



Observatoire des Sciences et Techniques

# Indicateurs de production S&T des établissements français

Journée IPERU

Paris, Université Paris Diderot et Hcéres – 10 décembre 2018

---

SYNTHESE

Mai 2019

# SOMMAIRE

## SEANCE PLENIERE

Accueil.....	3
Le contexte français et international.....	3
Le programme IPERU et ses évolutions.....	6
Conception, interprétation et usage responsables des classements d'universités.....	9
Archives institutionnelles, archives ouvertes : leur rôle dans le pilotage et l'évaluation des établissements.....	13

## LES ATELIERS

Base des publications et indicateurs bibliométriques .....	22
Base des brevets et indicateurs brevets.....	25
Base des projets du PCRI et indicateurs sur la participation aux projets H2020.....	28
Indicateurs sur la production scientifique : le cas des SHS.....	31
Méthodologie relative au calcul des indicateurs.....	36

# SEANCE PLENIERE

## ACCUEIL

---

*Michel COSNARD*  
*Président du Hcéres*

Mesdames et messieurs, soyez les bienvenus à cette conférence sur IPERU. L'acronyme désigne les Indicateurs de Production des Etablissements de Recherche Universitaire. Cette journée d'échanges se déroule en deux temps avec une session plénière ce matin et des ateliers cet après-midi. La matinée vise à présenter le contexte national et international. Les ateliers de l'après-midi permettront aux participants d'actualiser leurs connaissances et d'aborder les évolutions du programme IPERU.

Je laisse maintenant la parole à Alain Abécassis, chef du service de la coordination des stratégies de l'enseignement supérieur et de la recherche au ministère de l'Enseignement supérieur, de la recherche et de l'innovation (MESRI). Alain Abécassis est un peu le commanditaire d'IPERU ; il a suivi le programme quasiment depuis sa création. L'OST travaille pour la DGESIP dans le cadre des questions de contractualisation des établissements.

Je profite de la parole qui m'est donnée pour saluer la présence de Brigitte Plateau, directrice générale de l'Enseignement supérieur et de l'insertion professionnelle au sein du MESRI. Je la remercie de sa présence et de l'intérêt qu'elle porte à ces questions. Je remercie aussi l'Université Paris Diderot qui nous accueille dans ses locaux ce matin et nous permet de tenir cette séance plénière dans de très bonnes conditions.

## LE CONTEXTE FRANÇAIS ET INTERNATIONAL

---

*Alain ABÉCASSIS*

*Chef du Service de la coordination des stratégies de l'enseignement supérieur et de la recherche, Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation (MESRI)*

Mesdames et messieurs, je suis très heureux de me trouver parmi vous et d'avoir répondu à l'invitation que m'a faite Frédérique Sachwald pour intervenir au nom du ministère en ouverture de cette journée de réflexion et de travail.

J'ai eu certes la chance d'accompagner le programme IPERU, mais je ne le suis pas exactement depuis sa création. Ce programme a été mis en place à l'époque où le directeur de la DGESIP, Jean-Marc Monteil, souhaitait affirmer très fortement le rôle d'opérateur de recherche des universités. Il souhaitait ainsi identifier via des indicateurs et avec le concours de l'OST quels étaient les dominantes et les éléments de spécialisation des universités en matière de recherche au travers de leurs publications et des données que l'OST pouvait exploiter.

Le programme IPERU, de manière discrète – puisque je déplore qu'il ne soit pas suffisamment connu et valorisé du côté de la Conférence des présidents d'universités, comme du côté des indicateurs suivis par le ministère – a accompagné, voire suscité, les transformations importantes intervenues depuis dix-douze ans dans le paysage de l'enseignement supérieur et de la recherche.

Le programme IPERU est valorisé et utilisé sous des vocables divers dans le cadre du Programme des Investissements d'Avenir. A travers les contributions que l'OST a pu apporter aux travaux des jurys qui se sont succédé, les indicateurs produits par le programme IPERU ont constitué des éléments d'éclairage sur la compréhension et l'aide à la spécialisation. De manière discrète, il a aidé les établissements et leurs

directions à s'emparer de ces éléments pour définir des stratégies et identifier des éléments de spécialisation et de dominante.

Le Programme des Investissements d'Avenir a apporté de manière explicite, par les pouvoirs publics comme au niveau de chacun des porteurs de projet, des informations en vue d'assumer la différenciation et les stratégies de spécialisation qui peuvent être celles des établissements et leurs regroupements. Cette stratégie de spécialisation est devenue tellement claire qu'elle sert aussi d'inspiration et de fil conducteur à la réflexion qui est engagée entre la DGESIP et la DGRI pour refonder la démarche de contractualisation, à laquelle sont associés désormais plus étroitement les organismes de recherche. Il s'agit de faire en sorte que les contrats de site qui vont se développer, avec la nouvelle vague, prennent en compte et soutiennent les éléments de spécialisation qui peuvent être ceux des établissements et des regroupements. Dans ce cadre, IPERU aura un rôle clé à jouer pour aider à définir les éléments de spécialisation.

Au-delà des publications qui peuvent traduire la spécialisation en termes de recherche, les établissements ont vocation à établir des liens entre les politiques de formation et de recherche. De ce point de vue, le mouvement qui est engagé, en particulier avec les écoles universitaires de recherche, notamment *via* les laboratoires d'excellence, concerne évidemment la dimension recherche au travers des publications, mais elle concerne aussi le volet enseignement et le lien extrêmement fort qui doit être établi entre formation et recherche. Je voudrais souligner au passage un élément incident, mais qui démontre que ces démarches ne sont pas isolées mais participent d'un mouvement d'ensemble : la réflexion en cours et la concertation engagée sur la prise en compte des éléments de formation et de pédagogie dans les parcours des enseignants-chercheurs est importante pour valoriser la dimension recherche comme la dimension formation dans les carrières.

La contractualisation concerne non seulement les établissements, leurs sites et leurs regroupements, mais aussi les organismes de recherche afin de prendre en compte le rapprochement qui est à l'œuvre avec les universités. A vrai dire, ce rapprochement existe déjà depuis longtemps, *via* les unités mixtes de recherche, dont les activités quotidiennes ne font pas nécessairement la différence entre les institutions auxquelles elles sont rattachées. En définitive, les activités d'un laboratoire ou d'une unité mixte font valoir autant le rôle des universités que celui des organismes. Cela se lit en particulier au travers des politiques de signature qui ont considérablement évolué et qui font d'abord valoir la dimension de site et l'appartenance des membres d'une unité mixte à des laboratoires et à des unités. Le mouvement qui est en cours en France, notamment pour toutes les comparaisons internationales, est un mouvement qui ne distingue plus, au travers de la signature, l'appartenance institutionnelle à une université ou à un organisme. Le mouvement se dirige plutôt vers une signature commune pour que soient mises en avant la dimension du site et la politique menée en commun par les établissements sans opposer les organismes et les universités.

Dans une perspective internationale, les classements internationaux représentent un symptôme ou un indicateur. Depuis une quinzaine d'années, chacun sait que les discussions sont très vives à ce sujet. La contestation des différents types de classements n'a d'égales que la prolifération des classements et l'attention qui leur est portée par tous les responsables. Ces classements permettent d'opérer des comparaisons internationales, quelles que soient les critiques dont ils peuvent légitimement faire l'objet. De manière inévitable, délibérée, assumée et de mieux en mieux outillée, les établissements sont amenés ainsi à se comparer les uns aux autres, y compris à l'échelle internationale. Pour élaborer des stratégies d'établissements, des stratégies de spécialisation, des stratégies de site et des stratégies de différenciation, les classements sont des outils que les établissements peuvent mobiliser. Les établissements bénéficient en outre des rapports d'indicateurs du programme IPERU : ils offrent des éclairages mesurés et objectifs permettant aux établissements de se situer au niveau national et international (en choisissant le type d'établissements par rapport auxquels se comparer, le champ disciplinaire ou encore en optant pour un classement d'ensemble ou thématique).

Après avoir contesté et méprisé les classements pendant des années et après les avoir relativement ignorés dans un second temps, le ministère est désormais plutôt engagé dans un accompagnement des établissements pour comparer les classements. Il ne s'agit pas de jouer un rôle de prescripteur, mais de bien préciser les caractéristiques des classements et leurs effets afin d'éclairer au mieux les établissements.

Je souhaite ajouter quelques mots sur les démarches de science ouverte qui sont engagées dans une accélération et dans une effervescence internationale, qui sont parfois difficiles à maîtriser et qui font l'objet de beaucoup de débats et de tensions. Le ministère accompagne ces démarches de manière volontaristes. La semaine dernière, ont eu lieu les journées pour la science ouverte au ministère. Début juillet, la Ministre a engagé un plan pour la science ouverte. La France s'est, par ailleurs, associée au plan S lancé par la Commission européenne pour favoriser les démarches de science ouverte.

La production d'indicateurs sur les publications et leurs caractéristiques suppose de disposer de données bibliométriques. Aujourd'hui les bases de données bibliométriques permettant les comparaisons internationales sont très largement les bases commerciales éditées par Clarivate Analytics et Elsevier. N'étant pas spécialiste de la science ouverte, et les spécialistes me corrigeront, j'ai compris que des sources de données plus ouvertes se développent, comme CrossRef. L'accès aux données sur les publications s'inscrit ainsi dans la démarche plus générale de la science ouverte, partant du principe que les publications existantes à l'échelle mondiale ne doivent pas être uniquement ouvertes à quelques-uns et rester dans les mains de grandes entreprises multinationales qui commercialisent parfois l'accès à leurs bases à des tarifs très élevés. La pression pour l'ouverture et le partage est de plus en plus forte.

Dans ce contexte général, l'initiative du programme Bibliothèque Scientifique Numérique, le projet Conditor vise à recenser l'ensemble des publications des scientifiques français. Le baromètre de l'open access permet de lister les publications françaises disponibles en accès ouvert. Ces deux démarches menées doivent permettre d'accroître la visibilité sur l'ensemble des politiques de publications et de référencement. Pour l'OST et pour IPERU, les nouvelles informations qui doivent être issues de ces outils pourront constituer un enrichissement. Elles pourront sans doute contribuer à faire évoluer les indicateurs.

Au final, les évolutions de l'outil IPERU qui vont être évoquées lors de cette journée devraient lui permettre de mieux contribuer à l'élaboration des stratégies d'établissement.

Notre ministère poursuit l'accompagnement des acteurs dans l'élaboration des stratégies de site, au travers de la contractualisation ou au travers des programmes des investissements d'avenir, afin de renforcer la visibilité internationale de l'enseignement supérieur et de la recherche français.

Je vous remercie de votre attention et je vous souhaite de fructueux échanges durant cette journée.

# LE PROGRAMME IPERU ET SES EVOLUTIONS

---

Animation

Frédérique SACHWALD  
*Directrice de l'OST*

Hélène MECHOULAM  
*Responsable du programme IPERU, OST*

Françoise LAVILLE  
*Responsable du corpus Brevets, OST*

## Frédérique SACHWALD

Cette session présente le programme IPERU et ses évolutions récentes. Elle permet d'introduire la réflexion sur les évolutions du programme IPERU qui sera approfondie lors des ateliers de l'après-midi.

## Hélène MECHOULAM

 [Présentation en pdf](#)

La présentation décrit rapidement le programme IPERU avant d'aborder les interactions entre avec d'autres activités de l'OST. En effet les résultats du programme IPERU nourrissent différentes productions de l'OST et, réciproquement, de nouvelles analyses peuvent venir enrichir les productions d'IPERU.

Le programme IPERU a été lancé en 2006-2008 avec le soutien du MESRI pour fournir aux établissements du programme 150 de la LOLF des indicateurs d'aide au pilotage. Actuellement, suite aux scissions et fusions d'établissements, il concerne 130 établissements. Le programme produit des indicateurs de production à partir de trois corpus : les publications scientifiques, les brevets et les participations aux projets des programmes-cadres de recherche et d'innovation européens.

### **Le repérage institutionnel des publications, brevets et projets européens**

L'OST enrichit les bases de données sources par le repérage des établissements du périmètre IPERU. Le travail de repérage est le préalable au calcul d'indicateurs à la maille institutionnelle. La spécificité du programme consiste dans la validation de ce repérage par les établissements eux-mêmes. Les établissements sont considérés comme une somme de laboratoires et non pas comme une somme de chercheurs.

La mise en œuvre se déroule différemment suivant les corpus. Pour les publications scientifiques, l'OST propose une aide aux établissements pour effectuer le repérage des adresses institutionnelles de leurs publications par le biais d'un pré-cochage automatique dans la base de données. Si l'établissement le souhaite, il y a ainsi un travail en amont pour aboutir à une liste de mots clés. L'OST effectue un premier contrôle de cette liste de mots clés afin d'éviter les faux positifs qui aboutirait à un sur-repérage. Suit une étape de contrôle manuel pour garantir la qualité du repérage. Les résultats finaux sont validés par l'établissement. Pour les Brevets et les projets PCRI, les données se présentent avec une variété d'occurrences désignant l'établissement bien moindre, ce qui permet à l'OST d'effectuer un pré-repérage manuel sur la base des données existant dans les bases de données sources.

Pour le périmètre de repérage « strict », le travail de l'établissement consiste alors à valider ou contester le pré-repérage. Il peut en outre signaler son implication dans des brevets ou projets qui figurent dans la base mais que l'OST n'a pas pu identifier (ex : CNRS seul signataire de la convention de subvention pour un projet réalisé dans une UMR avec une université). Ce signalement qui génère un périmètre de repérage « large » et peut permettre d'apprécier plus complètement le nombre des brevets ou projets européens dans lesquels l'établissement est impliqué.

