

« MIEUX PARTAGER LES CONNAISSANCES »

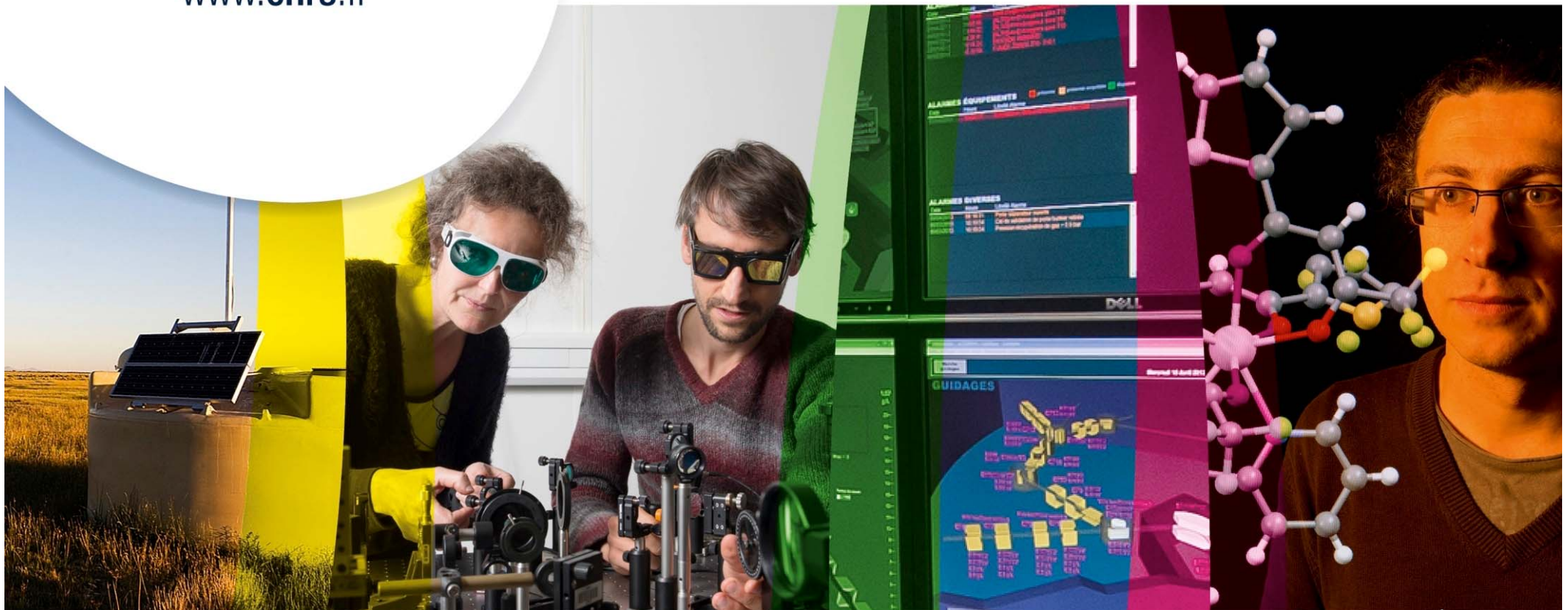


www.cnrs.fr

Pour une information scientifique et technique
interopérable, pérenne
fondée sur des ressources
évolutives, accessibles réutilisables, interopérables et pérennes.

30 ans de Démocrite - IN2P3

Direction de l'Information Scientifique et Technique
Laurence El Khouri,
Septembre 2017





INIST

DIST



LES CORRESPONDANTS IST DANS LES INSTITUTS

INC Brigitte Diers	INEE Joël Cuguen	INP Barendd Van Tiggelen Martina Knoop	IN2P3 Stéphane Plaszczyński Ursula Bassler	INSB Dominique Dunon-Bluteau	INS2I Patrice Bellot	INSIS Benoit Pier	INSHS Didier Torney	INSMI Frederic Helein	INSU Pascale Talour
------------------------------	----------------------------	---	---	--	--------------------------------	-----------------------------	-------------------------------	---------------------------------	-------------------------------



CEDRAM



NUMDAM



P. 3

DIST





EPRIST



DIST

- **Organisation**

- 10 instituts - Directions fonctionnelles DGDS – DGDR
- Unités d'IST CCSD – Persée – Inist – Cléo ...

