

Contenu archivé le 2024-04-23

En vedette - Sauvegarder les écosystèmes avec des données et des écosystèmes d'infrastructures ouvertes électroniques

Prenez des cartes de la biodiversité et comparez-les à des registres de capture de poissons; vous verrez clairement là où les stocks de poissons sont en danger. Ceci pourrait permettre de sauver les océans de la planète, mais il faudra pour de faire traiter et analyser une immense quantité de données complexes. Des chercheurs financés par l'UE sont en phase de résoudre le problème grâce à une approche innovante et inspirée par la nature aux infrastructures électroniques (les e-infrastructures) et en explorant les façons d'intégrer les initiatives de données ouvertes.



Les infrastructures électroniques utilisent l'informatique en grille et en nuage pour exploiter les fonctions de stockage, de traitement et de logiciel d'une multitude de ressources distribuées. Une infrastructure électronique peut être développée par un groupe de chercheurs en biologie par exemple pour étudier un problème spécifique. L'infrastructure électronique permet aux biologistes de créer un environnement de

recherche virtuelle (ERV) pour la collaboration tout en exploitant les ressources de l'informatique en grille pour traiter les informations d'une source et les analyser avec des logiciels d'extraction des données à partir d'une autre source. Mais que se passerait-il si pendant leurs travaux, ces derniers souhaitaient comparer leurs données avec des informations d'autres chercheurs à l'aide de données, logiciel et systèmes informatiques différents, ou même avec des ressources publiques de données?

«L'intégration de ressources dans différentes infrastructures électroniques est très complexe et extrêmement longue, et dans plusieurs cas nécessite le développement d'une nouvelle infrastructure électronique qui est, elle aussi, longue et coûteuse», explique Donatella Castelli, chercheur à l'Institut des sciences et des technologies de l'information du Conseil de recherche national d'Italie.

Si ces différentes infrastructures électroniques existent dans un écosystème où, comme dans la nature, elles sont conscientes les unes des autres et peuvent coopérer ou même entrer en compétition, le partage des ressources entre elles devient encore plus simple et bon marché. C'est cette vision qui a mené un consortium d'universités, d'instituts de recherche, de compagnies et un organisme des Nations Unies à lancer le projet [D4Science-II](#) («Data infrastructures ecosystem for science»). Soutenu à hauteur de 4,3 millions d'euros de financement par la Commission européenne, le projet a créé un cadre interopérable pour les infrastructures électroniques (un écosystème d'infrastructure électronique dans lequel les données, les ressources informatiques et logicielles appartenant à différentes infrastructures électroniques peuvent être partagées indépendamment de leur situation géographique, de la technologie, du format, de la langue, du protocole ou du flux de travail).

L'interopérabilité entre les infrastructures électroniques dans l'écosystème des connaissances D4Science-II apparaît de deux façons: par l'utilisation de normes communes entre infrastructures électroniques et plus important, par les «cadres de médiation». Ces cadres de médiation sont des logiciels qui traduisent et transforment les données et processus hétérogènes de telle manière qu'elles peuvent être utilisées par différentes infrastructures électroniques, rendant ainsi la coopération possible. La base du système est gCube, un cadre logiciel modulable qui permet l'interopérabilité et qui a été soumis à des tests par le partenaire hongrois du projet, 4D SOFT.

L'infrastructure électronique D4Science gère non seulement des ressources et les rend interopérables, mais les met également à disposition d'autres infrastructures électroniques, leur permettant d'accéder de manière dynamique aux données, aux logiciels et à la puissance informatique.

«Dans ce sens, les infrastructures électroniques dans l'écosystème peuvent être compétitive. Les chercheurs peuvent choisir les ressources disponibles qui leur conviennent le mieux à un moment donné», explique le Dr Castelli.

La force d'une telle approche se remarque dans les ERV et dans les applications gCube (ERV en accès ouvert) mis en place dans le projet D4Science-II et disponibles sur le [portal de D4Science](#).

«D4Science-II trouve son origine dans deux projets antérieurs, DILIGENT et

D4Science, projets pionniers dans le développement d'infrastructures pour les bibliothèques numériques reposant sur une infrastructure électronique à grille. Cependant, nous avons constaté que beaucoup d'infrastructures électroniques existent déjà pour des buts précis et avons réalisé qu'il vaut mieux utiliser les ressources dont elles disposent en collaboration plutôt que de reconstruire de nouvelles infrastructures électroniques à chaque fois. Notre centre d'intérêt dans D4Science-II est donc passé de l'habilitation des infrastructures électroniques à la construction d'un écosystème d'infrastructures électroniques», ajoute le Dr Castelli.

De la biodiversité et la pêche à la physique des hautes énergies...

