Sources et données

Pour produire des indicateurs, réaliser des études et des analyses, l'OST exploite différentes sources et parmi elles, des bases de données qu'il a préalablement enrichies.

Base OST-WoS

Les publications dans les revues scientifiques et les actes de colloques constituent un des principaux modes de diffusion des résultats de la recherche. Les bases de données bibliographiques contiennent des informations descriptives des publications sources (revue, auteurs, affiliations, pays), relatives à leurs contenus scientifiques (titre, mots-clés, résumé…), et à leurs liens avec d'autres publications (références bibliographiques ou citations).

Pour produire ses analyses bibliométriques, l'OST s'appuie sur la base multidisciplinaire Web of Science (WoS) de Clarivate Analytics qu'il enrichit avec des données complémentaires (géographiques, institutionnelles…). Le WoS est l'une des principales bases de données utilisées en bibliométrie : elle recense les revues scientifiques et les actes de colloques les plus influents au niveau international. Ils sont sélectionnés en fonction, à la fois, de leurs qualités éditoriales et de leur utilisation par les pairs mesurée à l'aide de citations. Le WoS est plus représentatif pour les disciplines bien internationalisées que pour les disciplines appliquées, de "terrain" ou à forte tradition nationale. Il est ainsi moins représentatif pour certains domaines des sciences pour l'ingénieur, des sciences humaines et des sciences sociales. Néanmoins, la couverture de la base évolue et de nouvelles revues y sont intégrées chaque année.

La classification en onze grandes disciplines dans la base OST résulte d'une agrégation des domaines de recherche établis par Clarivate Analytics. Ces disciplines sont :

- Biologie appliquée-écologie
- Biologie fondamentale
- Chimie
- Informatique
- Mathématiques
- Physique
- Recherche médicale
- Sciences de l'univers
- Sciences humaines
- Sciences pour l'ingénieur
- Sciences sociales

Les revues peuvent être rattachées à plusieurs domaines de recherche et donc, à plusieurs grandes disciplines. Les articles de revues multidisciplinaires (Nature, PNAS US et Science notamment) sont distribués dans les différentes grandes disciplines en fonction de leurs sujets.

Les indicateurs sont calculés en ne retenant que certains documents : les articles originaux (y compris ceux issus des actes de colloques), les articles de synthèse (reviews) et les lettres. Les documents pour lesquels une partie des informations est manquante (par exemple le pays) ne sont pas pris en compte.

> En savoir plus sur les indicateurs

Base Brevets

Le brevet d'invention est un titre de propriété qui confère à son titulaire ou à ses ayants droit, pour un temps et sur un territoire limités, un droit exclusif d'exploitation de l'invention. Pour être brevetable, une invention doit être nouvelle, impliquer une activité inventive substantielle et être susceptible d'application industrielle.

En échange du droit exclusif qui lui est accordé, le titulaire du brevet (appelé "déposant") accepte de rendre public son brevet. Sous peine de nullité, le brevet doit exposer l'invention de façon suffisamment claire et complète pour qu'elle puisse être réalisée par un homme de l'art. Le brevet est donc non seulement un titre juridique de droit de propriété mais également une publication technique.

Les indicateurs décrivant les activités technologiques sont calculés à partir de la base de données Patstat (EPO Worldwide Patent Statistical Database), qui a été créée par l'Office européen des brevets (OEB) avec l'appui d'autres organismes internationaux (OCDE, EUROSTAT…). Patstat contient les enregistrements des demandes de brevets dès lors qu'elles sont publiées, soit dix-huit mois après la date de dépôt, ainsi que des brevets délivrés. Cette base couvre environ 80 Offices de brevets nationaux et régionaux à travers le monde dont l'OEB, qui diffuse deux fois par an une version mise à jour de sa base.

Comme pour d'autres bases, l'OST améliore et enrichit les champs de la base par des traitements informatisés. Ces traitements portent notamment sur le nettoyage des adresses (inventeurs et déposants) avec un travail spécifique sur la régionalisation de certains pays (utilisation de la nomenclature NUTS d'EUROSTAT). Enfin un travail de catégorisation sectorielle est réalisé sur les déposants.

Ainsi les indicateurs sur les brevets, malgré certains biais, peuvent être considérés, avec les dépenses de R&D et les enquêtes sur l'innovation, comme une bonne approximation de la position technologique des pays, des régions, des secteurs et des firmes. Ces indicateurs sont déclinés selon une nomenclature technologique en cinq domaines et 35 sous-domaines, ces domaines reposant sur la classification internationale des brevets (CIB) et/ou sur la classification coopérative des brevets (CPC).

Base PCRI

Les programmes-cadres de recherche & développement (PCRD) sont utilisés par la Commission européenne pour développer la recherche européenne.

Succédant au 7ème PCRD (2007-2013), le 8ème programme-cadre ou Horizon 2020 (H2020), mis en place en 2014 pour sept ans, est le programme phare du financement des activités de R&D en Europe. Doté de près de 80 milliards d'euros et fortement axé sur l'innovation, H2020 regroupe les instruments de financement de la R&D et de l'innovation mis en œuvre par la Commission européenne, ses agences et ses partenariats publics-privés (en dehors des financements fournis dans le cadre des aides structurelles de la direction générale de la politique réginale (DG Regio). Ses trois priorités sont : l'excellence scientifique, la primauté industrielle et les défis sociétaux. La participation à Horizon 2020 est ouverte aux chercheurs du monde entier.

Les données relatives à H2020 générées par la Commission européenne sont mises à disposition du Hcéres par le ministère de l'enseignement supérieur de la recherche et de l'innovation.