## **Les rapports d'indicateurs IPERU**

Les indicateurs bibliométriques calculés sur les données de publication sont les plus développés, les rapports sont plus ou moins fournis selon le groupe auquel appartient l'établissement. Le nombre de publications et les parts de co-publications européennes, extra-européennes ou internationales sont fournis à tous les établissements, mais l'indice d'activité par classe de citations, par exemple, n'est fourni qu'aux établissements des groupes 1 et 2, ayant respectivement entre 300 et 500 publications en moyenne par an et plus de 500 publications. De même, seuls les établissements des groupes 1 et 2 sont positionnés entre eux sur les graphiques qui permettent de visualiser simultanément les indices d'impact et de spécialisation par discipline.

Pour les brevets et les projets du PCRI, des indicateurs étaient fournis à tous les établissements comme le nombre de brevets publiés, la part de brevets en collaboration internationale, la part des brevets français, le nombre de projets, les taux de participation et de coordination et la coprésence éventuelle d'autres pays dans les projets.

Ces indicateurs étaient seulement fournis sur le périmètre strict vérifié par l'OST ; les dénombrements sont fournis sur le périmètre large enrichi par l'établissement.

Plusieurs points sont acquis dans le cadre du programme IPERU. On utilise la même base de données, actualisée au même moment, ainsi que les mêmes modalités de repérage. En outre, la méthode de calcul est homogène. Ce choix identique pour tous les établissements constitue un préalable nécessaire pour garantir la comparabilité des résultats entre les établissements.

Le dispositif a évolué. Le périmètre a été adapté à chaque fois que possible, notamment pour prendre en compte le périmètre large. Indépendamment de remarques formulées par les établissements, le programme IPERU a bénéficié d'améliorations méthodologiques et sur les bases de données. Au cours des deux dernières années, les indices d'impact ont été normalisés au niveau des domaines de recherche (252 dans la base de publications, les indicateurs reposant sur un trop petit nombre de publications ont été grisés pour attirer l'attention des établissements et les co-publications européennes sont désormais définies de manière stricte (uniquement des pays de l'UE 28). En 2019, des améliorations seront apportées aux indicateurs portant sur les brevets et les PCRI qui n'avaient pas été produits au cours des dernières années. Enfin, l'OST réalise des explorations pour améliorer la couverture des données et leur traitement.

## **L'utilisation des enrichissements IPERU dans d'autres productions de l'OST**

L'OST fournit des indicateurs bibliométriques au MESRI pour alimenter les documents budgétaires (projet annuel de performance et rapport annuel de performance). Entre 2005 et 2015, la construction du périmètre du programme 150 de la LOLF donnait lieu à des pré-cochages automatiques à partir de mots clés. Depuis 2015, l'OST consolide les repérages des établissements IPERU pour déterminer le périmètre du programme 150.

Depuis 2017, l'OST produit es rapports d'indicateurs dans le cadre de l'évaluation des coordinations territoriales par le Hcéres. Ces rapports sont fournis en amont du processus pour que les évalués aient le temps de les exploiter dans leur rapport d'autoévaluation. Ils sont ensuite fournis au comité de visite. Un extrait par le département d'évaluation de la recherche du Hcéres pour réaliser les synthèses Recherche de site.

Les données issues du programme IPERU sont aussi utilisées dans d'autres productions régulières (rapports STRATER fournis aux MESRI) ou pour des productions spécifiques (pour des établissements ou l'évaluation de politiques publiques notamment).

Explorations et améliorations du programme IPERU

L'OST travaille de manière régulière sur les données issues du programme IPERU en mode exploratoire. Ces explorations permettent de répondre aux questions formulées par les établissements, d'approfondir la connaissance des données et de contribuer à des améliorations.

L'OST a ainsi comparé le nombre de publications des universités françaises présentes dans le classement de Leiden et le programme IPERU, alors même que les deux institutions travaillent sur la même base de données source, le Web of Science. Des écarts importants s'expliquent par le fait que le CWTS applique des filtres plus restrictifs sur les données de la base de publications et ne retient pas certains types de

documents et de revues qu'IPERU conserve (proceedings notamment). Les publications sont aussi filtrées par la langue, le classement de Leiden ne retenant que les publications en anglais.

Les écarts constatés tiennent aussi aux modalités de prise en compte des publications des CHU. Le CWTS considère qu'un acteur affilié à un établissement peut être une composante, un centre de recherche ou une organisation commune, ou encore une organisation associée. Les échanges entre l'OST et le CWTS, ont permis de comprendre que le CHU ou son équivalent à l'étranger n'est considéré dans le périmètre de l'université qu'en fonction de la qualification du lien. De plus, la coprésence des deux adresses dans les affiliations peut avoir un impact sur le nombre de publications en compte fractionnaire. Ce fait est d'autant plus important que le seuil d'entrée dans le classement de Leiden est justement un nombre de publications en compte fractionnaire.

Enfin, les écarts tiennent à la méthode de repérage. Le CWTS pouvait descendre à la maille du laboratoire mais uniquement lorsque la *main organization* de la base WoS ne concernait pas un établissement qui faisait l'objet d'un classement. Par ailleurs, le CWTS applique, comme IPERU, des critères géographiques : il ne recherche les adresses d'un établissement que sur une certaine liste de départements ou de villes.

L'OST travaille aussi à la comparaison de corpus de publications pour les disciplines SHS. Ces travaux ont été initiés en particulier pour répondre à une demande formulée par l'université de Bordeaux Montaigne. La présentation de ces travaux sera approfondie lors de l'atelier consacré aux SHS.

### **Frédérique SACHWALD**

Merci pour ce tour d'horizon. La suite de la journée doit permettre d'approfondir les différents modules du programme IPERU afin d'envisager des évolutions. Ces dernières dépendent à la fois des besoins qui seront exprimés par les établissements et les évolutions du contexte, que ce soit en matière de besoins ou d'opportunités, comme de nouvelles sources de données ou des ressources nouvelles le recensement des publications françaises dans HAL ou à terme par le projet Conditor.

Je vous propose de répondre à quelques questions dans le temps qui nous reste pour cette session.

### **Thomas COUDREAU, directeur du collège des écoles doctorales, Sorbonne Paris Cité**

Vous est-il possible de préciser le nombre de publications par doctorant de nos établissements ?

### **Frédérique SACHWALD**

C'est une évolution envisageable qui demanderait le développement d'un projet spécifique. Dans le cas présent, toutefois, l'information sera à relier au traitement des données personnelles. Il nous faudra disposer de la liste des doctorants afin d'identifier les publications auxquelles ils sont associés.

# RESPONSIBLE DESIGN, INTERPRETATION AND USE OF UNIVERSITY RANKINGS

---

## Conception, interprétation et usage responsables des classements d'universités

Animation

Frédérique SACHWALD,  
Directrice de l'OST

Martijn VISSER  
Centre for Science and Technology Studies (CWTS), Université de Leiden

 [Présentation en pdf](#)

Le classement de Leiden a été conçu comme une réponse aux classements existants : il est certes intéressant de critiquer les classements des autres mais il est évidemment plus constructif de publier son propre classement.

Parmi les différents classements ; le Times Higher Education (THE) ranking et le US News Best Global Universities sont les plus proches du classement de Leiden. Pour ce qui concerne le classement de Shanghai, la méthode utilisée n'est pas encore très transparente, ce qui rend les comparaisons précises difficiles. Le classement QS, pour sa part, se fonde essentiellement sur des enquêtes de réputation même si la méthode repose également sur des indicateurs bibliométriques. U Multirank enfin, est un classement auquel le CWTS participe et les remarques qui concernent le classement de Leiden s'appliquent aussi à U-Multirank.

Comment sont sélectionnées les universités qui entrent dans le classement de Leiden ?

Sont retenus les établissements qui ont produit au moins 1000 publications entre 2013 et 2016. Ce seuil est atteint par 938 universités réparties dans 55 pays. Le classement THE a un seuil similaire (1000 publications) mais s'appuie sur Scopus et non sur le Web of Science. Il classe 1258 universités de 86 pays. Les universités classées par le CWTS sont concentrées dans un plus petit nombre de pays que celles du classement THE, à l'exception du Japon et du Royaume-Uni qui sont plus largement représentés dans ce dernier. Pour figurer dans THE, les établissements doivent participer et donc fournir des données, ce qui n'est pas le cas dans le classement de Leiden réalisé à partir de la base de données de publications du CWTS.

Le classement US News, pour sa part, retient les 1 250 meilleures universités qui justifient de 1 500 publications au moins dans le Web of Science entre 2012 et 2016 ou qui figurent parmi les 250 meilleures universités du classement en matière de réputation utilisé comme référence (Clarivate Analytics).

Les classements Leiden et US News sont relativement proches car les sources de données sont similaires. Les deux classements sont donc cohérents à l'exception notable de la représentation de la Chine pour laquelle le classement de Leiden retient plus d'établissements que le classement US News. La position de la France est relativement stable dans l'ensemble de ces classements, autour d'une proportion de 3 % du nombre total d'institutions classées.

Pour figurer dans les classements Times Higher Education, QS et U-Multirank, les établissements doivent activement apporter leur participation. Les classements dépendent aussi de l'organisation des universités et des organismes de recherche (à travers la question du périmètre à retenir pour définir l'université) mais aussi de la concentration de la recherche dans ces institutions. Les variations constatées entre les classements dépendent également des bases de données exploitées. Pour le classement Leiden, le CWTS sélectionne les *core publications* du WoS et ne retient que les publications en langue anglaise.

Concernant les COMUE, les classements ont fait des choix différents. Le classement US News classe 13 COMUE depuis sa dernière édition. Le classement THE retient certaines COMUE, mais pas toutes. Le classement de Leiden ne s'appuie pas sur les communautés d'établissements ou regroupements mais sur

les universités qui en sont membres, en considérant que les universités fédérales ne sont pas comparables aux autres établissements.

La question des tutelles est très spécifique à la France. Il est souvent difficile pour les classements d'appréhender cette notion. Les tutelles sont souvent citées en même temps que les laboratoires. Par exemple, l'institut Pierre Simon Laplace est rattaché à dix tutelles, mais il est en fait une fédération de 9 laboratoires qui chacun dépendent de leurs propres tutelles. Dans le classement de Leiden, les tutelles ne sont repérées que lorsqu'elles sont explicitement mentionnées. L'organisation de certaines institutions peut aussi être difficile à appréhender, comme celle de l'Ecole des Ponts ParisTech.

Le positionnement des hôpitaux universitaires affiliés (academic hospitals) est une difficulté qui n'est pas spécifique à la France. Dans le classement de Leiden, les publications des CHU ne sont pas automatiquement prises en compte. La réponse varie selon que l'université est explicitement mentionnée dans les affiliations mais aussi du type de relation entre l'hôpital et l'université. En revanche, si l'université contrôle l'hôpital, nous considérons que la publication peut revenir dans le giron de l'université. En cas de difficulté, nous pouvons être amenés à étudier plus précisément un échantillon de publications de l'hôpital universitaire pour vérifier le statut de leurs auteurs pour déterminer s'ils sont rémunérés par l'université.

Selon les classements, différents indicateurs sont retenus : les publications, les citations mais aussi la notoriété, le nombre de Prix Nobel ou le nombre de post docs. Plusieurs classements (Shanghai, THE, QS et US News) utilisent un indicateur composite tandis que d'autres (Leiden, U-Multirank) proposent à l'utilisateur le choix de l'indicateur sur lequel il veut classer les établissements. D'un point de vue méthodologique, les indicateurs composites sont difficiles à interpréter. L'interprétation est encore plus complexe lorsque sont agrégés des indicateurs dépendants de la taille des universités et des indicateurs indépendants de la taille. Cela reviendrait à combiner des références au PIB d'un pays et au PIB par habitant de ce même pays : une telle comparaison n'aurait aucun sens. Il en est de même ici, quand certains classements (Shanghai et US News) mobilisent pour le calcul d'un indicateur composite le pourcentage de publications dans le décile des publications les plus citées et le nombre de publications dans cette classe de citations.

Le classement de Leiden porte sur les performances des universités en matière de recherche. Il distingue clairement les indicateurs dépendants de la taille des établissements d'intensité ou qualitatifs, indépendants de la taille de l'établissement. Il repose exclusivement sur des indicateurs bibliométriques sans données d'enquête fournies par les universités elles-mêmes. Sur le site de CWTS, vous pourrez trouver le classement des universités en fonction du nombre de publications, par défaut, mais l'utilisateur peut choisir un autre indicateur sur lequel classer les universités (par exemple, la proportion de publications dans le top 10% des publications les plus citées).

La méthodologie se veut robuste grâce à une approche qui réduit le poids des valeurs extrêmes, une classification fine par domaine de recherche, l'utilisation de comptes fractionnaires et l'unique prise en compte des *core publications* de la base de données. Les publications qui ne sont pas en langue anglaise étant généralement moins citées, elles sont exclues des calculs pour éviter de biaiser les indicateurs d'impact. La notion de compte fractionnaire renvoie au traitement des co-publications : pour le classement de Leiden, les co-publications sont partagées entre les auteurs tandis que les classements THE, QS mais aussi US News prennent en compte les co-publications de manière indivisible pour chacun des co-auteurs (compte entier). Pour le classement de Leiden, il fait sens de fractionner ces publications car ce fractionnement permet de ne pas favoriser les universités ayant une forte activité dans le domaine biomédical ou d'autres disciplines où le nombre d'auteurs par article est élevé.

Une vidéo insérée dans la présentation illustre la méthode employée par le classement de Leiden. En conclusion, il convient de rappeler que les classements, d'où qu'ils viennent, doivent être interprétés avec prudence. En définitive, les classements fournissent des informations utiles dès lors qu'ils sont conçus, interprétés mais aussi utilisés correctement. Les classements d'universités ne vont pas disparaître dans les années à venir mais nos analyses peuvent influencer la manière dont ils sont utilisés et à terme certains de leurs critères.

