

## La production scientifique en sciences humaines et sociales au prisme de l'évaluation par l'AERES (France)

Pierre Glaudes

Délégué scientifique coordinateur des Sciences humaines et sociales  
Agence d'évaluation de la recherche et de l'enseignement supérieur (AERES)

### Les principes

En guise de préliminaire, on voudrait indiquer quelques principes qui déterminent les pratiques d'évaluation à l'Agence d'évaluation de la recherche et de l'enseignement supérieur (AERES)<sup>1</sup>. Il convient en effet de rappeler que l'évaluation, telle que la comprennent les acteurs de l'Agence, doit être faite par les pairs et se fonder sur des processus raisonnés, transparents portant sur plusieurs critères : une évaluation pertinente doit être « multicritérisée » et mettre en jeu des batteries d'indicateurs. On considère en effet qu'il est vain de chercher l'indicateur universel, performant, discriminant<sup>2</sup>. L'AERES utilise donc trois types d'indicateurs pour l'évaluation des unités de recherche : les indicateurs de *caractérisation* (typologie des unités, ventilation des effectifs, surfaces disponibles, etc.) ; les indicateurs d'*environnement* (type d'établissement ; contexte socio-économique ; taille du site urbain) ; les indicateurs de *performance*, qui concernent le plus le sujet de ce colloque.

Ces derniers indicateurs se distribuent en quatre sous-catégories : des indicateurs de *production scientifique*, qui doivent être à la fois qualitatifs et quantitatifs (c'est en ce point que s'active l'approche bibliométrique) ; des indicateurs d'*impact*, de couplage aux demandes et aux enjeux sociaux (ce qui suppose que l'on trouve le moyen de préciser qui est touché par les activités de recherche d'une unité – les milieux académiques, les étudiants, de plus larges groupes sociaux – et comment il l'est) ; des indicateurs de *gouvernance* ; des indicateurs de *stratégie* et de *projet*.

Ces grands types d'indicateurs se déclinent en indicateurs secondaires, qui mettent en jeu un ensemble d'items. Par exemple, la gouvernance, qui relève des indicateurs de performance, prend en compte *trois indicateurs secondaires* (la démocratie interne, la vie de l'unité, la qualité de la communication interne et externe), qui se démultiplient en une série de critères d'évaluation. Pour la démocratie interne : la configuration de l'équipe de direction, composition du conseil de laboratoire, la fréquence de ses réunions, la procédure d'affectation de crédit, etc. Pour la vie de l'unité : les types de séminaires, leur fréquence, leur périmètre, l'existence ou non d'un séminaire spécifique pour les doctorants, le mode d'organisation de celui-ci, etc. Pour la qualité de la communication interne et externe : l'existence de comptes rendus du conseil de laboratoire, d'une newsletter, d'un site web, etc.

Cet ensemble d'indicateurs pour être opératoire est assujéti à quelques règles. Les indicateurs de l'AERES sont des instruments de mesure portant sur des *structures collectives* (et non pas sur des individus), qui n'ont de valeur qu'en fonction d'*un objectif* qui doit toujours être clarifié au préalable (importance de la question « pour quoi faire ? »). Ils sont considérés comme des mesures dont la durée de validité est *limitée*, car ils sont fonction de l'évolution des milieux évalués, chaque indicateur, lorsqu'il est intégré par une communauté donnée, étant aussitôt perçu comme un objectif à atteindre et pouvant, de ce fait, devenir plus ou moins rapidement peu discriminant, donc obsolète (cf. la situation actuelle du taux de produisant). Ils sont par principe susceptibles d'*évoluer* et d'*être renouvelés* (exemple : dans un récent séminaire, on s'interrogeait sur l'opportunité d'intégrer parmi les critères de gouvernance une mesure du bonheur, de la

<sup>1</sup> <http://www.aeres-evaluation.fr/>

<sup>2</sup> On voit bien à cet égard les limites de la « note globale » dont le maintien et/ou la transformation sont en cours de discussion au sein du conseil de l'Agence.

convivialité, de la qualité du vivre ensemble au sein des unités de recherche). Ils combinent évaluation *quantitative* et *qualitative*, et, de ce fait, s'appuient sur des données bibliométriques.

