

# geOrchestra, 10 ans après ! Modularité, Interopérabilité & Subsidiarité

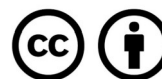
Une Infrastructure de Données Spatiales socle de la construction d'un écosystème de services numériques pour la science ouverte

Hervé Squvidant & Rodéric Béra – JDEV2020 – 9 juillet 2020

UMR SAS / Agrocampus Ouest – INRAE / Rennes - Quimper



CC BY 4.0





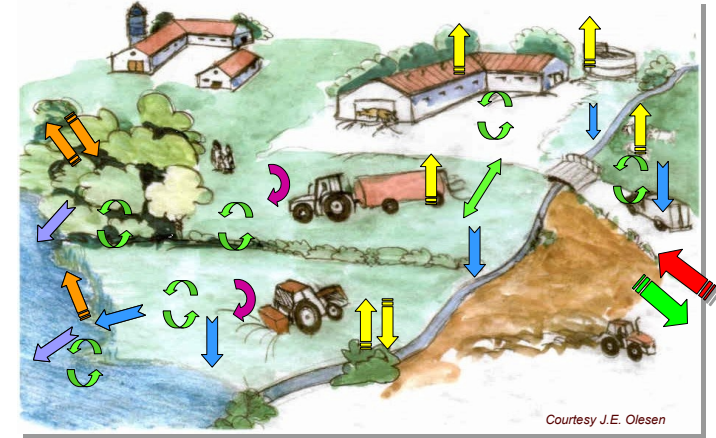
# Sommaire

- Contexte scientifique & historique
- Le choix de l'IDS geOrchestra
- GéoSAS une instantiation de geOrchestra
- Et le FAIR dans tout ça ?
- Perspectives / Conclusion

## ➤ Le contexte scientifique

# Contexte scientifique : thèmes ou objets de recherche à l'UMR SAS

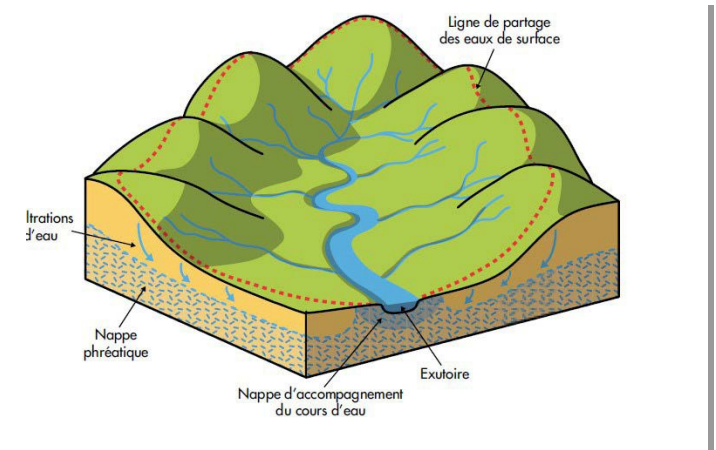
Systèmes Agricoles – Agronomie →



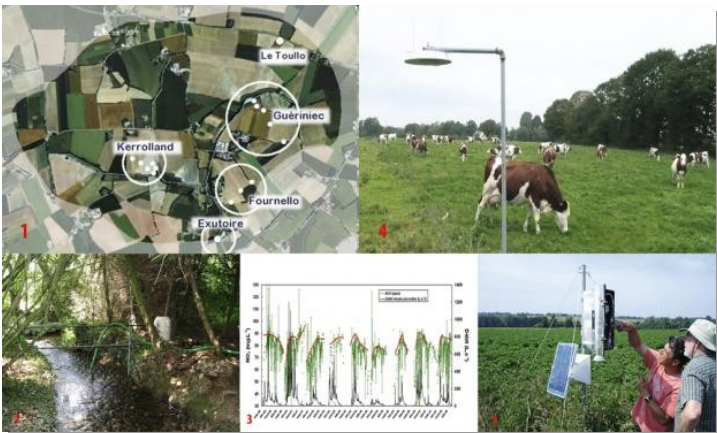
← Science du Sol – Pédologie



Hydrologie quantitative & qualitative →

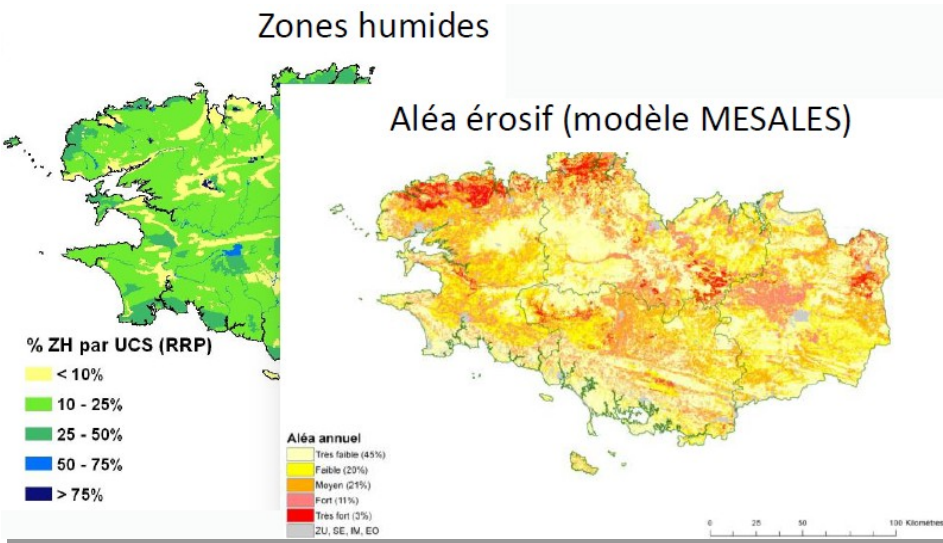
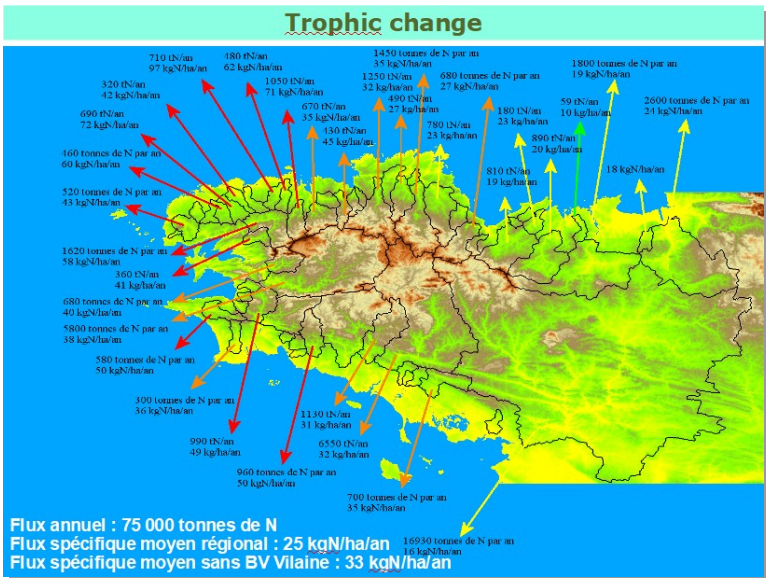


