

Biodiversité : les défis de la nature en open data

LE MONDE SCIENCE ET TECHNO | 08.02.2016 à 15h59 • Mis à jour le 14.02.2016 à 19h22 | Par Audrey Garric

Des centaines d'outardes canepetières – des oiseaux des plaines rares et protégés –, des grandes sauterelles vertes ou encore des papillons d'ianes. Voilà la moisson récoltée, parmi plus de 300 espèces de plantes, 150 d'oiseaux et 100 d'insectes, par le bureau d'études Biotope sur le chantier de la ligne à grande vitesse qui doit relier Montpellier à Nîmes en 2017. Comme pour tout projet d'aménagement, l'entreprise réalise des études d'impact chaque année afin d'évaluer les effets sur l'environnement du futur ouvrage et ses mesures de compensation. Mais, pour la première fois, elle devait verser ces informations, le 10 février, à l'Inventaire national du patrimoine naturel (INPN), la plus grande base de données française en matière de biodiversité, qui dépend du Muséum national d'histoire naturelle (MNHN).



Le pique-prune fait partie des espèces protégées en France. Robert Henno/ BIOSPHOTO

Cette démarche préfigure une forme d'«open data pour la biodiversité», une nouveauté intégrée par la ministre de l'écologie, Ségolène Royal, au projet de loi pour la biodiversité. L'article 3 *ter* du texte, voté par les sénateurs le 26 janvier et qui doit revenir en seconde lecture devant l'Assemblée mi-mars, prévoit que «*les maîtres d'ouvrage, publics ou privés, doivent contribuer à cet inventaire national par la saisie ou, à défaut, le versement des données brutes de biodiversité acquises à l'occasion des études d'impact des plans et programmes (...) et des projets d'aménagement*». Le texte précise que ces informations «*sont diffusées comme des données publiques, gratuites, librement réutilisables*».

«On va vers la démocratisation des données. Chaque citoyen pourra y avoir accès, ce qui améliore la connaissance de la nature et de son état de conservation», se réjouit Jean-Philippe Sibley, le directeur du Service du patrimoine naturel au MNHN. Objectif: tripler le nombre de données disponibles, pour atteindre près de 100 millions d'ici à 2020.

L'inventaire de 160 000 espèces

Aujourd'hui, l'INPN contient 35 millions d'entrées sur 160000 espèces de faune ou de flore françaises. La partie émergée de l'iceberg, c'est un site Internet qui affiche 120000 visites par mois et offre la possibilité à chacun de chercher des informations sur un programme, un habitat ou une espèce. Entrez «pic épeichette» – un oiseau que l'on entend parfois dans les jardins du muséum – et vous obtiendrez son nom latin (*Dendrocopos minor*), sa photo, sa classification dans l'arbre du vivant et sa répartition sur l'ensemble du territoire français.

La partie immergée, elle, est l'énorme machine qui agrège, standardise et rassemble toutes ces informations, à raison de 5000 nouvelles données par jour. «Notre travail n'est pas de découvrir de nouvelles espèces, précise Jean-Philippe Sibley. Nous les récupérons auprès des experts, des taxonomistes, et nous les validons scientifiquement.» Les données sont saisies grâce à des outils en ligne (tels que CardObs) par des milliers d'établissements publics, d'entreprises privées, d'associations naturalistes, de sociétés savantes ou bien de particuliers, qui acceptent de léguer leurs informations par voie de convention.

Mais, jusqu'à présent, les bureaux d'études, mandatés par les maîtres d'ouvrage, ne contribuaient guère. Pour l'essentiel, les informations restaient consignées dans des carnets, sans protocole méthodologique harmonisé. «Ce sont des trésors cachés, souvent inexploités. De sorte qu'il faut parfois tout réinventer», reconnaît Christian Caye, délégué au développement durable de Vinci, qui réalise une centaine d'inventaires chaque année sur ses chantiers d'autoroutes, de lignes ferroviaires ou d'aéroports. Des données qui ne sont pas encore versées à l'INPN.

