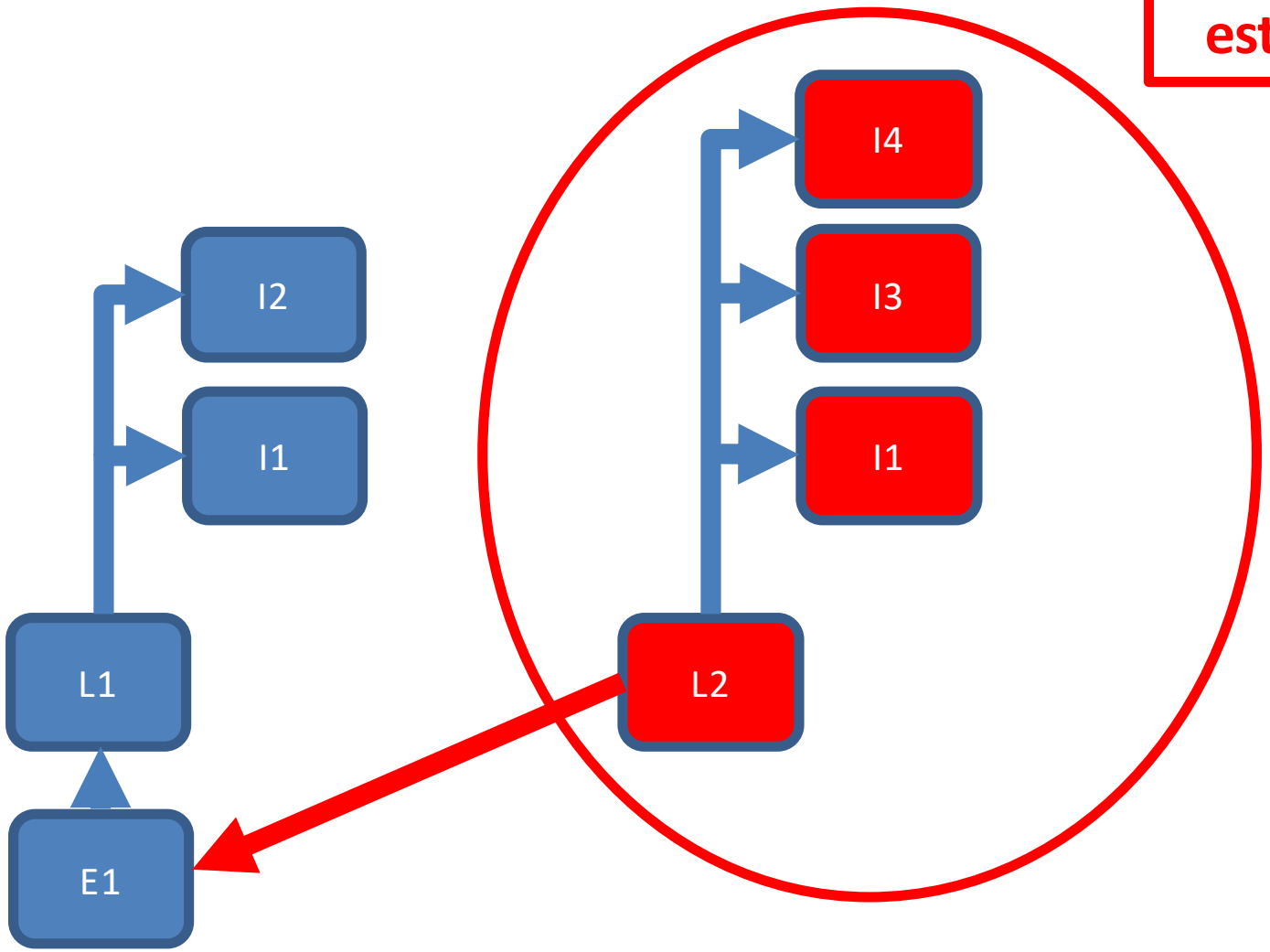
- Référent structure : action très limitée à sa structure (son équipe, son labo) : modification des métadonnées de sa structure

=> action principale sur les dépôts de sa structure

- Administrateur de portail : travail complet dans le référentiel : modification de métadonnées, remplacements de structures, verrouillage

REMPLACEMENT D'UNE STRUCTURE
1ER CAS

Toute la branche est remplacée



EXEMPLE 1

Dépôt ans HAL

Manifold-based mathematical morphology for graph signal editing of colored images and meshes