- **Documentation**

- BibCNRS
- ISTEEX
- PERSEE
- HAL
- Mais aussi : openedition.org, Huma-num, Numdam, ..., Inspire

- **Openscience**

- 4 piliers : Open access, open research data, opensource, altmetrics
- Loi pour une République numérique – Directives européennes –
- EOSC – GO FAIR
- Services : Acquisitions récentes par des acteurs du privé : Elsevier, Clarivate Analytics, Springer Nature,
- Interopérabilité, mutualisation, mutation des modèles économiques
- Nouveaux métiers du numérique en IST

- **Loi pour une République numérique**

- Article 30 - Article 38 - Et les autres ...



Quelques actions en cours

- Loi pour une République numérique
 - Article 30
 - Article 38 – décrets d'application en attente
- Directive Européenne pour le droit d'auteur DADVSI
 - Homogénéisation des traductions
 - Transposition française
- Actions de mutualisation
 - Stratégie
 - Enquêtes (UMR, personnels, copist...)
 - COPIST
 - Données de la recherche
 - RDA France
 - Formation - Doranum – DMPOpidor
 - GO FAIR EOSC
 - Nouveaux métiers
- projets transversaux
 - ISTEEX - ANR
 - Conditor - BSN/MESRI/DGRI/DGESIP
 - SCOAP3 - CERN



P. 6

Loi pour une République numérique

- Article 30
 - Information <http://openaccess.couperin.org/quoi/>
 - Automatisation ?
 - Données de recherche
- Article 38 – décrets d'application en attente
 - GT guide d'application de la loi

« The right to read is the right to mine.... »

<http://www.leru.org/index.php/public/news/the-right-to-read-is-the-right-to-mine/>



- Directive Européenne pour le droit d'auteur DADVSI
 - Homogénéisation des traductions
 - Transposition française



Loi pour une République numérique

- Harmonisation nécessaire d'un droit à l'open science

Projet de directive - Droit d'auteur dans le marché unique numérique
(COM(2016) 593 final) 14-9-2016



P. 8

- Définition de la notion de TDM :
 - Text and data mining means any automated analytical technique aiming to analyse text and data in digital form in order to generate information such as patterns, trends and correlations
- Création d'une exception droit d'auteur et droit du producteur de base de données en faveur du TDM
- Exception plus large que l'exception française (tous les types d'œuvres, y compris à des fins commerciales, dispositions d'ordre public)

Loi pour une République numérique

- Menaces sur le droit à l'open science

Projet de directive - Droit d'auteur dans le marché unique numérique (COM(2016) 593 final) 14-9-2016



P. 9

Appel SPARC EUROPE <http://sparceurope.org/copyrightreform/>



- retrait des articles 11 (droits voisins, citation) et 13 (contrôle des contenus a priori).
- amélioration de l'article 3 sur le TDM pour s'assurer de la capacité de mettre en œuvre celui-ci
- amélioration de l'article 4 (éducation)
- amélioration des articles 5 à 9 dans l'esprit de conforter les missions de conservation des bibliothèques.

Données de la recherche Loi Valter (2015) + Article 30

Ouverture des données de la recherche (open research data), mouvement plus large d'open data, ouverture des données publiques, mouvement pleinement applicable depuis la loi Valter de décembre 2015 aux « établissements et institutions d'enseignement et de recherche ».



