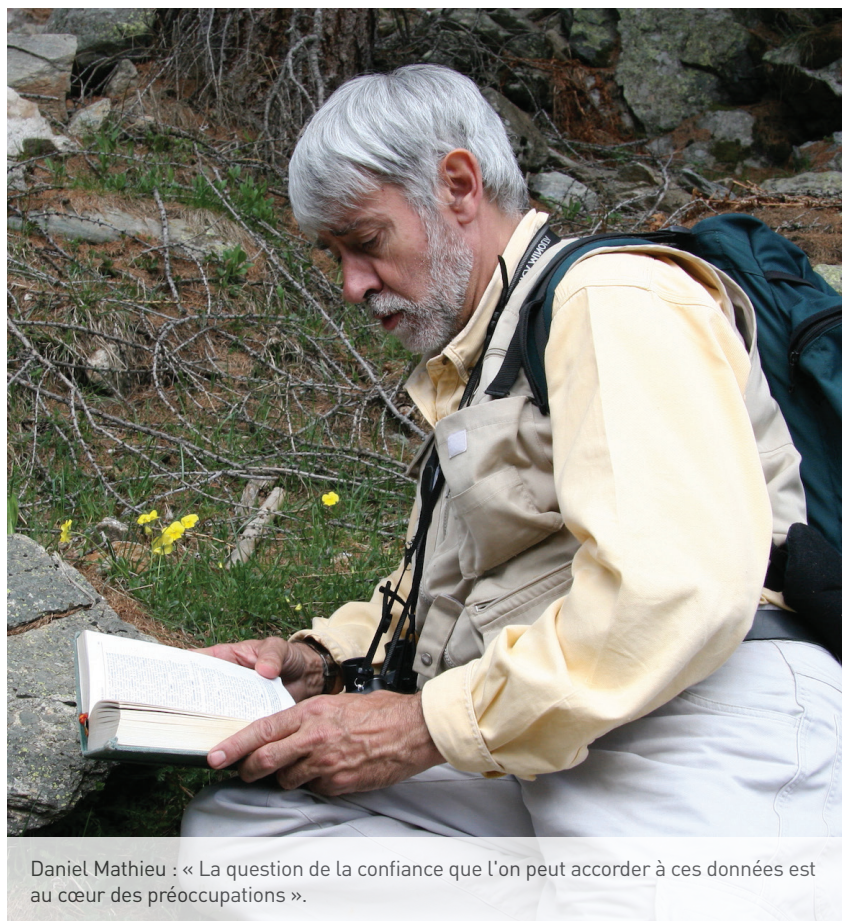


FIABILITÉ

Intégrer les données de réseaux amateurs

Des données d'amateurs ? Certains, comme Tela Botanica, y voient plutôt une richesse, d'autres pensent qu'intégrer des données ne venant pas d'experts peut polluer une base... Le traitement des données de sciences participatives ou de bénévoles pose en tout cas la question de la fiabilité.



Daniel Mathieu : « La question de la confiance que l'on peut accorder à ces données est au cœur des préoccupations ».

Le réseau Tela Botanica regroupe un panel important de botanistes de tous les niveaux qui fournissent des données dans le cadre de programmes de science participative et de relevés de terrain individuels. La question de la confiance que l'on peut accorder à ces données est au cœur des préoccupations de ceux qui les utilisent pour leurs travaux de recherche ou d'inventaires floristiques.

Deux cas se présentent pour les projets qui utilisent des données collectées par des non-spécialistes. D'une part des programmes de « science participative » qui sont issus de protocoles et conduits en partenariat avec des organismes de recherche, d'autre part des données non issues de protocoles fournies par des amateurs. Cette distinction est importante car elle conduit à des résultats et des traitements relativement différents.

Parmi les programmes animés par Tela Botanica disposant d'un protocole de collecte, l'Observatoire des saisons, conduit scientifiquement par le CNRS/CEFE, s'adresse à un très large public, notamment des scolaires. « Sauvages de ma rue » sensibilise également un large public, tout en collectant des don-

nées sur la biodiversité urbaine pour les chercheurs du MNHN. On peut citer aussi l'Observatoire des messicoles.

ADAPTER LE PROTOCOLE

Tous ces programmes de science participative sont de long terme et couvrent un très large espace géographique. La prise en compte, dans les protocoles, de l'incertitude liée au caractère amateur des contributeurs n'a pas posé à ce jour de réel problème en termes de fiabilité des données. Le nombre de plantes à identifier est relativement faible, ils sont accompagnés de guides de terrain et de fiches techniques limitant les erreurs d'identification ou d'interprétation et le service en ligne IdentiPlante permet de confirmer les déterminations à partir de photos. Concernant la collecte de données non issues de protocoles, la situation est différente. Prenons pour exemple la saisie des données réalisée dans le Carnet en ligne de Tela Botanica. Cette base de données collaborative cumule près d'un million d'observations et des centaines de milliers de photos de plantes. Que valent ces données amateurs ? Pour le savoir, le Conservatoire

botanique méditerranéen de Porquerolles a réalisé une étude comparative* des données du Carnet en Ligne de Tela Botanica avec celles de la base de donnée Silene en vue de leur intégration dans le SINP. Cette étude a permis de mettre en évidence l'apport non négligeable de ces données participatives à l'inventaire floristique du CBN-Med. Sur les quatre-vingt-mille observations enregistrées dans l'Hérault, 24 000 fourniraient un réel apport au SINP, car les observateurs amateurs et professionnels n'ont pas les mêmes centres d'intérêts, et sont donc complémentaires. Afin d'améliorer la fiabilité de ces données, une procédure est en train d'être mise en place avec le CBNMed pour filtrer automatiquement les observations et ne retenir que celles dont la qualité est conforme aux standards du SINP.

Reste la question de la gestion des données hors-normes. Enlever systématiquement les données qui paraissent aberrantes ? Et si c'était une information capitale de l'évolution de la flore ? • **Daniel Mathieu**, président de Tela Botanica

* Financée par la Dreal Paca.