Olivier Lézoray [1] AuthorId : 224109

Auteur

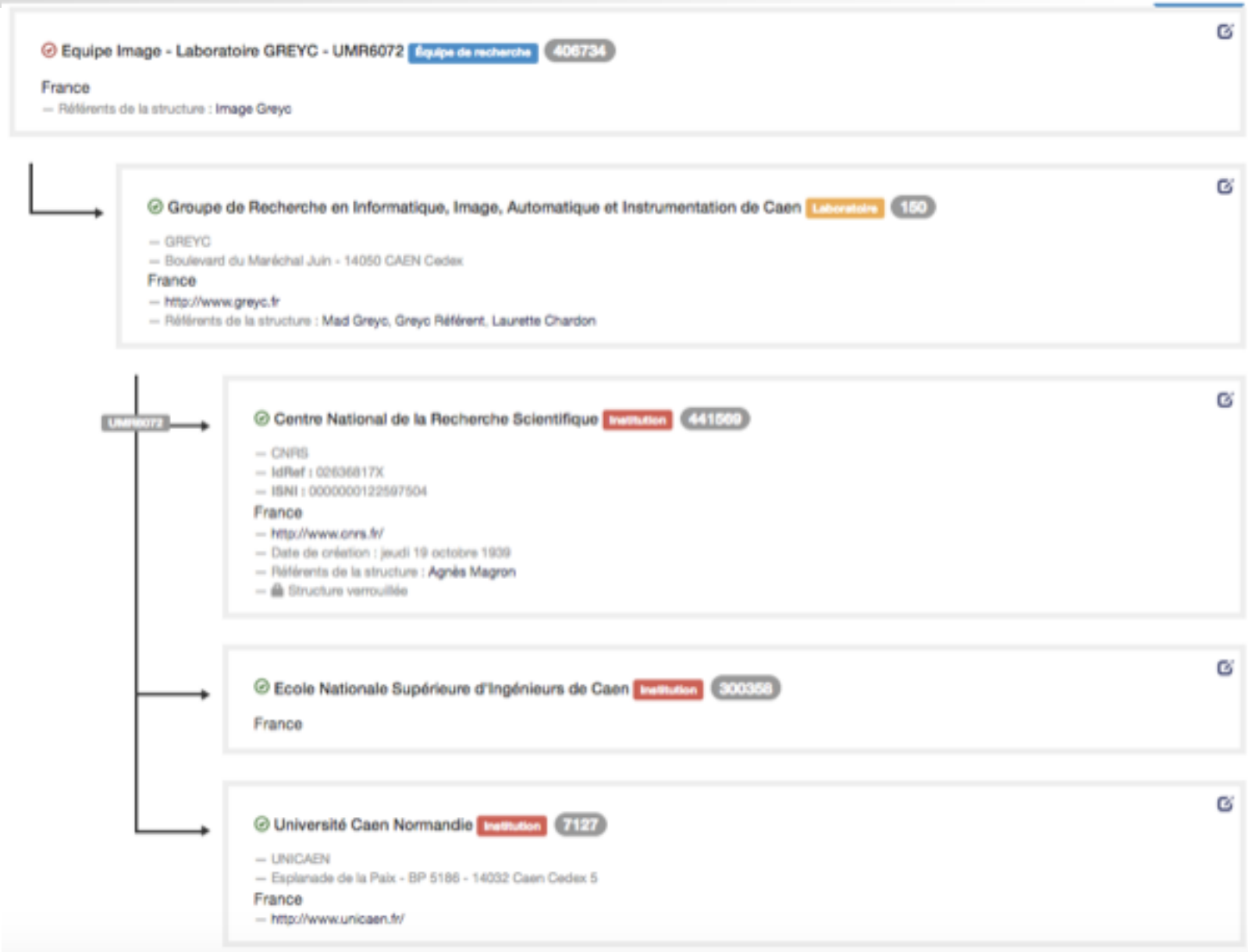
— IdHAL : olivier-lezoray

— ResearcherId : http://www.researcherid.com/rid/K-7326-2014

1 Equipe Image - Laboratoire GREYC - UMR6072 (France) StructId : 406734

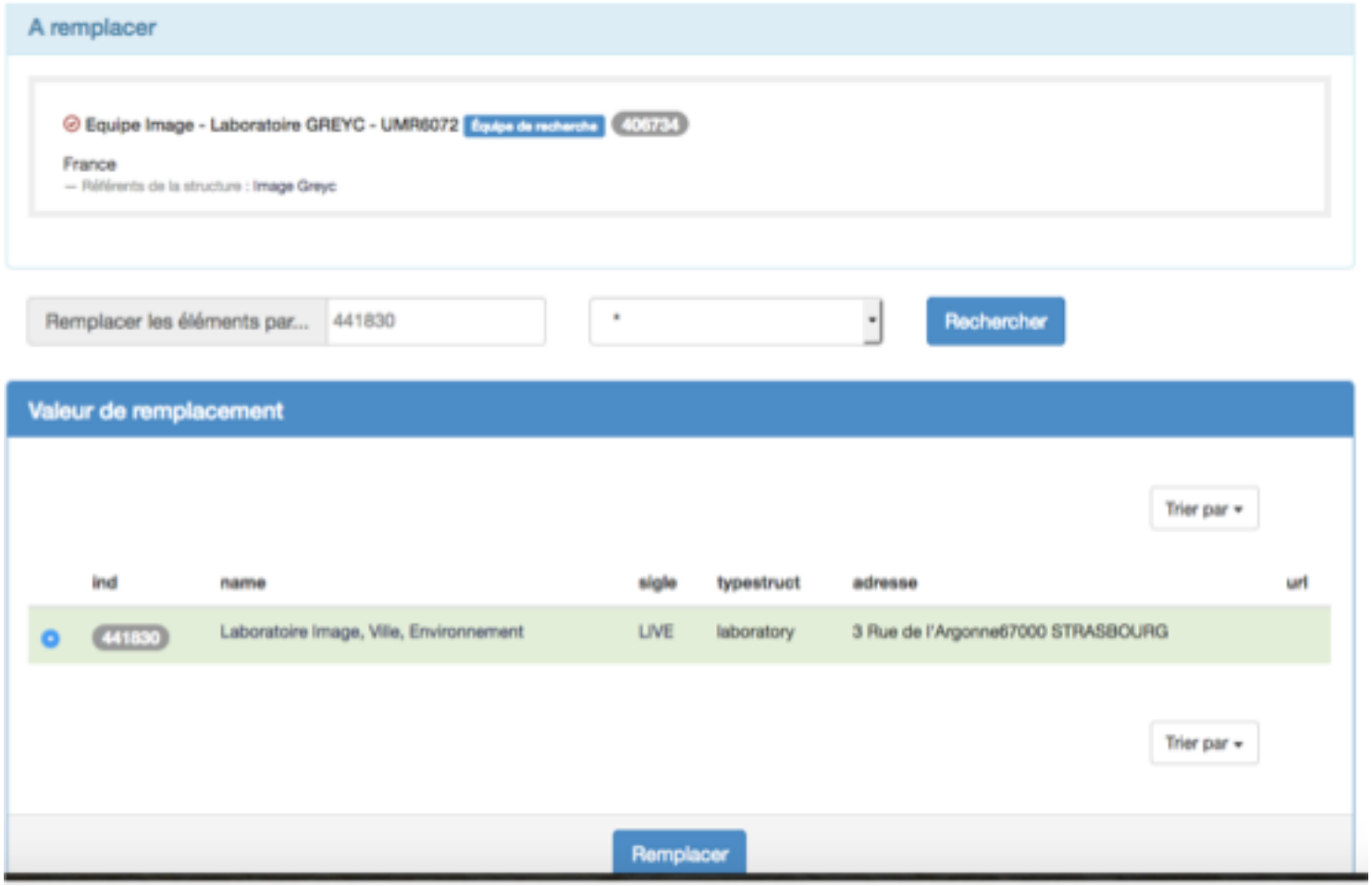
- GREYC - Groupe de Recherche en Informatique, Image, Automatique et Instrumentation de Caen (Boulevard du Maréchal Juin - 14050 CAEN Cedex - France) StructId : 150
  - CNRS - Centre National de la Recherche Scientifique : UMR6072 (France) StructId : 441569
  - Ecole Nationale Supérieure d'Ingénieurs de Caen (France) StructId : 300358
  - UNICAEN - Université Caen Normandie (Esplanade de la Paix - BP 5186 - 14032 Caen Cedex 5 - France) StructId : 7127