## Les listes de revues

Dans ce contexte, qu'en est-il de l'approche bibliométrique en Sciences humaines et sociales à l'Agence d'évaluation de la recherche et de l'enseignement supérieur ? Cette approche est à situer dans une réflexion globale sur la production scientifique. Deux facteurs doivent être pris en considération. Le premier est *l'objet de l'évaluation* par l'AERES. Celui-ci n'est pas les individus, on l'a dit, mais les groupes, parmi lesquels on distingue trois niveaux de « granularité », qui sont fonction de la taille : les petits groupes (5-15 personnes), soit le niveau des équipes réunies dans une unité de recherche ; les groupes moyens à grands (50 à 150 personnes), soit le niveau de l'unité de recherche ou du département au sein d'un établissement ; les très grands groupes (établissements, régions, pays). L'essentiel de l'évaluation faite par l'AERES porte sur les deux premiers niveaux, même si des rapports de vague et une synthèse quadriennale ont été tentés au niveau le plus large.

Le deuxième facteur à prendre en compte est *la visée de l'évaluation*. Dans l'évaluation des petits groupes et des groupes moyens à grands, l'AERES se fixe, en matière de production scientifique trois objectifs : produire des signaux d'alerte ; détecter l'émergence ; distinguer la consécration. C'est dans ce contexte que doit être située, parmi d'autres indicateurs de performance, l'approche bibliométrique, laquelle, en Sciences humaines et sociales, comme dans les autres domaines (Sciences et technologies, Sciences de la vie et de l'environnement) porte essentiellement sur les revues. Ce point soulève plusieurs problèmes.

Pour de nombreuses disciplines de Sciences humaines et sociales, notamment du côté des Humanités (lettres, langues, arts, philosophie), la production *d'ouvrages individuels* (ouvrages scientifiques, éditions de textes, traductions) ou *collectifs* (actes de colloque) reste la forme majeure de l'activité scientifique (à quoi s'ajoutent des spécificités liées à certaines disciplines : les productions artistiques théorisées pour les arts ; les rapports de fouille pour l'archéologie, etc.), dont il résulte une extrême diversité des productions scientifiques.

En l'absence de listes de revues internationales satisfaisantes pour l'ensemble des Sciences humaines et sociales en France – du fait notamment d'un déséquilibre, dans les listes existantes (*Web of Science, Scopus...*), jouant en faveur des revues anglo-saxonnes au détriment des revues de langue française –, l'AERES a considéré qu'il lui incombait d'intervenir dans ce domaine. Elle s'est fixé comme objectif, dès sa création en 2007, de rendre compte de la profusion internationale des périodiques, qui témoigne de la croissance de la communauté mondiale des chercheurs et de l'évolution dans la manière de publier les résultats de la recherche.

Elle a donc entrepris de dresser, par champ disciplinaire ou par domaine, des listes de revues scientifiques tenant une place importante dans la diffusion des résultats de la recherche à l'échelle nationale et surtout internationale, – listes qui ont été fondues dans une liste générale des revues en Sciences humaines et sociales publiée sur le site de l'Agence. Des commissions disciplinaires regroupant des représentants des sections du CoNRS et du CNU, des membres de l'IUF, des personnalités qualifiées étrangères et les délégués scientifiques de l'AERES ont été constituées en 2008 avec pour mission de tenter de parvenir à un consensus dans chaque communauté ; elles ont abouti à l'élaboration des premières listes, publiées sur le site de l'Agence. Ces listes sont actualisées chaque année en tenant compte du degré d'internationalisation, de l'organisation et des pratiques des milieux de chaque discipline : les commissions d'actualisation répondent aux sollicitations des périodiques, récemment créés ou non sélectionnés, et prennent en compte des revues en ligne.

## La disparité entre les listes

Les conditions dans lesquelles ces listes particulières et cette liste globale ont été constituées ont abouti à un résultat mitigé, dont on ne peut pleinement se satisfaire. Dans le contexte de la politique volontariste d'une Agence récemment créée et soucieuse d'être opérationnelle dans les meilleurs délais, chaque domaine (ou chaque discipline) a été invité à établir des listes, sans qu'une méthodologie commune ait été fixée ni qu'une réflexion collective suffisante sur les critères de définition d'une « revue scientifique » ait été considérée comme un préalable. De là des résultats d'une extrême disparité.