Des « problèmes à régler »

La construction d'une telle mémoire naturaliste, publique et transparente divise les experts. «Je suis favorable à la transmission de données, mais de manière encadrée. Cette mesure est absolument inapplicable en l'état. C'est une bombe», lâche Frédéric Melki, président-fondateur de Biotope, premier bureau d'études français dans l'environnement, avec 240 salariés et un millier d'études par an. Et d'énumérer la liste des «problèmes à régler»: à quel moment les données seront-elles fournies au Muséum? Quel sera leur degré de précision? «Si on n'attend pas la fin de l'instruction d'un dossier et la délivrance du permis de construire, on s'expose à la divulgation d'informations confidentielles aux concurrents, de même que si les données sont trop fines», poursuit-il. Enfin, reste à savoir qui va financer la diffusion de ces informations. «La création de bases de données coûte cher. On dépense entre 5000 euros et 1 million d'euros par inventaire.»

L'ouverture des données au grand public est «nécessaire» et favorisera une meilleure protection de la nature, jugent au contraire les ONG. «Dans dix ans, on pourra réutiliser des études d'impact pour voir si les mesures compensatoires proposées par un aménageur ont permis de restaurer un écosystème qui avait été détruit ou endommagé», avance Frédéric Jiguet, professeur au MNHN (CNRS) et ornithologue. «Les associations et les citoyens

pourront vérifier que les études ont été correctement réalisées, si tant est qu'on puisse s'assurer que les maîtres d'ouvrage versent bien les données», complète Emeline Bentz, de la Fondation Nicolas Hulot. Dernier exemple en date, rappelle-t-elle, le collectif des Naturalistes en lutte, opposé à la construction de l'aéroport de Notre-Dame-des-Landes (Loire-Atlantique), qui a affirmé fin 2015 avoir déniché cinq espèces rares et protégées «oubliées» lors de l'inventaire des bureaux d'études.

Contre-pouvoir

Alors, l'open data naturaliste, un contre-pouvoir citoyen? Le cas du petit scarabée baptisé pique-prune, qui a bloqué la construction d'une autoroute entre Alençon et Le Mans de 1996 à 2002, est resté dans les mémoires. *«C'est un outil pour les associations qui veulent protéger des sites: les données sont publiques, du coup nul ne peut prétendre les ignorer»*, juge Grégoire Lois, le directeur de Vigie-Nature, au MNHN, qui regroupe les programmes de sciences participatives. *«On est certain de voir se multiplier les recours d'opposants»*, renchérit Frédéric Melki. *«Nous sommes convaincus qu'il faut anticiper et privilégier la transparence dans tous les cas, rétorque Christian Caye, de Vinci. Notre intérêt est de prendre en considération le plus tôt possible les attentes de tous.»*

De manière étonnante, les réfractaires au partage des inventaires de faune et de flore ne se retrouvent pas forcément dans le secteur privé. Ainsi, au MNHN, seuls quatre des 17 programmes de sciences participatives protocolés de Vigie-Nature sont versés à l'INPN! L'établissement est partagé entre deux écoles, celle de l'open data contre celle de la protection des données. Sachant que les contraintes techniques et financières freinent tout le monde.

«Centraliser les données brutes, oui; les mettre en accès libre à tous, non», tranche Frédéric Jiguet, qui gère les observatoires oiseaux de Vigie-Nature. Le chercheur craint surtout *«les risques de détournement des informations par des groupes d'intérêt, comme nous l'avons déjà subi avec la Fédération nationale des chasseurs par exemple»*. Le danger réside également, ajoute-t-il, dans la course à la recherche à laquelle se livrent les grandes universités mondiales: *«Nous pourrions être "pillés" par des chercheurs concurrents, qui publieraient dans les revues scientifiques à partir de nos jeux de données.»*

«Il faut faire un gros travail pour changer les mentalités, juge Grégoire Lois. C'est un leurre que de croire qu'une donnée brute a une valeur monétaire. C'est son interprétation qui compte, et elle demande une expertise naturaliste.»

L'apport des amateurs

Les sciences participatives, qui font intervenir des non-scientifiques, enregistrent une forte croissance depuis une quinzaine d'années (recensements naturalistes, mesures, jeux sérieux en ligne, etc). Elles fournissent de nouvelles connaissances, mais suscitent des préoccupations (fiabilité et reproductibilité des données, rigueur des protocoles ou manipulation de la recherche). Un rapport sur le sujet, dirigé par François Houllier, PDG de l'INRA, a été remis à la ministre de l'éducation nationale, de l'enseignement supérieur et de la recherche le 4 février. Les auteurs recommandent d'organiser une communauté de pratiques, de soutenir ce domaine par des moyens techniques, financiers et réglementaires et d'associer les publics scolaires.