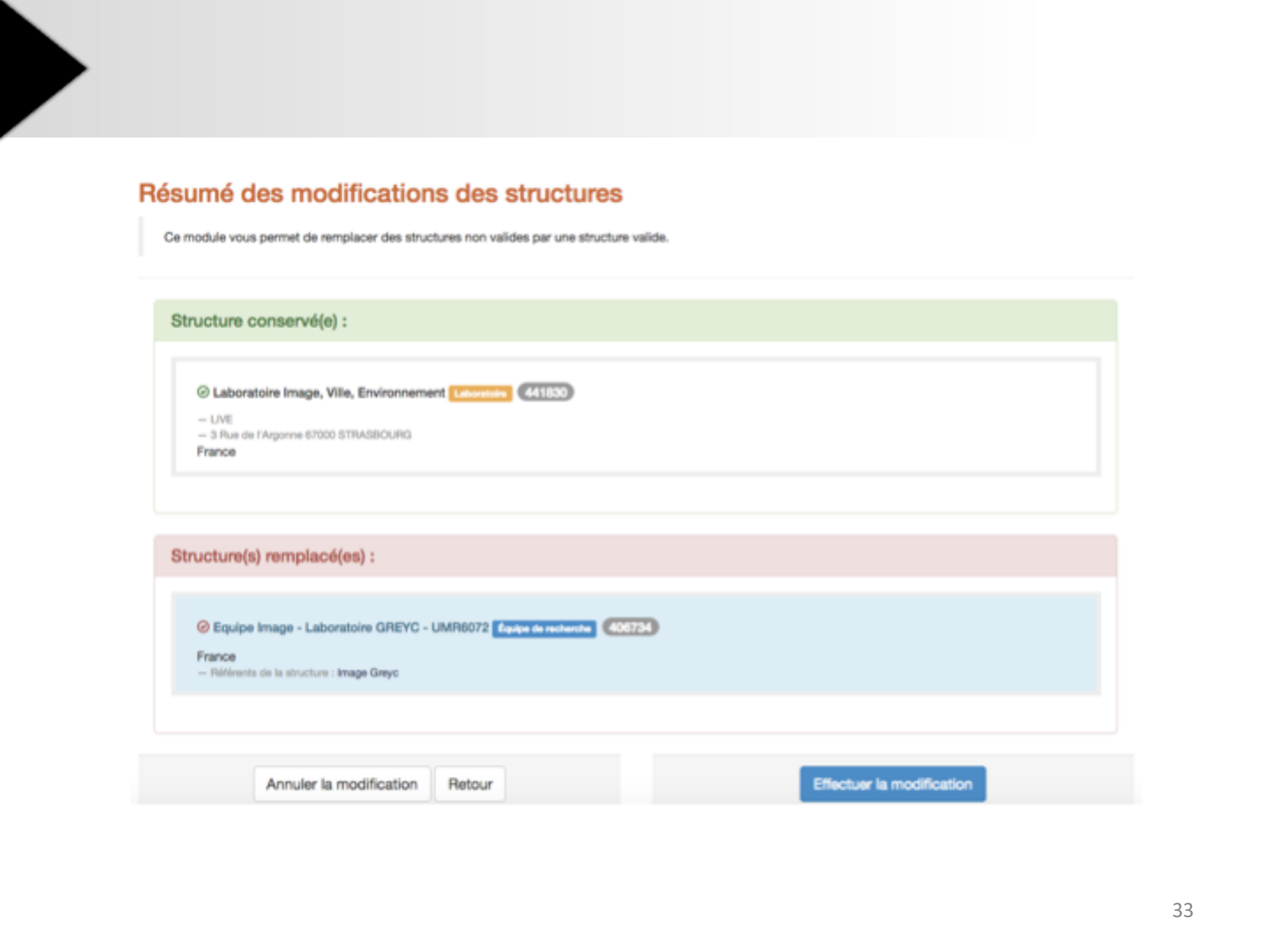
Structure dans le référentiel



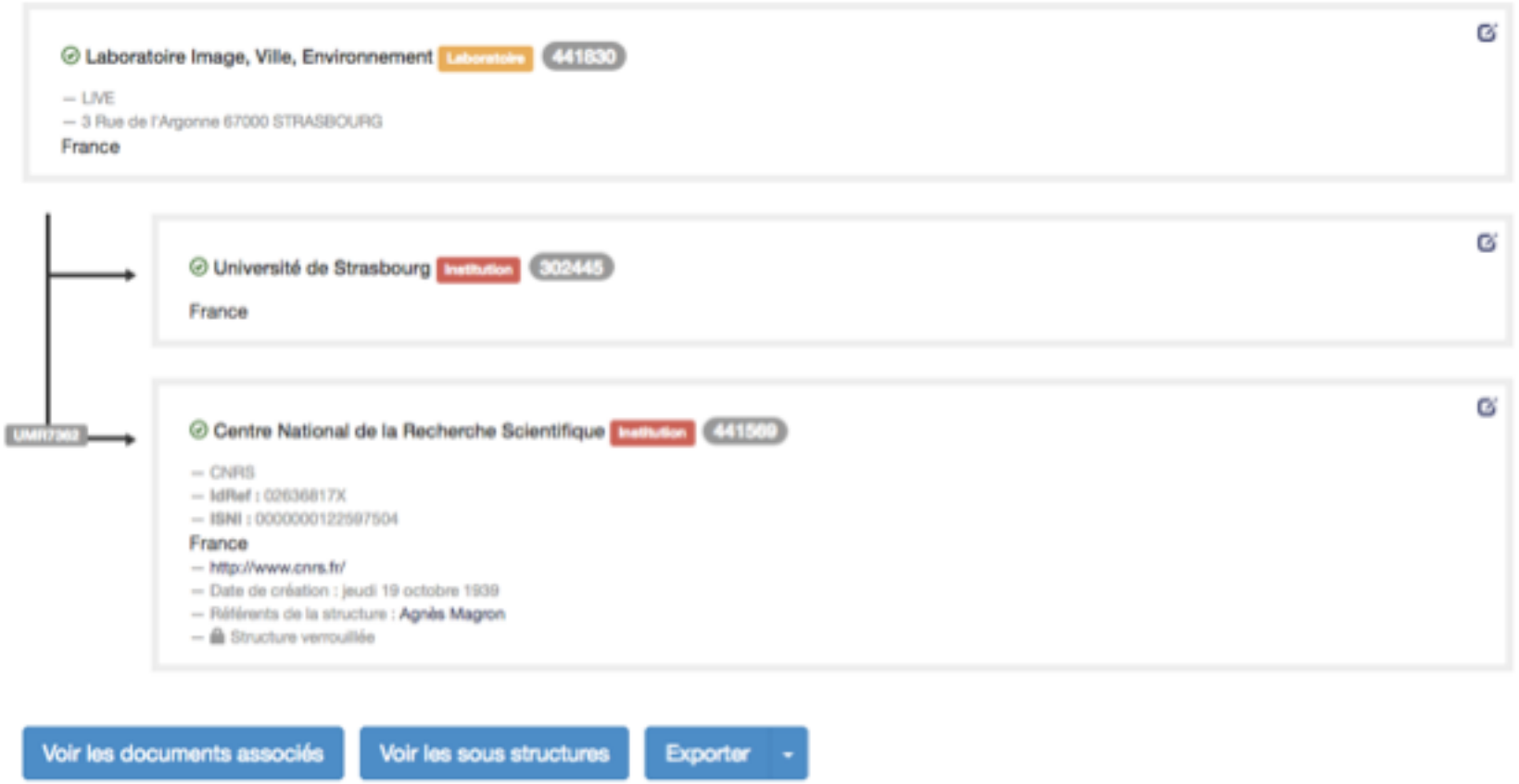
Remplacement dans le référentiel

	ind	name	sigle	typestruct	adresse	url	ACTIONS
☐	Tous/Aucun						Remplacer
	388984	Equipe AMACC - Laboratoire GREYC - UMR6072		researchteam			
	389385	Equipe Automatique - Laboratoire GREYC - UMR6072		researchteam		https://www.greyc.fr/fr/automatique	
	388594	Equipe CODAG - Laboratoire GREYC - UMR6072		researchteam			
	406687	Equipe Electronique - Laboratoire GREYC - UMR6072		researchteam			
	388300	Equipe Hultech - Laboratoire GREYC - UMR6072		researchteam			
	406663	Equipe MAD - Laboratoire GREYC - UMR6072		researchteam			
	150	Groupe de Recherche en Informatique, Image, Automatique et Instrumentation de Caen	GREYC	laboratory	Boulevard du Maréchal Juin - 14050 CAEN Cedex	http://www.greyc.fr	
☑	406734	Equipe Image - Laboratoire GREYC - UMR6072		researchteam			
☐	388352	Equipe Monétique & Biométrie - Laboratoire GREYC - UMR6072		researchteam			
☐	Tous/Aucun						Remplacer


A remplacer
Equipe Image - Laboratoire GREYC - UMR6072
Équipe de recherche
406734
France
— Référents de la structure : Image Greyc
Remplacer les éléments par...
441830
Rechercher
Valeur de remplacement
Trier par
ind
name
sigle
typestruct
adresse
url
441830
Laboratoire Image, Ville, Environnement
LIVE
laboratory
3 Rue de l'Argonne67000 STRASBOURG
Trier par
Remplacer


Résumé des modifications des structures
Ce module vous permet de remplacer des structures non valides par une structure valide.
Structure conservé(e) :
Laboratoire Image, Ville, Environnement