# Contexte scientifique : Données produites par l'UMR SAS



← Observatoires labellisés (ORE – SOERE PRO)

Référentiels pédologiques →



← Sorties de la modélisation hydrologique

## ➤ Le contexte historique

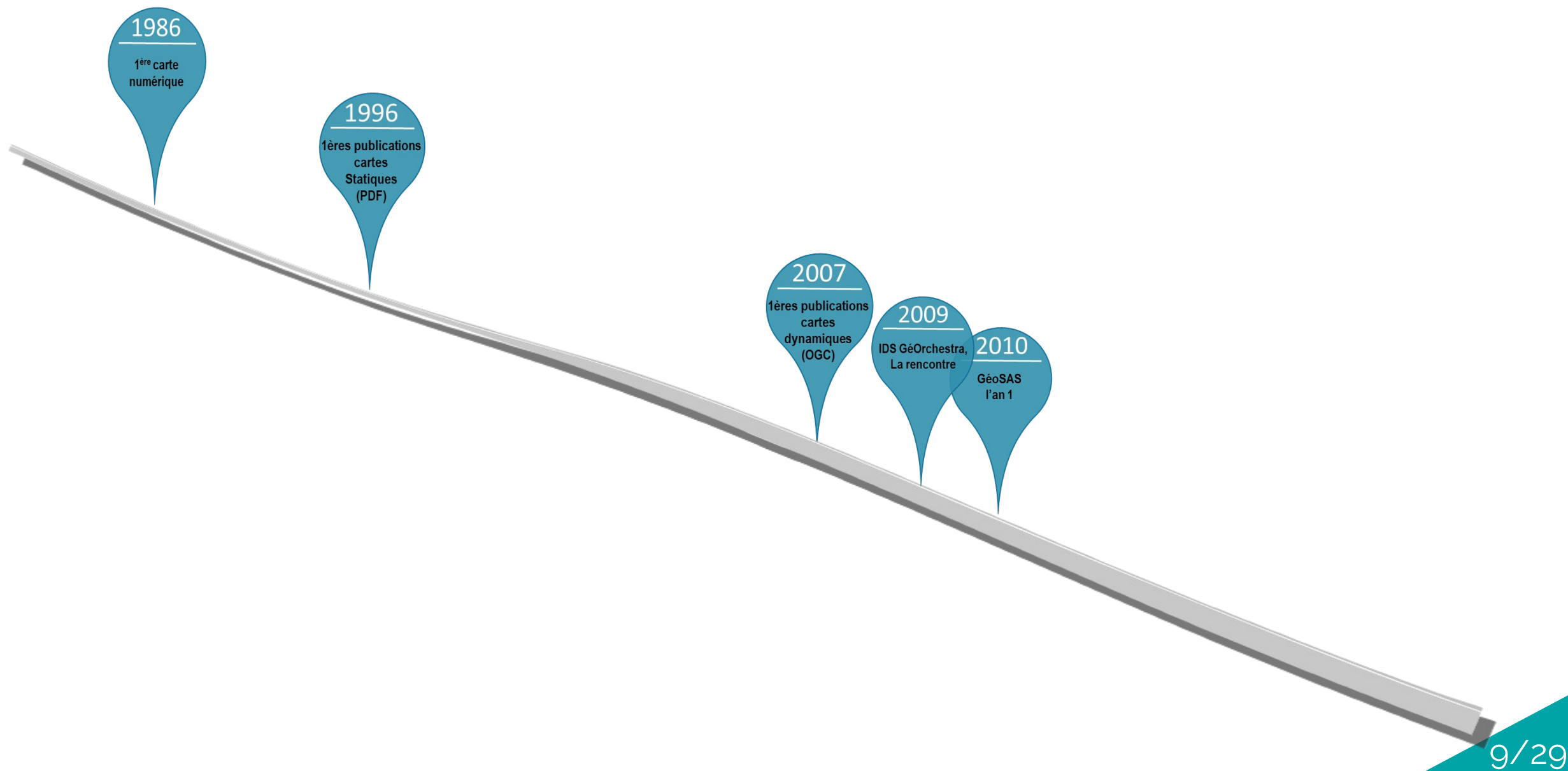
# Evolution du contexte institutionnel, législatif, politique