P. 10

- fourniture d'identifiants DOI
- aide à la création/gestion de plan de gestion de données (OPIDoR)
- formations à distance sur la gestion des données de recherche (DORANum)
- comment diffuser ses données : guide de lecture du cadre juridique de l'ouverture des données
- les services d'entrepôts institutionnels de données (Inra, Irstea,...)
- Participation aux WG de RDA (Best practices)

Collaboration, mutualisation, partage entre organismes, métiers

Objet décloisonnant, nécessite coopération entre IST/Science/techniques informatiques

Evaluation - Altmetrics

Les métriques de la science basées sur les seules publications montrent leurs limites. L'élargissement des objets pris en compte (jeux de données, data papers,...) et le développement de métriques d'impact (partage, réutilisation, valorisation) sont des chantiers à mener pour le développement de l'Open Science.



P. 11

Etude de modèles de l'édition scientifique innovants : nouveaux modes d'éditorialisation

Appel de Jussieu 5/07/2017 (joint) → accès ouvert : priorité à l'innovation ! (fait suite à l'appel d'Amsterdam (2016))

DORA Signatories, 2013, San Francisco declaration on research assessment

<http://www.ascb.org/dora/>

Leiden Manifesto (Hicks et al. 2015), qui cite 10 principes faisant évoluer les usages de bibliométrie

<https://www.nature.com/news/bibliometrics-the-leiden-manifesto-for-research-metrics-1.17351>

Responsible Metrics Forum : <https://responsiblemetrics.org/> forum de discussion

Rapport RISE (RISE High Level Group, 2017a),

<https://ec.europa.eu/research/openvision/index.cfm?pg=expert-groups>

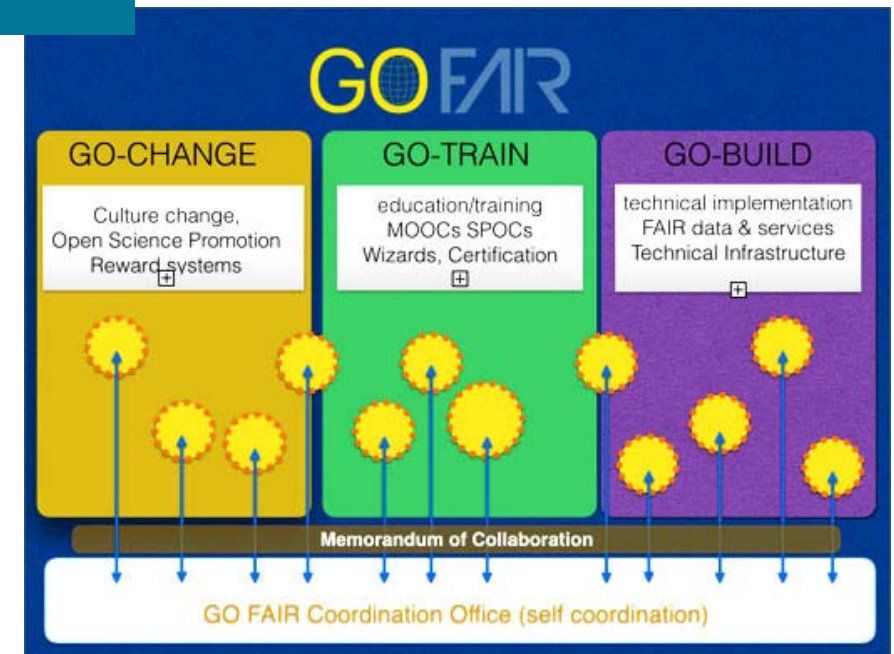
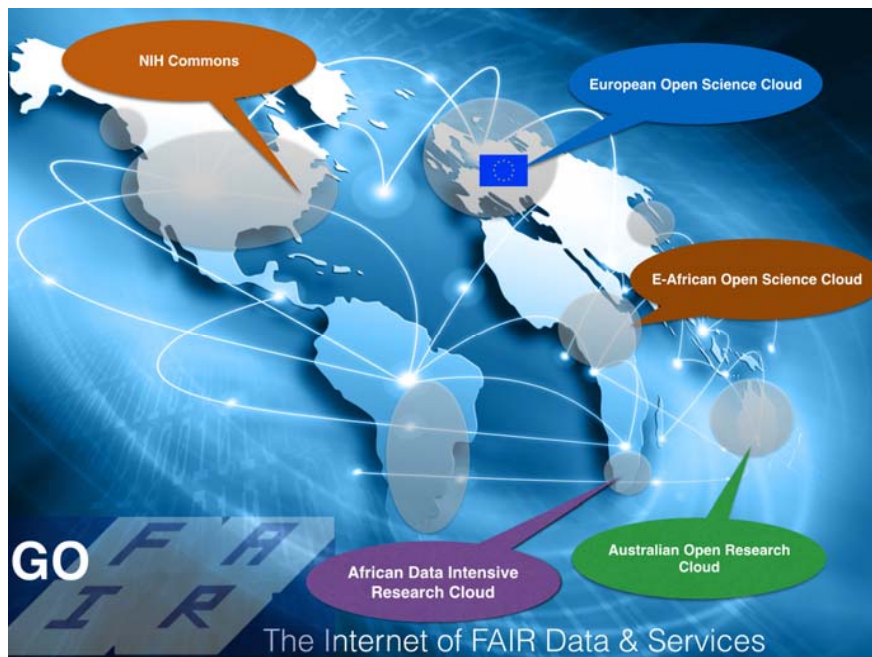
EU Altmetrics Report (European Commission Expert Group on Altmetrics, 2017)