## QUESTIONS DE LA SALLE

### Daniel EGRET, PSL

Pour procéder au repérage des universités, utilisez-vous la notion proposée par le Web of Science « d'organization enhanced », ou exploitez-vous les informations de l'adresse d'affiliation institutionnelle des auteurs indiquée dans les publications ?

### Brigitte PLATEAU, DGESIP, MESRI

Imaginons que je sois la présidente d'une université et que je sois classée entre le 100<sup>ème</sup> et le 200<sup>ème</sup> rang. Quels conseils me donneriez-vous pour améliorer mon positionnement ? Certes, une option serait d'améliorer la qualité de la science, mais un tel objectif prend du temps. Aussi quels conseils pourriez-vous me délivrer pour obtenir quelques *quick wins* ?

### Martijn VISSER

Le CWTS utilise ses propres systèmes pour analyser les données et ses propres filtres en fonction de la méthodologie développée et évoquée dans la présentation. Le CWTS cherche chaque année à améliorer cette méthodologie. Le classement de Leiden se distingue ainsi du classement US News qui utilise la base WoS de Clarivate Analytics sans aller plus loin dans l'analyse. Cette approche explique que le classement de Leiden ne recoure pas à l'information « enhanced organization » disponible dans les outils du WoS.

Concernant la deuxième question portant sur les moyens dont peut disposer une université pour améliorer son classement à court terme, il faut tout d'abord rappeler que la première étape consiste à figurer dans un ou plusieurs classements internationaux. La taille de l'établissement peut jouer dans la mesure où elle permet mécaniquement d'augmenter le nombre de publications, même si ce n'est pas de façon proportionnelle à l'augmentation du nombre d'étudiants ou du budget. Le nombre de collaborations peut aussi contribuer à améliorer le positionnement dans la mesure où elles donnent lieu à des co-publications et que ces dernières tendent à être plus citées que les publications d'une seule institution. Ceci dit, les collaborations avec des institutions prestigieuses dépendent en amont des qualités de l'établissement. Par ailleurs, certains indicateurs, comme ceux du classement de Leiden comptabilise les publications et les citations avec un système de fractionnement pour les co-publications.

### Michel COSNARD

Ma question portera sur le périmètre des institutions dans le classement de Leiden. En France, les écoles de commerce, mais aussi les beaux-arts ou encore les écoles d'architecture ne sont pas des universités. La notion de COMUE vise justement à permettre des collaborations plus fortes. Il en est de même pour liens entre les hôpitaux et les universités. Une piste consisterait à interroger les établissements pour leur demander sur quel périmètre ils doivent être considérés.

### Martijn VISSER

Cette ouverture peut conduire à accepter une définition très large de l'université or nous avons besoin d'objectivité pour éviter de créer des inégalités de traitement entre les établissements dans le cadre d'un classement.

### Frédérique SACHWALD

L'OST a mené un travail sur les dénombrements et les différences de méthodes. IPERU procède à un repérage automatique avant d'entrer dans une phase de consultation avec les établissements pour l'affiner. CWTS ne procède qu'à un repérage automatique ce qui peut s'expliquer par le fait que le classement de Leiden est effectué à l'échelle mondiale tandis qu'IPERU est circonscrit à la France. Par ailleurs, cette approche plus fine est possible en France car nous avons des liens avec les universités. CWTS veut pour sa part adopter l'approche la plus fiable et équitable possible. D'où l'utilise des informations harmonisées et vérifiables. Cependant, cette méthode ne permet pas au classement de Leiden d'entrer dans une analyse en profondeur. Par ailleurs, comme l'a souligné la présentation du classement de Leiden, le volume de publications n'est pas nécessairement corrélé aux indicateurs de qualité. C'est la raison pour laquelle le classement de Leiden compile des indicateurs sensibles à la taille et des indicateurs non liés à la

taille. Le fait que le classement de Leiden ne retienne que les publications rédigées en anglais constitue un filtre sélectif de fait, même si cette option correspond à un objectif d'égalité en termes de potentiel de citation pour les publications retenues.

### Questions

- Lors des Open Science Days qui se sont tenus récemment, Vincent Larivière, chercheur à Montréal, a présenté une étude sur l'impact des classements et des évaluations sur la recherche. Il a notamment souligné que les classements tendent à en encourager, voire forcent les chercheurs à publier en anglais pour accroître leur visibilité à l'échelle internationale. Cette tendance peut toutefois conduire les chercheurs à négliger des sujets de recherche d'intérêt local qui auraient une moindre résonance à l'international. Un autre intervenant des Open Science Days a affirmé que 50 % des articles scientifiques ne seraient jamais lus et que 90 % ne seraient jamais cités. En outre, cet intervenant indiquait que 80 % des articles cités ne seraient en fait pas lus.
- Il est demandé aux chercheurs de publier toujours davantage. Le nombre de publications ne cesse d'enfler. Certaines universités semblent même mettre en place des encouragements financiers pour faire croître le nombre de publications, comme certaines universités chinoises qui récompenseraient un article dans *Nature* ou *Science* avec des primes pouvant aller jusqu'à 200 000 dollars.

### Martijn VISSER

Je ne peux pas réagir en séance sans avoir pris connaissance des études évoquées dans la question et des conditions dans lesquels les chiffres cités ont été obtenus. La question des incitations auxquelles font face les chercheurs est aussi très débattue et fait l'objet d'analyse. Ces incitations dépendent non seulement des indicateurs bibliométriques, mais aussi de leur utilisation et des méthodes d'évaluation dans les différentes institutions et pays. L'analyse précise des biais demande ainsi d'approfondir le contexte.

### Frédérique SACHWALD

J'irais dans le même sens pour dire que l'analyse des citations et de leurs déterminants font l'objet d'études dans la littérature en bibliométrie avec des estimations diverses. De même concernant les effets pervers des classements : ils peuvent exister, mais ce ne sont pas les seuls déterminants de l'augmentation du nombre de publications dans le monde, ou encore des modalités d'évaluation des chercheurs.

### Question

Comment évaluez-vous l'*open science* qui ne répond pas aux mêmes critères que la recherche classique ? Comment entendez-vous changer les règles pour que l'*open science* puisse jouir d'une place plus importante ? Si l'*open science* devient un indicateur supplémentaire, son impact peut-il s'avérer positif pour le positionnement des universités ?

### Martijn VISSER

Le CWTS n'a pas le projet de faire de l'*open science* un indicateur. Nous considérons que l'*open science* est une politique de l'établissement. Cependant, si l'*open access* aux publications conduit de fait à accroître la visibilité des universités qui le promeuvent, son développement devrait conduire à constater un meilleur positionnement de ces universités.

### De la salle

L'*open access* sera-t-il aussi un futur indicateur pour IPERU ?

### Frédérique SACHWALD

Une partie des publications en accès ouvert est déjà incluse dans nos indicateurs. Pour aller plus loin dans l'analyse, nous pourrions suivre les écarts entre les publications en accès ouvert et les autres. En revanche, le respect d'objectifs de politiques publiques qui sont qualitatives ne fait pas partie des indicateurs IPERU. Je suppose que le ministère mettra en place des indicateurs pour étudier dans quelle mesure ces politiques sont suivies.

# ARCHIVES INSTITUTIONNELLES, ARCHIVES OUVERTES : LEUR ROLE DANS LE PILOTAGE ET L'EVALUATION DES ETABLISSEMENTS

---

## Animation

Michel COSNARD, président du Hcéres

- Jean-François BALAUDE, président de l'Université Paris Nanterre
- Christine MUSSELIN, directrice de recherche CNRS, Centre de Sociologie des Organisations-Sciences Po et CNRS
- Christophe PERALES, directeur du Service commun de la documentation de l'Université Paris Diderot et président de l'Association des directeurs et personnels de direction des bibliothèques universitaires et de la documentation (ADBU)
- Michel ROBERT, directeur du département d'évaluation des établissements du Hcéres

### Michel COSNARD

Jean-François Balaudé, président de l'Université de Nanterre, ne peut pas nous rejoindre car il est bloqué – pour des raisons que je ne vais pas commenter – dans son établissement.

### Christophe PERALES

La thématique qui nous réunit soulève notamment la question de la définition de l'*open access*. De quel *open access* parlons-nous ? Certaines politiques d'*open access* visent principalement à ouvrir les contenus et conduisent notamment à ce que les publications soient davantage connues et citées, et que la production scientifique stimule le développement technologique et économique. L'*open access* peut prendre aussi s'inscrire dans une perspective plus radicale dès lors que l'on considère que l'*open access* peut également avoir un effet transformant sur l'écosystème des publications et conduire à modifier la manière dont l'évaluation de la recherche peut être conduite. Il me semble que cette table ronde peut s'inscrire dans cette seconde perspective puisqu'elle lie archives institutionnelles et archives ouvertes au pilotage et à l'évaluation des établissements.

Le pilotage et l'évaluation recouvrent aussi des réalités très diverses. Sans être exhaustifs, nous pouvons distinguer l'évaluation qualitative qui concerne le recrutement et la carrière des chercheurs mais qui permet également de dresser des cartographies disciplinaires et d'identifier les thèmes de recherche les plus dynamiques, ou les critères des programmes européens (H2020) qui incitent à publier en *open access* pour bénéficier des financements. L'évaluation peut être aussi plus quantitative, basée sur des métriques avec l'évaluation des publications et la production scientifique. Indirectement, comme cela a été souligné précédemment, l'*open access* a déjà un impact sur les métriques puisque plusieurs études aux résultats aujourd'hui bien établis indiquent que les publications libres d'accès reçoivent plus de citations que les autres.

Le sujet est donc pluriel et il faudrait pouvoir s'interroger à ces différents niveaux.

La question qui nous réunit aujourd'hui prend aussi place dans un contexte qui a énormément évolué. Depuis une vingtaine d'années, nous vivons un changement de paradigme avec l'essor du *web* en particulier. Nos méthodes d'évaluation ne sont plus quantitatives uniquement mais aussi qualitatives. Nous pouvons désormais exploiter d'immenses bases de données pour produire des cartographies. Nous constatons aussi un élargissement des objets qui sont pris en compte dans un certain nombre d'outils visant à cartographier la recherche (SciVal, Incites) : des publications mais aussi d'autres objets (bases ou jeux de données, et les algorithmes pour les manipuler, etc.), voire des signaux plus faibles (avec des traces laissées sur les réseaux sociaux de la recherche comme Academia ou ResearchGate, ou sur Twitter).

Le mouvement de fond actuel, dans le contexte de l'*open access*, c'est la volonté de rompre avec le biais fondamental des métriques classiques, qui date de l'instauration même de l'*impact factor* dans les années 1950 : un article vaut par l'*impact factor* de la revue qui le publie, alors qu'en réalité, une revue est d'autant meilleure qu'elle publie de bons articles, et non l'inverse. D'où la critique des métriques *review centric*, reposant sur l'*impact factor*, et la promotion, sous des formes diverses, de métriques alternatives (*altmetrics*), qui ont toutes en commun d'être *article centric*. Les initiatives sont foisonnantes et diverses (*epijournals*, F1000, etc.), pas encore *mainstream*, mais elles nourrissent indéniablement la réflexion en questionnant les métriques traditionnelles au niveau de leur biais le plus fondamental.

La question dépend aussi des communautés, puisque les pratiques sont très variables suivant les communautés, ce qui complexifie encore la question. Les monographies ont un poids plus important dans les humanités. Dans les sciences médicales, les métriques restent de format plus classique. Les communautés peuvent se structurer y compris à un niveau international pour baser leur travail d'évaluation sur les archives ouvertes qui sont, cette fois, thématiques. Elles peuvent alors passer en revue les *working papers* et *pre-prints* qui sont des éléments pris en compte dans l'évaluation de la recherche, comme le fait par exemple à un niveau international l'archive RePEc, en sciences économiques et de gestion.

### **Christine MUSSELIN**

Mon propos sera centré sur les archives institutionnelles. L'archive institutionnelle de Sciences Po, dénommée Spire, est connectée à HAL (archives ouvertes) et à RePEc car certains économistes ne veulent déposer leurs archives que dans RePEc. Les archives institutionnelles présentent un intérêt lorsqu'elles sont connectées les unes aux autres. Dans le cas contraire, il existe évidemment le risque d'un repli sur soi.

A quoi servent ces archives institutionnelles et comment pouvons-nous les utiliser ? Leur premier intérêt est de rendre visible. Déposer ses publications ou annoncer une publication sur une archive institutionnelle, c'est donner de la visibilité individuelle au chercheur mais aussi de la visibilité collective à l'institution. Lors d'une recherche sur internet, ce sont d'ailleurs ces archives qui apparaissent en premier en permettant l'accès à une notice bibliographique ou directement au texte.

Ces archives institutionnelles permettent aussi l'accès à des documents puisqu'elles permettent d'intégrer les avant-dernières versions des articles dès lors que la dernière version n'est pas encore publiable en raison de barrières mobiles. Cependant, si ce procédé fonctionne pour des articles, il est moins évident pour d'autres types de publications, notamment pour les livres qui restent un support important en SHS. Sur Google Scholar ou d'autres moteurs de recherche, je ne recherche pas uniquement des articles mais aussi des ouvrages, mais force est de reconnaître qu'il est plus difficile de déposer des chapitres de livres que des articles.

