



Suivi de la reproduction des brochets dans une frayère en Seine-et-Marne.

© Laurent Mignaux - Terra

ENJEUX

## Donner ses données

Le partage des données dans le domaine environnemental est aujourd'hui inscrit dans la culture commune. Mais il continue de soulever des questions qu'il convient de bien cerner et prendre en compte afin d'en poursuivre sereinement la production et la diffusion.

Présence et effectif d'une espèce animale ou végétale dans un espace, date de passage d'un oiseau en migration sur un territoire, période de fauche d'une prairie, niveaux d'un plan d'eau en fonction de la période de l'année... La connaissance et le suivi des milieux naturels mobilisent depuis fort longtemps une galaxie d'informations très diverses et importantes en quantité, produites par une communauté étendue qui va du naturaliste amateur jusqu'au scientifique. L'apparition et la diffusion des outils numériques à la fin du XX<sup>e</sup> siècle a profondément bouleversé les conditions de production de ces informations, devenues des « données », rendant ces processus plus efficaces, ouvrant de larges champs d'investigation. Cette révolution technologique a par ailleurs tissé, par les réseaux qui devenaient possibles, de nouvelles possibilités de partage.

Pour autant, tous les acteurs concernés par la production de données environnementales ne sont pas de farouches défenseurs de sa diffusion sans limite. Un panorama des idées en présence s'impose.

**Partager une donnée, c'est aussi se donner les moyens de la faire valider.**

De quoi parle-t-on tout d'abord lorsque l'on parle de données ? Historiquement, les informations enregistrées et archivées dans le monde naturaliste concernaient surtout les espèces en tant que telles (présence, âge, état de la population, etc.), mais les pratiques s'élargissent aujourd'hui à l'ensemble des données qui peuvent enrichir les travaux sur ces espèces : état et évolution des

milieux, pratiques de gestion, perturbations accidentelles, etc. Qui produit des données ? Le citoyen qui repère un triton dans son jardin et le signale à une association locale, les membres des associations environnementalistes, les gestionnaires d'espaces naturels, les services de l'État, les collectivités, les scientifiques, des professionnels au cours de leur activité (agriculteurs, pêcheurs, forestiers, aménageurs). Les données qu'ils fournissent, volontairement (on ne parle pas ici des données sournoisement collectées par les acteurs d'Internet...), ont donc des statuts divers. Une partie d'entre elles est publique, et a donc pour vocation à être partagée (cf. encadré p. 26), l'autre est privée, son accès est donc conditionné à l'obtention de l'accord de son propriétaire. Le débat sur le partage des données porte donc principalement sur les données privées.

## UN CHAMP CONTINU ?

Pourquoi a-t-on intérêt à partager ses données, autrement dit à les mettre à disposition d'autrui ? Pour nombre d'interlocuteurs interrogés, le partage fonde la force d'une donnée. Seule, celle-ci n'a qu'une portée très courte. Mise en perspective avec d'autres informations, elle s'enrichit d'un sens précieux. Pour René Lalement, directeur adjoint de l'appui aux politiques publiques à l'Agence française pour la biodiversité, il convient d'ailleurs de parler de LA donnée (et non pas des données) au même titre que l'on parle de l'énergie, du vivant. « La donnée n'a pas de frontières », résume-t-il avant de poursuivre : « une donnée individuelle, ce n'est pas intéressant, comme la constante de Planck seule n'a pas d'intérêt, ce qui compte c'est de l'associer à d'autres. Si on ne s'intéresse qu'à une espèce, on ne comprend pas tout. Alors que lorsqu'on considère par exemple les liens trophiques qui la font intervenir, l'état de sa population, alors là, les choses s'éclairent. » Partager une donnée, c'est aussi se donner les moyens de la faire valider. « Une donnée qui n'est pas partagée n'est pas une donnée, car elle ne peut être critiquée », estime sur ce point René Lalement.

Grâce à la force de l'action collective, le partage des données améliore par ailleurs la préservation de l'environnement, et en particulier la conservation des espèces. « Nous échangeons volontiers nos données car nous savons qu'elles peuvent servir à d'autres et qu'en retour, nous avons besoin des données extérieures pour nourrir nos propres travaux de gestion », témoigne ainsi Francis Meunier, directeur adjoint du Conservatoire des espaces naturels de Picardie. « Par exemple, nous sommes à l'affût de toute information nous permettant de repérer dans la région des "points chauds" de biodiversité qui demeureraient encore dépourvus de statut de préservation. La mutualisation d'informations au sein de l'observatoire régional auquel nous participons aux côtés d'associations, d'organismes publics, nous aide clairement dans ce sens. »

Pour Laurent Poncet, directeur adjoint de l'unité mixte de service PatriNat (MNHN/AFB/CNRS), « plus on partage les données existantes et plus la nature peut être prise en compte dans l'élaboration des politiques. La connaissance

sur la biodiversité peut ainsi peser dans la décision publique. » La diffusion des données contribue d'autre part à créer un débat serein autour des politiques publiques, « on éviterait des climats tels que celui autour du projet d'aéroport à Notre-Dame-des-Landes si l'on partageait davantage les données nécessaires à l'évaluation. La nouvelle loi Biodiversité va dans ce sens », poursuit-il.

“  
La diffusion d'une donnée,  
sa mise en commun, peut  
contribuer à sa pérennité.”