<https://ec.europa.eu/research/openscience/pdf/report.pdf>

EOSC – European Open Science Cloud



https://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf#view=fit&pagemode=none





- Une vaste équipe pluridisciplinaire :
 - Une cinquantaine de titulaires qui interviennent sur le projet
 - Une contribution totale de 42 CDD et 2 CDI en 2016
 - Dont 16 CDD au CNRS

- Un projet transversal qui nécessite l'expertise de plusieurs directions du CNRS :
 - Direction des Affaires Juridique (DAJ)
 - Direction de la Stratégie Financière, de l'Immobilier et de la Modernisation (DSFIM)
 - Délégation
 - Service Financier et Comptable (SFC)
 - Service des Ressources Humaines (SRH)

- Un accès pérenne aux ressources acquises

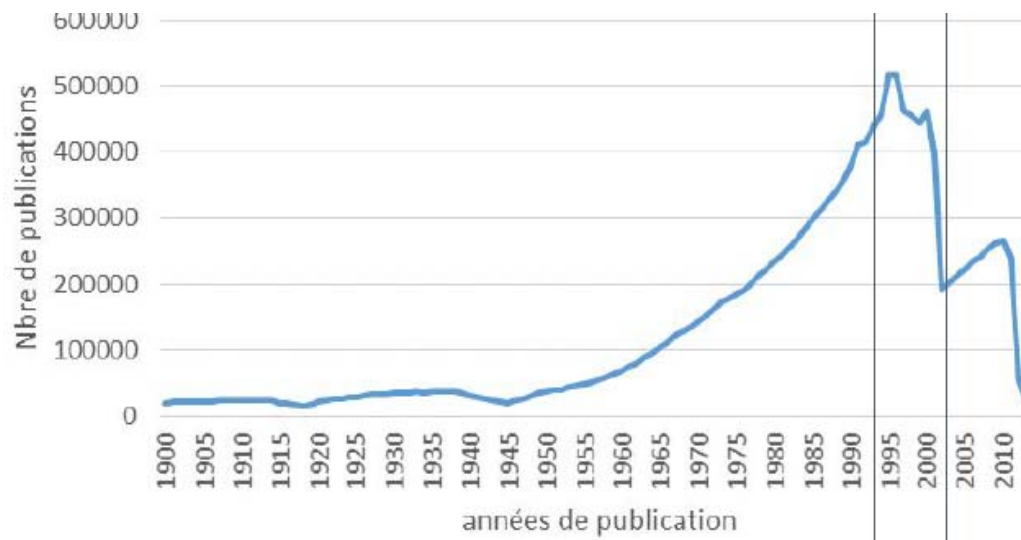
- Plus de 18 millions de ressources
- Plus de 7500 revues et 13000 E-books
- Ressources généralistes mais aussi de niches
- Formats variés : E-books, livres, articles, revues