L'avantage des archives institutionnelles en termes de science ouverte est qu'elles permettent aussi de faire connaître d'autres types de publications que les articles et les livres. Par exemple, sur l'archive de Sciences Po – Spire – nous pouvons annoncer la création ou la publication de logiciels en accès ouvert ou encore annoncer la constitution de bases de données. D'ailleurs, le Hcéres s'est montré pionnier en la matière puisque la dernière évaluation nous a encouragés à lister les logiciels libres que nous avons créés ainsi que les bases de données. Pour les SHS, en particulier, cette ouverture du champ est enrichissante en rendant plus visibles des productions autres que les publications.

Il apparaît plus difficile d'utiliser les archives institutionnelles pour le pilotage et l'évaluation. Tout d'abord, les archives institutionnelles dépendent de la bonne volonté des chercheurs de déposer leurs productions. Il est certes possible d'encourager ces saisies. Certaines institutions vont jusqu'à pénaliser ceux qui ne le font pas. Personnellement, je ne suis pas favorable à cette méthode, mais c'est une piste exploitée par d'autres.

Nous nous heurtons aussi à un problème de fiabilité des données qui sont déposées. A Sciences Po, nous avons une référente Spire qui se rend dans les laboratoires pour encourager les dépôts et qui fait preuve de pédagogie sur la manière de procéder, preuve s'il en est que le mouvement n'est pas naturel ni

spontané. Il est nécessaire de mettre en place une infrastructure humaine pour que les données soient fiabilisées et qu'un contrôle soit effectué sur les dépôts.

Par ailleurs, sur Spire comme sur HAL, les modules statistiques sont très peu développés. De gros progrès ont été fait pour SPIRE, mais les statistiques disponibles sont encore souvent peu exploitables et nécessitent un important travail manuel pour extraire des informations réellement utiles en termes de pilotage. Par exemple, je peux connaître le nombre d'articles publiés par Sciences Po dans des revues à comité de lecture et même avoir l'information par laboratoire. En revanche, les outils statistiques ne me permettent pas de disposer d'informations sur une population particulière, alors que c'est ce dont nous avons besoin si nous souhaitons entrer dans une démarche de pilotage. Nous devons cependant reconnaître que ces archives institutionnelles n'ont pas été conçues pour ce pilotage fin. Il nous revient donc de mener un important travail de développement informatique afin de pouvoir répondre à de nouvelles utilisations. Ces archives permettent cependant de dresser une cartographie et de visualiser sur quoi travaillent les chercheurs d'une institution à un moment donné. Il est plus difficile d'aller plus loin et d'interpréter ces données. Ces informations permettent toutefois d'identifier de nouvelles thématiques en train d'émerger, que nous pouvons alors renforcer via des recrutements. Elles pourraient aussi permettre d'identifier les thématiques en train de disparaître alors qu'il pourrait exister un intérêt à continuer à les soutenir via les recrutements ou les thèses.

### **Michel ROBERT**

Il existe quatre niveaux d'évaluation au Hcéres : l'évaluation des formations, l'évaluation de la recherche, l'évaluation des établissements et l'évaluation des coordinations territoriales. Pour ces évaluations, nous utilisons des référentiels. Il existe par exemple un seul référentiel pour l'ensemble des établissements (écoles et universités) en France. Pour la recherche, l'évaluation est menée par les pairs, comme pour les établissements, mais ces pairs sont disciplinaires (SHS, santé, sciences, etc. ). Dans ce cadre, on y parle de produits de la recherche, de faits observables et c'est au laboratoire de s'approprier tout degré de liberté. Pour citer mon exemple, lorsque j'étais directeur de laboratoire, j'insistais davantage sur l'usage des brevets, et donc sur la création de valeur, que sur le nombre des brevets. Il existe certes certains indicateurs standardisés – comme nous l'avons entendu ce matin – mais chaque communauté peut ajuster ce contexte à sa réalité.

Il me semble qu'il existe un fil conducteur essentiel qui va du chercheur, de l'équipe et du laboratoire jusqu'à l'établissement. Ce fil conducteur est la clé de l'évaluation en Italie, ce qui n'est pas le cas en France où l'on distingue l'évaluation du chercheur en tant que personne et l'évaluation de l'établissement. Pour l'évaluation de l'établissement, ce sont des comités de pairs qui se prononcent, composés de français et d'étrangers qui n'entreront pas dans le détail des composantes de l'établissement ni dans des analyses disciplinaires. En revanche, le rapport d'auto-évaluation permet de coupler la recherche et la formation. Ce rapport doit permettre de décrire comment l'établissement formalise et suit ses partenariats, d'évoquer l'organisation et le pilotage de l'activité de recherche et la mise en œuvre opérationnelle de la politique de recherche. Il explique comment l'établissement est en capacité de suivre, d'analyser et de qualifier son activité de recherche et d'en diffuser les résultats.

Pour ce qui nous concerne, nous ne préconisons pas l'utilisation d'indicateurs : les Indicateurs sont ceux du contrat entre l'établissement évalué et le ministère et les Indicateurs mis en place par l'établissement lui-même pour qualifier ses résultats. Ainsi, un établissement qui affichera une priorité sur un thème donné devra préciser si l'indicateur choisi concernera la recherche, la formation ou les deux volets. Il devra aussi préciser la mesure de ce positionnement et comment l'établissement se compare à d'autres établissements dans son domaine au niveau national comme international.

A l'occasion des évaluations qui ont lieu tous les cinq ans, le Hcéres observe de grands progrès dans l'appropriation d'outils comme IPERU. Ces outils permettent aux établissements de se situer sur une échelle, mais ils peuvent aussi souligner des limites dans la comparaison.

## **Michel COSNARD**

Quelles archives doivent être privilégiées ? Quelles sont les bonnes pratiques en la matière ? Par ailleurs, quel est le rôle de l'archive nationale HAL ? Plus généralement, comment est-il possible de passer d'archives institutionnelles à des archives nationales et internationales qui sont très souvent disciplinaires ?

## **Christine MUSSELIN**

Spire est une archive pour les publications. Pour déposer les données de la recherche, il faut utiliser une autre archive spécifique. Celle-ci s'inscrit dans le cadre de la science ouverte en rendant possible l'accès à des données qui ont été collectées dans le cadre de projets de recherche.

Spire permet d'identifier plus facilement les publications qui sont propres à l'établissement. Cependant, il reste nécessaire d'assurer une connexion à HAL qui a une vocation nationale. S'il est pertinent que chaque établissement dispose de ses propres archives, il existe aussi un intérêt à ce que celles-ci soient connectées aux archives nationales.

Nous avons fait le choix de nous connecter à l'archive internationale RePEc à la demande des économistes de Sciences Po, mais il est évident que cette connexion est plus difficile si de nombreuses disciplines sont représentées au sein d'un même établissement, ce qui n'est pas le cas pour Sciences Po.

## **Michel COSNARD**

Je renvoie donc la question à Christophe puisque Paris VII accueille de nombreuses disciplines.

## **Christophe PERALES**

Je souhaite tout d'abord rejoindre le propos précédent de Christine pour appuyer le fait que les archives ouvertes n'ont pas été conçues à l'origine pour le pilotage et l'évaluation. L'articulation des différents niveaux d'archives (institutionnelles, thématiques internationales, nationales) nécessite d'insister fortement sur la fiabilisation des données et sur la mise en place de référentiels car une archive, par principe, est moissonnée. Or moissonner sur une base qui n'est pas propre posera de vraies questions et ne permettra pas de dédoubler les données pour calculer des indicateurs correctement.

Nous assistons aujourd'hui à une multiplication des métriques et des indicateurs possibles, ou à l'émergence de méthodes plus qualitatives et plus ouvertes, le tout dans le sillage de l'*open access*. Pour les communautés comme pour les établissements, c'est aussi un moyen de s'approprier les évaluations. Le but ultime n'est toutefois pas que chaque établissement crée ses propres métriques car le besoin d'aligner et de normaliser les références afin de pouvoir se comparer aux autres établissements demeure. Cependant, cette relative liberté permet dans une certaine mesure d'adapter les indicateurs en fonction des disciplines et du profil des établissements.

Ainsi, même si les archives ouvertes n'ont pas été conçues à des fins d'évaluation, elles offrent un grand potentiel. Certes, comme l'a souligné Christine Musselin, il convient d'en assurer l'exhaustivité et la fiabilité, mais ces objectifs ne sont pas hors de portée et posent surtout des questions techniques, en termes de fiabilisation du signalement, ou, si l'on préfère utiliser un vieux vocable, de catalogage. C'est donc une possibilité supplémentaire, potentiellement très intéressante, d'évaluer la recherche.

## **Christine MUSSELIN**

Au-delà des archives institutionnelles, nationales ou thématiques, il existe aussi des archives d'éditeurs, comme Academia ou Research Gate par exemple, qui font du moissonnage, interrogent directement le chercheur et qui entrent ainsi dans une concurrence plus frontale avec les archives institutionnelles. Les chercheurs d'une institution peuvent faire le choix de ne pas déposer dans les archives institutionnelles de leur établissement ou dans l'archive nationale dès lors qu'ils sont répertoriés dans des archives d'éditeurs qui souvent leur mâchent beaucoup plus le travail que les archives de nos établissements. Pour autant, ces archives d'éditeurs posent les mêmes problèmes de fiabilité, sans compter les problèmes de privatisation puisque ces outils proposent des services payants.

## **Christophe PERALES**

Il est en effet très intéressant de regarder ce que font ces acteurs privés comme Academia ou ResearchGate, car ils sont très réactifs et très au fait des besoins des chercheurs. Ils se situent cependant dans un modèle de *net economy* qui est très éloigné des archives ouvertes, et pas fondamentalement différent de celui des grands acteurs « éditoriaux » comme Elsevier, avec lesquels ils sont du reste en concurrence.

## **Michel ROBERT**

L'évaluation de la recherche dans les laboratoires est contrainte par l'observation des productions. Au niveau des établissements, c'est en revanche peu contraint et il faut se saisir de cette opportunité pour se poser les bonnes questions. Par exemple, lors de sa dernière évaluation, le CNRS, en utilisant à la fois le WoS et Scopus, a été en capacité de consolider toute sa production pour la comparer à celle d'autres pays. Le comité de pairs a pu alors se focaliser sur l'analyse posée par l'établissement lui-même pour se situer dans son contexte particulier.

## **Michel COSNARD**

Que contiennent réellement les archives ouvertes ? Qu'est-ce qu'une publication : est-ce uniquement un texte, ou l'ensemble d'une présentation textuelle, enrichie par des tableaux reposant sur des données qui pourraient être publiées (voire la publication des traitements) ? Et comment s'assurer de l'intégrité scientifique des publications ? Rappelons que dans le cadre de l'archive institutionnelle, c'est le président de l'université ou le directeur de l'école qui publie sous sa responsabilité. N'y a-t-il pas un intérêt à co-publier systématiquement les données qui ont permis de produire la publication et les algorithmes qui ont permis les traitements, voire d'autres supports comme des vidéos si des enregistrements ont été nécessaires, pour valider la qualité et la pertinence des données ?

## **Christophe PERALES**

Je crois que la question de l'accès ouvert se pose différemment pour la publication et pour les données. Il est certes important de déposer les données pour les archiver, et ainsi éviter qu'elles ne demeurent sur un disque dur local qui pourrait être détruit ou de ne plus pouvoir les lire car le logiciel en capacité de le faire a changé de version. Mais la nécessité d'assurer la conservation pérenne des données n'implique pas qu'elles aient vocation à entrer dans le même cycle de visibilité et d'ouverture que les publications. Les chercheurs y sont d'ailleurs réticents, de manière très compréhensible. Rappelons que la collecte et la production de ces données primaires représentent la mobilisation de moyens souvent plus conséquents que la production de la ou des publications correspondantes. C'est souvent ce qui prend le plus de temps dans le processus de recherche. Par ailleurs, ces données primaires permettent de publier plusieurs articles : souvent ces données fournissent des matériaux pour la recherche pour plusieurs années. Ceci explique que les chercheurs soient réticents à faciliter l'accès à leurs données. Elles peuvent certes être archivées à des fins de vérification et de preuve scientifique si jamais un doute apparaît. Par contre, je ne pense pas que nous pourrions les ouvrir complètement. Outre les données, les algorithmes en permettant le traitement devraient aussi être archivés, mais c'est au chercheur de décider quand il faut donner accès aux données primaires de ses travaux.

## **Christine MUSSELIN**

Je partage ce point de vue d'autant plus que, comme cela vient d'être dit, les données produites dans le cadre d'une recherche sont beaucoup plus larges que celles utilisées pour une seule publication. Déjà certaines revues mettent en ligne les données qui ont servi de base aux articles, avec le problème de la privatisation de ces données par l'éditeur. Il est donc important de proposer des archives institutionnelles de publication qui incluent le dépôt de logiciels, d'algorithmes ou de logiciels et de posséder des archives de bases de données qui seront accompagnées de métadonnées permettant ainsi de les réutiliser. Pour ma part, je ne peux pas ouvrir les entretiens que je mène, car ils contiennent des informations confidentielles. Mais ces données ne peuvent de toute façon pas être exploitées de manière brute : pour les interpréter, il faudrait disposer d'informations de contexte (pourquoi la recherche a été faite,...).

### **Michel ROBERT**

Dans certains domaines des sciences et technologies, il existe des benchmarks avec des données qui font partie de la publication. Souvent on a tendance à vouloir séparer les revues des conférences mais je considère que, dans mon domaine, les conférences sélectives avec actes sont incontournables car ce sont des lieux de confrontation directe avec la communauté.

