



Données et algorithmes en recherche publique, jusqu'où (est-on et) peut-on aller avec geOrchestra?



GEOCOM – Perros-Guirec - 16 mai 2022 Hervé Squividant

> Plan

- Enjeux
- L'écosystème « Open Data » Recherche
- Algorithmes & appui aux politiques publiques
- Conclusion / Perspectives



Enjeux



Enjeux intra et extra-académiques de la publication des données de la recherche

Au sein de la communauté « Recherche » :

- ➤ Capitalisation des productions numériques → Pérennité
- Valorisation scientifique : 1 data paper = 1 publication
- Pari : Ouvrir le champ des possibles en recherche

En dehors du monde académique :

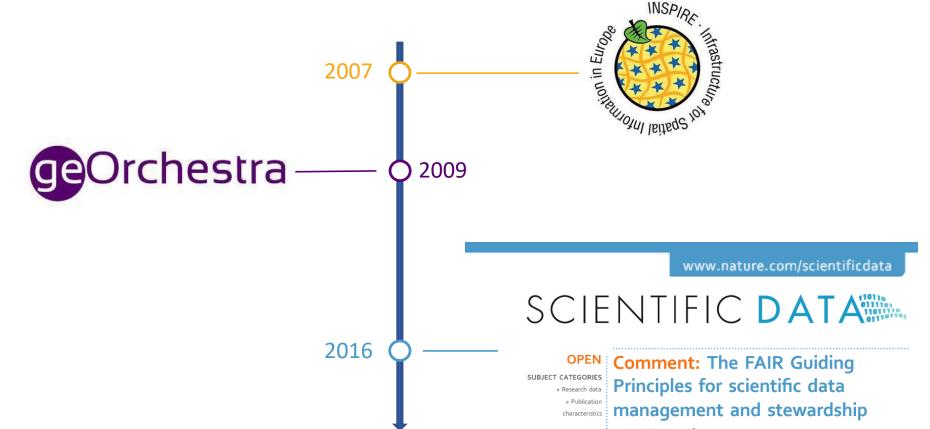
- Rôle sociétale de la recherche
- ➤ Avènement de l'Open Science → FAIR data



L'écosystème « Open Data » Recherche



Naissance d'un écosystème



- La directive INSPIRE, la bonne afFAIRe!*
- geOrchestra, pierre angulaire de l'écosystème

Received: 10 December 2015 Accepted: 12 February 2016 Published: 15 March 2016 Mark D. Wilkinson et al.*

There is an urgent need to improve the infrastructure supporting the reuse of scholarly data. A diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers—have come together to design and jointly endorse a concise and measureable set of principles that we refer to as the FAIR Data Principles. The intent is that these may act as a guideline for those wishing to enhance the reusability of their data holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting its reuse by individuals. This Comment is the first

formal publication of the FAIR Principles, and includes the rationale behind them, and some exemplar



implementations in the community.

> Un Ecosystème IDG « Recherche » à maturité





... Extension à d'autres services de données

- Application de la philosophie « libre, interopérable et modulaire »
- Mobilisation de certaines briques de geOrchestra (Sécurité, annuaire, catalogue ...)
- Développement et déploiement de nouveaux services de données interopérables autonomes
- → Data as a Service (DaaS)



SensorThings **OGC API**

Données temporelles

-> Observatoires

Services spécifiques au TYPE de la donnée

WPS OGC API SOL

Données pédologique

-> DONESOL RRP. UCS. UTS, propriétés, arbre de décision....

WPS OGC API Histoire Culturale

Occupation du Sol

-> RPG 2006 à 2019. reconstitution de l'H.C. nomenclature, requêtes spatiales

Autres Services de données

Autres TYPES, autres THEMES...

Ex: Données SHS,

Services spécifiques au THEME de la donnée



Algorythmes & appui aux politiques publiques

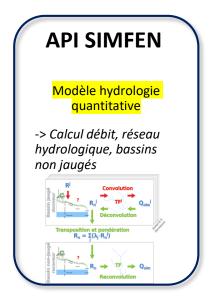


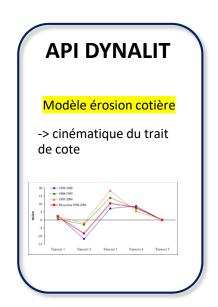
... vers une approche « services » pour la modélisation

- Application de la philosophie « libre, interopérable et modulaire » aux calculs, aux modèles produits par la recherche en environnement
- Développement et déploiement de nouveaux services de « modélisation élémentaire » sous forme d'APIs
- → Software as a Service (SaaS)

Quelques exemples entre terre et mer :

API MNTSurf Modéle hydrologie spatiale -> MNT, modèle de drainage, variables dérivées, indice de Beven...





Autres
Services de
calcul

Autres modèles : météo,
agronomie, pédologie,
hydrologie, ...



