

La valorisation de HAL par les laboratoires de recherche

Joachim Schöpfel¹, Florence Thiault²

Stéphane Chaudiron¹, Bernard Jacquemin¹, Eric Kergosien¹, Hélène Prost³,

¹Université de Lille, Gériico

²Université de Rennes 2, URFIST Rennes, PREFICS

³CNRS, Gériico

(1) Le projet HAL/LO

Valorisation sur **HAL** de la production des Laboratoires dans l'environnement de la science **Ouverte**

Financement par GIS « Réseau Urfist »

HAL comme objet, outil et ressource



Les objectifs du projet HAL/LO

Le projet HAL/LO s'intéresse aux laboratoires comme acteurs-clés de l'activité scientifique, porteurs de projets de recherche et objets l'évaluation par le Hcéres

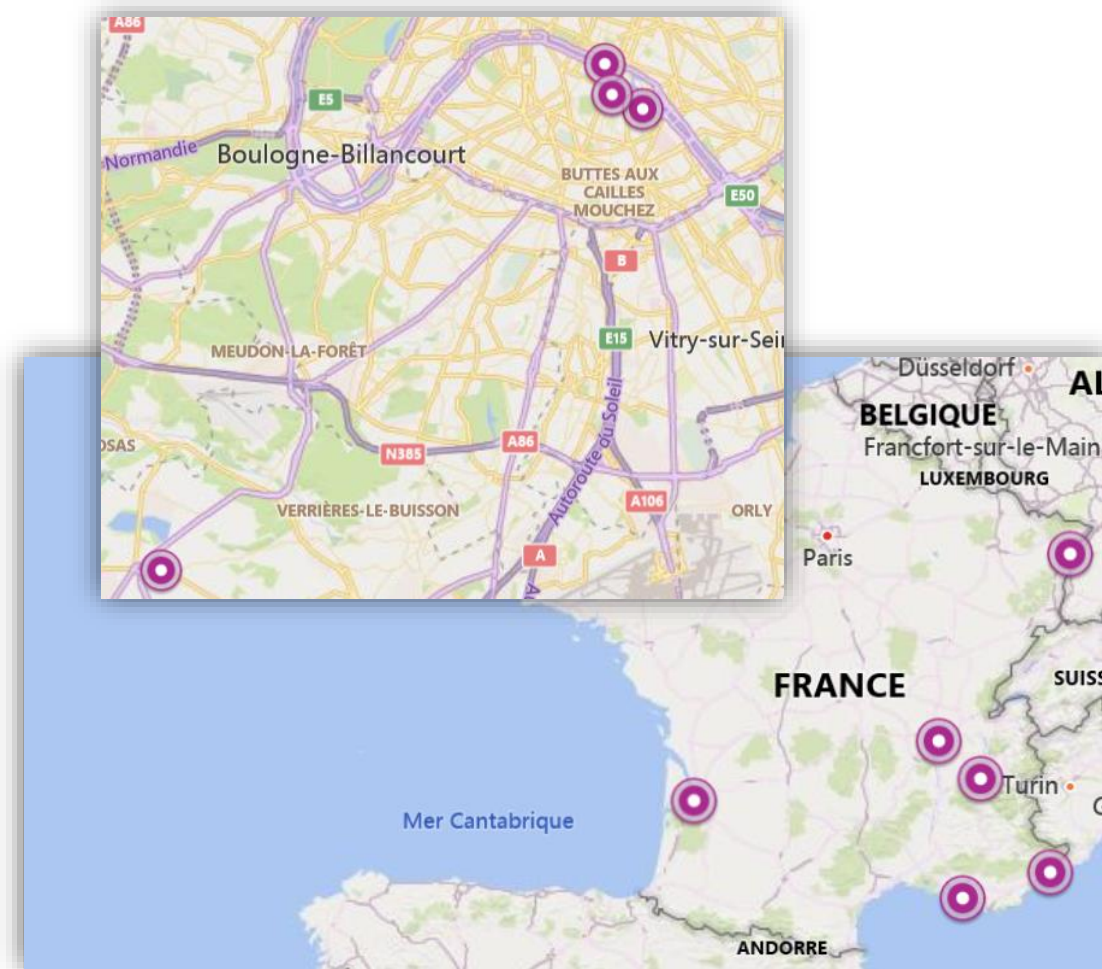
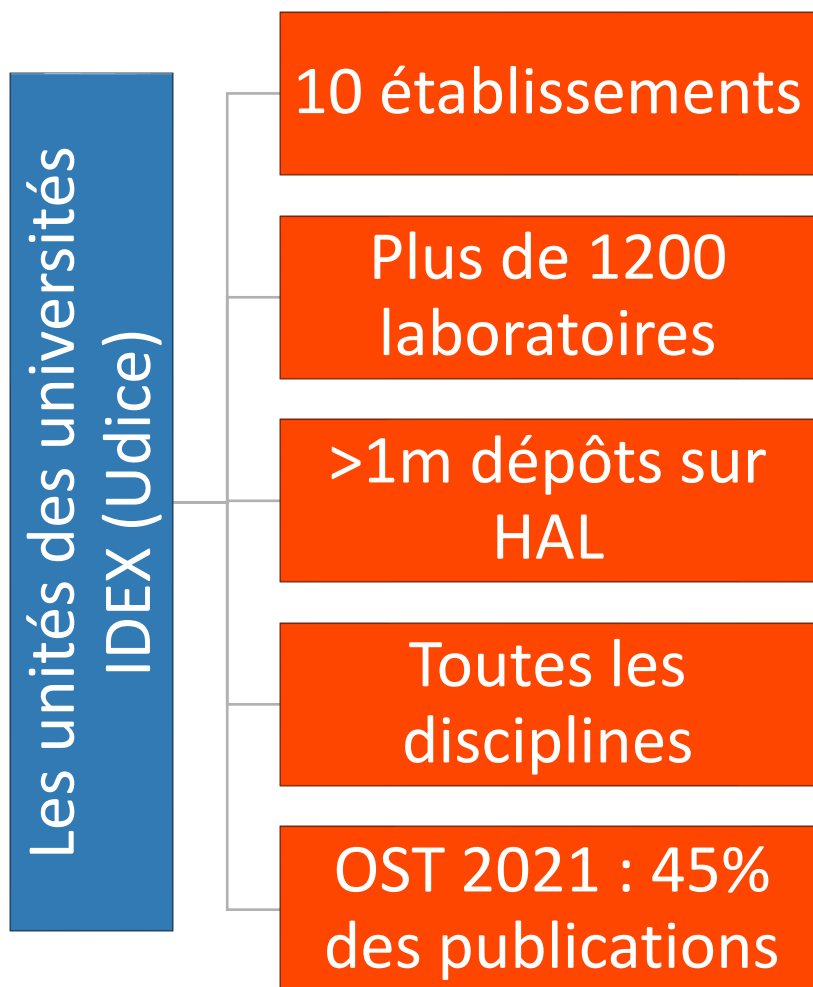
Comment gèrent-ils et valorisent-ils leurs productions scientifiques sur la plateforme HAL ?

Quel est l'intérêt et le potentiel de HAL pour les laboratoires ?