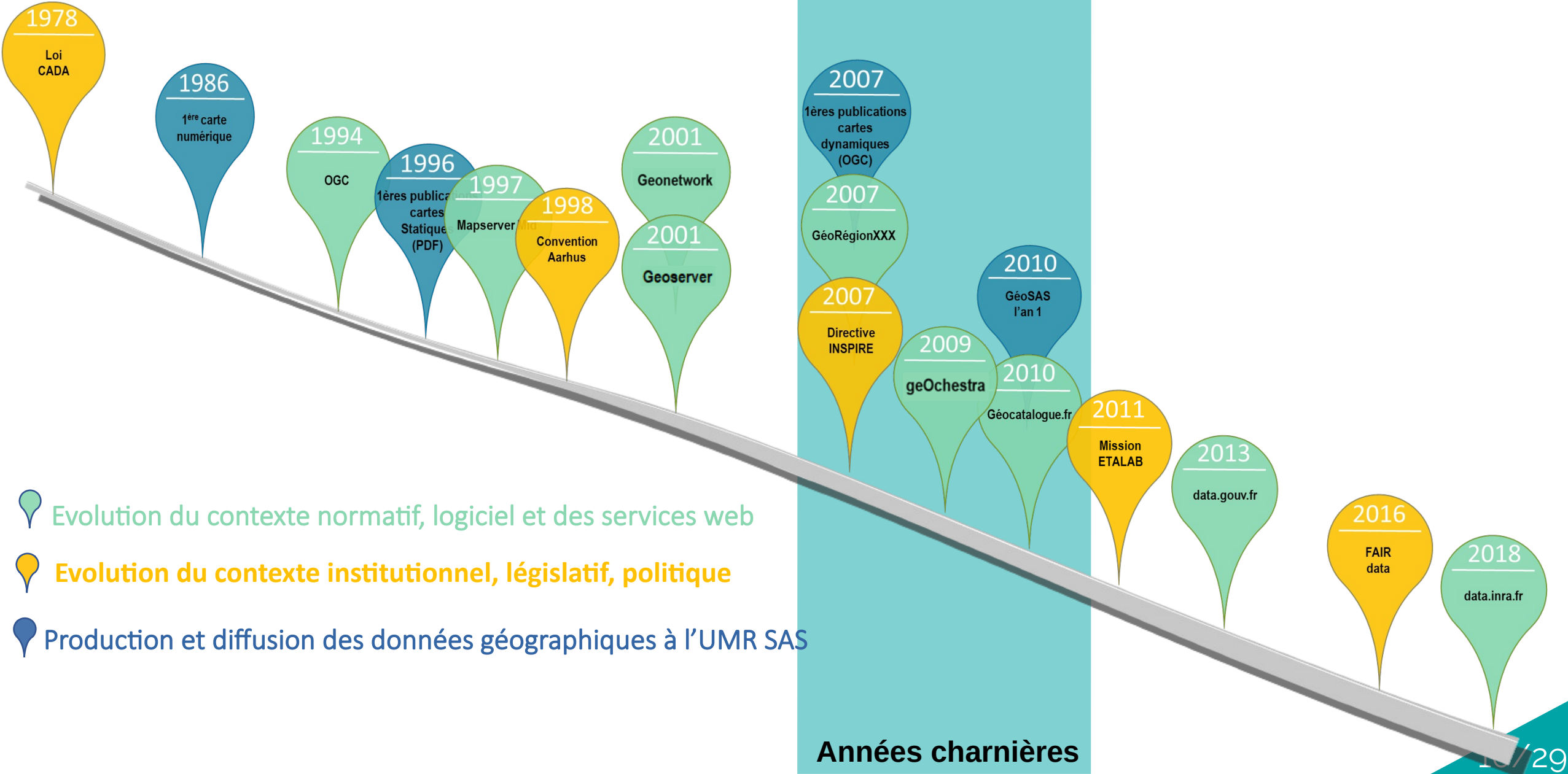
# Evolution du contexte normatif, logiciel et des services web



# Contexte historique : de la production de données géographiques à leur diffusion à l'UMR SAS



# Convergence du contexte institutionnel et technique avec les besoins croissants de diffusion des données de l'UMR SAS



## ➤ Le choix de l'IDS geOrchestra

# Le choix de l'Infrastructure de Données Spatiales geOrchestra

## Un hasard

- Rencontre avec l'équipe GéoBretagne en 2009
- 1<sup>ers</sup> échanges autour de la directive INSPIRE

## Une communauté

- GEOCOM : Rencontre annuelle des utilisateurs, développeurs, administrateurs, ...
- Mailing lists :
  - [georchestra@googlegroups.com](mailto:georchestra@googlegroups.com)
  - [georchestra-dev@googlegroups.com](mailto:georchestra-dev@googlegroups.com)
- Code source :
  - [Github.com/georchestra](https://github.com/georchestra)
- PSC : Comité de pilotage

## Des principes

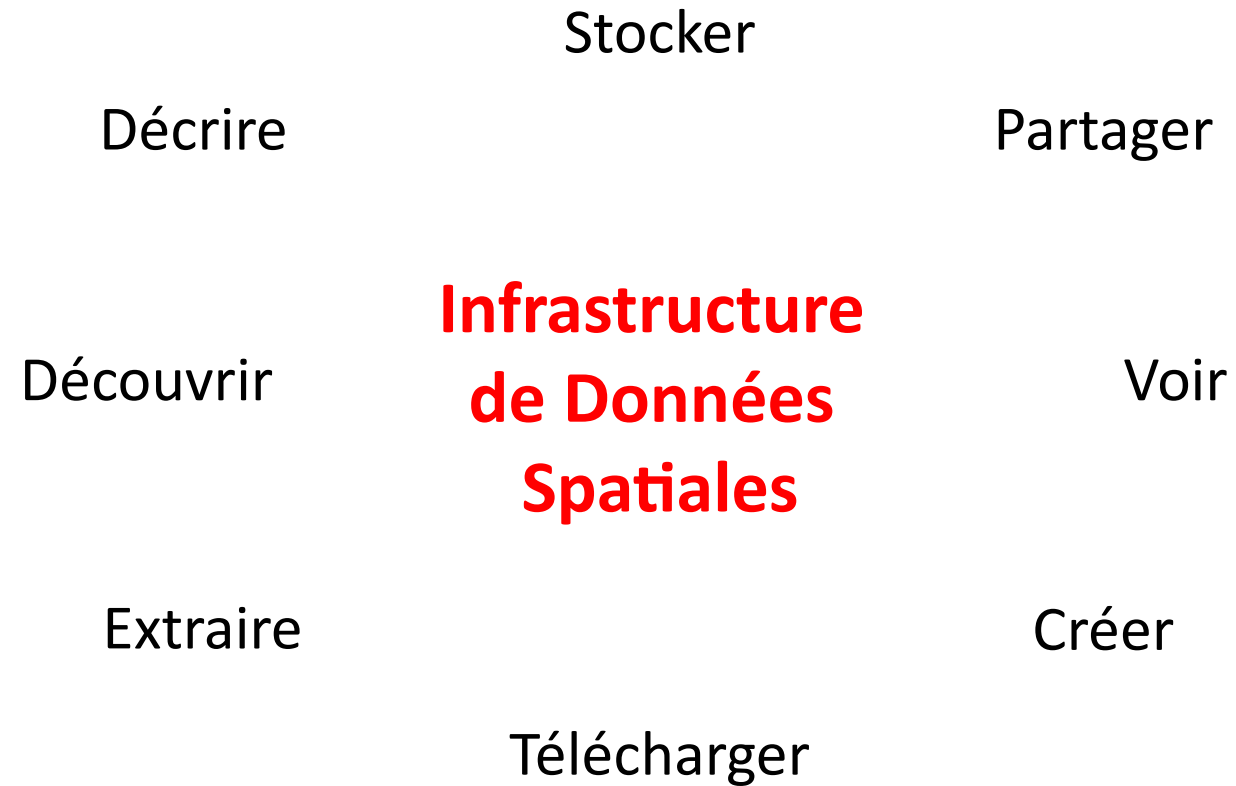
- Libre – Licence GPL
- Modulaire – Plus de 10 modules disponibles
- Interopérable – Service OGC + APIs REST
- Sécurisé – Support https + MAJ continue