Plus trivialement, la diffusion d'une donnée, sa mise en commun, peut contribuer à sa pérennité. Si l'annotation sur un carnet de terrain n'est pas nécessairement plus risquée que l'enregistrement dans un fichier numérique en termes de durabilité (un disque dur, ça se dégrade aussi...), l'archivage dans des plateformes communes aboutit à la capitalisation d'une multitude d'informations qui risqueraient d'être oubliées au fil du temps. Elles viennent alors abonder un patrimoine commun. Avec près de quarante millions de données d'occurrence sur les espèces, l'Inventaire national du patrimoine naturel illustre bien cette dimension du partage mais renvoie d'emblée à un problème de fond : celui de la mise à jour. « Il faut être vigilant dans les processus de centralisation de l'information », prévient ainsi Gérard Largier, directeur du Conservatoire botanique national des Pyrénées et de Midi-Pyrénées. « Une fois qu'une donnée est archivée

au Muséum, il faut s'organiser de telle sorte qu'elle puisse être révisée facilement en relation avec les sources. Car une donnée naturaliste est, par nature, sujette à discussion : une espèce peut être confondue avec une autre... »

## CONTRÔLER ET FINANCER LA DONNÉE

Mais les réticences et les freins au partage généralisé des données se révèlent multiples. Qu'il s'agisse d'un ornithologue amateur ou d'un chercheur professionnel, celui qui déniche une pépite a d'abord envie de faire savoir qu'il en est à l'origine, au moyen d'un article publié par exemple. Il est donc respectueux de conserver la « primeur » d'une information à celui qui s'est démené. Cette paternité s'accompagne parfois de la revendication d'une expertise exclusive : « les organismes producteurs de données ont parfois peur de la mauvaise interprétation qui pourrait être faite à partir de leurs données estimant être les seuls à pouvoir les lire dans leur contexte », explique Laurent Poncet. Il convient donc de bien documenter les informations que l'on partage, tant sur leurs sources que sur le contexte dont elles sont issues et, poursuit Laurent Poncet « comme dans toute démarche "académique", d'être prêt à la controverse scientifique. » Il serait par ailleurs nécessaire que les producteurs de données soient informés des projets qui risquent d'exercer une pression sur les sites qu'ils suivent. Il s'agirait là d'une vraie avancée dans la construction d'une relation de confiance entre producteurs et utilisateurs.



Observation sur la Réserve naturelle nationale, étang de Cousseau.

© Laurent Mignaux - Terra



Fiche d'inventaire de la flore.

© Patrick Folliet - PN-Vanoise

Sur un plan plus technique, produire des données dans la perspective de les diffuser nécessite du travail. Il faut en effet vérifier l'information, la présenter sous une forme normée, et se doter pour cela d'équipements informatiques, disposer d'un temps de travail suffisant. Autant de choses qui ont un coût. À la Ligue de protection des oiseaux, « sept millions de données sont générées chaque année par les associations, témoigne Dominique Aribert, sa directrice du pôle Conservation de la nature. La saisie seule représente mille équivalents temps plein, assurés à 95 % par des bénévoles. Mais la validation, l'organisation des informations sont pris en charge par des salariés. On ne peut pas tout demander aux bénévoles... » Or, alors que de telles associations fournissent beaucoup de données de qualité aux services de l'État, leur contribution n'est pas toujours suffisamment subventionnée en retour. Si les discussions entre État et associations naturalistes avancent sur ce point, « aucun compromis satisfaisant sur le financement pérenne des travaux conduits par les associations n'est encore trouvé », juge Dominique Aribert, et les réseaux naturalistes revendiquent que « leur expertise soit reconnue au-delà du seul recueil de données. »

### SECRET BIEN GARDÉ

Toute donnée n'est toutefois pas bonne à dévoiler sans précaution. « C'est le cas notamment des données publiques dont la divulgation pourrait nuire à la sauvegarde d'une espèce sensible », pointe à ce sujet Laurent Poncet. En effet, les défenseurs d'espèces fragiles peuvent craindre du dérangement, voire des actes malveillants. Emilie Leblond, qui travaille sur le Système d'informations halieutiques (SIH) de l'Ifremer, doit faire face quant à elle à l'inquiétude des pêcheurs auprès desquels le SIH obtient une partie des informations (effectifs et caractéristiques des espèces pêchées) qui alimentent les travaux de recherche et d'expertise de l'institut. « On manipule là des données qui relèvent en partie du secret industriel, les zones de pêche,

les techniques employées, etc. », explique-t-elle. Pour favoriser leur participation, l'Ifremer leur explique qu'une information solide contribue à faire des estimations pertinentes, agrège un minimum de données, et lorsque c'est possible, convient avec eux de ce qui est diffusable ou pas.

**Sept millions de données sont générées chaque année par les associations.**

À son échelle, Emilie Leblond illustre un enjeu transversal à l'ensemble des filières de production de données environnementales : la concertation entre toutes les parties prenantes au processus. • **Christophe Trehet**

### REPÈRE

La mise à la disposition des citoyens des données publiques est fondée sur la loi n° 78-753 du 17 juillet 1978 relative au droit d'accès aux documents administratifs. Cette loi définit ces données comme celles produites ou reçues « dans le cadre de leur mission de service public, par l'État, les collectivités territoriales ainsi que par les autres personnes de droit public ou les personnes de droit privé chargées d'une telle mission. » Progressivement, des évolutions législatives ont institué la diffusion obligatoire et gratuite de ces données publiques et ce, sous des formats qui permettent la réutilisation des données. Depuis 2012, l'État s'est engagé dans un vaste chantier pour mettre en place « l'open data » public, avec notamment la création de la plateforme data.gouv.fr où un peu plus de 21 000 séries de données sont renseignées actuellement. Pour autant, d'importants progrès restent à faire, car l'État rechigne encore à partager certaines de ses données, y compris avec le secteur de la recherche publique...