Laboratoire
441830
– LIVE
– 3 Rue de l'Argonne 67000 STRASBOURG
France
Structure(s) remplacé(es) :
Equipe Image - Laboratoire GREYC - UMR6072
Équipe de recherche
406734
France
– Référents de la structure : Image Greyc
Annuler la modification
Retour
Effectuer la modification

Fiche d'une structure



Remplacement dans HAL

Manifold-based mathematical morphology for graph signal editing of colored images and meshes

Olivier Lézoray [1] AuthorId : 224109

Auteur

— IdHAL : olivier-lezoray

— ResearcherId : http://www.researcherid.com/rid/K-7326-2014

1 LIVE - Laboratoire Image, Ville, Environnement (3 Rue de l'Argonne 67000 STRASBOURG - France) StructId : 441830

- Université de Strasbourg (France) StructId : 302445
- CNRS - Centre National de la Recherche Scientifique : UMR7362 (France) StructId : 441569

Manifold-based mathematical morphology for graph signal editing of colored images and meshes

Olivier Lézoray [1] AuthorId : 224109

Auteur

— IdHAL : olivier-lezoray

— ResearcherId : http://www.researcherid.com/rid/K-7326-2014

1 Equipe Image - Laboratoire GREYC - UMR6072 (France) StructId : 406734

- GREYC - Groupe de Recherche en Informatique, Image, Automatique et Instrumentation de Caen (Boulevard du Maréchal Juin - 14050 CAEN Cedex - France) StructId : 150
  - CNRS - Centre National de la Recherche Scientifique : UMR6072 (France) StructId : 441569
  - Ecole Nationale Supérieure d'Ingénieurs de Caen (France) StructId : 300358
  - UNICAEN - Université Caen Normandie (Esplanade de la Paix - BP 5186 - 14032 Caen Cedex 5 - France) StructId : 7127