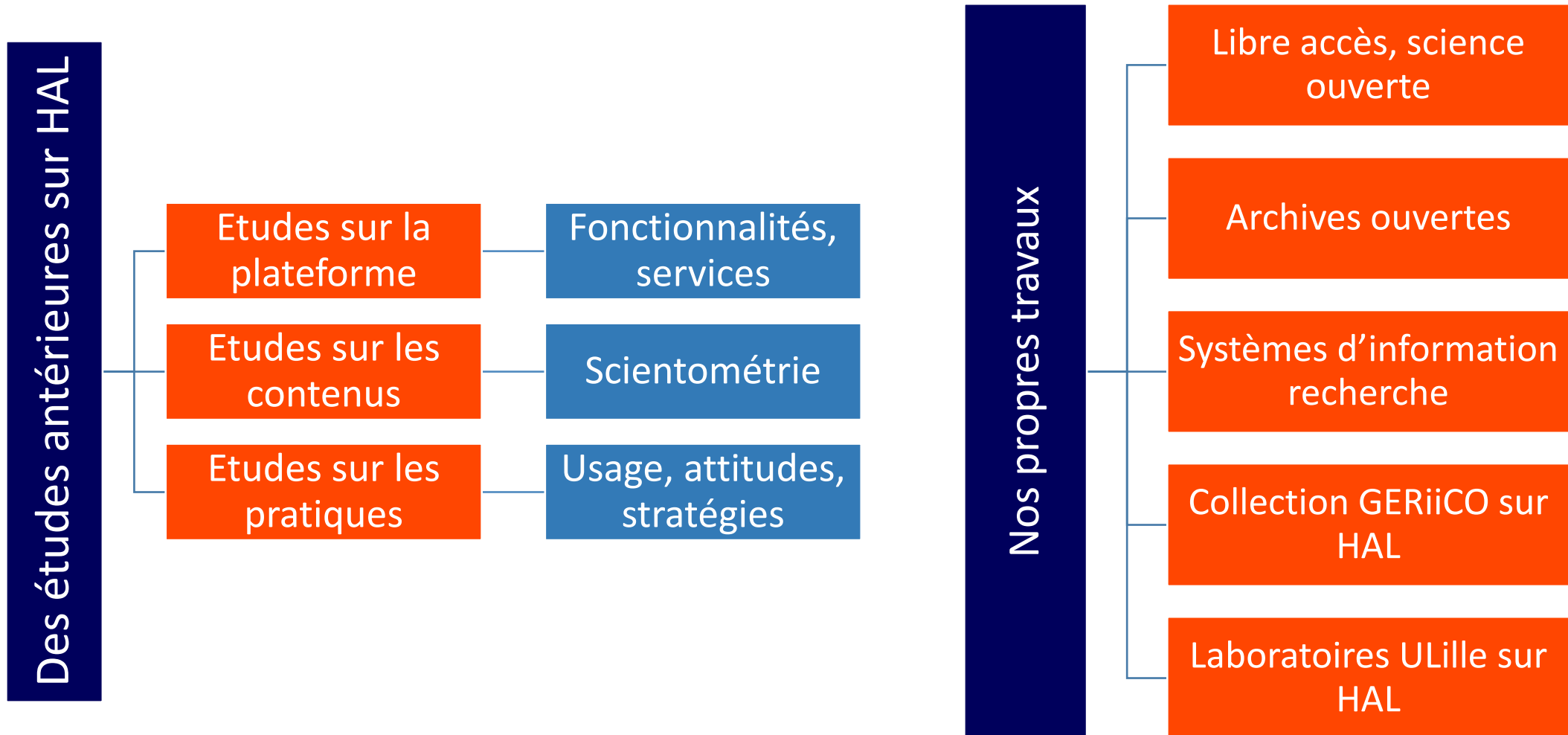
Focus sur les stratégies collectives

Analyse des facteurs favorables et défavorables

L'échantillon



La genèse du projet



Les participants

Lille

- Joachim Schöpfel (resp)
- Stéphane Chaudiron
- Bernard Jacquemin
- Eric Kergosien
- Hélène Prost

Rennes

- Florence Thiault (resp)
- Marie-Laure Malingre

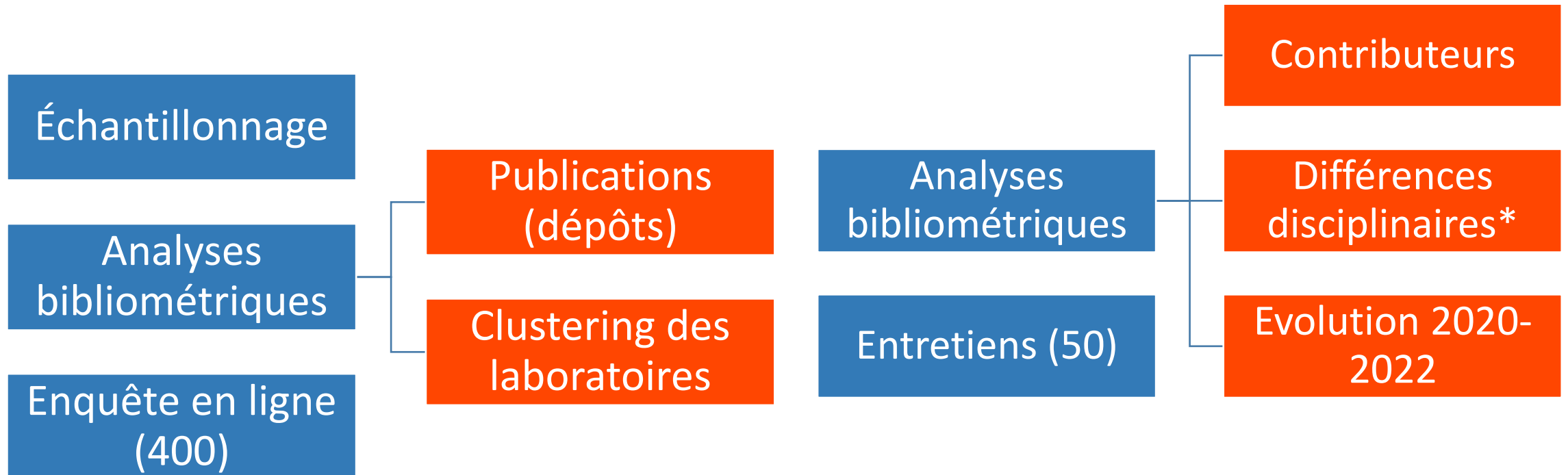
Stagiaires et vacataires

- *Arezki Achouri*
- *Julie Barrié*
- *Aline Benvegnu dos Santos*
- *Marine Chochoy*
- *Chamseddine Hammouda*
- *Odelin Jousset*
- *Ibrahima Ndiaye*

La méthodologie

2019-2021

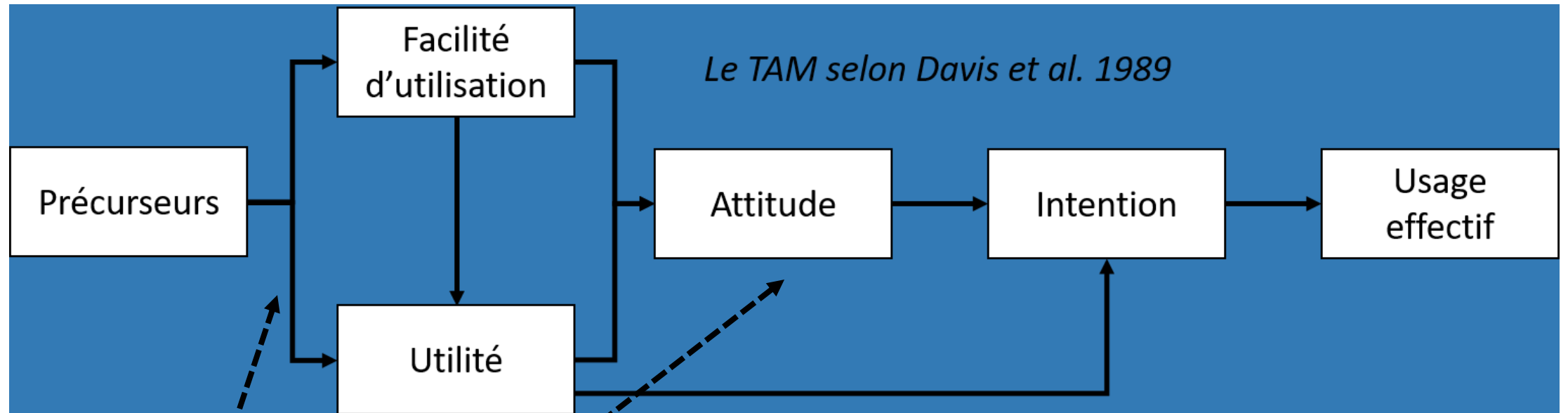
2022-2023



* Analyse avec Maxence Larrieu (UVSQ, Saclay ; UGA)

(2) Opinions

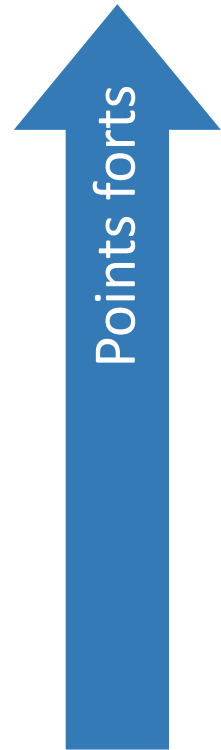
L'acceptation d'une nouvelle technologie



Influence sociale
Culture organisationnelle
Incitation/obligation

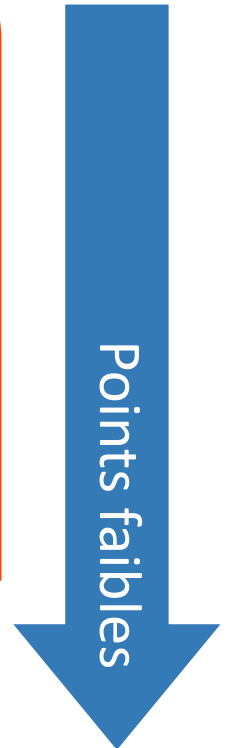
12 juin 2024

Des avis partagés

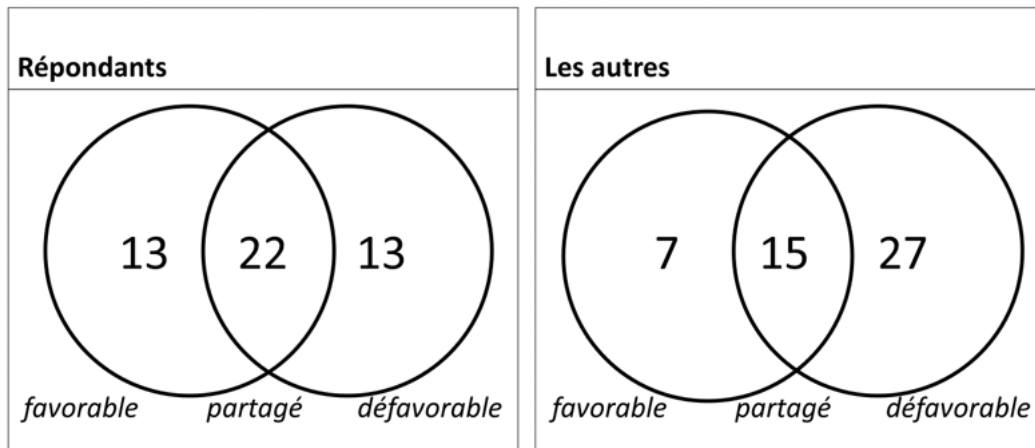


Visibilité des publications
Valorisation
Suivi, bilans et rapports
Pérennité
Economie
Légalité