P. 14

- Des ressources enrichies, plus accessibles et exploitables

- Du contenu en Xml enrichi, Full Text, PDF, XML structuré...
- Des ressources OCRisées à la demande



Corpus ▾		16
<input type="checkbox"/> elsevier		6015851
<input type="checkbox"/> wiley		4982769
<input type="checkbox"/> springer		2304783
<input type="checkbox"/> oup		1450642
<input type="checkbox"/> sage		804286
<input type="checkbox"/> bmj		689840
<input type="checkbox"/> iop		394495
<input type="checkbox"/> nature		379969
<input type="checkbox"/> rsc-journals		294765
<input type="checkbox"/> degruyter-journals		242278
<input type="checkbox"/> ecco		207414
<input type="checkbox"/> emerald		159955
<input type="checkbox"/> brill-journals		130257
<input type="checkbox"/> eebo		124221
<input type="checkbox"/> rsc-ebooks		23879
<input type="checkbox"/> brill-hacco		11177

ISTEX consolide les performances de sa plateforme...

L'excellence documentaire pour tous

- Disponibilité optimale de la plateforme ~ 99,9%

- Développement de services :

- De bases :

- Annotation de termes et de leurs variantes
- Extraction d'entités nommées
- Reconstruction des références bibliographiques...

- À valeur ajoutée :

- **CILLEX** - moteur de réponse offrant des outils de classification automatique
- **LorExplor** - création d'une bibliothèque Open Source de composants XML d'exploitation du corpus ISTEX
- **ISTEX-R** - caractérisation de l'évolution des recherches et des connaissances dans le temps grâce à la construction de cartes diachroniques
- **ISTEX-SNU** - intégration de la plateforme ISTEX dans les Environnements Numériques de Travail des Universités

- Actions de perfectionnement :

- Environnement de pré-production
- Tests d'intégration continue garantissant la non-régression
- Augmentation de la puissance de calcul de l'API et du moteur de recherche



P. 15

... et travaille sur ses bancs d'essais

- Chantiers d'usage :

- Projet **Biosystémique** : Recherche de résultats expérimentaux dans les publications scientifiques concernant la biologie systémique ;
- Projet **ALPAGE** : annotation de corpus ISTEEX et codage en TEI ;
- Projet **Terre-ISTEEX** : Identification et analyse des TERRains d'Études dans les corpus ISTEEX ;
- Projet **3ST** : Surligneur Sémantique de Textes Scientifiques ;
- Projet **OTELO** : chantier thématique sur les Terres rares, de l'exploration à l'impact environnemental ;
- Projet **NEOTEX** : Exploitation de documents textuels d'un domaine par un néophyte ;
- Projet **FULLAB** : Comparaison fulltext et abstract en sciences environnementales ;
- Projet **ReITEX** : Usage du corpus ISTEEX pour l'exploitation de méthodes d'extraction de connaissances à partir de textes

- Séminaire ISTEEX du 06 et 07 juin 2017 - Nancy : Les résultats et l'avancement des projets soutenus en 2015 et 16 seront bientôt disponibles (captation video)