### **Aurore CARTIER, Université Paris Descartes**

L'université Paris Descartes possède un portail institutionnel sur l'archive ouverte HAL. Je pense pour ma part que les archives ouvertes ont un rôle à jouer dans le pilotage et l'évaluation de la recherche en tant que prérequis de mise à disposition des données. C'est aussi une mémoire que l'on dépose (des métadonnées, des articles) et c'est parfois la seule mémoire qui va rester. Dans les sciences humaines et sociales, plusieurs chercheurs sont venus nous voir cette année car leurs éditeurs avaient fait faillite et que leurs contenus n'étaient plus accessibles. Par ailleurs, les plateformes comme InCites ou SciVal sont des outils commerciaux. Dans cet environnement, comment garantir que nous aurons des données à disposition pour des politiques de pilotage et d'évaluation à l'échelle de l'établissement ? A mon sens, les archives ouvertes ont un rôle à jouer car elles sont la mémoire d'un établissement, en particulier pour les sciences humaines et sociales qui ne sont mal répertoriées dans les outils commerciaux. En outre, nous pourrions nous désabonner demain de ces outils de manière volontaire ou contrainte pour des motifs budgétaires.

### **Solenn BIHAN, Université de Lille**

Comment peut-on articuler les messages contradictoires qui sont envoyés aux chercheurs, aux laboratoires et aux établissements ? A chacun l'on demande d'être visible, mais cette visibilité pour un chercheur passe d'abord par une activité sur les réseaux sociaux académiques. Ce sont aussi les outils les plus ergonomiques sur le plan individuel. De plus, la visibilité d'un laboratoire n'est pas celle de son établissement, puisque le laboratoire cherche à mettre en avant son activité quelle que soit la tutelle alors que chaque tutelle veut apparaître, ce qui entraîne des tensions pour des questions de signature unique entre autres. Le plan national pour la science ouverte aborde-t-il ces questions ?

### **Christophe PERALES**

A ma connaissance, aucun projet ne porte sur une simplification possible du dépôt dans les archives, permettant notamment qu'un seul dépôt alimente d'autres circuits. Le plan national pour la science ouverte n'aborde pas ces thématiques, mais vous posez de vraies questions.

### **Christine MUSSELIN**

Pour encourager à déposer, j'ai souvent utilisé la technique du *naming and shaming* en indiquant, lors des réunions mensuelles avec les directeurs des unités, qu'il manquait des informations dans tel laboratoire. Une autre tactique est de ne citer que les publications déposées dans Spire dans les rapports annuels de Sciences Po.

Je souhaite également insister sur le fait que ces outils participent de l'évaluation mais qu'il ne faut surtout pas réduire l'évaluation au nombre de publications. Il convient absolument de remettre du qualitatif dans l'évaluation et revenir aussi sur la valeur intrinsèque des publications. Le nombre de publications doit aider à repérer des problèmes ou des évolutions : par exemple, une baisse des publications, une augmentation des publications en anglais, etc. Toutefois, ces seules informations ne suffisent pas à jauger la qualité de la science.

### **Christophe PERALES**

Au-delà des incitations ou encouragements que vous avez cités, il est aussi possible de développer des archives institutionnelles avec des services à valeur ajoutée pour le chercheur, comme ORBi, l'archive de l'université de Liège (<https://orbi.uliege.be/?locale=fr>). Il existe certaines réussites en la matière et ce sont des voies possibles mais aujourd'hui peu nombreuses. Ces expériences nécessitent de s'appuyer sur une volonté politique affirmée, à travers des incitations au dépôt très fortes.

### **Michel ROBERT**

La question posée renvoie à la nécessité de mettre en cohérence les différents acteurs à la maille d'un établissement. Il existe cependant ici une spécificité française à prendre en compte : en effet, on sépare en France l'évaluation des chercheurs et l'évaluation des structures. Lorsque j'étais responsable d'une unité, j'avais expliqué que ce qui ne serait pas référencé dans HAL ne serait pas cité dans le rapport d'activité du laboratoire. Cette méthode certes radicale permet de sensibiliser les acteurs sur l'importance de la visibilité de nos actions. Par ailleurs, il faut savoir que les chercheurs CNRS doivent renseigner annuellement un petit rapport d'activité, obligation qui n'est pas opposable aux enseignants chercheurs. Le rôle du responsable d'établissement, c'est aussi d'être en capacité de fédérer les acteurs et d'être incitatif.

### **Michel COSNARD**

Je vous remercie toutes et tous pour votre participation. Je vous invite à faire une pause avant de participer aux ateliers de l'après-midi.



# ATELIERS

Base des publications et indicateurs bibliométriques .....	22
Base des brevets et indicateurs brevets.....	25
Base des projets du PCRI et indicateurs sur la participation aux projets H2020.....	28
Indicateurs sur la production scientifique : le cas des SHS.....	31
Méthodologie relative au calcul des indicateurs.....	36

## BASE DES PUBLICATIONS ET INDICATEURS BIBLIOMETRIQUES

---

Au cours du tour de table, les participants ont été invités à faire part de leurs questions et suggestions relatives au sujet de l'atelier. Les points soulevés ont servi de base à la discussion.

Les échanges peuvent être synthétisés en trois thèmes : le positionnement des rapports IPERU publications dans le contexte de l'ESRI français, les usages du rapport IPERU publication par les établissements ; les améliorations souhaitables.

### 1. Contexte et interprétation des rapports IPERU publications

Des interrogations portent sur l'interprétation des indicateurs des rapports IPERU : la présentation de cas pratiques permettrait d'améliorer la politique de communication que les établissements mettent en œuvre auprès de la communauté scientifique. C'est important dans un contexte où certains chercheurs perçoivent une « course aux indicateurs ». Loin d'être un objectif en eux-mêmes, les indicateurs devraient constituer un levier pour mobiliser davantage la communauté scientifique. Parmi les réserves que celle-ci exprime figure la question de la couverture des disciplines des SHS par les bases de données : faudrait-il envisager de nouveaux indicateurs qui leur soient adaptés et qui soient basés sur des sources qui seraient reconnues comme plus représentatives des publications dans ces disciplines ?

Dans cette perspective, dans quelle mesure le programme IPERU pourrait-il travailler aussi à partir d'autres sources de données, et notamment de bases en Open Access ? Certaines bases en accès ouvert ont une couverture potentiellement beaucoup plus large dans la mesure où elles prennent en compte divers types de productions, tels que les ouvrages, les documents de travail... En revanche, elles n'assurent pas la couverture systématique d'un périmètre donné. Elles ne permettent pas non plus de travailler de façon systématique sur la dimension citationnelle, ce qui rend la construction et la contextualisation d'indicateurs bibliométriques délicates. L'OST explore néanmoins des bases de données non commerciales, notamment l'archive HAL ou d'autres sources dans la mesure où le domaine est en évolution.

Pour différents établissements, il serait par ailleurs souhaitable d'améliorer la fiabilité des affiliations indiquées par les auteurs lorsqu'ils signent des articles de revue ou d'actes de colloque. Cependant, il apparaît difficile pour les établissements de contrôler la bonne application des chartes de signature lorsqu'elles existent.

Certains participants s'interrogent sur l'impact du projet Conditor évoqué lors de la plénière de la journée IPERU. L'interrogation porte notamment sur l'apport de Conditor en matière de recensement des publications des établissements. Un lien est établi avec le développement de l'utilisation de l'identifiant ORCID et l'amélioration du recensement des publications. L'OST a rappelé sa participation au projet Conditor d'entrepôt des métadonnées de l'ensemble de la production scientifique française qui à terme peut simplifier le processus de repérage IPERU. L'utilisation d'identifiants pour les auteurs, ORCID ou autre, est un axe structurant de Conditor.

### 2. Usages des rapports IPERU publications

#### *La notice méthodologique*

La notice méthodologique qui accompagne le rapport IPERU relatif aux publications est consultée par les établissements. Ceux-ci estiment qu'elle est claire et permet de s'approprier les indicateurs, malgré sa densité et sa technicité.

#### *Les indicateurs « préférés » du rapport IPERU*

Si tous les indicateurs calculés par les rapports IPERU présentent un intérêt, il ressort de la discussion que le nombre de publications de l'établissement, ainsi que les indicateurs d'impact et de co-publications font l'objet d'une attention particulière. Les comparaisons permises par le volume 2 des rapports, relatif au positionnement national et international de l'établissement selon les disciplines, sont particulièrement appréciées. Il n'existe en effet que peu d'outils et de sources de données fiables pour conduire ce genre d'exercice.

### *Les usages actuels du rapport IPERU*

Les établissements ont des usages variés du rapport IPERU. Des informations issues du rapport sont souvent intégrées au dossier d'auto-évaluation dans le cadre des évaluations du Hcéres. Les indicateurs présentés dans le rapport aident à la décision et au pilotage des établissements, à la fois en permettant le suivi des performances académiques et le positionnement de l'établissement par rapport aux autres. Le rapport est enfin utilisé pour la communication interne et externe, que ce soit pour une présentation au conseil de la recherche, ou pour les chiffres-clés auprès de différentes instances.

Cependant, la question est soulevée de l'usage éventuellement risqué, voire « contre-productif », de ces indicateurs auprès de la communauté scientifique du fait de la couverture lacunaire des bases de données pour certains champs disciplinaires.

### *Développement des usages des repérages IPERU*

Certains établissements se sont approprié le rapport en développant des usages spécifiques et originaux. Par exemple, les adresses extraites de l'interface IPERU ont servi à contrôler et corriger, le cas échéant, les affiliations dans les bases de données bibliométriques Web of Science ou Scopus. Par ce moyen les établissements s'assurent de la bonne application de la charte de signature et peuvent nourrir ainsi un dialogue avec les chercheurs.

Par ailleurs, dans la mesure où le repérage IPERU des publications d'un établissement est considéré comme exhaustif, certains établissements ont calculé le taux de couverture des archives institutionnelles pour le périmètre des publications concernées. Ils ont ainsi pu instaurer de nouveaux processus en vue d'assurer la bonne mise en œuvre de la politique de dépôt des publications dans leur archive institutionnelle.

### *Indicateurs complémentaires*

Le souhait est émis de pouvoir rapporter les données de production scientifique des rapports IPERU (output) au nombre de chercheurs (input). L'OST n'a pas à ce stade de réflexion sur la définition qu'il conviendrait d'utiliser pour faire ce type de calcul. A l'échelle internationale, la notion d'un ETP pour une université varie, et en France il n'y a pas de source simple.

En interne, les établissements disposent de données d'effectifs qu'ils peuvent exploiter, mais la comparabilité inter-établissement risque d'être limitée.

## 3. Améliorations

### *Source de données*

Les améliorations attendues portent d'abord sur la source des données utilisées. La couverture de la base WoS étant limitée dans le cas des SHS et des sciences de l'ingénieur, il serait utile d'explorer l'apport que pourraient représenter des sources complémentaires, telles que Scopus et HAL. Des comparaisons permettraient d'apprécier les questions de représentativité et de valeur ajoutée des différentes sources.

### *Le repérage et la signature*

En ce qui concerne le repérage, une granularité plus fine permettrait l'élaboration de stratégies pour les « équipes d'accueil ». À cette fin, les établissements souhaiteraient effectuer le repérage au niveau du laboratoire. Les *patterns* venant du champ « Organization Enhanced » dans le WoS pourraient être utilisés, et les repérages seraient alors fournis directement par les établissements ayant nettoyé le champ comme informations pour le calcul des indicateurs par le programme IPERU.

Les *patterns* de repérage devraient être élargis afin de disposer de repérages au niveau des laboratoires, faute de quoi le reliquat à traiter manuellement resterait très important. Le cas échéant, il faudrait avoir la possibilité de contrôler les faux positifs.

### *Interface, réutilisation des données*

Les établissements souhaiteraient pouvoir réutiliser les données fournies par le programme IPERU. Un format exploitable (Excel) pour les indicateurs agrégés serait ainsi apprécié. Dans le mail de fin de repérage, par ailleurs, un accès automatique à un export de l'ensemble des publications pourrait être présenté à l'acteur et refléter le dernier état de son repérage.

Par ailleurs, les établissements apprécieraient que le délai d'ouverture de l'interface soit plus long et qu'un accès aux résultats soit disponible après la fermeture de l'interface.

### *Contenu du rapport*

Pour le rapport IPERU, certains établissements pensent qu'un niveau de détail plus grand pourrait leur être utile ; par exemple, le détail par pays des partenariats au niveau européen. Une granularité des indicateurs plus fine, allant jusqu'au niveau des laboratoires, ainsi qu'une nomenclature disciplinaire plus fine, sont

également suggérées. Ces questions de granularité doivent prendre en compte le nombre de publications et la capacité à calculer des indicateurs statistiquement significatifs.

Des fenêtres de citation plus longues (3-5 ans) seraient utilement explorées pour le calcul des indicateurs d'impact dans les disciplines SHS et les mathématiques. Des fenêtres plus longues sont pertinentes dans certaines disciplines ; suite à des tests, l'OST a par exemple retenu une fenêtre de 5 ans pour analyser la discipline mathématique dans le rapport sur la France publié en 2018. Il y a un arbitrage à effectuer avec la fraîcheur des données. Une fenêtre de citations à 5 ans entraîne en effet un décalage temporel important entre la diffusion des rapports et l'année la plus récente pour laquelle des indicateurs seraient proposés (cela correspond à la durée d'un cycle de contractualisation entre les opérateurs de l'enseignement supérieur et leur ministère de tutelle).

#### *Diffusion du rapport*

Une diffusion plus large que la pratique actuelle du programme IPERU apparaît souhaitable.

## 4. CONCLUSION

Les établissements rencontrent des difficultés pour communiquer autour des rapports et faciliter leur appropriation en interne. En particulier, les questions de couverture des bases de données dans certaines disciplines restreignent l'utilisation des rapports.

Les établissements ont des usages variés des rapports IPERU ; au pire, ils ne sont pas utilisés par manque de temps pour se les approprier.

Le volume 2 – positionnement national et international de l'établissement suivant la discipline – est clairement ce qui intéresse le plus les établissements, leur permettant de « se comparer ».