## Un certain succès

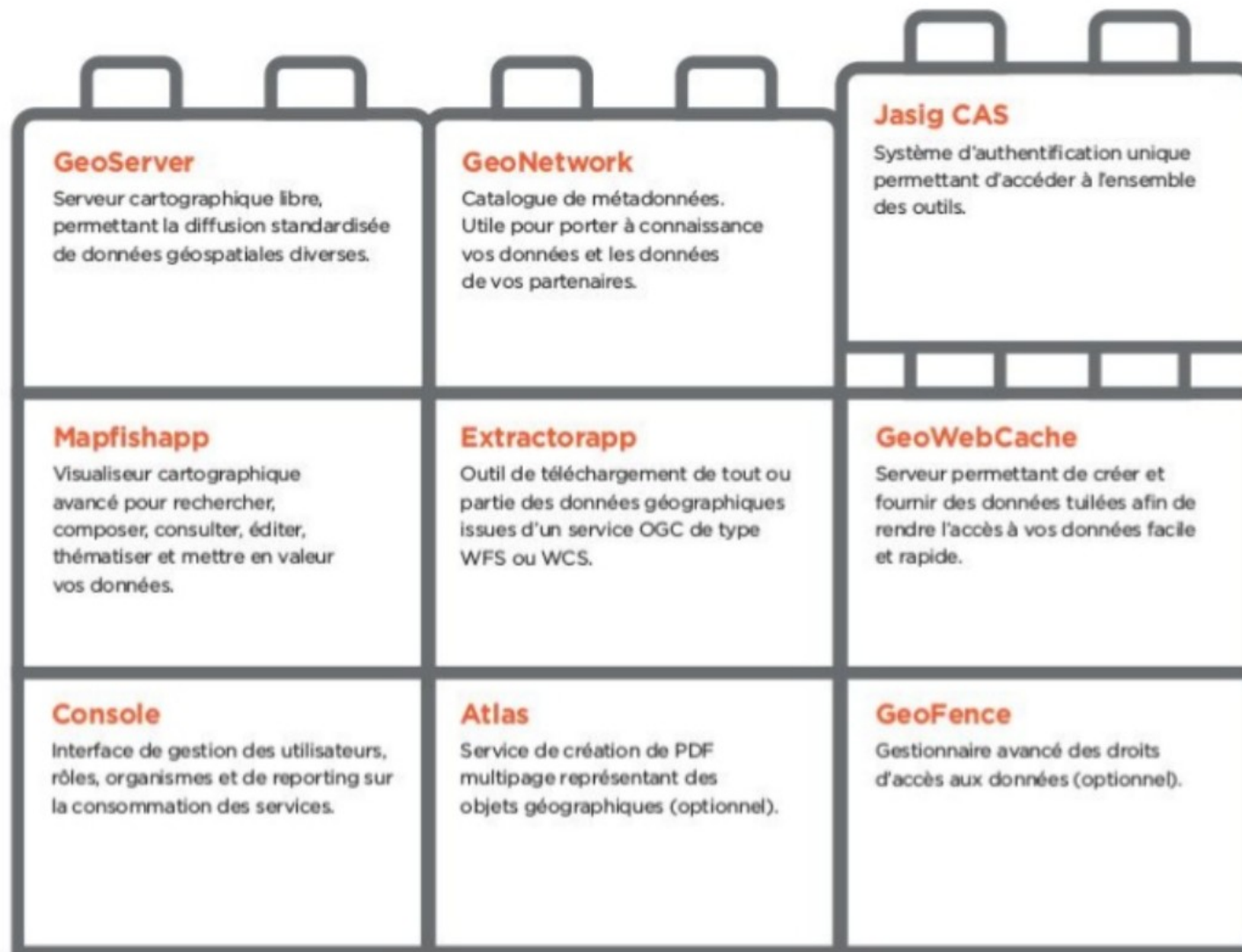
- Des états
- Des régions
- Des agglos
- Des organismes ou unités de recherche
- Des ONG
- Des entreprises



# Une IDS, pour quoi faire ?



# geOrchestra, une IDS modulaire, à la carte

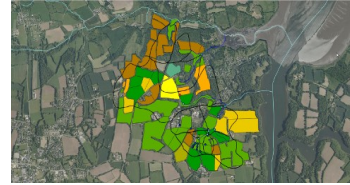


➤ GéoSAS, une instanciación de geOrchestra

# GéoSAS, l'IDS de l'UMR SAS

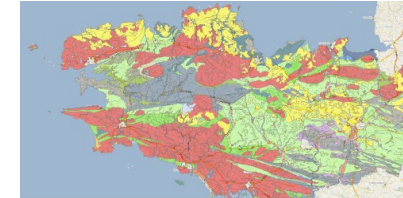
## Mise en production en 2010

- Publication des 1ères données géographiques et métadonnées

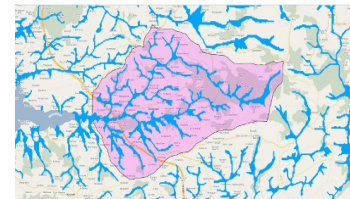


Observatoires labélisés (ORE - SOERE PRO)

Référentiel pédologique régional



Modélisation hydrologique



## Bilan en 2020

- 10 projets de recherche
- 5 observatoires
- 2 services de modélisation
- Une centaine de jeux de données ouverts
- 11 000 utilisateurs / an et 1700 utilisateurs / mois
- 7 000 000 requêtes / an et 600 000 / mois
- 15 000 téléchargements / an



# Choix et stratégie autour de GéoSAS

- Positionnement de l'IDS comme un maillon d'un écosystème multi-institutionnel, multi-thématique, multi-échelle géographique.
  - Le pari de la subsidiarité
- Démarche pro-active dans différentes communautés :
  - En interne à l'UMR SAS
  - Communauté « Recherche », institutionnelle et thématique
  - Communauté geOrchestra
  - Communauté régionale - GéoBretagne
  - Enseignement : Master TELENVI Rennes 2 - Agrocampus, Moodle Agrocampus IWDS
- L'IDS comme base de R & D pour bâtir un écosystème de services numériques dans une démarche d'« Open Science »