<http://www.istex.fr/seminaire-technique-les-6-et-7-juin-2017/>





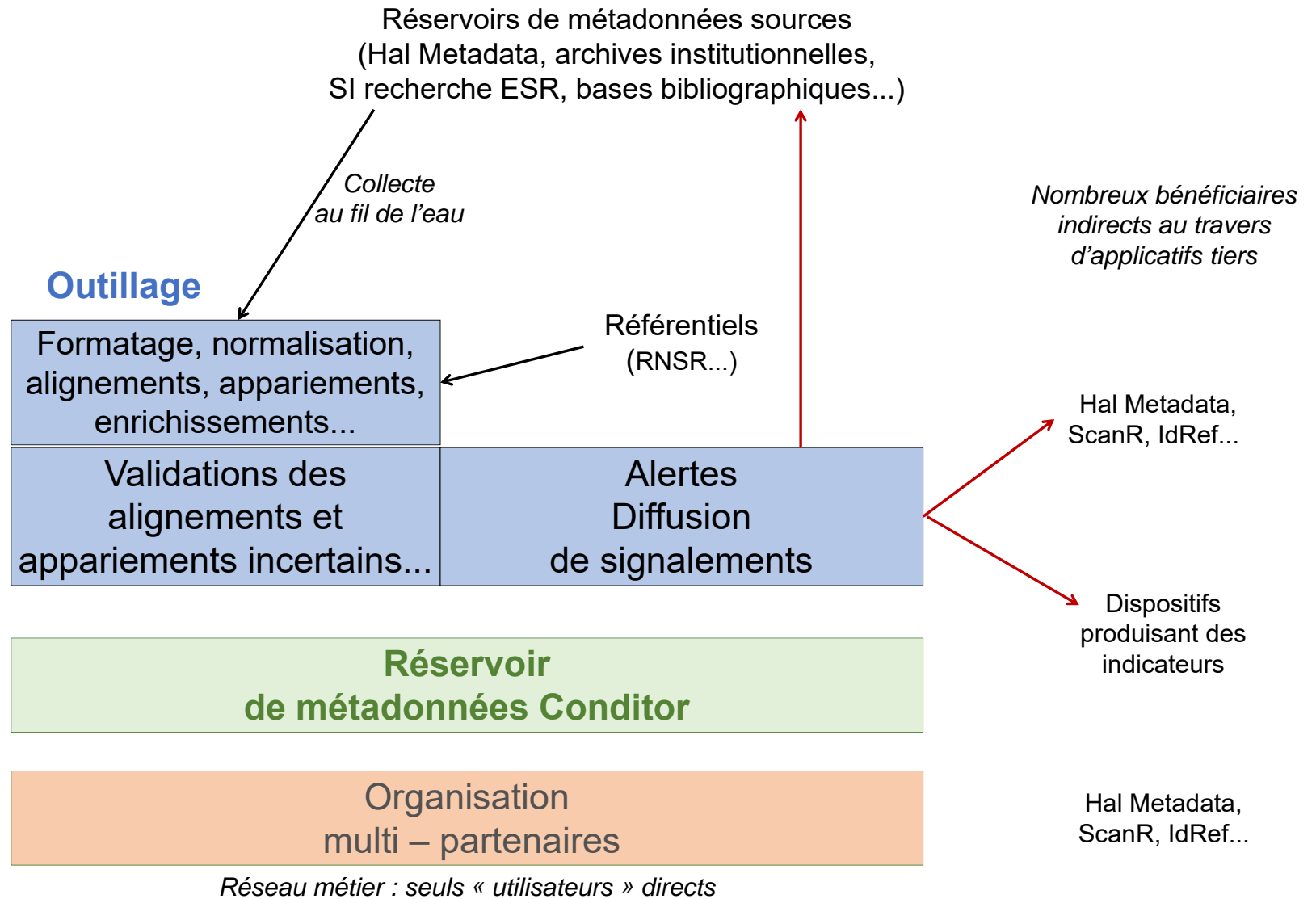
P. 17

Une recommandation de BSN3 (signalement)
« Construire un *référentiel* de la *production scientifique française* en s'appuyant sur des *référentiels communs* aux universités et organismes de recherche (unités, thématiques, auteurs,...) »