Ergonomie (interface)
Performance
Trop franco-français
Finalité (évaluation)
Règles compliquées (droits, embargo)
Discours mieux adapté aux STM



La perception des autres



- Les interlocuteurs - DU, chargés de HAL, documentalistes etc. - semblent moins critiques, moins négatifs que les autres membres du laboratoire
- Interrogés sur l'avis de leurs collègues, les réponses sont significativement différentes, dans la mesure où l'avis négatif prévaut
- *“Les autres membres du laboratoire, les chercheurs, voient surtout les problèmes techniques”*
- *“Ils ne sont pas aussi enthousiastes que moi sur HAL”*

(3) Collections

Les collections

60% ont une collection, et 7% ont le projet

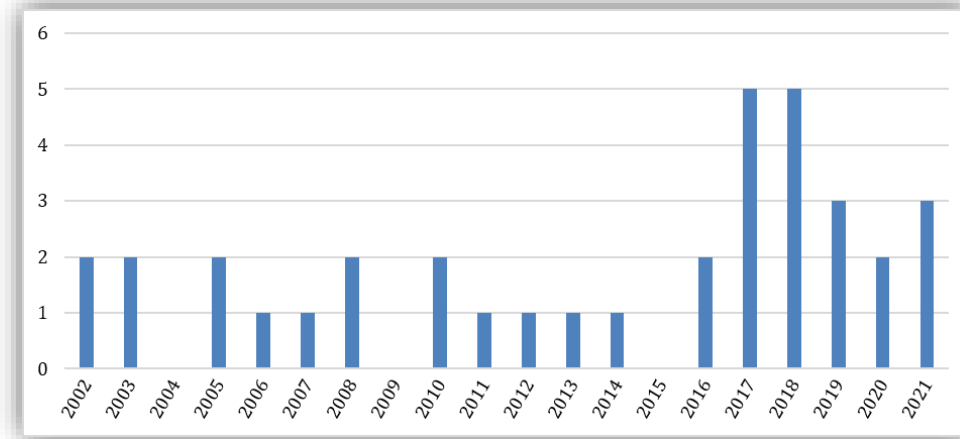
Une gestion par le personnel scientifique, infodoc, ou administratif

Plusieurs objectifs

Visibilité, impact

Suivi, évaluation, bilan

Diffusion des résultats



Nuage de mots des réponses libres sur le but de la collection (N=76)

L'impact significatif des disciplines

- Utilisation de HAL**
 - STIC, mathématiques, physique
- Typologie des dépôts**
 - SHS : livres
- Le degré d'ouverture**
 - DEG et ALLSHS moins ouvertes
- Les collections**
 - Moins de collection en SDV santé

Discipline	Moyenne
Sciences et technologies de l'information et de la communication	1521
Mathématiques et leurs interactions	1496
Physique	1313
Sciences pour l'ingénieur	1177
Sciences de la terre, de l'univers et de l'espace	1086
Sciences de la société	980
Chimie	972
Sciences agronomiques et écologiques	915
Sciences humaines et humanités	787
Biologie, médecine, santé	306

Nombre moyen de dépôts par laboratoire et discipline

(4) Alimentation

Une diversité de situations

Alimentation systématique

- **Médiation**
 - Pour tous les chercheurs
 - En complémentarité, partiellement
- **Auto-archivage**
 - Alimentation « disciplinée », par tous les chercheurs

Alimentation non-systématique

- **Médiation**
 - Aide ponctuelle
- **Auto-archivage**
 - Choix individuel, par certains chercheurs

Différents types de contributeurs

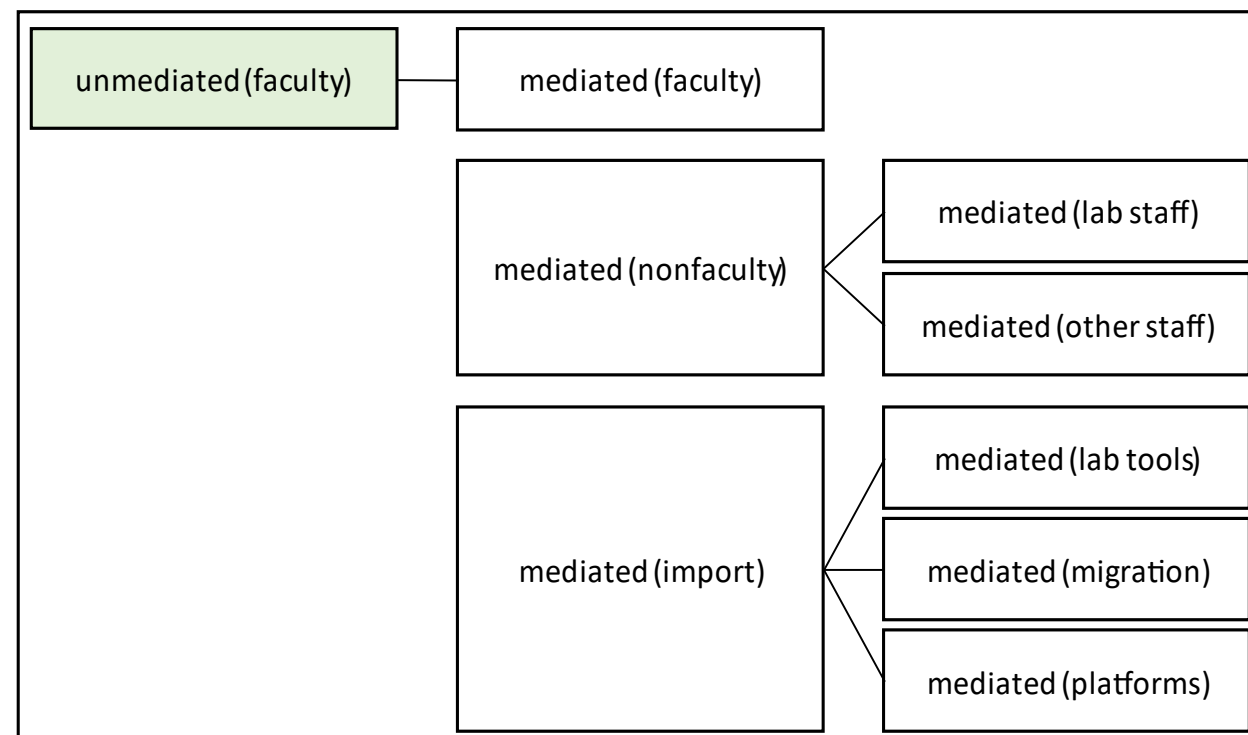
L'auto-archivage est un cas d'alimentation parmi d'autres

Des dépôts par d'autres

- Des collègues
- Des ITA/BIATSS du laboratoire
- D'autres personnels en dehors du laboratoire

Import de notices

- Import via des outils de laboratoire
- Import via des flux de migration
- Import via d'autres plateformes (éditeurs...)

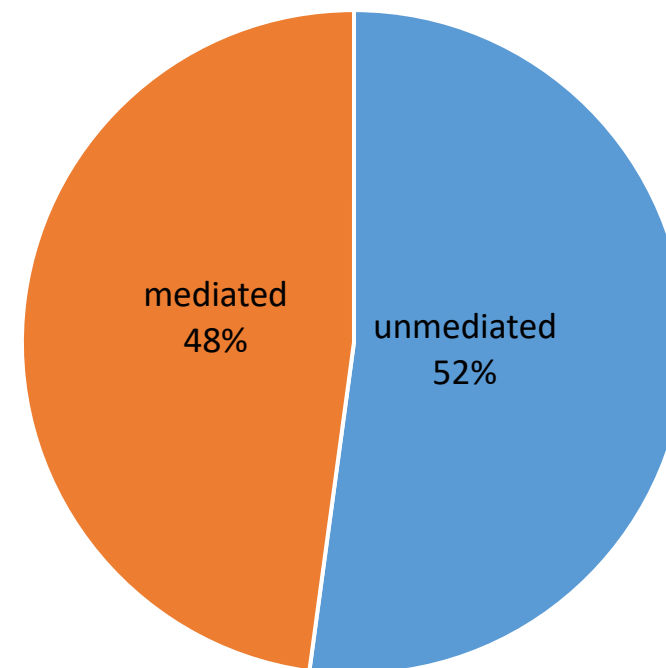


L'importance relative de l'auto-archivage

La part des comptes de contributeurs « médiatisés » peut être estimée à 1,2 %

Ce faible nombre de contributeurs représente 78 510 dépôts, soit 48% des 165 070 dépôts effectués en 2020

En d'autres termes, l'auto-archivage sur HAL correspond à un peu de la moitié de tous les dépôts (chiffres de 2020)



Les communautés de pratiques

Âge

- Jeunes chercheurs (doctorants, post-docs)
- Les anciens

- 50% : tout le monde fait pareil
- 50% : tout le monde ne fait pas pareil
- Communautés vs choix individuel

Statut

- Chercheurs (CNRS, INRAE, INRIA...)
- Enseignants-chercheurs

Discipline

- Équipes, pôles
- Communautés (méthodes, équipements...)