# Le pari de la subsidiarité

- 2 grandes options possibles dans le positionnement de l'IDS

Guichet unique d'accès  
aux données

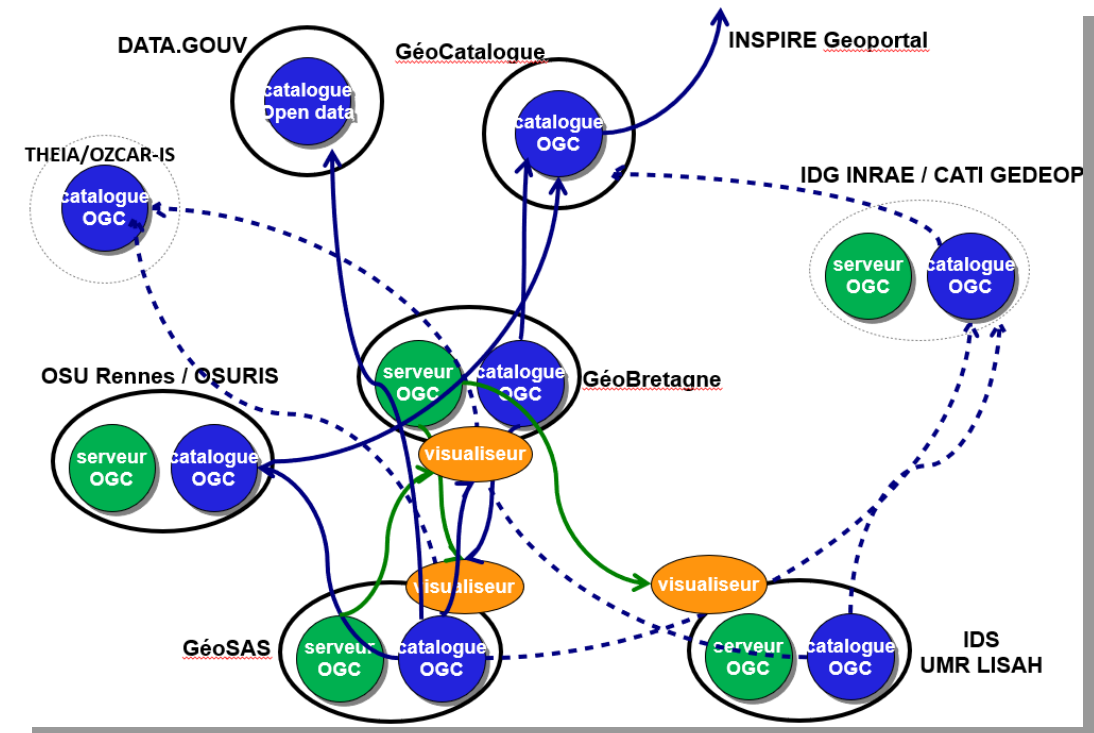
Maillon d'un écosystème de services



- Choix de l'éco-système multi-institutionnel, multi-thématique, multi-échelle géographique reposant sur le principe de **subsidiarité**.

Wikipédia : Le **principe de subsidiarité** est une maxime politique et sociale selon laquelle la responsabilité d'une action publique, lorsqu'elle est nécessaire, revient à l'entité compétente la plus proche de ceux qui sont directement concernés par cette action.

Application à l'Open Data : La donnée est publiée au plus près de son producteur.



- MAJ simplifiée
- Base à l'innovation & Socle de l'écosystème interne à l'UMR
- Acteur de l'écosystème global
- Support de formation / Evangélisation

- Cout humain/infra
- Approche horizontale au sein de tutelles adeptes de la verticalité

# Démarche pro-active en interne à l'UMR & boucle de l'Open Science

➤ Question : Comment inciter et accompagner les collègues de l'UMR à publier les « connaissances numériques » qu'ils produisent ?

## 1 Sensibilisation à L'Open Science

- Valorisation de la recherche par la publication des données (data paper, ...),
- Rôle sociétal par la mise à disposition de modèles ou d'outils d'aide à la décision,
- Open Science comme Thème de recherche en soit
- Notion de service public
- ...

## 2 Lancement du projet

- R.H. dédiées au projet
- Co-financement de l'infra
- Question de l'externalisation ?



## 3 Conception

- Des données,
- Chaîne de traitement automatisée
- SI dédié au projet, (BdD, ...)
- Développement de service de modélisation (scripts, API, ...)

## 6 Reporting

- Reporting sur l'utilisation des services du projet
- A destination de l'équipe projet, des financeurs, ...

## 7 Valorisation scientifique

- Publication de data papers
- Publication d'article scientifique "thématico-informatique". Ex : Hydro-informatique...