Les améliorations souhaitées concernent différentes étapes du processus IPERU : la mobilisation d'autres sources de données, le repérage des publications, le calcul de certains indicateurs, l'accès aux publications repérées, ou le format de restitution des indicateurs. Les conclusions des différents ateliers seront diffusées et l'ensemble des points évoqués examinés par l'OST pour préciser la faisabilité et établir des priorités. Cette perspective transversale permettra aussi de prendre en compte les points de convergence, par exemple entre cet atelier et l'atelier consacré aux SHS.

## 5. Annexe

### *Animateurs*

Lesya Baudoin, Mathieu Goudard, Esther Lardreau

### *Participants*

- CARDOVILLE Régis, Muséum national d'Histoire naturelle
- CHROSTOWSKA Anna, Université de Pau et des Pays de l'Adour
- COUTHURES-IDRIZI Anna, Institut d'Études Politiques de Paris
- DIAZ Pierre, Institut National Polytechnique de Toulouse
- DUFAU Carole, Université Paris Sud
- GOURDON Guillaume, Université Paris-Est Marne-la-Vallée
- HOMONT Isabelle, Université de Versailles - Saint-Quentin en Yvelines
- JAUNY Isabelle, Université de Versailles - Saint-Quentin en Yvelines
- JORRY Hélène, Université de Cergy-Pontoise
- KELLER Hélène, Muséum national d'Histoire naturelle
- LE BIGOT Ludovic, Université de Poitiers
- MERIGAUD Françoise, Université de Limoges
- MICHOU-SAUCET Philippe, Collège de France
- MIRABEL Pierre, Université de Strasbourg
- PELLÉ Anne, Université Paris 13 Nord
- PRETRE Véronique, Centrale Supélec

## BASE DES BREVETS ET INDICATEURS BREVETS

---

L'atelier se donne pour objectif de recueillir les besoins des établissements en matière d'indicateurs sur les brevets et d'échanger sur des évolutions du programme IPERU pour répondre à ces besoins.

L'OST présente le programme IPERU dans son ensemble pour que tous les participants soient informés des données et indicateurs auxquels chaque établissement peut avoir accès. L'OST présente ensuite les évolutions de sa base de données brevets au cours des trois dernières années durant lesquelles le module brevets n'a pas été fourni.

Les indicateurs peuvent être produits désormais à partir des données de quatre offices de propriété intellectuelle (INPI, OEB, USPTO et OMPI) et pour les Etats-Unis, les demandes américaines sont désormais incorporées (et pas uniquement les délivrances de titres). L'OST travaille actuellement sur des indicateurs portant sur les familles de brevets, les demandes prioritaires, sur une extension du projet antérieur sur les inventeurs académiques et sur une nouvelle catégorisation institutionnelle des déposants.

Les questions abordées lors de l'atelier à partir des problématiques soulevées par les établissements participants, peuvent être synthétisées en trois thèmes.

### 1. Utilisation des données et indicateurs IPERU

Le programme IPERU n'est pas connu de tous les participants et une meilleure diffusion de l'information apparaît utile aux participants.

Pour l'utilisation par la direction en charge du pilotage dans les établissements, il faudrait pouvoir apparier les bases et donc disposer au moins d'un identifiant commun. Un établissement a souhaité avoir accès à la base complète avec les numéros des brevets en format Excel car la base de l'établissement est sous ce format. Cette demande a été soutenue par les autres établissements.

Il ressort des discussions que l'utilisation du rapport IPERU pourrait être utile pour disposer de plus d'informations. Mais dans cette perspective, les établissements auraient besoin d'être accompagnés pour l'interprétation des résultats et le commentaire des indicateurs. Cependant pour le moment, en dehors d'un format Excel pour les données de repérage aucun représentant n'a émis de souhait précis concernant les modalités d'accompagnement.

### 2. Processus de repérage des brevets

Un consensus s'établit sur le fait que le repérage des seuls titulaires et/ou déposants ne suffit pas à caractériser la production technologique d'un établissement. Les participants souhaitent une prise en compte des données « inventeurs », car, dans de nombreux cas selon eux, l'établissement ne figure pas en tant que déposant.

Il a ainsi été souligné l'intérêt de bien identifier les affiliations des inventeurs dans un contexte (le RGPD) où leurs adresses personnelles pourraient ne plus être systématiquement indiquées. L'OST a proposé la mise en place d'un groupe de travail sur les inventeurs académiques dont les modalités seraient définies courant 2019.

Certains participants évoquent les difficultés de rapprochement de leurs données internes avec celles fournies dans l'interface OST. Le (pré)repérage par l'OST est effectué au niveau de l'établissement et utilise le numéro de publication des brevets comme identifiant. Or les numéros de publication ne sont pas toujours normalisés dans les données des établissements. De plus, il serait utile de connaître les années de présence des chercheurs au sein des laboratoires.

Les participants souhaiteraient pouvoir disposer d'une fenêtre de validation du pré-repérage plus longue et d'un export au format Excel afin de pouvoir modifier les éventuels changements dans la vie du brevet. Cette demande fait écho à celle de pouvoir disposer de la base complète formulée précédemment. L'OST souligne que l'objectif du repérage est le dénombrement des demandes produites par l'établissement à une date donnée afin de pouvoir produire les indicateurs dans le cadre d'un calendrier précis.

### 3. Bases/Indicateurs

L'intérêt des participants est centré sur la production de leur établissement, donc le dénombrement de leurs brevets, sans demande particulière concernant des points de comparaison / benchmark. Ils souhaiteraient pouvoir suivre la vie de leur portefeuille d'inventions et connaître leurs partenaires publics et privés.

Les établissements présents à l'atelier n'ont pas mentionné la nécessité de disposer d'indicateurs brevets par rapport à une référence, ni de spécialisation par technologie.

La discussion a également fait ressortir la demande de la maintenance et de la gestion d'une base unique, définie selon certains critères acceptés par une majorité d'établissements, intégrant l'ensemble des dépôts (et leurs extensions) faits par l'ensemble des institutions et entreprises françaises.

#### Synthèse des points de discussion

- a. Une même invention peut donner lieu à plusieurs dépôts dans le monde. C'est pourquoi l'OST travaille par office ; il est possible d'ajouter des indicateurs sur les brevets prioritaires ou les familles pour éviter les doubles comptes.
- b. Différencier les brevets actifs des brevets abandonnés demanderait un travail exploratoire dans la base de données. Les participants souhaitent une distribution par année, technologie et laboratoire de leur nombre de brevets.
- c. Il pourrait être intéressant de disposer d'informations sur le devenir des brevets, ce qui supposerait des développements spécifiques.
- d. Indicateurs de co-dépôts avec des entreprises ou des partenaires publics.
- e. Concernant la quote-part d'inventivité et de copropriété, l'OST n'a pas accès aux données relatives aux déclarations d'invention. De même, l'OST n'a pas accès aux informations relatives aux licences pour l'ensemble des établissements. Des données utilisables sur ces questions demanderaient de mettre au point un processus de collecte spécifique.
- f. La discussion a fait ressortir la demande de la maintenance et de la gestion d'une base unique intégrant les dépôts faits par les institutions et entreprises françaises sous un format propre utilisable immédiatement. Il a aussi été jugé utile de disposer d'indicateurs à des niveaux plus fins (laboratoire). Pour cette demande les participants ont mentionné l'apport potentiel des SATT. Le dispositif pourrait être le suivant :
  - création d'une base unique qui serait alimentée par le mandataire et accès aux données pour les établissements co-proprétaires ;
  - les SATT, et notamment la SATT Nord, pourraient aider à nettoyer/consolider la base de données interne des établissements.

## 4. CONCLUSION

En conclusion, deux besoins nouveaux ont été exprimés lors de cet atelier.

1. L'intérêt de bien identifier les inventeurs affiliés à telles ou telles institutions, alors que du fait du RGPD leurs adresses personnelles pourraient ne plus être systématiquement indiquées. L'OST a proposé la mise en place d'un groupe de travail sur les inventeurs académiques dont les modalités seraient définies courant 2019.
2. L'intérêt de la création d'une base unique concernant les établissements et organismes de recherche que l'OST pourrait gérer/maintenir. Cette demande devra être traitée plus amont car n'est pas forcément du ressort unique de l'OST.

L'analyse des comptes rendus des cinq ateliers permettra d'établir une liste d'ensemble des besoins ou demandes exprimés. L'OST établira ensuite des priorités dans le cadre de son programme de travail et en lien avec son Conseil d'orientation scientifique.

## 5. Annexe

### *Animateurs*

Françoise Laville, Wilfriedo Mescheba et Justin Quemener

### *Participants*

- BOURBON Natacha, Université Savoie Mont-Blanc
- FRITSCH Florian, Université Strasbourg
- ITEMAN Isabelle, Sorbonne Université
- MOUQUET-NOPPE Christine, ENS de Lyon
- ROBERT Michel, HCERES-DEE
- TIRMAN Cyrielle, Université d'Artois
- TOURNOIS Jennifer, Université de Haute-Alsace

## BASE DES PROJETS DU PCRI ET INDICATEURS SUR LA PARTICIPATION AUX PROJETS H2020

---

Du début du programme IPERU à la fin du FP7, des indicateurs ont été fournis sur les participations au programme-cadre de recherche et de développement de l'Union européenne.

Depuis le début de H2020, le format des données livrées évolue dans le sens d'une plus grande complétude des informations. En 2019, l'OST prévoit de fournir des indicateurs dans le cadre du programme IPERU en exploitant les nouvelles données disponibles.

Le but de cet atelier est de présenter ce qui est prévu pour 2019, ainsi que de recueillir les suggestions et demandes des établissements en matière d'indicateurs sur les PCRI.

L'atelier débute par un tour de table où chacun se présente et précise ses attentes pour la suite d'IPERU PCRI. Les questions abordées lors de la discussion portent sur quatre thématiques.

### 1. Production IPERU à partir de la base PCRI

Seuls quelques participants à l'atelier ont un souvenir du dernier exercice IPERU PCR

D, plutôt concernant l'interface de repérage que les rapports d'indicateurs. L'une et l'autre sont présentés rapidement pour l'information des autres participants.

A partir de ces éléments, les participants considèrent que des indicateurs sur les PCRI peuvent avoir un triple intérêt pour leur établissement :

- faciliter l'évaluation ;
- faciliter le pilotage ;
- permettre à l'établissement de mieux communiquer sur ses réussites.

### 2. Sources de données utilisées par les établissements sur les projets européens

La plupart des participants utilisent le « participant portal » (qui deviendra le « Funding & Tenders Portal »).

Certains mentionnent « Oscar » (Organisation et Suivi des Contrats et des Activités de Recherche).

Data ESR et ScanR sont peu utilisés par les participants.

### 3. SYNTHÈSE DES PROJETS D'ÉVOLUTION ET DES DEMANDES D'AMÉLIORATION

Les participants sont sollicités pour formuler des critiques, suggestions, ou demandes d'indicateurs répondant à leurs besoins. L'OST précise que la faisabilité des demandes sera analysée plus en détail ultérieurement mais que compte tenu du calendrier de production prévu pour 2019, il ne sera probablement pas possible de mettre en œuvre toutes les demandes dès le prochain exercice. Néanmoins, on essaiera d'implémenter au plus vite tout ce qui peut l'être. Par ailleurs, sur certains points, les détails de la mise en place seront renvoyés à l'examen par un groupe de travail.

Il en ressort, globalement, les points suivants :

- Demande de faire un repérage à la maille la plus fine possible (labo), lorsqu'elle est disponible. L'attention des établissements est attirée sur le fait que pour faciliter le repérage, ces données devraient être saisies (dès le stade de la réponse à l'appel à proposition) de la manière la plus complète et harmonisée possible, afin de faciliter, autant que faire se peut, le rapprochement avec le RNSR. Actuellement, les données ne permettront pas forcément un repérage exhaustif ou sans erreur.
- Enrichir la nomenclature des types d'implication des établissements sur un projet
- Revoir l'ergonomie de l'interface de repérage pour que son utilisation prenne moins de temps et fournisse davantage d'informations. L'OST rappelle que les données qui ne sont pas visibles sur Cordis ne seront pas visibles dans l'interface, à moins de concerner l'établissement.
- Il n'est pas utile de maintenir la nomenclature OST des domaines d'application : des indicateurs calculés selon la structure du programme-cadre suffisent.
- Ajouter des indicateurs et autres éléments de contexte statistique pour faciliter le diagnostic de l'établissement face à ses indicateurs.
- Demande de calcul du taux de succès par établissement. L'OST précise que cela suppose un repérage sur la base des « PROPOSALS » et pas seulement sur la base des projets financés et ne pourrait être qu'un test, pour commencer. De surcroît, la question du périmètre n'est pas résolue.
- Indicateurs sur les montants (global au niveau du projet ; au niveau des établissements)

#### 4. REMARQUES DIVERSES

Les établissements ne sont pas tous prêts à partager leurs indicateurs.

A plus long terme, il serait intéressant d'essayer d'établir un lien entre la base des PCRI et les publications, notamment pour pointer celles qui sont en Open Access..

#### 5. CONCLUSION

Un groupe de travail avec les établissements intéressés pour étudier les éventuelles propositions de l'OST sera créé.

L'atelier donnera lieu à un compte-rendu qui sera transmis aux participants.

L'analyse des comptes rendus des cinq ateliers permettra d'établir une liste d'ensemble des besoins ou demandes exprimés. L'OST établira une synthèse et une priorisation.