Olivier Lézoray [1] AuthorId : 224109

Auteur

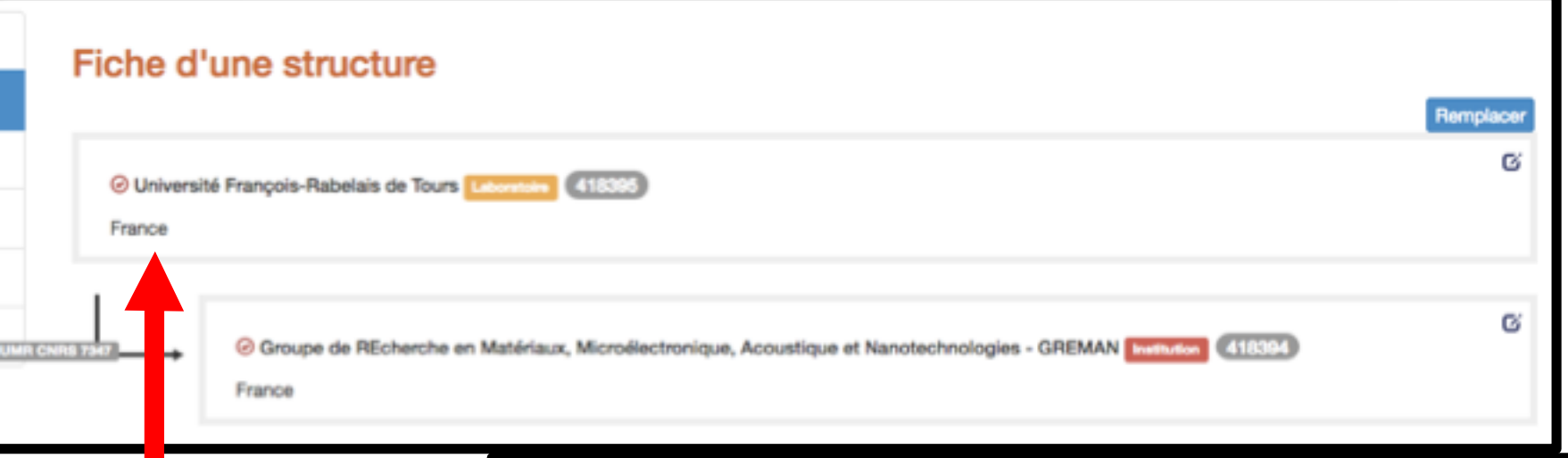
— IdHAL : olivier-lezoray

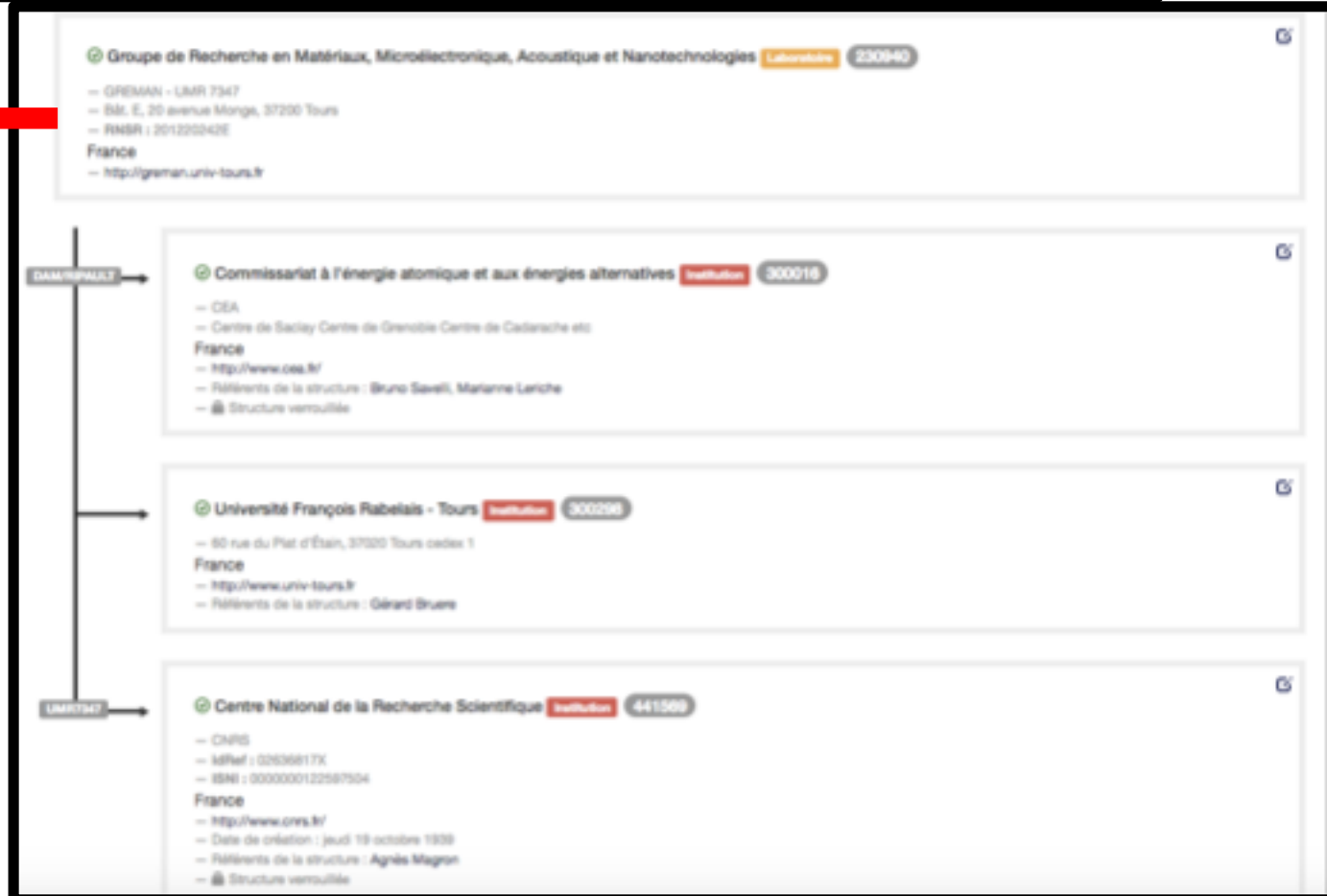
— ResearcherId : http://www.researcherid.com/rid/K-7326-2014

1 LIVE - Laboratoire Image, Ville, Environnement (3 Rue de l'Argonne 67000 STRASBOURG - France) StructId : 441830

- Université de Strasbourg (France) StructId : 302445
- CNRS - Centre National de la Recherche Scientifique : UMR7362 (France) StructId : 441569