Un projet multi-partenaires soutenu par le MESRI

Abes, Amue, CNRS (Dist/Inist, InSHS, SAPPS), IAVFF – Agreenium, Inra, Inria, Irstea, MESRI (RNSR-ScanR), UMS CCSD et Huma-Num, universités de Bordeaux, Grenoble, Paris Dauphine, Paris Diderot, UPMC

Vision du produit Conditor



Une démarche progressive

- 
- P. 19
- Phase projet lancée en décembre 2016 pour un service opérationnel début 2019 :
 - Un corpus de métadonnées sur plusieurs années
 - L'outillage pour le gérer
 - L'équipe pour le faire vivre

 - Un impact sur l'ensemble de l'ESR d'ici 5 ans :
 - Une mutualisation des travaux de recensement (saisie et repérage)
 - Une intégration aux systèmes d'informations de l'ESR

Des apports potentiels multiples



Amélioration de la **qualité** et la **complétude** des réservoirs ESR existants

Mutualisation optimale des recensements au sein de l'ESR
Allègement des tâches des **chercheurs**

Amélioration de la **qualité** des **référentiels** ESR (RNSR, IdRef...)








Contribution à la mise en place de **référentiels communs** ou à **l'interopérabilité** de référentiels existants

Meilleure **connaissance** de la production de la communauté ESR

Indicateurs de la science Fiabilisés



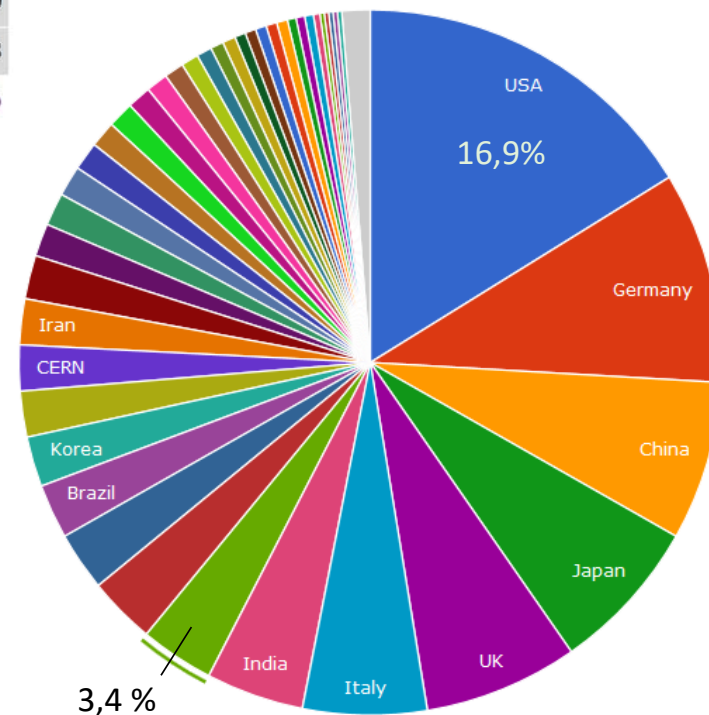
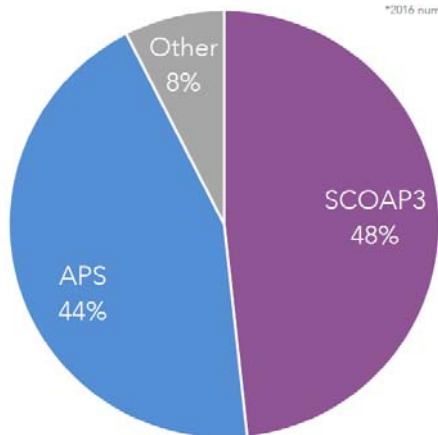
SCOAP³ – Sponsoring Consortium for Open Access Publishing in Particle Physics