## 6. Annexe

### *Animateurs*

Eric HAUET, David SAPINHO et H el ene MECHOULAM

### *Participants*

- Antoine BOURGET, Universit  d'Angers
- Rapha l BRETIN, Universit  de Bourgogne
- El anor FONTAINE, ENSAM
- Aur lie KASPRZAK, Observatoire de Paris
- Carole MATTHIA, Universit  Toulouse III
- Gabrielle MINOT, Universit  Paris-Est Marne-la-Vall e
- S bastien RESPINGUE-PERRIN, Universit  d'Evry Val-d'Essonne
- St phanie ROSSARD, Universit  de technologie de Compi gne
- Caroline SAINT-GIRONS, MESRI (DGESIP)
- Sandrine SCHOTT, Universit  de Strasbourg
- H el ene VEILLARD, Observatoire de Paris

## INDICATEURS SUR LA PRODUCTION SCIENTIFIQUE : LE CAS DES SHS

---

L'atelier SHS a porté sur la problématique des indicateurs de production scientifique dans les disciplines des sciences humaines et sociales (SHS). Il a notamment abordé le choix d'une ou de plusieurs bases de données en fonction de l'objectif visé (recension des productions, calculs d'indicateurs à des fins de comparaisons inter-institutionnelle ou internationale...). La présentation de l'atelier donnait de façon indicative le type de questions qui pouvait être abordé.

- Comment constituer un/des corpus représentatifs de la production dans les disciplines SHS ?
- Comment identifier les corpus pouvant permettre le calcul d'indicateurs pertinents ?
- Quels approfondissements sont nécessaires pour explorer de nouvelles sources ?

L'atelier s'est ouvert avec un tour de table où chacun des participants a indiqué ses intérêts et questions sur le sujet. Il s'est poursuivi avec la restitution par l'OST des résultats de l'exploration de différentes sources de données pour analyser la production de l'université Bordeaux Montaigne, établissement « spécialisé » en SHS.

Les questions abordées lors de la discussion avec les établissements participants (voir annexe) ont porté sur les besoins en matière d'indicateurs de pilotage et de valorisation des recherches en SHS, ainsi que sur les solutions pour atténuer les biais résultant des sources et méthodes actuelles pour ces disciplines. Les débats sont synthétisés ci-dessous.

### 1. Des indicateurs pour le Pilotage des établissements

Les participants ont exprimé des besoins en matière d'indicateurs de pilotage de leurs établissements et de valorisation des recherches en SHS. L'évaluation par le Hcéres et le positionnement par rapport aux classements internationaux induisent un besoin régulier d'indicateurs et d'études bibliométriques.

Les participants répondent à ces besoins par l'utilisation d'outils et d'archives internes, par la participation à IPERU ou des dépôts dans HAL. Dans certains établissements, le repérage IPERU et le dépôt dans HAL sont assurés par le service commun de documentation (SCD), ce qui constitue une « assurance » de qualité des données.

Malgré l'existence de différents outils, l'usage des indicateurs bibliométriques reste problématique pour les disciplines SHS. Les participants ont exprimé leur perplexité voire leur insatisfaction par rapport aux indicateurs produits dans le cadre du programme IPERU pour les SHS malgré les précautions méthodologiques que prend l'OST. Ils considèrent que ces indicateurs se fondant sur des données qui ne représentent qu'une partie de la production en SHS, donnent une image tronquée de la production des établissements spécialisés en SHS. Il y a typiquement un écart important entre les dénombrements issus des sources internes et ceux des rapports IPERU.

### 2. Des bases de données plus appropriées

La discussion a abordé le choix de bases de données qui seraient plus appropriées que les bases internationales comme le WoS ou Scopus pour calculer des indicateurs de pilotage dans les disciplines des SHS. Elles devraient mieux prendre en compte les ouvrages et plus généralement les autres productions que les articles dans des revues à comité de lecture à diffusion internationale. Elles devraient aussi permettre de mieux prendre en compte la diversité des langues de publication.

Les universités incitent leurs chercheurs à déposer dans l'archive nationale HAL. HAL peut constituer une source de données pour les SHS françaises, mais ne permet pas de réaliser des comparaisons internationales ou de calculer des indicateurs bibliométriques. Certains participants ont suggéré que les rapports IPERU pourraient évoluer en intégrant une comparaison entre le WoS et HAL afin de mieux prendre en compte les SHS. Cependant, la production française en SHS n'est que partiellement déposée dans HAL et il est actuellement difficile de savoir la part que représentent ces dépôts, ou leurs caractéristiques comparées au total de la production française en SHS.

D'autres participants estiment que le WoS relève de l'ancien monde car de plus en plus de données sont disponibles en Open Access. Le DOAJ (Directory of Open Access Journals), par exemple, est un répertoire international de revues avec une couverture bien plus large que le WoS. Les éditeurs ont intérêt à y être.

Le DOABook (Directory of Open Access Books) permet de créer des BDD ouvertes. Dans le but de créer un écosystème en dehors des BDD privées (comme le WoS ou Scopus), Openedition, qui pilote le DOABook en Europe, fait en sorte d'y référencer les livres et de créer un label. Dans ce contexte, la question peut se poser de savoir comment développer des outils opérationnels de pilotage basés sur l'écosystème de l'Open Access d'ici 3 à 5 ans.

La première piste évoquée porte sur étude par l'OST des nouvelles bases de données, comme le DOAJ ou DOAbook pour savoir quel est leur niveau de couverture et leur audience. Dans ce sens, l'OST explore déjà des nouvelles sources comme CrossRef qui fait la promotion de l'initiative pour l'Open Citation et plaide pour des données ouvertes sur les citations. L'OST échange également avec ScienceMetrix<sup>1</sup> qui développe l'finder, une base de données bibliométriques en développement et qui devrait privilégier l'accès ouvert. Toutefois, certaines limites de cette piste ont été pointées.

Premièrement, obtenir un DOI peut être considéré comme coûteux par certains éditeurs. Un soutien financier pourra encourager les éditeurs à se faire référencer dans CrossRef (avec un abonnement national par exemple pour obtenir des DOI moins chers). Certains éditeurs académiques s'organisent en réseau ou en association et l'accompagnement peut aussi les concerner, en complément des chercheurs.

Deuxièmement, le problème des affiliations demeure dans les bases de données en accès ouvert. Il est possible de faire des analyses par thématique ou par discipline dans ces bases, mais, pas au niveau des établissements dans la mesure où l'information est manquante. Les éditeurs disposent des informations sur les affiliations, mais, ils ne font pas le travail de transformation en métadonnées pour que les bases en accès ouvert soient utilisables pour des travaux de bibliométrie. Ces difficultés peuvent être signalées au Ministère de l'Enseignement Supérieur et de la Recherche.

Par ailleurs, Le temps de publication est plus long dans certaines disciplines SHS et les carrières sont morcelées. Ce qui pose des problèmes d'affiliation lorsque la publication intervient alors que l'auteur a quitté l'établissement, comme c'est le cas notamment pour les doctorants ou les post-doc. Dans la base SPIRE de Sciences-Po, par exemple, il y a 11.000 auteurs dont 50% n'ont pas d'affiliation. Enfin, il n'y pas d'existence juridique des unités de recherche, ce qui rend difficile le repérage des auteurs d'une unité. Que faut-il repérer : l'unité de recherche, le chercheur, ou l'établissement? Au-delà des outils, il faudrait disposer de la définition juridique ou administrative des unités de recherche pour améliorer le repérage des publications, même si les chercheurs n'utilisent pas forcément les signatures normalisées.

La deuxième piste concerne la collaboration entre l'OST et les groupes de travail ou comités de pilotage de l'écosystème de la science ouverte. Le but de cette collaboration serait de définir les données et les métadonnées dont l'OST a besoin pour faire des indicateurs dans 5 ans à partir de bases en open access. L'OST a fait différents travaux à partir de HAL et explore la base de données l'finder en cours de constitution. Elle est en effet plus riche que Scopus et le WoS grâce à l'intégration de nombreuses revues, notamment en accès ouvert.

D'autres pistes ont également été avancées, notamment pour mieux rendre compte des thématiques de recherche. Il s'agirait d'explorer des archives thématiques ou les réseaux sociaux scientifiques. Des analyses en termes de cartographie des réseaux et des thématiques de recherche devraient permettre de rendre compte de la production des SHS d'un point de vue thématique plutôt que disciplinaire.

### 3. Publications dans les BDD internationales versus produits de la recherche en SHS

Une des préoccupations majeures des établissements spécialisés en SHS est le décalage important qui existe entre la production que leurs équipes de recherche déclarent et le faible volume de publications indexées dans les bases bibliométriques internationales comme le WoS. Les représentants de ces établissements à l'atelier ont exprimé le besoin de trouver des solutions pour réduire ce décalage.

L'OST a présenté l'étude de caractérisation de la production de l'université Bordeaux Montaigne (UBM), qui dans le domaine des SHS est plutôt spécialisée en sciences humaines (voir annexe). L'analyse montre que les productions déclarées dans les dossiers Hcéres ou les dépôts dans HAL ne peuvent pas être comparés globalement aux publications à diffusion internationale indexées dans les bases bibliométriques. Les dépôts dans HAL comme les dossiers d'autoévaluation fournis au Hcéres comporte outre les différents types de productions académiques (y compris les contributions dans des colloques sans actes), des

---

<sup>1</sup> ScienceMetrix a depuis été racheté par Elsevier (décembre 2018).

productions relatives à la diffusion des connaissances notamment. Elles couvrent ainsi l'ensemble des activités des unités de recherche et pas uniquement les publications scientifiques. Le périmètre de comparaison avec les bases de publications scientifiques doit donc être ajusté aux publications évaluées par les pairs. La comparaison entre les trois corpus de l'UBM (dossiers Hcéres, HAL et WoS) montre que l'écart se réduit significativement lorsque l'on compare uniquement les publications académiques dans des revues à comité de lecture. Dans le cas d'UBM, le ratio entre les publications dans le WoS et le total des productions déclarées au Hcéres est de 1 à 24. Il n'est plus que de 1 à 6 quand il s'agit uniquement de publications scientifiques avec revue par les pairs (yc les contributions à des colloques avec actes). Dans le cas de HAL, le ratio est de 1 à 8 entre les publications dans le WoS et les dépôts dans HAL. Il est de 1 à 2,3 lorsque le périmètre est limité aux publications académiques.

A la lumière de ces résultats, certains participants se sont demandés si des ratios de ce type pourraient être établis sur un échantillon plus important de façon à pouvoir déduire la production académique en SHS d'un établissement à partir des publications dans une base bibliométrique internationale et inversement. Faudrait-il affiner de tels ratios entre les sources (Hcéres, HAL, WoS) en fonction des disciplines et donc de la spécialisation des établissements ? Des ratios moyens par discipline pourraient contribuer à analyser les marges de progression d'un établissement à partir de différents indicateurs bibliométriques. L'idée est que, pour un service en charge du pilotage dans une université, l'objectif est de se positionner même à grosse maille plutôt que de calculer tous les indicateurs à partir d'un recensement exhaustif des publications. Ce recensement peut être effectué de façon complémentaire par ailleurs et la comparaison établie régulièrement entre les sources de façon à être en mesure d'en apprécier la pertinence.

#### 4. Visibilité des productions en SHS

La visibilité des résultats de la recherche en SHS est une question qui se pose avec acuité aux représentants des établissements présents à l'atelier. Comment rendre des disciplines comme l'histoire, les langues ou les sciences juridiques « lisibles » ou « visibles » ?

Face à cette question de visibilité, certains établissements ont une politique à court terme et une politique à plus long terme. A court terme, dès qu'il y a une évaluation, l'établissement cherche à augmenter sa visibilité. Mais, le travail des agents continue en interne et la visibilité augmente à long terme. Certains participants considèrent que c'est aux instances politiques et administratives des établissements d'apporter les supports administratifs et les moyens pour prendre en charge une meilleure diffusion des productions scientifiques et l'augmentation de la visibilité.

D'autres participants estiment que certains chercheurs peuvent ne pas tenir à être plus visibles. A titre d'exemple, les philosophes ou les juristes ne cherchent pas à être visibles à travers leurs publications. Dans le secteur juridique, les chercheurs ont de nombreux étudiants en master et en doctorat. Ils ne se réfèrent pas aux différentes métriques car ils se connaissent entre eux et se considèrent suffisamment connus à l'étranger. Le discours de l'établissement sur la visibilité peut être inaudible auprès de ces enseignants-chercheurs. Cependant, cela risque d'évoluer avec l'apparition des COMUE et leurs positionnements respectifs par rapport à d'autres disciplines.

La visibilité des SHS varie selon les disciplines et leurs pratiques en matière de publication. Certaines disciplines comme la géographie, l'économie ou la psychologie publient plus d'articles et plus souvent en anglais. Elles sont ainsi plus présentes dans les bases internationales. D'autres disciplines, certaines sciences humaines notamment, sont moins visibles parce qu'elles publient plus en français et plus d'ouvrages, de rapports ou autres produits considérés comme non académiques et non recensées dans ces bases. Les rapports de fouille sont par exemple considérés comme une production scientifique par les archéologues ainsi que les recensions ou les éditions critiques en littérature alors qu'ils ne sont pas visibles dans les bases de données internationales.

Ainsi, la visibilité des SHS ne dépend pas que de la politique incitative menée par les établissements (signature commune, recensement exhaustif...), mais aussi des artefacts des bases bibliométriques. L'étude de l'OST sur la visibilité des SHS françaises dans le WoS, en comparaison avec d'autres pays européens non-anglophones, montre que l'Espagne a bénéficié de l'arrivée de nouvelles revues dans le WoS. Par ailleurs, les éditeurs espagnols, mais aussi italiens, ont présenté plus de revues pour entrer dans le nouvel index du WoS, « Emerging sources citation index ». Ce qui augmente leur visibilité.