EXEMPLE 2



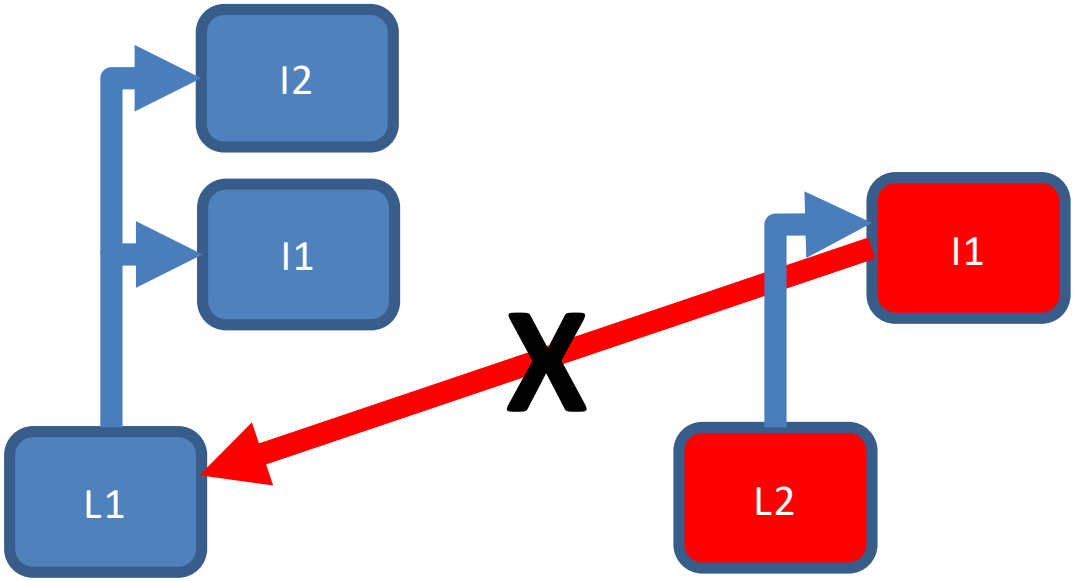


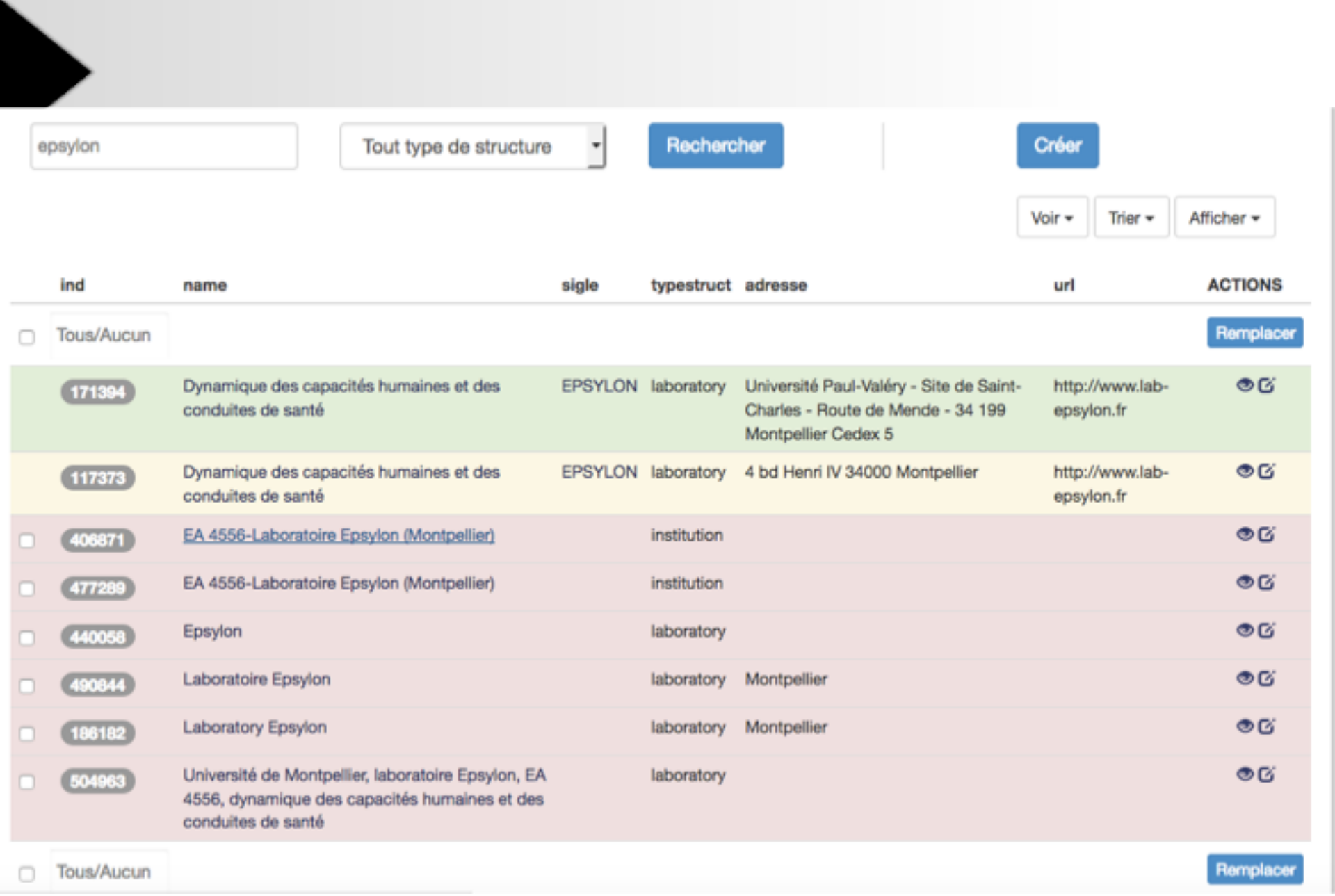
Remplacement possible

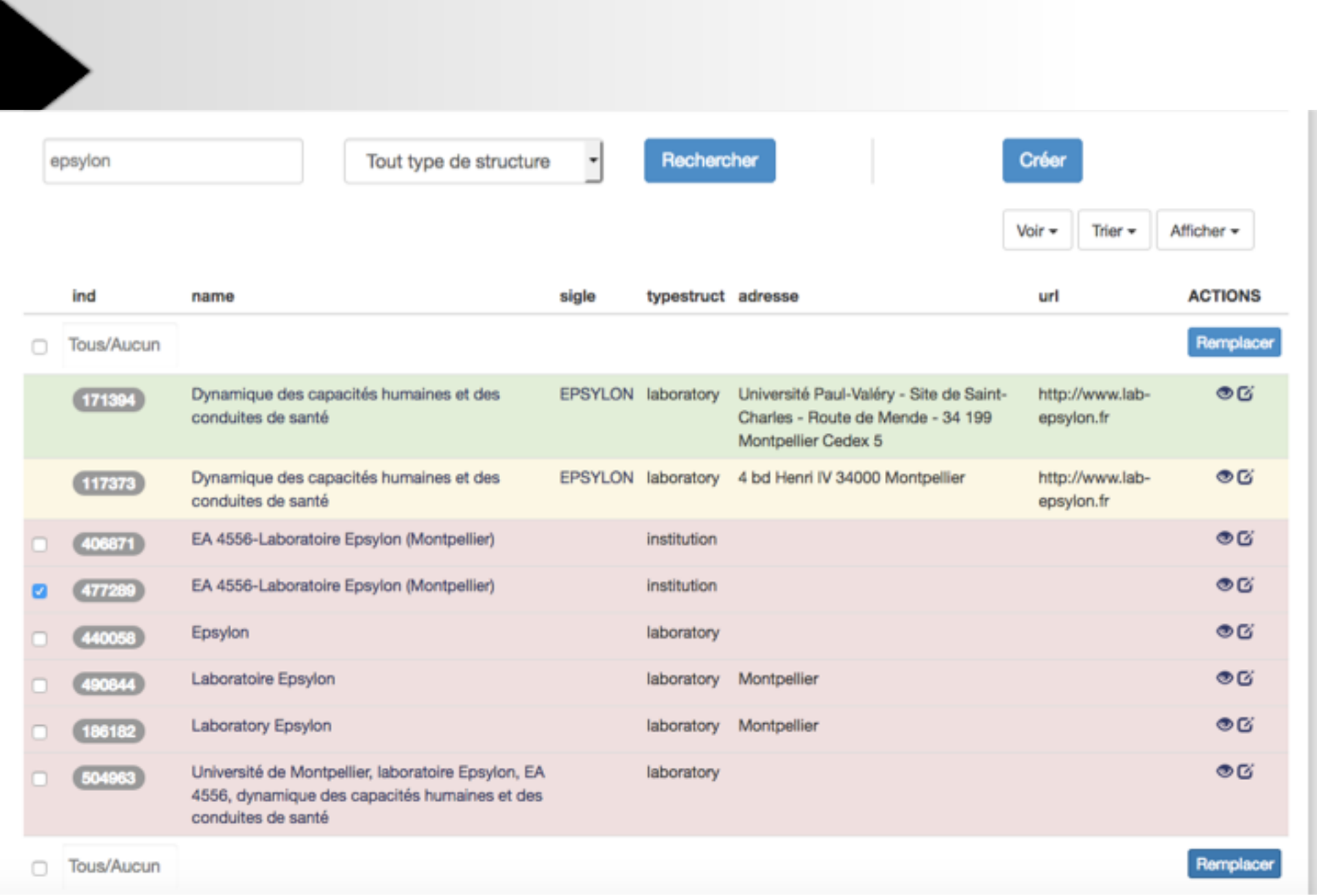
REMPLACEMENT D'UNE STRUCTURE 2E CAS

- Impossibilité de remplacement

STRUCTURES : REMPLACEMENT IMPOSSIBLE




epsylon
Tout type de structure
Rechercher
Créer
Voir
Trier
Afficher
ind
name
sigle
typestruct
adresse
url
ACTIONS
Tous/Aucun
Remplacer
171394
Dynamique des capacités humaines et des conduites de santé
EPSYLON
laboratory
Université Paul-Valéry - Site de Saint-Charles - Route de Mende - 34 199 Montpellier Cedex 5
http://www.lab-epsylon.fr
117373
Dynamique des capacités humaines et des conduites de santé
EPSYLON
laboratory
4 bd Henri IV 34000 Montpellier
http://www.lab-epsylon.fr
406871
EA 4556-Laboratoire Epsylon (Montpellier)
institution
477289
EA 4556-Laboratoire Epsylon (Montpellier)
institution
440058
Epsylon
laboratory
490844
Laboratoire Epsylon
laboratory
Montpellier
186182
Laboratory Epsylon
laboratory
Montpellier
504963
Université de Montpellier, laboratoire Epsylon, EA 4556, dynamique des capacités humaines et des conduites de santé
laboratory
Tous/Aucun
Remplacer


epsylon
Tout type de structure
Rechercher
Créer
Voir
Trier
Afficher
ind
name
sigle
typestruct
adresse
url
ACTIONS
Tous/Aucun
Remplacer
171394
Dynamique des capacités humaines et des conduites de santé
EPSYLON
laboratory
Université Paul-Valéry - Site de Saint-Charles - Route de Mende - 34 199 Montpellier Cedex 5
http://www.lab-epsylon.fr
117373
Dynamique des capacités humaines et des conduites de santé
EPSYLON
laboratory
4 bd Henri IV 34000 Montpellier
http://www.lab-epsylon.fr
406871
EA 4556-Laboratoire Epsylon (Montpellier)
institution
477289
EA 4556-Laboratoire Epsylon (Montpellier)
institution
440058
Epsylon
laboratory
490844
Laboratoire Epsylon
laboratory
Montpellier
186182
Laboratory Epsylon
laboratory
Montpellier
504963
Université de Montpellier, laboratoire Epsylon, EA 4556, dynamique des capacités humaines et des conduites de santé
laboratory
Tous/Aucun
Remplacer

A remplacer

EA 4556-Laboratoire Epsylon (Montpellier) Institution 477289