L'écosystème est depuis longtemps utilisé pour soutenir les ERV dans des domaines tels que la physique des hautes énergies, la biodiversité, la pêche et les ressources en aquaculture. Il a permis d'ouvrir de nouveaux domaines de recherche entre eux et s'étend maintenant à de nouveaux domaines.

AquaMaps, un projet visant à créer des cartes de distribution mondiale des espèces du monde marin, utilise les ressources des infrastructures électroniques des données et à grille par le biais d'un ERV défini sur l'infrastructure de D4Science.

La génération de cartes haute résolution montrant la distribution des espèces de poissons est une tâche intense sur le plan computationnel: concevoir une carte multi-espèces nécessite 125 millions de calculs. Sans infrastructure électronique à grille, la génération de la collection de cartes nécessaire pour soutenir une activité de recherche pourrait prendre plusieurs jours; avec l'informatique à grille, il suffit de quelques heures. Au sein de l'écosystème D4Science, trois ERV distincte bien que liés travaillant sur des données sur la pêche ont pu utiliser les informations et les ressources fournies par différentes infrastructures électroniques différentes (GENESI-DEC pour les données d'observation de la planète, GBIF pour les données sur la biodiversité et FIGIS pour les informations sur la pêche). Avec cette capacité, ils ont pu mener des procédures d'analyse statistique innovante qui étaient jusqu'à présent tout simplement impossibles, associant des informations sur les espèces de poissons et les situations géographiques des captures à des données environnementales et géospatiales par exemple.

«Nous rassemblons des statistiques de toutes sortes d'exploitations piscicoles de plusieurs pays et d'une grande diversité en terme de qualité des données. D4Science nous a aidé à rassembler toutes ces données», explique Anton Ellenbroek du département des pêches et de l'aquaculture du FAO à Rome, en Italie. «Il s'agit d'une infrastructure réellement importante...elle nous permet d'analyser les statistiques de manière encore impossible jusqu'à présent et de partager d'autres environnements de recherche virtuelle.»

La FAO (Organisation des Nations Unies pour l'alimentation et l'agriculture) a

organisé un atelier dans le cadre du projet «Digital Repositories - Linked Open Data» afin d'examiner les solutions possibles pour la publication des référentiels numériques en tant que données ouvertes à l'aide d'outils avancés tels que les ERV de D4Science.

Le succès des ERV traitant des exploitations piscicoles et de la biodiversité dans D4Science-II a inspiré deux projets de suivi dans ce domaine.

Dans i-Marine, les chercheurs appliquent une approche écosystémique à la gestion des exploitations piscicoles et la conservation de l'environnement marin à l'aide d'une plateforme ouverte basée sur l'infrastructure D4Science pour travailler avec une série de connaissances et de sources de données bien plus vastes que celles utilisées dans la gestion traditionnelle des exploitations piscicoles.

Et dans le projet [EUBrazilOpenBio](#) («EU-Brazil open data and cloud computing e-Infrastructure for biodiversity»), des chercheurs européens et brésiliens utilisent l'approche écosystémique aux infrastructures électroniques pour créer une plateforme en accès ouvert intégrant des infrastructures électroniques européennes et brésiliennes ainsi que des ressources pour la science de la biodiversité.

«La coopération au sein des infrastructures électroniques ouvre de toutes nouvelles possibilités et domaines de recherche. Nous pouvons analyser des données scientifiques par rapport aux statistiques économiques par exemple, pour développer une perspective entièrement nouvelle qui n'existait pas jusqu'à présent», explique le Dr Castelli.

Liens utiles:

- [Site web du projet «Data infrastructures ecosystem for science»](#)
- [Fiche du projet D4Science-II sur CORDIS](#)
- [Site web du projet «EU-Brazil open data and cloud computing e-Infrastructure for biodiversity»](#)
- [Fiche du projet EUBrazilOpenBio sur CORDIS](#)
- [Discours sur les données de la Commissaire européenne Neelie Kroes](#)

Articles connexes:

- [D4Science-II contribue à la révolution des ressources électroniques en sciences](#)
- [Les travaux sur l'infrastructure de grille paneuropéenne passent au niveau supérieur](#)
- [Le calcul distribué pour lutter contre la maladie d'Alzheimer](#)
- [EELA-2 élargit l'utilisation de la grille de calcul mise en place en Amérique latine par le projet EELA de l'UE](#)

Projets connexes



ARCHIVED

Data Infrastructure Ecosystem for Science

D4Science-II

2 Août 2019

PROJET

Cet article apparaît dans...

MAGAZINE RESEARCH*EU



À la loupe: moteurs de recherche, données et aventure sémantique

Dernière mise à jour: 27 Janvier 2012

Permalink: <https://cordis.europa.eu/article/id/87946-feature-stories-saving-ecosystems-with-open-data-and-einfrastructure-ecosystems/fr>

European Union, 2025