## 4 Publication

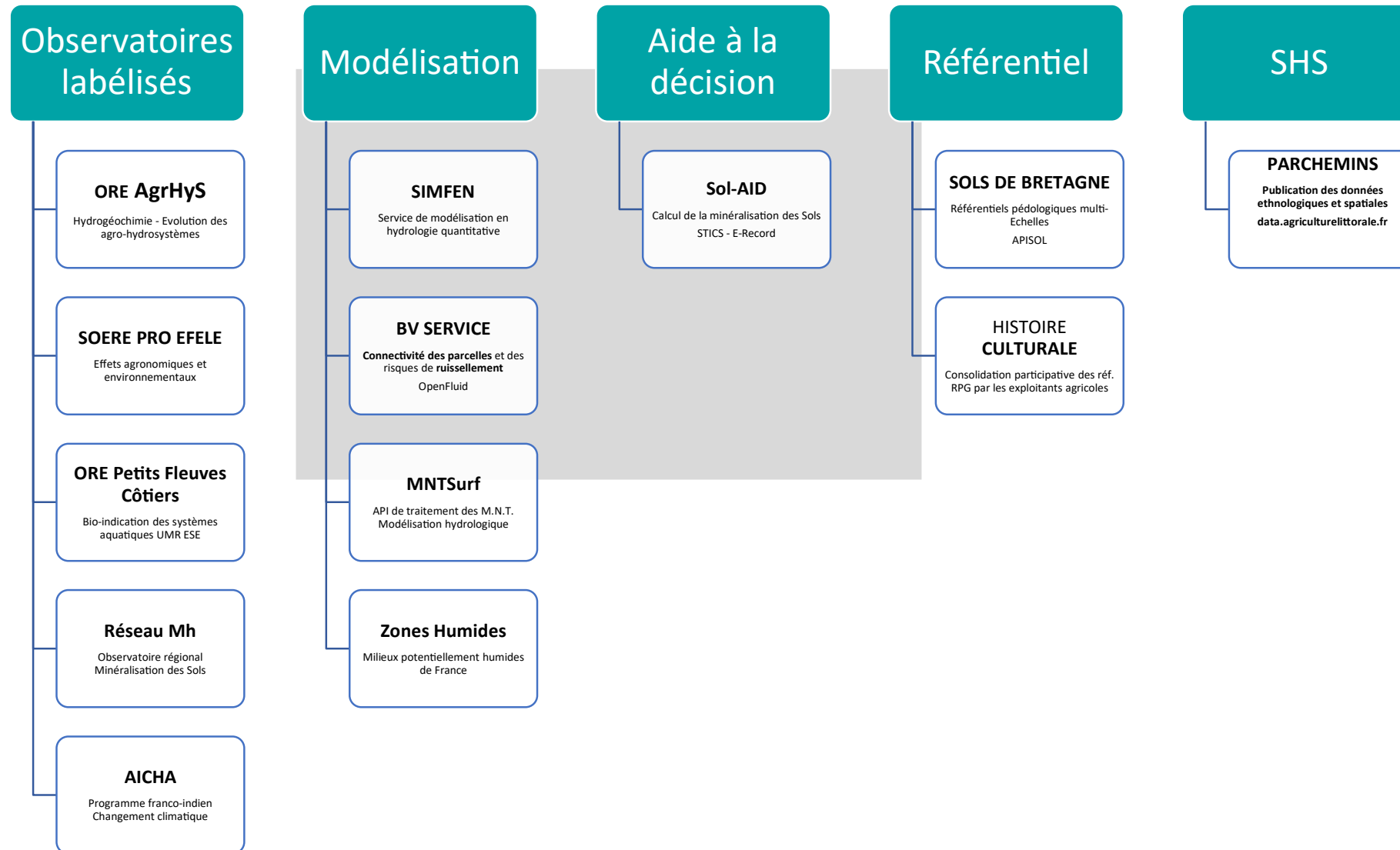
- Des données spatiales et des métadonnées via l'IDS
- Déploiement API de données
- Déploiement API de modélisation
- Métadonnée de service dédiée au projet

## 5 Réutilisation

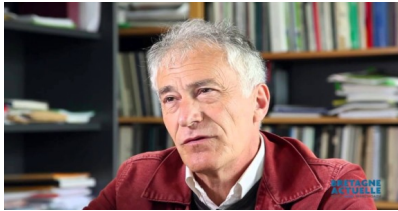
- Développement du portail dédié au projet
- Design de service – externalisation
- Favoriser, encourager tout autre réutilisation
- Réutilisations inconnues...



# Les projets modulaires, ouverts etinteropérables de GéoSAS



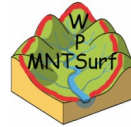
# Exemple n°1 : Sorties de modèle



*Recherche Hydrologique*

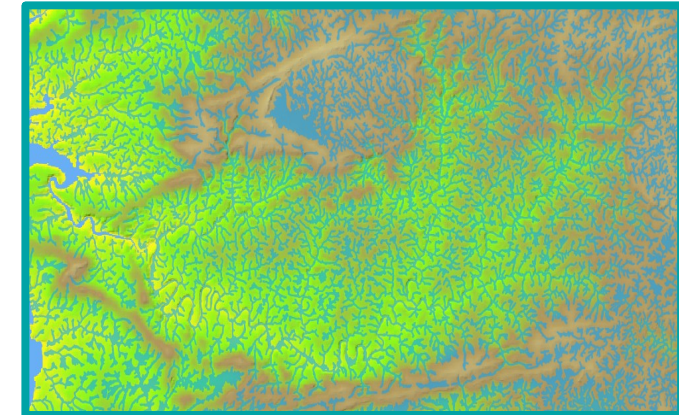
Merot, et al 2003. Testing a climato-topographic index for predicting wetlands distribution along an European climate gradient. Ecological Modelling.

## Conception



*Développement géo-informatique*  
*Logiciel MNTSurf*

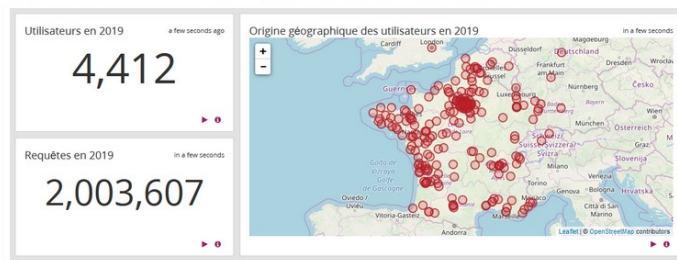
**Cartes vecteur des zones humides potentielles**



## Reporting

- Reporting sur l'utilisation des services du projet à destination des scientifiques

Rapport d'utilisation de la couche « Milieux potentiellement Humides » pour l'année 2019



## Réutilisations

- Portail dédié au projet



- Autres réutilisations ++
  - [reseau-zones-humides.org](http://reseau-zones-humides.org)
  - IDS Ministère de l'Ecologie
  - GéoBretagne