> Enjeux sociétaux et politiques publiques

- De nouvelles possibilités pour les outils de diagnostic et d'aide à la décision sur des enjeux sociétaux :
 - Réchauffement climatique (Submersion marine, Sécheresse, ...)
 - Environnement (Marées vertes, Bilan d'Azote,)
- Mobilisant les services numériques préexistants :
 - De **données** (géographiques ou pas) émanant des plateformes nationales thématiques, régionales et scientifiques,
 - de traitements, calculs, de modélisation,





LEGO de Services



Exemple appliqué à la lutte contre les marées vertes

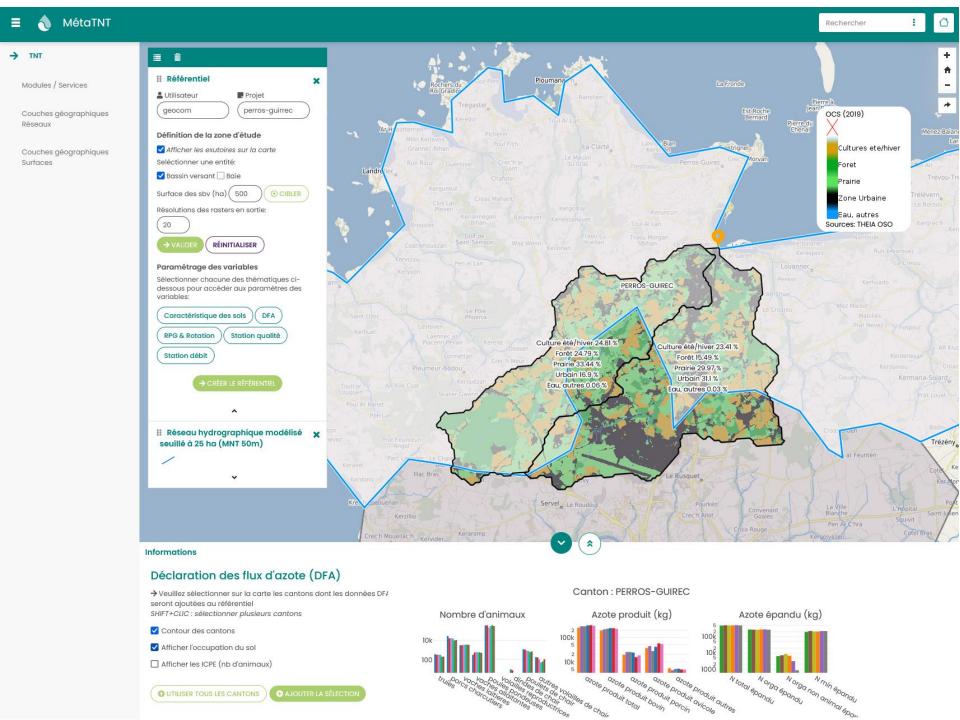
Projet MétaTNT²

- Objectif: Lutte contre les algues, Modélisation de l'azote, Scénarios prospectifs
- Financement : Etat PITE
- Services tiers mobilisés :
 - Données: Hub'Eau (débit, qualité des cours d'eau), Agri4cast (météo), GéoBretagne (Cultures/RPG, Déclaration Flux d'Azote), GéoSAS (Sol, Altimétrie, cours d'eau)
 - Calculs/Modèles : API MNTSurf (BV, Variables hydro), API SIMFEN (Débit)
- Développement et déploiement
 - d'une nouvelle API de modélisation dédiée = Service OGC WPS
 - d'une interface cartographique dédiée (Mviewer + extensions)

« Ce que permet MétaTNT² en quelques minutes, il nous fallait des mois pour le réaliser! »

Patrick Durand, hydrologue INRAE.





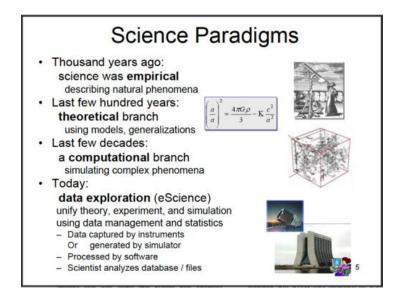
Conclusion

- > Naissance d'un écosystème de services numériques pour la science ouverte :
 - > pour le partages des données et des algorithmes produits par la recherche,
 - > Parfois concomitante mais le plus souvent précédée par le mouvement INSPIRE,
 - ... inspiré par les choix techniques et philosophiques de geOrchestra.
- Cet écosystème s'entremêle avec, complète et se nourrit des écosystèmes de plateformes des « administrations »,
- → Pour former un écosystème de plateformes et de services numériques publiques.
- De nouvelles possibilités et des économies d'échelle ENORMES pour le diagnostic et l'aide à la décision



Perspectives

4 paradigmes...



- ➤ 2 approches : Big data VS circuit court de la donnée
- > et 1 nouvelle question de recherche :

L'ouverture des données et des modèles contribue-t-elle à ou modifie-t-elle la recherche? *



