



# Normalisation des données

## sur les abeilles et l'apiculture

*Pouvez-vous me parler des objectifs du groupe de travail n°15 d'Apimondia sur la normalisation des données relatives aux abeilles et à l'apiculture ?*

On recueille plus de données sur les abeilles et l'apiculture que jamais auparavant. C'est une bonne chose. Cependant, la plupart des données sont collectées et enregistrées de différentes manières, ce qui rend difficile la fusion et l'harmonisation des ensembles de données. Cela limite notre capacité à tirer profit d'outils tels que l'apprentissage automatique (*machine learning*) qui nécessitent un grand nombre de données pour fonctionner efficacement. L'objectif de ce groupe de travail est de faire le travail nécessaire pour permettre le partage des données afin que ceux qui sont prêts à partager puissent le faire efficacement.

*Pourquoi est-il si important de partager les données ?*

Le partage des données permet d'appliquer la science des données à l'apiculture à grande échelle. L'apiculture est un passe-temps ou une profession pratiquée par beaucoup de merveilleuses personnes et d'organisations. La plupart de ces organisations sont relativement petites si l'on parle de science des données (*data science*). Dans ce domaine, on est plutôt habitués à des entreprises disposant de nombreux téraoctets ou exoctets de données.

En outre, les besoins en matière de science des données pour les techniques d'intelligence artificielle à appliquer à l'apiculture sont beaucoup plus importants que ce que nous constatons avoir dans de nombreuses applications commerciales. Cela s'explique par le fait que de nombreuses applications actuelles des grandes bases de données tournent autour de systèmes fermés ne comportant qu'une poignée de variables nécessaires pour fabriquer un produit ou gérer un client.

Cependant, lorsqu'il s'agit d'organismes vivants touchés par le climat, les maladies/ pathogènes, la génétique, les pratiques de gestion, etc., le problème est beaucoup plus complexe. Plus il y a de variables ayant un impact sur le problème, plus il faut de données à analyser efficacement. Par conséquent, vous avez besoin de plus de données que de nombreuses grandes applications de données actuelles et nous en avons actuellement beaucoup moins. La seule façon de tirer pleinement profit de la science des données dans le monde apicole est de partager les données, sinon nous n'en aurons jamais assez.

*Quels sont certains des avantages probables de l'application de la science des données à l'apiculture ?*

La plupart des lecteurs de cet article auront entendu parler d'une ruche intel-

ligente. Une ruche intelligente est une ruche qui peut vous informer de sa situation. Le poids, la température, l'humidité, les sons, les images et les compteurs d'abeilles peuvent vous en dire beaucoup sur ce qui se passe dans une ruche. C'est intéressant, mais pour être utile, vous devez ensuite savoir que faire de ces informations.

C'est là qu'intervient la science des données. La science des données consiste à donner un sens à de grandes quantités de données. Dans ce cas, elle nous aiderait à passer d'une ruche intelligente qui vous parle d'elle-même à une ruche géniale qui peut vous dire ce dont elle a besoin pour être meilleure. Avec suffisamment de données, nous pouvons identifier les meilleures pratiques de gestion personnalisées pour optimiser vos ruches et orienter les actions vers une

Groupe de travail BeeXML - Dernière réunion phy.



meilleure apiculture. Nous pouvons suivre les implications politiques en temps réel et prédire ce qui se passera ensuite, et bien plus encore.

Cela prendra du temps, mais si nous ne commençons pas dès maintenant à ouvrir la voie au partage des données, cela ne se produira jamais.

*Comment ce groupe de travail facilite-t-il le partage des données ?*

Le groupe de travail a identifié trois domaines principaux dans lesquels nous pouvons travailler pour permettre le partage des données. Ces domaines sont les suivants :

- Convenir de formats standard pour le stockage des données relatives aux abeilles et à l'apiculture et formuler des recommandations à cet égard.

- Encourager et publier des études d'harmonisation visant à trouver des moyens de mieux agréger de grands ensembles de données sur la mesure d'éléments similaires. Par exemple, harmoniser plusieurs méthodes de mesure du varroa afin que ces données, mesurées différemment, puissent être agrégées avec un certain degré de confiance statistique.

- Formuler des recommandations de meilleures pratiques en matière de collecte, de stockage ou de sécurisation des données afin de faciliter le partage.

*Je vous ai entendu utiliser le terme beeXML en référence à ce groupe, pouvez-vous expliquer ce terme ?*

La plupart de nos lecteurs connaissent le HTML (*Hyper Text Markup Language*) qui est un moyen de normaliser la présentation des données dans un navigateur web afin que les documents soient similaires quel que soit l'ordinateur ou le navigateur utilisé.

Le XML (*eXtensible Markup language*) est un jumeau du HTML qui se concentre sur la signification des données. En utilisant des balises définies par l'utilisateur similaires à celles du HTML, les groupes peuvent baliser leurs données et les partager grâce à cette technologie. Il s'agit d'un cadre et d'un langage permettant de créer une norme qui, une fois adoptée,

permet de partager facilement des données qui ont été utilisées des milliers de fois, du partage de documents à beeXML. C'est de là que vient le terme beeXML.

L'un des principaux avantages du XML par rapport aux autres langages standard est qu'il est lisible par l'homme. Nous voulions que le langage standard soit accessible à tous, ce qui donne au XML un net avantage sur les autres options.

*Comment les gens peuvent-ils aider ?*

Ils peuvent rejoindre le groupe de travail en m'envoyant un courriel à [cazierja@appstate.edu](mailto:cazierja@appstate.edu), adopter la norme dans leur organisation, donner leur avis sur les normes proposées ou faire un don à notre centre de recherche universitaire à [care.appstate.edu](http://care.appstate.edu) pour soutenir cet effort.

visique en Allemagne en 2019



#### MOTS CLÉS :

Big Data, informatique, ruche connectée, matériel

#### RÉSUMÉ :

les objectifs et perspectives du groupe de travail Bee XML sur les échanges de données appliqués à l'apiculture