## Publication

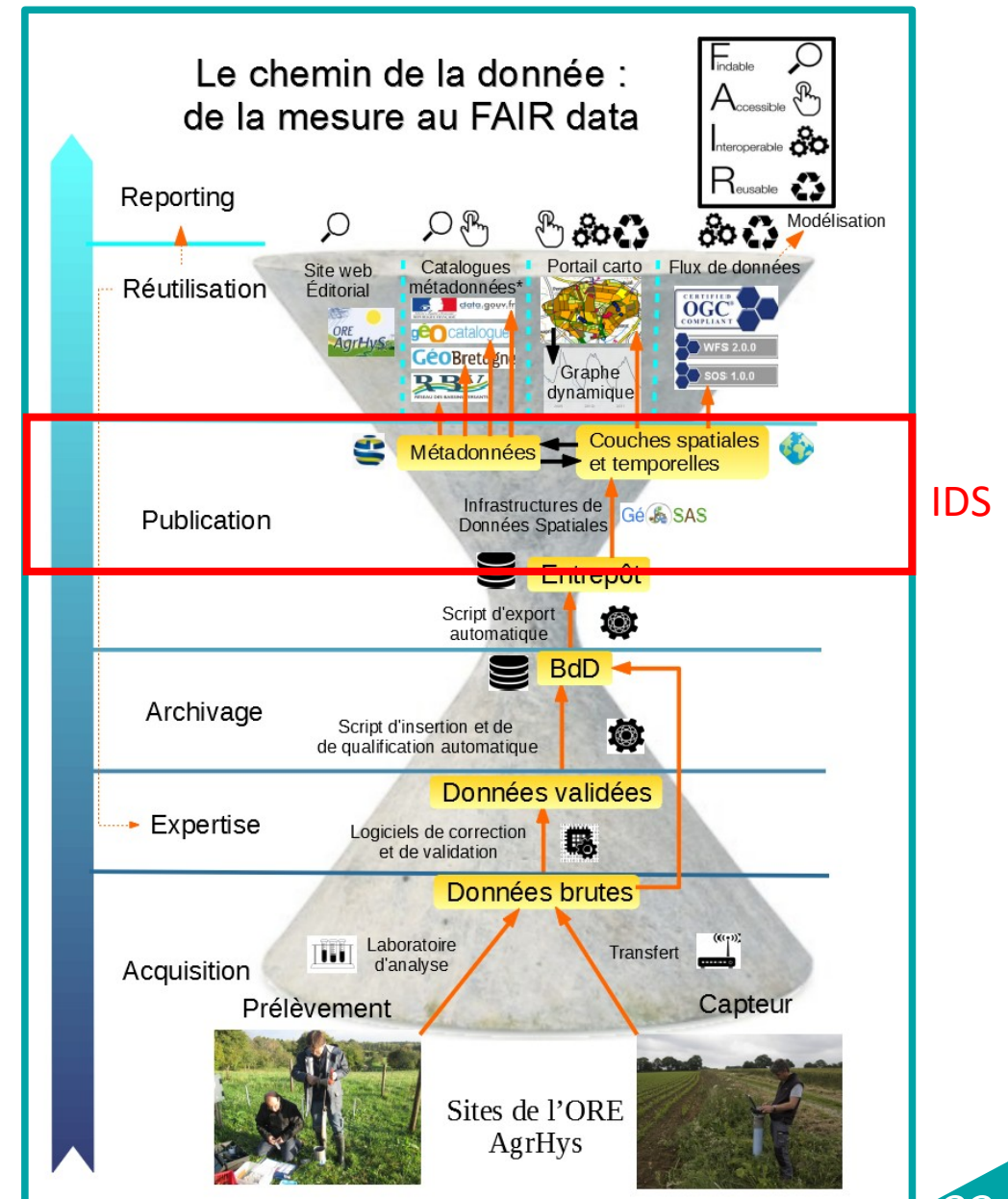
- Des données spatiales et des métadonnées via l'IDS



# Exemple n°2 : L'IDS comme maillon du SI des Observatoires

Dispositifs d'acquisition et de publication des données des observatoires ORE AgrHyS, SOERE PRO EFELE, ORE PFC → Approche générique

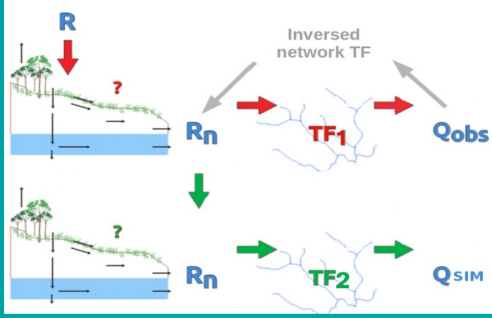
- Extension des capacités de l'IDS aux données temporelles → Standards OGC SOS & WFS
- Interface générique aux Time Series (Highcharts)
- Evolution vers les API GraphQL
- **Geneviève Le Hénaff et Al. 2018. De la mesure environnementale à sa diffusion : mise en place d'une chaîne de traitement modulaire et générique pour les données de l'Observatoire de Recherche en Environnement AgrHyS.** n° spécial du Cahier des Techniques de l'INRA consacré aux Données de la recherche à l'INRA
- [Page web du projet](#)



# Exemple n°3 : Projet SIMFEN - Modélisation ouverte en hydrologie

## Recherche fondamentale

- **20 ans de recherche** en hydrologie. Pr C. Cudennec
- **Une Thèse** A. Delavenne. Modélisation Hydrologique à base Géomorphologique de bassins versants non jaugés par Régionalisation et transposition d'hydrogramme
- **Un modèle :**



## Valorisation scientifique

- **Dallery, D. et Al. 2020.** An end user friendly hydrological Web Service for hydrograph Prediction in Ungauged Basins. Hydrological Sciences Journal (sous presse).

## Lancement du projet

- Financement C.R. + AELB
- CDD 3 ans
- Co-financement infra GéoSAS

## Conception

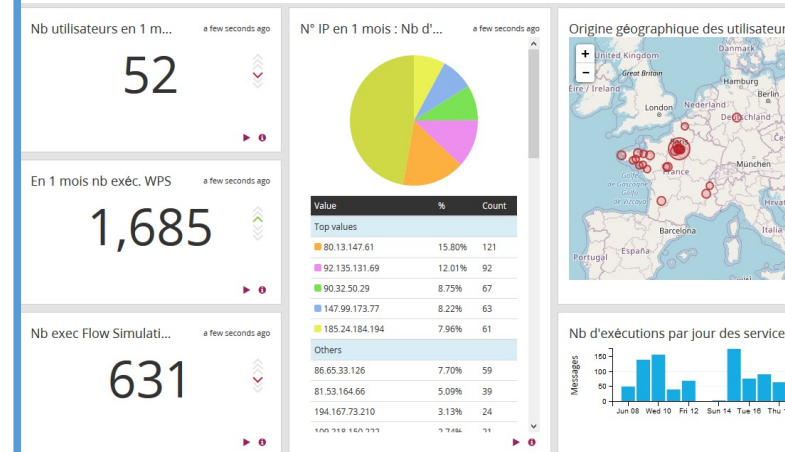
- Dév. Du module TransfR sous R
- Dév. du Service OGC WPS sous PyWPS
- BDD Postgres
- Forge github

## Publication

- Déploiement en production API de modélisation standard OGC WPS
- Métadonnée de service dédiée au projet