France

Remplacer les éléments par... epsylon

Rechercher

Valeur de remplacement

Trier par

ind	name	sigle	typestruct	adresse	url
171394	Dynamique des capacités humaines et des conduites de santé	EPSYLON	laboratory	Université Paul-Valéry - Site de Saint-Charles - Route de Mende - 34 199 Montpellier Cedex 5	http://www.lab-epsylon.fr
117373	Dynamique des capacités humaines et des conduites de santé	EPSYLON	laboratory	4 bd Henri IV 34000 Montpellier	http://www.lab-epsylon.fr

Structure conservé(e) :

Dynamique des capacités humaines et des conduites de santé Laboratoire 171394

— EPSYLON
— Université Paul-Valéry - Site de Saint-Charles - Route de Mende - 34 199 Montpellier Cedex 5
France
— http://www.lab-epsylon.fr

Structure(s) remplacé(es) :

Structure(s) non remplacé(es) :

EA 4556-Laboratoire Epsylon (Montpellier) Institution 477289

France

Modifications non autorisées

Les structures suivantes ne seront plus valides car elles dépendent de la structure remplacée (Structure):

- 477289

OUTILS EN DÉVELOPPEMENT (SCRIPTS)

- Remplacer dans un lot de dépôts une structure par une autre
  - Fermeture d'une institution `I1` date t
  - Ouverture d'une nouvelle institution `I2` > t
  - Pour le lot de dépôts avec date de publication > t remplacer `I1` par `I2`
- Dupliquer toutes les sous-structures d'une structure donnée :
  - Anciennes sous-structures = **old**
  - Nouvelles sous-structures créées = **valid**

Fiche d'une structure

CRLCC Paul Papin Institution 300545

France

Voir les documents associés | Voir les sous structures | Fermer et remplacer par | Exporter

Historique

Transfert d'une structure vers une autre + Fermeture

Ce module vous permer de remplacer une structure A par une autre structure B. Les structures filles A seront fermées et dupliquées pour être associées à la structure B.

A remplacer

CRLCC Paul Papin Institution 300545

France

Sous structure(s) fermée(s)/dupliquée(s)

Interactions cellulaires et applications thérapeutiques Laboratoire 1807

— ICAT
— Bâtiment Monteclair 4 rue Larrey 49033 Angers Cedex 1
France
— http://www.univ-angers.fr