Autre

- Chercheurs étrangers

(5) Outils

Un écosystème d'outils

Des outils
externes
développés
pour HAL

- En amont pour alimenter HAL
- En aval pour les extractions

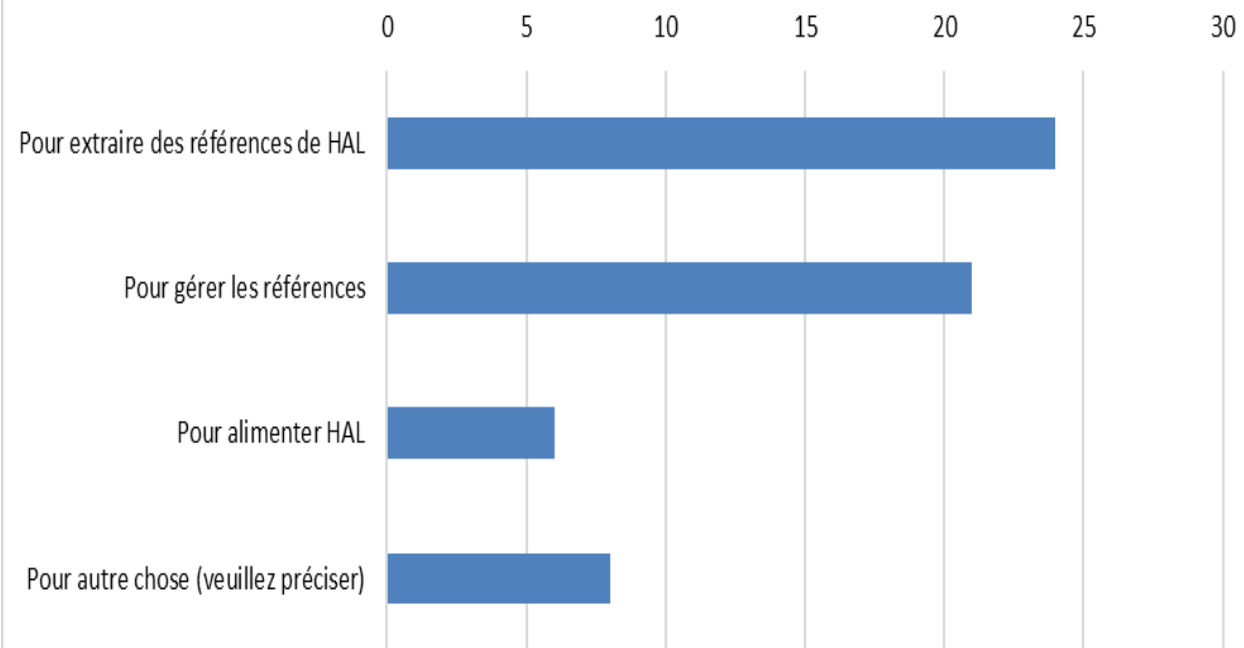
D'autres outils
externes

- Bases de données, moteurs de recherche etc.
- Gestionnaires de références

Des outils
développés
en interne

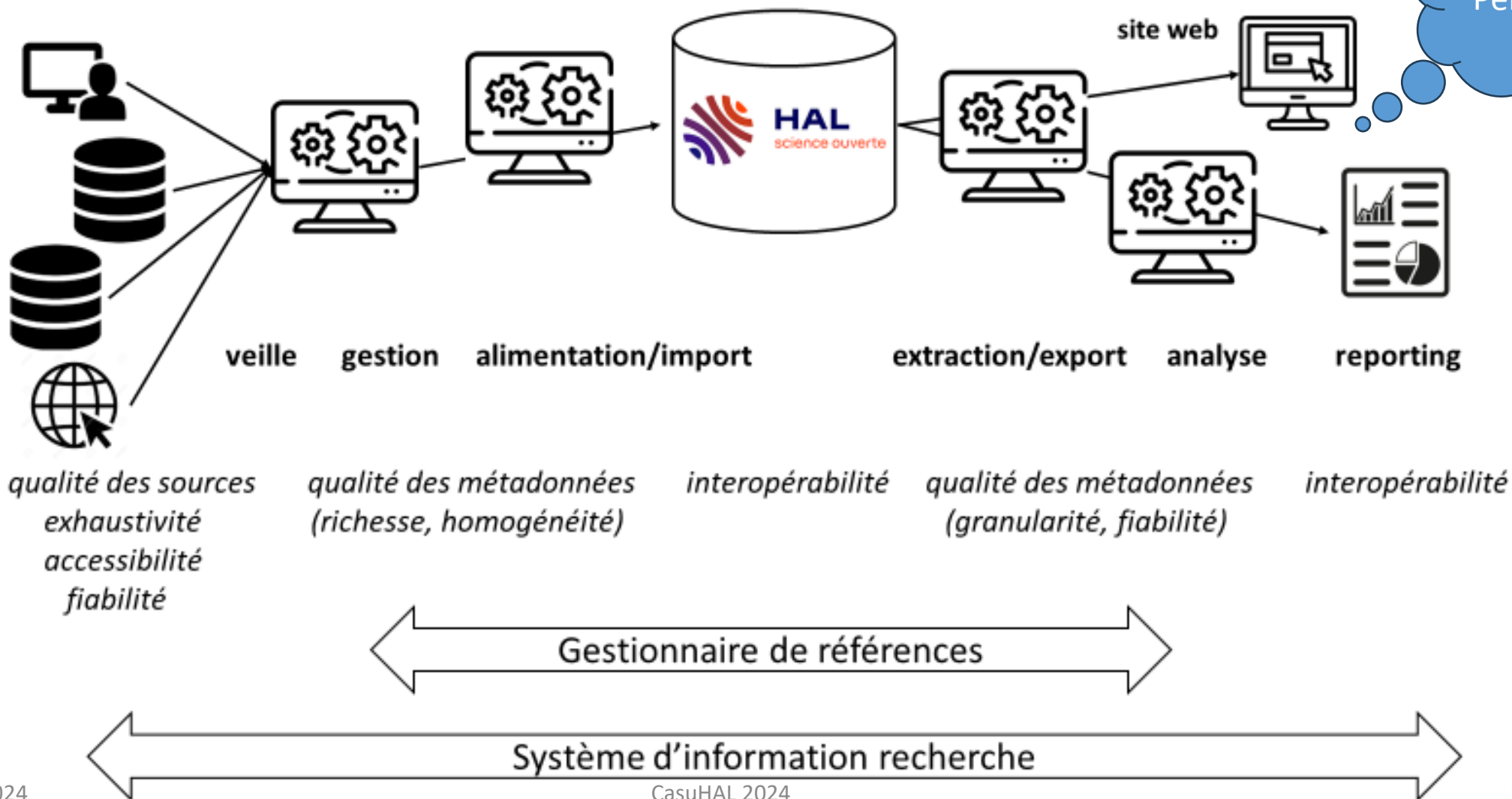
- Des bases de données
- Des outils bibliométriques

Avez-vous des outils particuliers pour préparer les bilans ?



Périmètres et enjeux de cet écosystème

Qualité ?
Pérennité ?



Les outils

Enjeux

- La veille : la qualité des sources exploitées et la fiabilité de leurs données, leur exhaustivité et aussi leur accessibilité ;
- La gestion : la qualité des métadonnées, leur richesse et aussi leur homogénéité, par exemple quand elles proviennent de sources différentes ;
- L'import/export : l'enjeu majeur est l'interopérabilité des formats, par exemple la typologie des documents, la nomenclature des domaines scientifiques, les identifiants des personnes et des organisations etc. ;
- L'analyse : comme pour la gestion, l'enjeu principal est la qualité des métadonnées, leur fiabilité et aussi leur granularité (axes, projets, équipes...) ;
- Le reporting : le problème est surtout l'interopérabilité avec les systèmes de suivi et d'évaluation des établissements et organismes et du Hcéres.

Observations

- Le périmètre des dispositifs :
 - Certains laboratoires se limitent à la gestion des références de leur production scientifique, avec des outils de gestion, d'import et d'export, où HAL joue essentiellement le rôle d'une base de données bibliographiques, parfois en remplacement d'une base interne.
 - D'autres laboratoires vont plus loin, avec des dispositifs qui, à l'instar de LabMetry, couvrent au moins partiellement les fonctions d'un système d'information sur la recherche, notamment pour assurer le suivi des projets scientifiques et fournir des informations nécessaires aux bilans, aux rapports d'activité et aux campagnes d'évaluation
 - Très peu de laboratoires semblent utiliser des logiciels fonctionnels et adaptés, comme SAMPRA ou QUASAR.
- La qualité des outils :
 - Sont-ils provisoires ou temporaires (rustines), ou représentent-ils une solution plus durable, avec une certaine stabilité ?
 - Qui en assure la maintenance et le développement ?
 - Existe-t-il une sorte de marché, une offre clairement identifiée ?
 - A quel degré sont-ils ouverts, standardisés, interopérables avec d'autres systèmes ; ou pour utiliser un terme actuel, à quel degré sont-ils FAIR ?

