

Cantharella, une base de données pour capitaliser les substances naturelles

Sylvain Petek et Adrien Cheype

Dès lors que l'on souhaite étudier les substances naturelles²⁴, quelle qu'en soit l'origine, on se retrouve très rapidement confronté à gérer un important volume de données, de différentes natures et origines, concernant :

- les sites de prélèvement : pays, localité, point GPS, inventaire des espèces, information sur le milieu/biotope, etc. ;
- l'identification des organismes (taxonomie) échantillonnés, leur abondance, leurs caractéristiques physiques, génétiques, etc. ;
- les protocoles chimiques mis en œuvre, les molécules identifiées ;
- les tests d'activités biologiques réalisés.

De plus, ces travaux sont par nature pluridisciplinaires et nécessitent l'intervention de nombreux collaborateurs spécialisés, souvent géographiquement éloignés.

Au final, seule une partie de ces informations figureront dans des articles et seront ainsi pérennisées. À la longue, il y a par conséquent un risque que les données « brutes » deviennent inexploitables ou disparaissent alors qu'elles peuvent constituer un historique et ainsi servir de base à de nouveaux projets.

Pour finir, au fil du temps, l'hétérogénéité des supports papier ou informatiques, des formats de fichier, ou dans la manière de structurer les données compliquent très rapidement la réutilisation des informations.

Cantharella (PETEK et CHEYPE, s.d.), une base de données dédiée à l'étude des substances naturelles a été conçue pour apporter une solution aux différentes problématiques posées par ces données, en termes :

- d'accès et de partage entre collaborateurs ou de restitution auprès des collectivités ;
- d'analyse et d'actualisation ;
- de pérennisation sur le long terme.

Cet outil de travail collaboratif, accessible en ligne, développé à partir de briques logicielles « libres », permet, au travers de quatre modules spécialisés, de capitaliser l'ensemble des données allant de la récolte des organismes en passant par les tests biologiques jusqu'aux molécules identifiées.

Par ailleurs, dans le cadre de la démarche APA (Accès et partage des avantages), l'outil permet également d'effectuer une restitution des travaux auprès des collectivités prospectées, et ainsi d'avoir un suivi des études menées sur leur biodiversité.

Pour les universités ou les laboratoires qui souhaiteraient l'utiliser, le logiciel est mis à disposition gratuitement sous licence libre.

Concernant l'instance de l'IRD, opérationnelle depuis 2010, elle capitalise les données de nombreux projets essentiellement dans le Pacifique (environ 700 stations de prélèvement, environ 950 espèces, plus de 7 700 résultats de tests biologiques).

Référence bibliographique

PETEK S., CHEYPE A. *Cantharella : Base de données pharmacochimique des substances naturelles* : (cantharella.ird.nc).



²⁴ Remerciements : financements IRD pour les programmes Spirales (DDUNI, Direction pour le développement des usages numériques innovants, IRD) et Maturation de projets innovants (SIV).

Petek Sylvain, Cheype A.

Cantharella, une base de données pour capitaliser les substances naturelles [encadré 26].

In : Payri Claude (ed.), Moatti Jean-Paul (pref.). Nouvelle-Calédonie : archipel de corail. Marseille (FRA), Nouméa : IRD, Solaris, 2018, p. 208.

ISBN 978-2-7099-2632-4