Certaines disciplines, comme l'économie et la gestion, *ont simplement repris et combiné des listes établies ailleurs* (liste de la section 37 du CNRS, la liste *Econlit : Expanded Journal List* et la liste *Harzing : Journal Quality List*) : la pratique se limite ici à un simple référencement à des objets constitués à partir d'une méthodologie qui n'a pas été examinée par les délégués scientifiques de l'AERES et soumise à la démarche de recherche de la qualité qui est un des principes de l'évaluation de l'Agence. D'autres disciplines, comme le droit, ont constitué leur propre liste, à partir de *critères empiriques de notoriété*, fixés par les autorités représentatives des disciplines juridiques (présidents des sections 01, 02, 03 du CNU, président de la section 36 du CoNRS, DSA en charge du droit au sein du CNRS, président du conseil national du droit, président de la conférence des doyens des facultés de droit, présidents d'université à dominante juridique, personnalités qualifiées). D'autres disciplines encore ont établi des listes *en fonction d'indicateurs*, sans que ceux-ci soient unifiés pour l'ensemble du secteur Sciences humaines et sociales.

Pour la philosophie, quatre critères ont été utilisés : 1) publication comportant principalement des articles de recherche en philosophie ; 2) présence effective d'un comité scientifique et d'un comité de lecture distincts et comportant des chercheurs de la communauté internationale ; 3) lecture des articles proposés à la revue en « double aveugle » (y compris pour les « dossiers » avec contributions commandées) ; 4) sélectivité. Pour les sciences de l'information et de la communication, ce ne sont pas moins de sept critères qui ont permis d'établir la liste : 1) existence d'un comité de rédaction ouvert (représentativité des membres, non concentration sur un seul lieu de recherche, appel à des collègues étrangers) ; 2) sélection rigoureuse des articles (relecture croisée, en aveugle, commentaires critiques pouvant aboutir à des demandes de réécriture ou des refus) ; 3) ouverture internationale (présence d'auteurs étrangers, résumé en langue étrangère) ; 4) origine institutionnelle de la rédaction (association scientifique, laboratoires et équipes de recherche reconnues, soutien d'une maison d'édition reconnue) ; 5) exigences internes de publication (régularité annuelle des parutions, nombre minimal d'articles parus, part limitée d'autopublication des membres de la rédaction, taille moyenne des articles conséquente) ; 6) liens institutionnels et scientifiques avérés avec la discipline (origine disciplinaire des auteurs, problématiques soulevées, contribution aux débats de la discipline, apports à la diffusion des savoirs par l'existence de notes de lecture) ; 7) indexation dans les grandes banques de données documentaires internationales (Francis, ISI, etc.).

À ces disparités de méthode s'ajoutent des différences notables dans la nature des listes obtenues. Certains secteurs disciplinaires (droit, sciences de l'éducation, sociologie-démographie, philosophie, arts, histoire-histoire de l'art-archéologie, etc.) ont limité leur travail à l'établissement d'un périmètre scientifique. D'autres disciplines (Psychologie, STAPS par exemple) non seulement ont établi un périmètre scientifique, mais se sont engagés de surcroît dans la voie du « ranking », en procédant à un classement, sans qu'aient été fixées pour autant des normes de classement communes aux disciplines des Sciences humaines et sociales. En psychologie, on dispose d'une échelle à 5 niveaux : modéré, assez bon, bon, très bon, excellent. En STAPS, on est passé d'une échelle à deux niveaux (A+ et A) à une reconfiguration qui ne prend en compte que les revues dites « de rang A » (*i.e.* prise en compte dans le décompte des « producteurs »), parmi lesquelles on distingue désormais trois niveaux (1, 2, 3), en fonction de la qualité scientifique et du rayonnement. Enfin, un dernier groupe de disciplines (la langue et la littérature française, la littérature comparée, les langues et littératures anciennes, les langues, littératures et civilisations

étrangères) ont refusé d'établir des listes en vue d'une évaluation bibliométrique, considérant que cette approche n'étant pas pertinente. (Ces disciplines sont en train d'évoluer : les lettres ont par exemple accepté le principe d'un périmètre scientifique permettant d'établir une liste de revues).