(6) Finalités

Trois types de discours

Un discours administratif qui met l'accent sur l'incitation institutionnelle, sur la demande de la part de la tutelle

- *“Cela devient une obligation pour les chercheurs CNRS”*

Un discours politique voire idéologique qui souligne l'intérêt d'utiliser HAL pour une recherche meilleure, plus responsable et plus ouverte à la société

- *“Une démarche très romantique, dans le sens où c'est un bel idéal (...) un peu utopique pour l'instant”*

Un discours utilitaire et pragmatique qui met en exergue l'usage de HAL pour le bon fonctionnement et le développement de l'unité

- *“Pour les bilans annuels du laboratoire. Nous avons besoin de l'information sur le nombre de publiants car notre budget en dépend en partie. On utilise HAL aussi pour assurer le suivi de l'utilisation d'un certain nombre de ressources”*

Pourquoi utiliser HAL ?

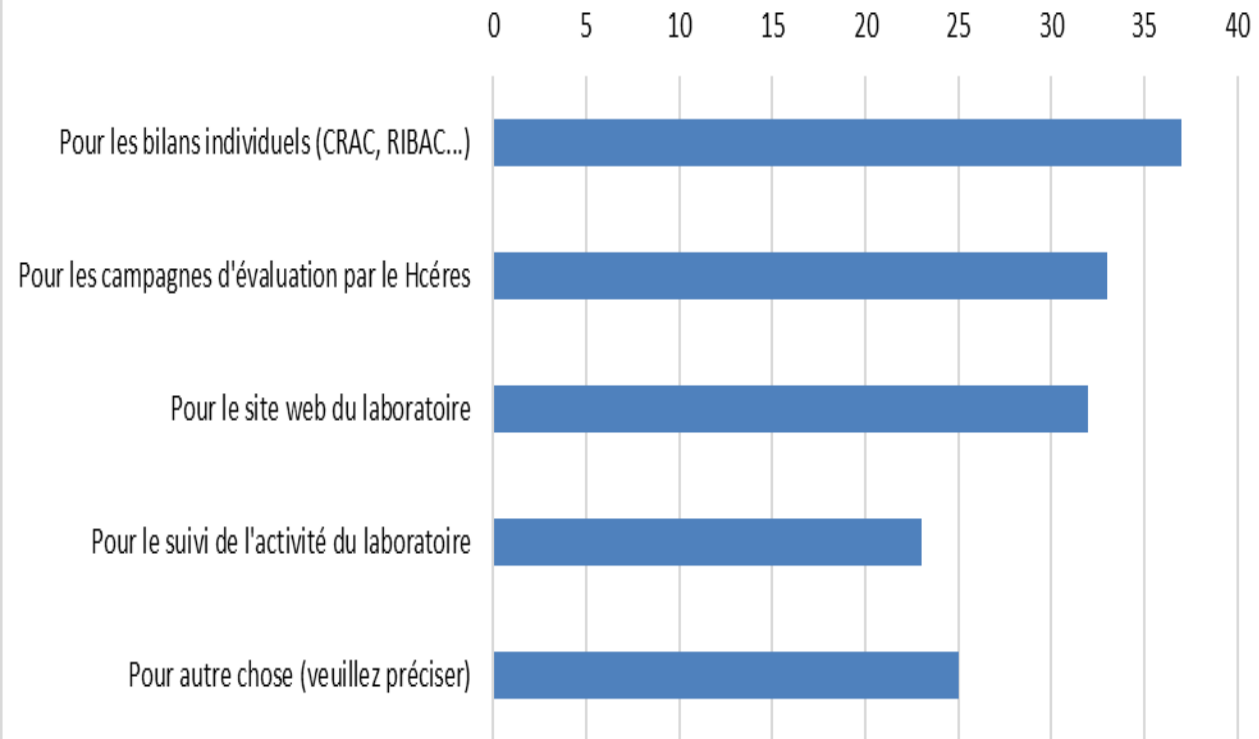
Pratiques individuelles vs collectives

Souvent plusieurs finalités

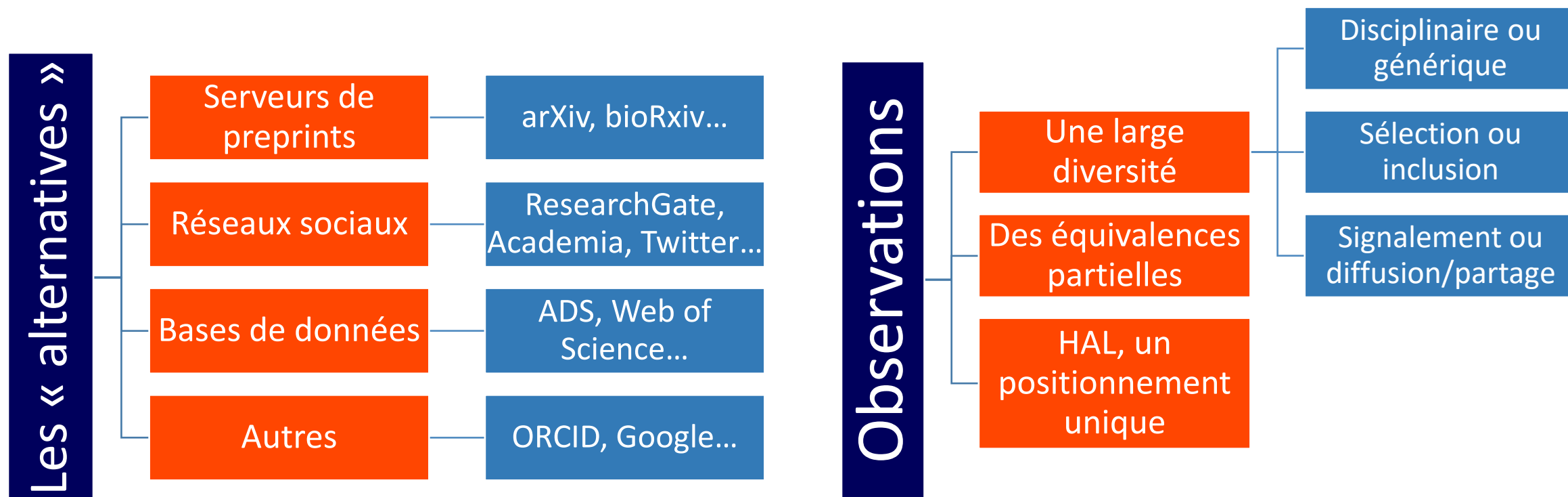
Pour certains, surtout pour la diffusion des résultats

Pour d'autres, outil indispensable de gestion

Comment utilisez-vous HAL dans votre laboratoire ?



Les équivalents fonctionnels



(7) Accompagnement

Une approche collective

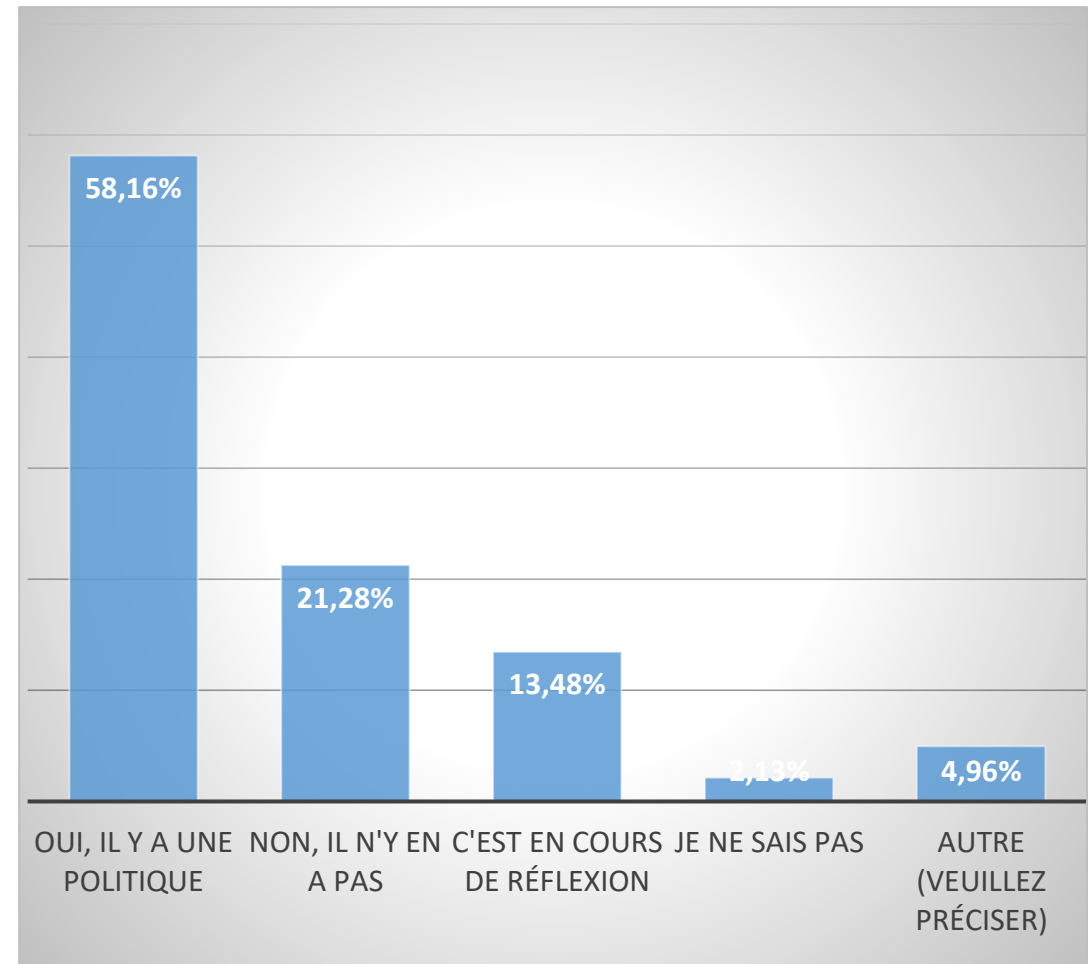
80% des unités ont discuté de HAL

La plupart des laboratoires ont mis en place une politique HAL

Souvent en AG, sinon en conseil ou en bureau, en séminaire, en réunion d'équipe ou en atelier de formation, ou d'une manière plus informelle