Nous pouvons conclure que, pour être visible dans ces bases, il faut publier plus d'articles dans des revues (à la différence des ouvrages) et en anglais ou au moins avec un résumé en anglais. Cependant, certains participants estiment qu'il est très difficile de contraindre les chercheurs à publier dans des revues 'visibles' et qu'il y a peut-être un travail en amont avec les éditeurs des revues SHS pour être visibles. Il y a un rapprochement à avoir avec les éditeurs, très largement subventionnés, pour qu'ils demandent aux chercheurs un résumé en anglais et se portent candidats pour le WoS.

## 5. Développer de nouveaux indicateurs

Par ailleurs, suivant l'idée de « publier moins mais mieux », l'impact sociétal des SHS est à prendre en compte. Les participants à l'atelier suggèrent que l'OST travaille sur les « altmetrics » qui ne concernent pas que l'impact académique. A terme, et lorsque les nouvelles sources et les nouveaux indicateurs seront fiabilisés, les mesures de l'impact sociétal pourraient constituer une voie d'amélioration de la visibilité des SHS.

## 6. Conclusion

Plusieurs pistes ont émergé pour améliorer la visibilité des publications dans les disciplines des SHS et développer des indicateurs de pilotage pertinents pour les établissements spécialisés. Certaines de ces pistes pourraient faire l'objet d'un groupe de travail.

A court-moyen terme, l'OST pourrait étudier l'enrichissement des rapports IPERU pour les établissements SHS avec une comparaison entre les publications dans le WoS et les dépôts dans HAL, notamment ceux qui correspondent à des publications académiques. Ces comparaisons seraient surtout pertinentes pour les établissements qui déposent une part substantielle des productions en SHS dans HAL. La démarche pourrait être étendue aux établissements disposant d'une archive institutionnelle systématique pour les SHS.

Les nouvelles sources de données en Open Access (DOABook, DOAJ) sont à explorer en réalisant des comparaisons avec les données disponibles dans les dossiers Hcéres, le WoS et HAL, à l'instar de l'étude réalisée par l'OST sur l'Université Bordeaux Montaigne. La visibilité dans les bases de données internationales dépend des disciplines, y compris au sein des SHS. La définition de ratios entre l'ensemble des productions d'une discipline et les publications à diffusion internationale peut être approfondie.

La visibilité des SHS peut également être améliorée, selon les participants, en lien avec l'évolution de l'écosystème de l'édition : augmentation de la présence des revues françaises dans les bases internationales, systématisation des résumés en anglais, nettoyage des affiliations au sein des établissements et avec les éditeurs qui disposent des métadonnées.

En ce qui concerne les indicateurs, les participants à l'atelier ont suggéré que l'OST travaille en lien avec l'écosystème de l'Open Access sur les données et les métadonnées nécessaires pour produire des indicateurs bibliométriques à partir des archives ouvertes à un horizon de 5 ans.

Enfin, certains participants à l'atelier<sup>2</sup> se sont déclarés intéressés par l'organisation d'un groupe de travail qui serait animé par l'OST pour explorer les alternatives aux indicateurs standards pour les SHS, comme la cartographie des thématiques de recherche ou l'analyse des réseaux.

## 7. Annexe

### *Animateurs*

Aouatif de La Laurencie, Anne Glanard, Abdelghani Maddi

### *Participants*

#### **Services de documentation :**

- PIÉJOS Albane, Chargée de valorisation, Conservatoire national des arts et métiers.
- MUSNIK Noémie, Bibliothécaire - chargée de mission Archive Ouverte et responsable SPIRE, Sciences Po.
- ASHTA Emmanuelle, Directrice adjointe du Pôle Documentation, chargée du numérique et des services à la recherche, université de Bourgogne.

---

<sup>2</sup> Cécile Sebban, U. Rennes 2, Solenn Bihan, U. Lille, Jean-Marc Llorens, U. Côte d'Opale, Virginie Fiesinger, U. Haute-Alsace, Stéphane Kojayan, U. Toulouse Capitole, Pierre Yves Steunou, U. Lyon II.

- BIHAN Solenn, Direction Valorisation de la recherche et membre de Lillometrics (plateforme lilloise de bibliométrie), université de Lille.
- SIMON Caroline, Ingénieur de recherche, cellule Ingénierie de Projets, direction de la Recherche et de la Valorisation, Université Polytechnique Hauts-de-France.
- SEBBAN Cécile, Responsable du Département Recherche, service commun de documentation, université Rennes 2.

### **Service Recherche / pilotage**

- BEDOS Elsa, Responsable du suivi des activités scientifiques, Sciences Po Paris.
- FIESINGER Virginie, Directrice de la recherche, université de Haute-Alsace.
- LLORENS Jean-Marc, Directeur recherche et valorisation, université du littoral Côte d'Opale.
- ZYLA Eric, Direction de la recherche et de la valorisation, université Paris 1 Panthéon-Sorbonne.
- KOJAYAN Stéphane, Directeur de la Recherche, de l'Observatoire et du Pilotage, université Toulouse Capitole.
- STEUNOU Pierre-Yves, Directeur service des études statistiques, Univ. Lyon 2.

## METHODOLOGIE RELATIVE AU CALCUL DES INDICATEURS

---

L'atelier porte sur des questions de méthode relatives à la production des indicateurs des rapports IPERU livrés aux établissements. La présentation de l'atelier donnait de façon indicative le type de question qui pouvait être abordé, donner lieu à discussion et éventuellement aboutir à des propositions d'évolution des rapports IPERU et de leur utilisation.

- L'usage des rapports IPERU est-il gêné dans certains établissements par des questions de compréhension ? La lecture de la note méthodologique IPERU : la rédaction pourrait-elle être améliorée pour en faciliter l'interprétation des indicateurs et l'appropriation par les établissements ?
- Le périmètre des publications utilisé par l'OST pour produire les rapports IPERU est-il satisfaisant : base de données utilisée, mailles disciplinaire ...
- Le type de positionnement des institutions IPERU dans la sphère de la production scientifique mondiale est-il satisfaisant ? pourrait-il être complété ?

Les questions abordées lors de l'atelier à partir des besoins des établissements participants (voir annexe) portent sur quatre thématiques. Les débats sont synthétisés ci-dessous.

### 1. Appropriation et Usage des rapports IPERU par les établissements

Le rapport IPERU est généralement diffusé dans les établissements au-delà des correspondants de l'OST, mais son usage pour le pilotage et la communication avec les partenaires de l'université est variable. Les informations fournies pourraient être davantage exploitées par les établissements.

Les représentants des établissements soulignent que la note méthodologique IPERU est bien rédigée, claire et complète. La lecture de la note facilite l'appropriation du rapport IPERU. Néanmoins, le format du rapport et la richesse des informations fournies ne permettent pas de communiquer directement avec différentes instances de l'université ou avec des partenaires comme la région.

La présentation à des instances de l'université et l'utilisation pour le pilotage supposent un travail de sélection des indicateurs les plus pertinents et plus généralement, une présentation synthétique des éléments contenus dans les différents volumes du rapport IPERU. Les établissements ne sont pas tous en mesure de réaliser ce travail d'appropriation. Ce travail au sein des établissements pourrait être facilité par l'identification de quelques formats de rapport selon les publics visés, par exemple certains conseils des universités ou les interlocuteurs de l'établissement auprès de la région. La question de l'appropriation, en lien avec le format des rapports IPERU et de potentielles versions plus ciblées, pourrait faire l'objet d'un groupe de travail avec des établissements intéressés.

L'appropriation et l'usage des informations fournies pourraient aussi être stimulés par une analyse complémentaire des liens avec les classements internationaux pour certains établissements. L'OST pourrait envisager de conduire ce type d'analyse en fonction des demandes exprimées par les établissements intéressés.

## 2. Méthode de calcul des indicateurs bibliométriques et perimetre pris en compte

La notion de compte fractionnaire est bien présentée dans la note méthodologique. Cependant, dans la mesure où les notions de type de compte sont spécifiques au domaine de la bibliométrie, il serait utile de bien mettre la définition en exergue et de l'illustrer en dehors même de l'annexe méthodologique. Cette démarche soutiendrait l'effort d'appropriation que font les correspondants IPERU dans les établissements.

La fenêtre 2 ans utilisée pour les indicateurs relatifs aux citations peut s'avérer trop courte pour refléter la réalité disciplinaire de certains établissements. L'ajout de la fenêtre 3ans et/ou sans fenêtre pourra être envisagé.

Certains établissements souhaiteraient mieux comprendre l'impact que peuvent avoir leurs recommandations en termes de signature des publications sur le repérage des adresses dans le cadre d'IPERU. D'autres s'interrogent sur la possibilité de repérer le périmètre des COMUEs. Ce type de questions pourrait être abordé dans le cadre du programme IPERU, et notamment des sessions de formation qui sont proposées par l'OST chaque année avant le lancement de la campagne de repérage.

## 3. Couverture bases de données et open access

Les établissements utilisent parfois d'autres bases pour analyser leur production et comparer avec les informations fournies dans le rapport IPERU. Certains ont fait remarquer que les nombres de publications du rapport IPERU sont parfois différents de ceux issus d'autres bases, en particulier dans les disciplines Maths, Informatique, Astronomie/Astrophysique et SHS.

Pour les Maths, la remarque est basée sur l'usage de MathSciNet : le typedoc 'livre' de MathSciNet n'est pas dans le WoS, ce qui pourrait expliquer la différence.

Pour l'Informatique, la base IEEE (Institute of Electrical and Electronics Engineers) comporte plus d'informations que le WoS : les nouveaux proceedings (nouvelles revues) non encore pris en compte dans le WoS peuvent être à l'origine de cette différence.

La base de données de références bibliographiques en Astronomie, Astrophysique (ADS : NASA Astrophysics Data System) comporte moins de publications que le WoS. Il serait notamment utile de vérifier le périmètre disciplinaire de la base ADS qui pourrait expliquer la différence.

Pour les SHS, il n'existe pas de base qui serait réellement satisfaisante pour produire des indicateurs de type IPERU. L'OST mène des travaux de comparaison de corpus, notamment avec HAL. HAL comporte des dépôts relatifs à des documents de types très divers, notamment des ouvrages, mais aussi des documents de travail et des mentions d'interventions dans des colloques sans actes. HAL n'est cependant pas une source exhaustive et pour de nombreux établissements repose sur le volontariat. L'atelier de la Journée IPERU consacré aux SHS aborde spécifiquement les questions relatives à ces disciplines et la restitution de la journée établira des liens entre les deux ateliers.

Les participants évoquent l'utilisation des publications en « open access » (OA) comme complément des sources actuelles sur les publications. Il est suggéré en particulier d'utiliser les informations sur les publications en OA disponibles dans la base WoS pour souligner la production en OA des établissements et éventuellement calculer des indicateurs sur ces publications. L'OST a un projet en cours pour exploiter le format XML des données WoS qui lui permettra d'exploiter les informations sur les OA. Il ne dispose pas des capacités de traiter les données des bases ouvertes qui demandent un travail important de mise en qualité.

## 4. Etudes thématiques et interdisciplinarité

Les participants à l'atelier ont exprimé leur intérêt pour le sujet de l'interdisciplinarité et des indicateurs qui seraient pertinents. Les établissements se sentent démunis lorsqu'il s'agit de se positionner par rapport à ce sujet, notamment dans le cadre de projets européens où cela peut être demandé. L'apport de l'OST en termes d'indicateurs est souhaité.

L'OST a travaillé sur un indicateur d'interdisciplinarité, notamment avec des établissements qui ont permis de calculer cet indicateur. Afin de passer à une phase de production, il manque une étape de validation avec le Conseil d'orientation scientifique à laquelle l'OST peut travailler en 2019 dans la mesure où il y a une demande dans le cadre d'IPERU. Un groupe de travail avec les établissements intéressés pourrait ensuite permettre d'analyser et de tester plus avant les propositions de l'OST.

Certaines universités pourraient être intéressées par des études thématiques. Ce type d'étude suppose de constituer un corpus spécifique dans la mesure où les frontières disciplinaires (et donc les nomenclatures usuelles) ne sont pas pertinentes. L'OST mène ce type de travaux, mais en dehors du cadre d'IPERU qui fournit des rapports standardisés.

## 5. CONCLUSION

L'atelier donne lieu à un compte rendu transmis aux participants.

L'analyse des comptes rendus des cinq ateliers permettra d'établir une liste d'ensemble des besoins ou demandes exprimés. L'OST établira une synthèse et des priorités dans le cadre de son programme de travail et en lien avec son Conseil d'orientation scientifique.

## 6. Annexe

### *Animateurs*

Frédérique Sachwald et Agénor Lahatte

### *Participants*

- Agnès BINET, université de Pau et des Pays de l'Adour
- Jérémie CASTELLO, MESRI DGEIP B1-1
- Lauranne CHAIGNON, Ecole normale supérieure de Paris
- Georgette DAL, université de Lille
- Florence FAVRE-BONTÉ, université Clermont Auvergne
- Marjolaine GAUTRET, Sorbonne Université
- Perrine HELLY, université de Bretagne Occidentale
- Tifenn JÉGO, Muséum national d'Histoire naturelle
- Félicien LACAN, université Paris Diderot
- Sylvie LEBORGNE, Institut National des Sciences Appliquées de Rennes
- Kareen LOUEMBE, université Lyon 2
- Thomas ROGAUME, université de Poitiers
- Frédéric SACONNET, Observatoire de Paris
- Alain SARFATI, université Paris Sud
- Malika ZENATI, université Le Havre Normandie



2 rue Albert Einstein  
75013 Paris, France  
T. 33 (0)1 55 55 60 10  
hceres.fr

Haut Conseil de l'évaluation de la recherche et de l'enseignement supérieur