Dans ce contexte contrasté surgissent inévitablement des problèmes, comme des disparités d'indexation ou de classement liées en particulier à la pluridisciplinarité. Il peut arriver qu'une revue du fait de son extension pluridisciplinaire soit indexée dans un domaine, mais pas dans un autre ; ou qu'elle soit diversement classée selon les domaines de spécialité où elle a été retenue. Cela n'est pas sans incidence sur la cohérence de la liste globale des revues en Sciences humaines et sociales produite par l'Agence.

## Le premier bilan

Après cinq années d'existence, l'AERES est à l'heure d'un premier bilan et s'oriente donc vers une reprise méthodique de la question des indicateurs. Alors qu'elle vient de changer de président, l'Agence, profitant de l'allègement de la prochaine vague C (du fait au passage à un rythme d'évaluation quinquennal), a fixé comme l'une de ses priorités la question des indicateurs en Sciences humaines et sociales. Cette question sera abordée en se fondant une fois encore sur quelques principes. Un travail sur les indicateurs a pour préalable la définition des *éléments de culture partagée*, de langage commun, en ayant le souci de distinguer l'essentiel de l'accessoire. S'agissant en particulier des Sciences humaines et sociales, il est important d'abord de *relativiser leur exceptionnalité* en la rapportant à des facteurs qui sont peut-être moins constitutifs (*i.e.* relevant d'un fait de « nature » ou d'une « essence ») qu'épistémologiques, historiques et institutionnels. La coexistence des Sciences humaines et sociales au sein de l'AERES avec les Sciences de la vie et de l'environnement et les Sciences et technologies apprend à ceux qui en font l'expérience qu'il y a des clivages au sein de ces grands domaines qui clivent aussi les Sciences humaines et sociales : recherche finalisée *vs* recherche fondamentale, recherche collective *vs* recherche individuelle, grosses structures de recherche *vs* petites structures, etc. En conséquence, il y a des similitudes entre certains domaines des Sciences humaines et sociales et certains domaines des Sciences et technologies et des Sciences de la vie et de l'environnement : la recherche individuelle est de première importance en mathématique comme en philosophie ou en lettres, par exemple ; en biologie, la tendance est aux petites équipes, comme en droit, etc.

Il faut ensuite *définir ce qui constitue les Sciences humaines et sociales comme telles* (la question du sens ? celle de la valeur ? la combinaison, variable selon les disciplines, de la modélisation et de l'herméneutique ? la typologie des productions ? les modalités de la recherche ? etc.), non pas négativement (par exclusion, retranchement, parti pris d'insularité) mais positivement (par affirmation du caractère propre). Il convient enfin de poser en principe qu'elles sont *commensurables* aux autres disciplines du savoir, qu'il existe donc des indicateurs pertinents pour objectiver leurs performances (et rendre compte de l'utilisation des crédits dont elles bénéficient), tout en tenant compte de leurs spécificités.

Un autre préalable à la réflexion consistera à tenir compte des méthodes et pratiques concurrentes dans le domaine des Sciences humaines et sociales : travailler en concertation avec les autres parties prenantes : DGRI, OST, CoNRS, DSA de l'INSHS, sections concernées du CNU, alliance ATHENA considérer les démarches d'évaluation et les indicateurs retenus par les institutions d'évaluation étrangères équivalentes à l'AERES (ERIH, HEFCE en Grande-Bretagne, NWD aux Pays-Bas, etc.).

## Les perspectives

C'est dans ce contexte général qu'il faut situer l'intérêt de l'Agence d'évaluation de la recherche et de l'enseignement supérieur pour la bibliométrie, pour les listes de revue et pour les

bases de données indexant ces listes. Dans ce domaine, somme toute limité mais crucial, le travail sur les indicateurs programmé à l'Agence consistera à se doter *a posteriori* de critères communs en vue de la constitution de périmètres scientifiques pour les revues des différents domaines ou disciplines.