Souvent incitation

Impact des tutelles (bilans, dotation budgétaire)



Définition d'une stratégie HAL (N=141)

L'accompagnement dans les unités

Facilitation	Aider à faire • <i>Conseil, aide au dépôt</i>
Médiation	Apprendre à faire • <i>Formations collectives, sensibilisation, promotion de HAL</i>
Coproduction	Faire avec • <i>Formations individuelles, conseil</i>
Sous-traitance	Faire à la place • <i>Gestion de la collection, saisie des références</i>

- 2/3 oui, différents types
 - Facilitation (aider à faire)
 - Médiation (apprendre à faire)
 - Coproduction (faire avec)
 - Sous-traitance (faire à la place)
- 1/3 non, pas d'accompagnement
 - Pas besoin
 - Pas de ressources
 - Pas systématique
 - Délégation/externalisation
 - SCD, CCSD

La formation, l'accompagnement

- **Présentation moments collectifs (AG/JE)**
- **Formation des nouveaux EC, des doctorants**
- **Formations courtes**
- **Accompagnement individualisé (création IdHal)**
- **Autoformation (webinaire, tutoriels...)**

NÉCESSITÉ D'UN ACCOMPAGNEMENT

- *« Parce qu'il y a le discours que c'est de la communication scientifique directe entre chercheurs et qu'ils font tout tous seuls ; mais beaucoup d'entre eux ont besoin d'une médiation »*

INTÉRÊT MÉDIATION DE PROXIMITÉ

- *« Les gens de la bibliothèque sont allés voir les chercheurs dans les laboratoires sur place, et ils ont beaucoup apprécié »*
- *« Nous faisons aussi des rappels régulièrement par courriel, pendant les AG et c'est très efficace quand la référente montre les chiffres, les graphiques et les commente en AG, car cela stimule les collègues ; on fait aussi circuler les rapports faits à partir des extractions de HAL et cela incite les chercheurs à participer et être actifs sur HAL (quand ils voient le risque de ne pas être pris en compte dans le bilan, par exemple) »*

Médiation et gestion documentaire

- Mise à jour / corrections / validation des affiliations
- Archivage pérenne / rétroconversion
- Identifiant chercheur / tamponnage
- Versement dans d'autres AO
- Ajout du texte intégral

RÔLE DU RÉFÉRENT HAL

« Elle reçoit une alerte dès qu'il y a un dépôt d'un des chercheurs des trois laboratoires... Elle vérifie l'intégralité des métadonnées, l'ordre des auteurs, les affiliations auteurs / labos des dépôts par rapport aux affiliations de la publication, ajout de l'abstract ou de la communication entière, vérifie le titre, le DOI »

MÉMOIRE SCIENTIFIQUE

« Cette collection contient aussi des notices : il s'agit de garder une trace du passé scientifique du laboratoire, même sans document. »

AFFICHAGE PUBLICATIONS AUTEUR

« C'est regrettable qu'après tant d'années, HAL n'a pas mis en place une sorte d'API qui affiche la liste de publications de tout auteur sur n'importe quel page Web »

« On peut regretter que HAL n'aille pas moissonner automatiquement les nouvelles publications, et que les auteurs ne soient pas juste sollicités pour le texte intégral. »

Les acteurs de la gestion et du dépôt dans HAL

Quelle catégorie de personnels ?

- Référent HAL (administratif / chercheur)
- Documentaliste (labo)
- Bibliothécaire SCD
- Auteurs
- Contrat doctorant (rattrapage biblio HCERES)

« Le directeur dit qu'il ne va jamais déposer ou demander aux chercheurs de faire les dépôts ; c'est le travail des bibliothécaires »

Pour quels rôles ?

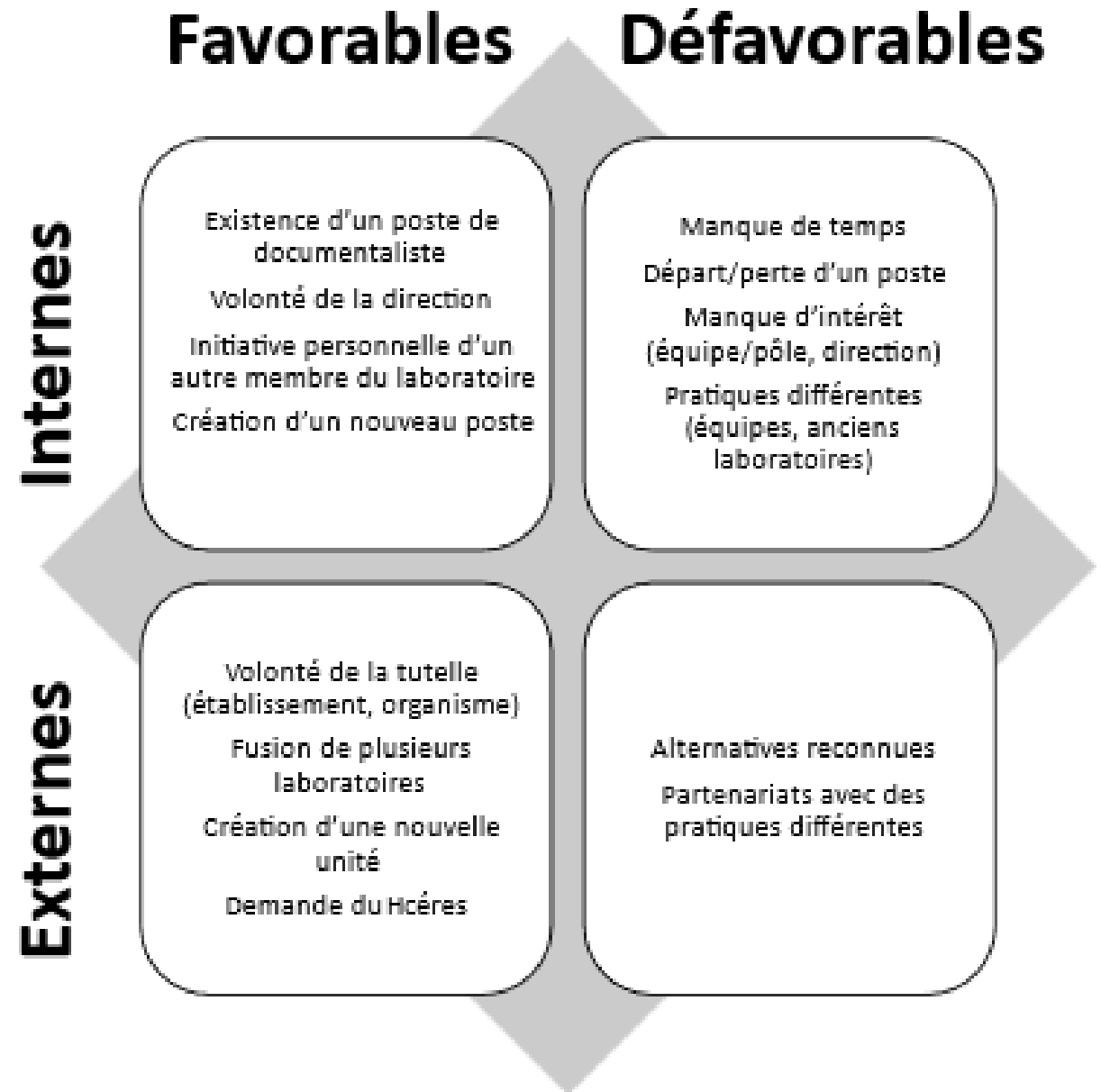
« La saisie dans HAL est du ressort du chercheur. Notre documentaliste est là pour gérer les collections, pour homogénéiser les signatures, pour faire des statistiques et pour avoir des indicateurs plus ou moins automatiques pour l'HCERES. Elle n'est pas là pour saisir les publications des chercheurs »

Critique

« Quand l'université leur [aux EC] demande de déposer sur HAL, c'est perçu comme un ajout de tâches de type administratif »

Facteurs favorables et défavorables d'une pratique de HAL dans les laboratoires


Pour la plupart des laboratoires, il s'agit d'une sorte d'acceptation sans acculturation.