Departement de Chirurgie Oncologique Laboratoire 96424

— 2, rue Moll 49000 Angers
France

Remplacer les éléments par... CRLCC

*

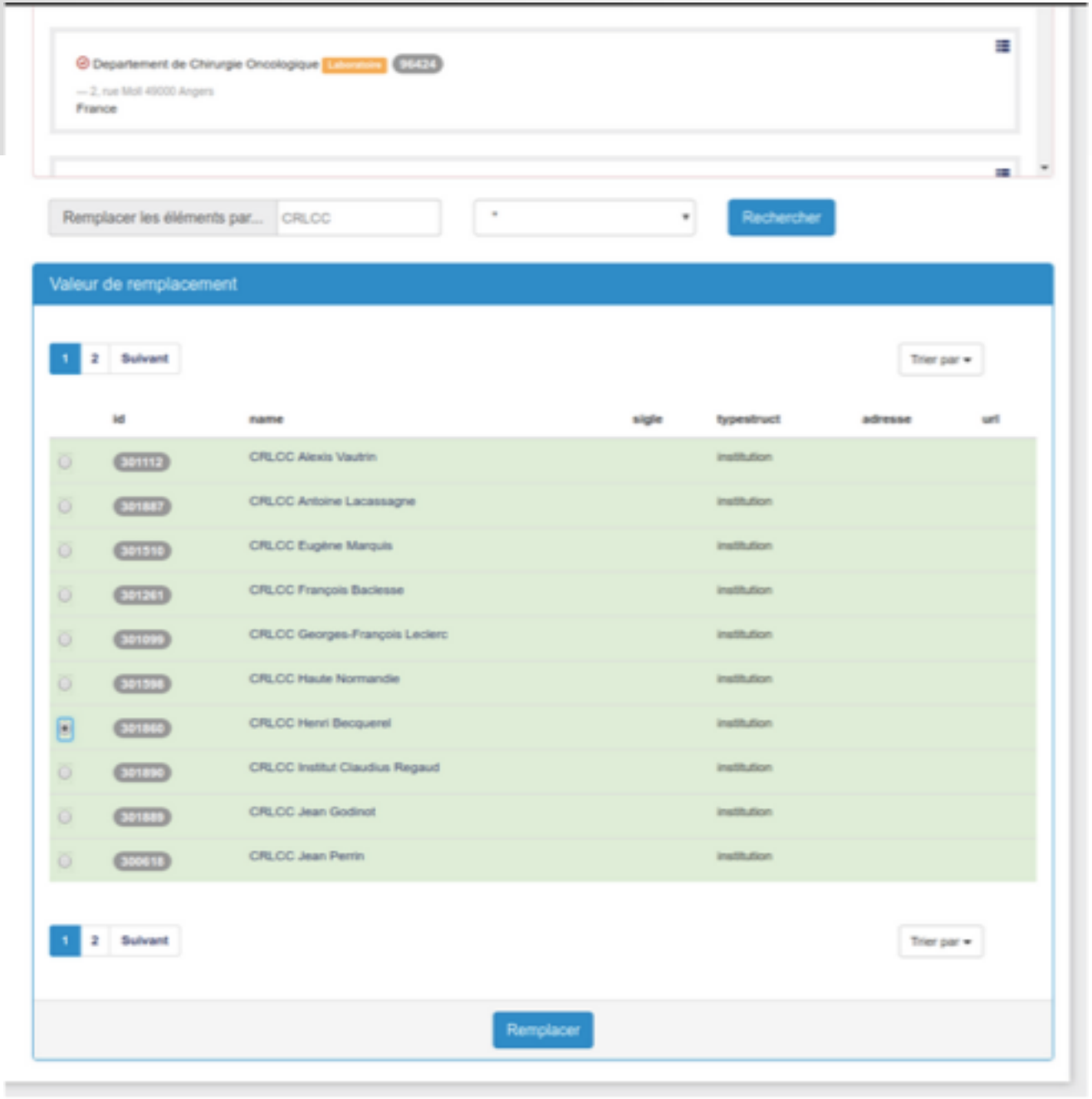
Rechercher

Valeur de remplacement

1 2 Suivant

Trier par

id	name	sigle	typestruct	adresse	url
301112	CRLCC Alexis Vautrin		institution		


Département de Chirurgie Oncologique
96424
2, rue Moll 49000 Angers
France
Remplacer les éléments par...
CRLCC
Rechercher
Valeur de remplacement
1
2
Suivant
Trier par
id
name
sigle
typestruct
adresse
url
301112
CRLCC Alexis Vautrin
institution
301887
CRLCC Antoine Lacassagne
institution
301510
CRLCC Eugène Marquis
institution
301261
CRLCC François Baclesse
institution
301099
CRLCC Georges-François Leclerc
institution
301598
CRLCC Haute Normandie
institution
301860
CRLCC Henri Becquerel
institution
301890
CRLCC Institut Claudius Regaud
institution
301889
CRLCC Jean Godinot
institution
300618
CRLCC Jean Perrin
institution
1
2
Suivant
Trier par
Remplacer

Résumé du transfert d'une structure

Ce module vous permer de remplacer une structure A par une autre structure B. Les structures filles A seront fermées et dupliquées pour être associées à la structure B.

Structure conservé(e) :

CRLCC Henri Becquerel Institution 301860

France

Structure(s) remplacé(es) :

CRLCC Paul Papin Institution 300545

France

Annuler la modification | Retour | Effectuer la modification

Résultat du transfert d'une structure

Ce module vous permer de remplacer une structure A par une autre structure B. Les structures filles A seront fermées et dupliquées pour être associées à la structure B.

Structure(s) fermée(s) :

CRLCC Paul Papin Institution 300545

France

Structure conservée :

CRLCC Henri Becquerel Institution 301860

France

Sous structure(s) fermée(s)/dupliquée(s)

Interactions cellulaires et applications thérapeutiques Laboratoire 1807

— ICAT
— Bâtiment Montéclair 4 rue Larrey 49033 Angers Cedex 1
France
— http://www.univ-angers.fr

Département de Chirurgie Oncologique Laboratoire 96424

— 2, rue Moll 49000 Angers
France

Oncology Surgical Department Laboratoire 102260

France

Institut de cancérologie de l'Ouest - Nantes Laboratoire 104568

— ICO Nantes
— Nantes
France
— http://www.centregauducheau.fr/

Structure(s) créée(s)/dupliquée(s) :

Interactions cellulaires et applications thérapeutiques Laboratoire 534450

— ICAT
— Bâtiment Montéclair 4 rue Larrey 49033 Angers Cedex 1
France
— http://www.univ-angers.fr

Département de Chirurgie Oncologique Laboratoire 534451

— 2, rue Moll 49000 Angers
France

Oncology Surgical Department Laboratoire 534452

France

Institut de cancérologie de l'Ouest - Nantes Laboratoire 534453

— ICO Nantes
— Nantes
France
— http://www.centregauducheau.fr/

RECOMMANDATIONS

- Normalisation de l'écriture, bonnes pratiques
  - Code unité : attention, propre à une institution
  - Structures étrangères
  - Fermeture / ouverture de structures
- Mutualisation des structures => concertation
- Pragmatisme : préférer un nom significatif
  - Exemple : 'équipe informatique'
- Toujours penser à l'impact sur le fonctionnement des portails et des collections

EXEMPLES

- Université Antilles-Guyane
- Devient Université de Guyane + Université des Antilles
- Fermer les labos
- Ouvrir de nouveaux labos
- Changements de tutelles = collections ou portails modifiés

- Labo X : tutelle Université Antilles-Guyane jusqu'en 2016

- A partir de 2016 tutelle = Université des Antilles
- Les dépôts ne doivent pas être tamponnés dans le portail Université de Guyane à partir de 2016
- => fermeture du laboratoire X (Old)
- + Ouverture d'un nouveau laboratoire X2 avec tutelle = Université des Antilles
- Même démarche pour les labos de l'Université de Guyane

ATTENTION : COLLECTIONS

- Les modifications et remplacements ne sont pas pris en compte dans les requêtes des collections
- Collections : éviter de construire la requête à partir des identifiants des référentiels, préférer le texte (texte1 OR texte2)
- Lors des remplacements dans les référentiels, évaluer l'impact sur d'éventuelles collections (laboratoires), voire portails