Plus précisément, il s'agira d'abord de poser une définition de ce que l'on entend par « revue scientifique » (au-delà de la simple présence d'un numéro ISSN). Il faudra ensuite définir les indicateurs pertinents permettant d'évaluer ce type de revues en Sciences humaines et sociales et, le cas échéant, de les classer. En l'état, embryonnaire, de la réflexion, voici un premier repérage d'indicateurs secondaires et les items afférents, qui restent à valider :

- *Mode de publication et de diffusion* :
  - Existence d'un diffuseur ;
  - Nombre d'exemplaires diffusés par numéro ;
  - Existence d'une publication en ligne ;
  - Modalités de cette publication (*open access*, péage) ;
- *Présentation matérielle de la publication* :
  - Présence de notes de bas de page dans les articles ;
  - Présence d'une bibliographie par article ;
  - Existence d'abstracts en langue anglaise ;
- *Politique éditoriale*
  - Existence d'une ligne éditoriale identifiable ;
  - Adéquation des contenus aux objectifs scientifiques de la publication ;
  - Diversité des auteurs d'articles (extension locale, nationale, internationale) ;
  - Diversité des éditeurs scientifiques/numéro (extension locale, nationale, internationale) ;
  - Taux d'acceptation ;
  - Existence d'un site web (en français seul ; bilingue : anglais/autre langue) présentant la revue et ses procédures de sélection.
- *Notoriété*
  - Présence au catalogue des bibliothèques régionales, nationales, internationales
  - Niveau de citation des articles dans les bases disponibles ;
  - Niveau de citation des éditeurs scientifiques dans les bases disponibles ;
  - Évaluation par les pairs.

Ce n'est qu'au moment où l'AERES aura établi et testé la batterie d'indicateurs permettant de dresser un état des lieux pour les revues en Sciences humaines et sociales (mais aussi pour les ouvrages individuels et collectifs, notamment les actes de colloque) qu'elle pourra reprendre et affiner les périmètres scientifiques qu'elle a dessinés pour chaque domaine disciplinaire, qu'elle pourra aussi avancer dans la voie d'un classement des revues (en affectant un coefficient à chaque indicateur et en définissant des paliers permettant de passer du niveau inférieur au niveau supérieur de l'échelle ?), qui est son objectif à terme. Deux hypothèses sont envisagées.

La première est la constitution d'un périmètre scientifique large (recensement le plus complet possible des revues d'un champ scientifique, nationales et internationale) et, à l'intérieur de ce périmètre, la distinction des revues « à étoile » (i.e. des revues ayant atteint, au vu des indicateurs retenus, un certain niveau de qualité). Quatre indicateurs, à première vue, sont considérés comme

prioritaires dans la liste dressée plus haut : l'existence d'une publication en ligne ; l'existence d'abstracts en langue anglaise ; la diversité des auteurs d'articles ; le taux d'acceptation. La seconde hypothèse porte sur la constitution d'un périmètre restreint d'emblée à des revues remplissant un ensemble de conditions d'accès à la liste de l'AERES (conditions toujours fixées en fonction des indicateurs déjà mentionnés).

Il ne s'agit là que de pistes à peine ébauchées qui seront affinées de septembre à décembre 2010 dans un séminaire hebdomadaire qui réunira les délégués scientifiques de l'AERES et des spécialistes de l'évaluation. La priorité est à la constitution, au sein de ce groupe de travail, d'*un référentiel* : un corpus d'indicateurs communs à toutes les disciplines des Sciences humaines et sociales. Ceux-ci pourront être différemment pourvus de coefficients, voire, pour certains, activés ou désactivés selon les disciplines, pour tenir compte des spécificités de chaque domaine, la question d'indicateurs fondamentaux, qu'il est impossible de neutraliser, étant évidemment posée.

Une fois ce corpus établi, les délégués scientifiques en Sciences humaines et sociales mèneront auprès de leurs communautés respectives une consultation, en s'appuyant sur les représentants du CNU et du CoNRS concernés, sur les DSA du CNRS, ainsi que sur des personnalités qualifiées. Cette consultation aura pour but de légitimer la démarche et de s'assurer de son acceptabilité, l'objectif visé, dont on espère qu'il n'est pas trop chimérique, étant celui d'une adhésion largement majoritaire. L'AERES, on le voit, est engagée dans une démarche de rationalisation méthodique de ses indicateurs, qu'elle souhaite fonder sur une exigence de qualité tenant compte de la diversité des pratiques en Sciences humaines et sociales et permettant, en même temps, de caractériser avec cohérence, rigueur et finesse les productions issues de ces disciplines.