(8) Écosystème

L'environnement

44 % des laboratoires se trouvent au sein d'un établissement explicitement favorable au libre accès, avec une politique affichée et une archive institutionnelle sur HAL et/ou locale.



L'impact de cet environnement sur les choix des laboratoires est significatif mais ne détermine pas tout.



Même dans un tel environnement, certains laboratoires n'ont ni stratégie, ni collection; mais cela reste marginal.



Le lien entre l'environnement et la valorisation par les laboratoires est statistiquement significatifs (pour les bilans, évaluation Hcéres, site web).

Quelle politique institutionnelle ?

Nationale

- Science ouverte : PNSO, financeurs, Loi pour une République numérique
- Évaluation HCERES (bilan bibliométrique, visibilité)

« En 2018, a eu lieu l'évaluation HCERES et c'était obligatoire de déposer au moins les notices (car l'évaluation ne prend en compte que ce qui est sur HAL) ; à ce moment-là il y a eu un grand mouvement pour le dépôt sur HAL... »

Établissement

- Politique tutelles obligation de dépôt : CNRS (CRAC compte-rendu annuel d'activités), INRIA...
- Archives ouvertes locales : OSKAR (Bordeaux), UnivOAK (Strasbourg)
- BONUS-MALUS : *« Il existe un bonus performance, et pour déclencher ce bonus performance et obtenir la gratification de l'université (max. 10% de la dotation annuelle initiale), ... il faut au moins un tiers de dépôts avec texte intégral pour l'année N-1 pour déclencher le bonus »* (mis en place en 2017 par l'AMU)
« Malus sur budget si non-respect de la charte des signatures et si pas assez de texte intégral dans HAL. »

(9) Perspectives

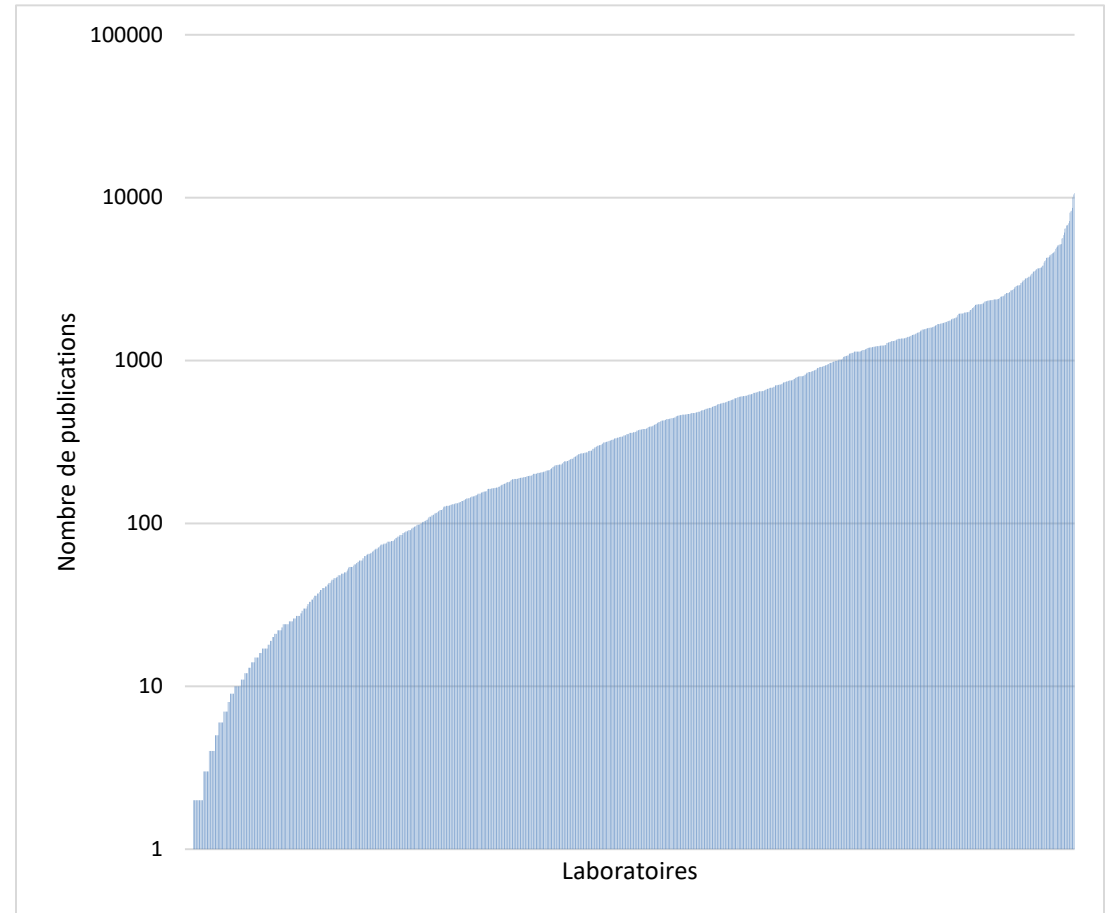
9.1 - Un usage généralisé : Le début d'une longue traîne

99% des laboratoires ont des publications sur HAL

Le nombre des dépôts varie de 1 à 10645 (moyenne 843, médiane 359)

Le début d'une longue traîne

- 20% dépôts \approx 70% unités
- 20% unités \approx 67% dépôts



Nombre de publications par laboratoire (N = 1228 laboratoires)

La typologie des dépôts

L'article est de loin
le contenu le plus
important

Trois types de
documents sont
déposés par au
moins 80% des
laboratoires :
communications,
thèses et chapitres

	% laboratoires	% dépôts
Articles	99%	52%
Communications	87%	23%
Thèses	86%	4%
Chapitres	80%	8%
Preprints	68%	3%
Posters	62%	1%
Ouvrages	58%	2%
Direction ouvrages	55%	2%
Rapports	51%	2%

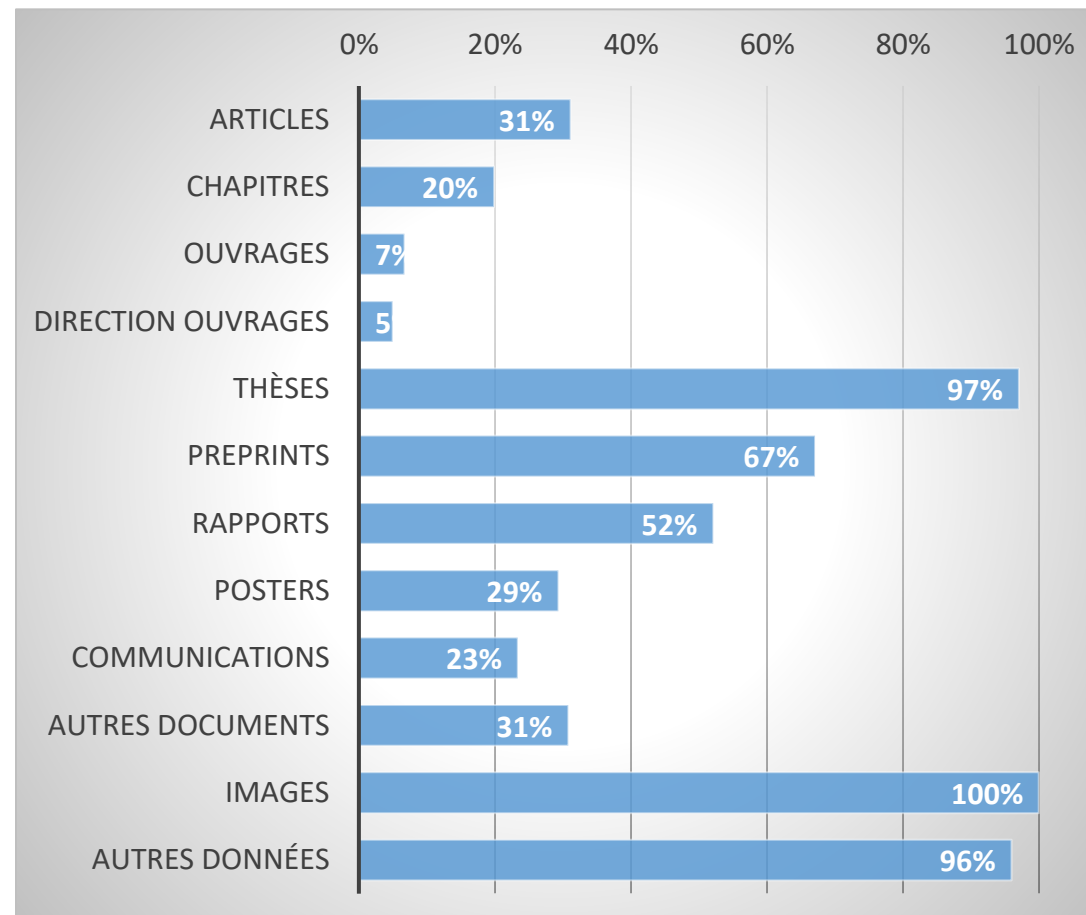
Une ouverture toute relative

Le taux d'ouverture est de 32%

Moins ouvert : articles,
chapitres, livres,
communications,
posters...