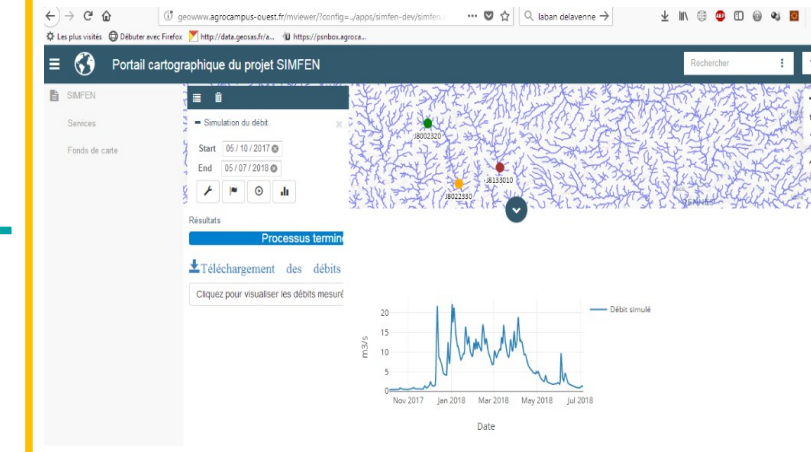
## Reporting

- Reporting sur l'utilisation des services du projet
- A destination de l'équipe projet, des financeurs, ...



## Réutilisation

- Portail cartographique proposant l'exécution du modèle. MViewer – JQuery – OpenLayers. Design de service.
- Public visé : Gestionnaire de l'eau, recherche, autres ?



# L'IDS comme base de R & D : Valorisation scientifique

Quelques publications sur le thème de l' « hydro-informatique » & Open Science...

- Dallery, D. et Al. 2020. *An end user friendly hydrological Web Service for hydrograph Prediction in Ungauged Basins*. Hydrological Sciences Journal (sous presse),
- Braud, I., et Al. 2020. *Building the information system of the French Critical Zone Observatories network: Theia/OZCAR-IS*, Hydrological Sciences Journal, DOI: 10.1080/02626667.2020.1764568,
- Bera, R. et Al. 2015. *GéoSAS : A modular and interoperable Open Source Spatial Data Infrastructure for research*, PIAHS, 368, 9–14, <https://doi.org/10.5194/piahs-368-9-2015>,
- Squidant, H., et Al. 2015. *Online watershed boundary delineation: Sharing models through Spatial Data Infrastructures*. PIAHS, 268, 144–149, <https://doi.org/10.5194/piahs-368-144-2015>.

... et quelques autres dans les cartons : « Agro-informatique » (Sol-AID), « SHS-informatique » (Parchemins)

- Morvan, T. et Al. Le web service Sol-AID, un exemple d'une instantiation d'un écosystème de services.
- Revelin F. et Al. « Ce que l'intention de mise en partage des données fait aux pratiques des chercheurs en anthropologie »

# Et le FAIR Data dans tout ça ?

- Avènement des principes du FAIR data, Wilkinson et al., 2016. « The FAIR Guiding Principles for scientific data management and stewardship »
- Guidelines on FAIR Data Management in Horizon 2020
  - Le bâton d'INSPIRE n'a pas marché, maintenant voilà le carotte du FAIR \$\$\$\$ ! ☺
- Compatibilité INSPIRE et FAIR data ?
  - Trouver, Voir, Télécharger = Findable, Accessible, Interoperable, Reusable ?
  - La directive Inspire est globalement compatible avec les différents critères d'évaluation du FAIR data. Le respect de ces critères peut être le fait de différents éléments parfois concomitants :
    - directement par application des préconisations de la directive **Inspire** (*F2, F3, A1.1, I2, R1.1, R1.2*),
    - par l'implémentation de normes comme les **standards ISO19115** et **ISO19139** (*I1, R1.2, R.1.3*),
    - par le déploiement d'outil comme les **IDG** (*A1.2, I3*)
    - par le respect de **bonnes pratiques** non imposées mais permises par la directive Inspire (*F1, F4, A2, I2*).
- Article de blog sur GéoSAS : [La directive INSPIRE, la bonne affaire !](#)



## The FAIR Guiding Principles

To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

To be Accessible:

- A1. (meta)data are retrievable by their identifier using a standardized communications protocol
  - A1.1 the protocol is open, free, and universally implementable
  - A1.2 the protocol allows for an authentication and authorization procedure, where necessary
- A2. metadata are accessible, even when the data are no longer available

To be Interoperable:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
- I2. (meta)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

To be Reusable:

- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
  - R1.1. (meta)data are released with a clear and accessible data usage license
  - R1.2. (meta)data are associated with detailed provenance
  - R1.3. (meta)data meet domain-relevant community standards

# Perspectives

- Continuer cette démarche pro-active en interne et en externe à l'UMR SAS facilité par :
  - L'entrée « science ouverte » reconnue dans la nouvelle organisation de l'UMR
  - Le projet d'IDG institutionnelle de l'INRAE,
- Interopérabilité & Modélisation
- Elargir l'approche « spatiale » à d'autres types de données
  - Données à entrée spatiale
  - Données a-spatiales, ex SHS / Projet Parchemins → test concluant CKAN
  - En route vers un [data.umrsas.fr](http://data.umrsas.fr) ?

# Conclusions

- Arrêtons de parler de BdD ! Parlons de SERVICES.
- Attention à l'interopérabilité galvaudée : « interopérable avec » = compatible  $\neq$  interopérable
- Une IDS, on en fait ce qu'on en veut (Modularité VS Plateforme) ou ce qu'on peut (Disponibilité VS Coût)
- Merci Mr Wilkinson, merci au FAIR
- Les étoiles semblent alignées pour la construction d'un écosystème de service numériques pour la « Science ouverte » !

# Merci aux humains de l'écosystème !

... de la communauté geOrchestra



... de la recherche (INRAE, IDS West, OZCAR, DEVLOG , ...)



... de l'UMR SAS



# Ressources

- GeOrchestra
  - Site : <https://www.georchestra.org/>
  - Forge : <https://github.com/georchestra>
- GéoSAS
  - Site : <http://geowww.agrocampus-ouest.fr>
  - Forge : <https://github.com/geosas>
- Formation
  - Moddle "Infrastructures et Web de Données Spatiales" : <https://tice.agrocampus-ouest.fr/course/view.php?id=887>