Publisher	Journal	Articles 2014 – 2016*
	Nuclear Physics B	997
	Physics Letters B	2,693
	Advances in High Energy Physics	509
	Chinese Physics C	89
	Journal of Cosmology & Astroparticle Physics	648
	New Journal of Physics	27
	Acta Physica Polonica B	59
	Progress of Theoretical & Experimental Physics	221
	European Physical Journal C	1,820
	Journal of High Energy Physics	6,283

Total number of articles funded during Phase 1*: 13,346

*2016 numbers extrapolated based on the articles published as of November 1st 2016


- phase 1 : 3.8% → 210k€/an (420€/article)
- phase 2 : 3.4% → 183k€/an (360€/article)



Digitalisation de la science



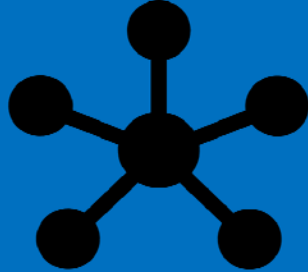
Nouveaux volumes :
 $1+1 = 3$
le téraoctet



Nouveaux supports et
démarches :
plateformes



Nouvelles formes :
métaconnaissances...



Nouveaux usages :
recommandation,
captation contrôlée

Des attentes partagées entre les acteurs de l'IST une dynamique pour la science publique ouverte



Etre formés à (et informés de) l'interaction avec l'IST

A green rounded square containing a black silhouette of a human head in profile, facing right. Inside the head, there are three interlocking gears of different sizes, colored in a light green. Below the icon, the text 'Etre formés à (et informés de) l'interaction avec l'IST' is written in a white, sans-serif font.

Accéder à de larges fonds documentaires et outils d'analyse au meilleur coût

A blue rounded square containing three black icons: a stack of three books, a film reel, and a microscope. Below the icons, the text 'Accéder à de larges fonds documentaires et outils d'analyse au meilleur coût' is written in a white, sans-serif font.

Unifier l'offre d'outils aux utilités superposées

An orange rounded square containing two black icons: a funnel and a target with an arrow hitting the bullseye. Below the icons, the text 'Unifier l'offre d'outils aux utilités superposées' is written in a white, sans-serif font.



- [Stratégie IST 2014](#)
- [Livre blanc et guide stratégique d'application](#) (« Une Science ouverte dans une République numérique : études et propositions en vue de l'application de la Loi »)
- Enquête COPIST ([premiers résultats d'octobre](#), [version du 14 mars 2016](#))
- [Enquête auprès des DU](#) (« Mieux partager l'information scientifique », résultats de mars 2015 de l'enquête sur les usages et besoins d'IST des Unités de Recherche, réalisée auprès des Directrices et Directeurs d'Unités du CNRS)
- Actes du Colloque de Meudon ;
 - « [Dynamiques de l'édition scientifique, de l'industrie, de l'information, de la documentation](#) » (4 et 5 novembre 2014)
 - « [Publication scientifique, innovation et services à la recherche](#) » (9 et 10 novembre 2015)
 - « [La Loi Numérique : et après ?](#) » (9 et 10 novembre 2016)
- [Rapport de mission](#) sur les systèmes de recommandation (Elsevier, 6 avril 2017)
- Schéma opérationnel (Panorama de l'écosystème IST 2017)
- [Ouverture des données de recherche : cadre juridique](#) (INRA)
- [Rapport 2017 au MESR](#) « Les enjeux du numérique concernant les ressources documentaires en SHS »
- Rapports d'activité des segments 4 et 7 de la BSN (2016)
- [Travaux sur les indicateurs IST](#) des EPST et des Epic (séminaire ADBU 2012, Dominique Cavet)



www.cnrs.fr

Merci pour votre attention !