Plus ouvert : rapports,
preprints, thèses

Les thèses sont un cas particulier :
obligation de déposer le document
(tout comme les images et autres
données)



La part des dépôts avec fichiers (% *open access*) (N = 1035612 dépôts)

HAL, un partenaire obligé et essentiel

HAL est devenue une infrastructure incontournable de la politique en faveur de la science ouverte.

Les chercheurs utilisent HAL, peut-être moins pour chercher de l'information mais surtout pour signaler et diffuser leurs travaux.

Les laboratoires jouent un rôle important pour la mise en œuvre de cette politique.

Ce rôle varie d'un laboratoire à l'autre, en fonction des ressources et personnels, des engagements et initiatives, de l'environnement institutionnel et des tutelles et des traditions propres aux unités et aux communautés.

On l'apprécie ou on ne l'apprécie pas. Mais (presque) tous connaissent et utilisent HAL, plus ou moins.

HAL est sorti de sa phase « de jeunesse » pour devenir un partenaire obligé et essentiel des chercheurs et de leurs unités de recherche.

9.2 – La transformation de HAL : De la « voie verte » au SI recherche

Nos données empiriques révèlent à grande échelle la transformation de HAL d'une archive ouverte (comme arXiv) en une plateforme ouverte avec des publications et des métadonnées provenant de différentes sources.

La contribution médiatisée nécessite (et reflète) un soutien institutionnel, avec trois objectifs :

- Le développement du libre accès et de la communication scientifique directe;
- la préservation à long terme des ressources (CINES);
- le développement d'une infrastructure pour le suivi et l'évaluation de la production scientifique des chercheurs, des laboratoires, et des tutelles.

Grâce au soutien institutionnel et aux contributions des laboratoires et des universités, HAL est devenu une sorte de vitrine de leur production scientifique; il fournit également des données à l'Open Science Monitor français.

La contribution médiatisée est une partie importante du fonctionnement normal et permanent du dépôt (pas temporaire).

De nouveaux enjeux

Cette transformation n'est pas spécifique pour la France

Les nouveaux enjeux sont ceux des infrastructures de la recherche et en particulier, ceux des SI sur la recherche

La qualité des métadonnées

La FAIRness des données et métadonnées

La transformation de l'évaluation scientifique (DORA, CoARA...)

Les questions éthiques liées à l'évaluation

La transformation de HAL

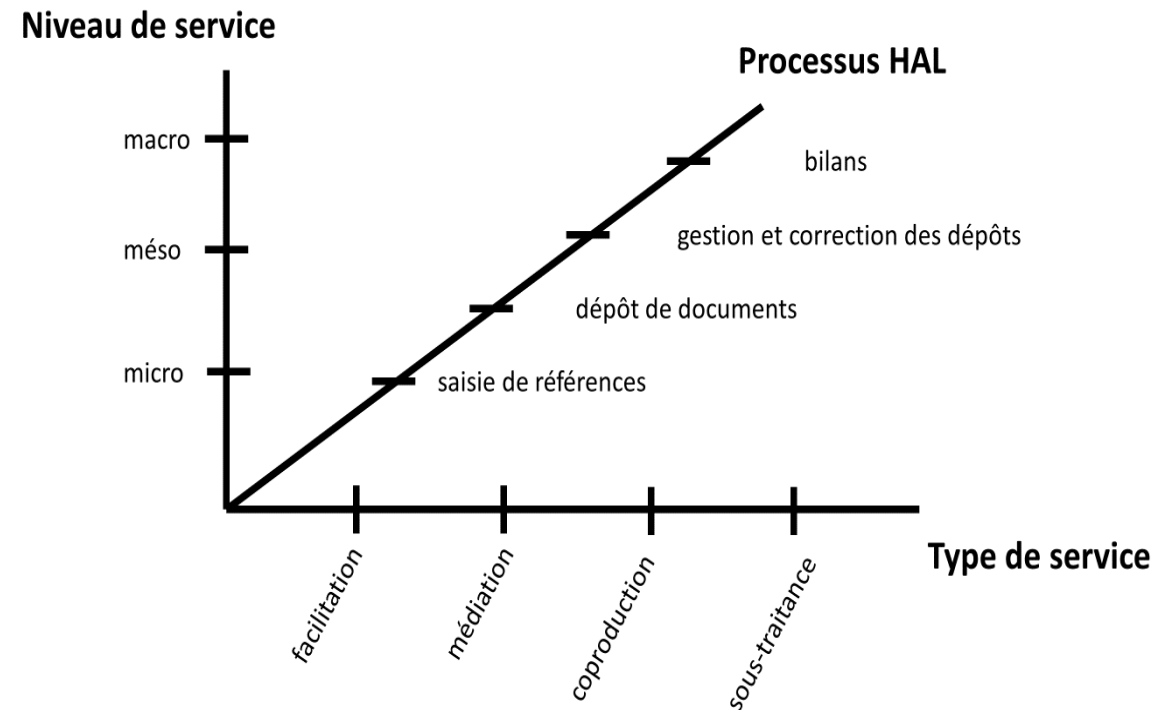
- *“Chaque fois qu’on a besoin d’un indicateur, on utilise HAL”*
- Intégration de nouveaux services et fonctionnalités dans HAL ?
 - Remplacement des outils mis en place dans les laboratoires
 - Externalisation (délégation) de certaines fonctions (CCSD, SCD...)
 - Risque d’une perte de contrôle et de compétences par les laboratoires
- Ouverture des données de HAL pour développer d’autres outils ?
 - Par qui : communautés de pratique, secteur privé ?
- L’enjeu de la qualité des métadonnées
- L’enjeu éthique

9.3 - Des pratiques en évolution

- Les entretiens révèlent des pratiques en évolution, dans des situations dynamiques dont les variables indépendantes se situent à plusieurs niveaux :
 - Organisation : création ou fusion de structures (unités, établissements...)
 - Institution : incitation et/ou soutien (formation...)
 - Hcéres : nouvelle campagne d'évaluation
 - Ressources humaines : départ ou arrivée d'une nouvelle personne (direction, informatique, documentation)
 - Dispositifs : création ou disparition d'outils, de fonctionnalités ou de services
- Ces différents facteurs sont liés
- Les trajectoires sont plus ou moins longues
- Ce que les entretiens révèlent surtout, c'est la capacité d'adaptation des laboratoires afin de relever les défis internes et externes

Esquisse d'un modèle des pratiques

- Niveau des actions
 - Micro/méso/macro
- Type d'accompagnement
 - Médiation/facilitation/...
- Processus scientifique
 - Evaluation/publication/...



Analyse stratégique

- Une large diversité de situations
- Possibilité d'identifier des facteurs favorables et défavorables
- Des trajectoires dynamiques
- Parfois, des stratégies assumées
- Souvent, une adaptation aux opportunités et contraintes
- La capacité d'adaptation comme facteur principal ?

9.4 – Pistes de recherche

- Etude longitudinale de l'utilisation de HAL
 - Le signalement des publications sur les sites web des laboratoires
- Etude longitudinale de l'évolution de HAL
- Communautés de pratique
- Qualité des métadonnées
- L'impact du Plan S (*rights retention strategy, responsible publishing*)
- L'impact de DORA/CoARA (*responsible research assessment*)

- *Science studies* (SIC, STIC, sociologie...)

Bibliographie

- Larrieu, M., & Schöpfel, J. (2022). Différences disciplinaires en contexte de Science ouverte. Étude avec les publications de l'archive ouverte HAL. *DocSoc 2022. 8e Conférence Document Numérique & Société. 23-24 Juin 2022, Liège (Belgique)*. <https://hal.science/hal-03760316>
- Schöpfel, J., Kergosien, E., Prost, H., & Barrié, J. (2022). « Pas si simple que ça... » : Une enquête sur l'usage de HAL par les unités de recherche des universités IdEx. *I2D - Information, Données & Documents*, 2(2), 150–183. <https://doi.org/10.3917/i2d.222.0150>
- Schöpfel, J., Chaudiron, S., Jacquemin, B., Kergosien, E., Prost, H., & Thiault, F. (2023). The Transformation of the Green Road to Open Access. *Publications*, 11(2), 29. <https://doi.org/10.3390/publications11020029>
- Schöpfel, J., Thiault, F., Prost, H., Jacquemin, B., & Kergosien, É. (2023). L'utilisation de HAL par les laboratoires de recherche. *Balisages*, (6). <https://doi.org/10.35562/balisages.1166>
- Schöpfel, J., Prost, H., Jacquemin, B., Kergosien, É., & Thiault, F. (2024). La valorisation de HAL : Finalités, outils et process. *JDMDH Preprint*. <https://hal.univ-lille.fr/GERIICO/hal-04450898v1>
- Schöpfel, J., Kergosien, É., Prost, H., Jacquemin, B., & Thiault, F. (2024). L'utilisation de HAL : Pratiques, avis et alternatives. *JDMDH Preprint*. <https://hal.univ-lille.fr/GERIICO/hal-04523829v1>

Merci !

joachim.schopfel@univ-lille.fr

florence.thiault@univ-rennes2.fr

Remerciements : Le projet HAL/LO a été financé par le GIS « Réseau Urfist ».

