



Transformation des pratiques scientifiques à l'ère du numérique, pour une science ouverte

Fred Garcia

03/05/16, Ecole-chercheur Connaissances



Transformation des pratiques scientifiques à l'ère du numérique, pour une science ouverte

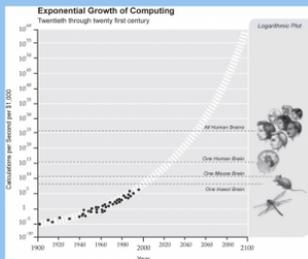
Révolution numérique

Science 2.0

- Mutation des pratiques de recherche
- Mutation du travail et de l'organisation collective

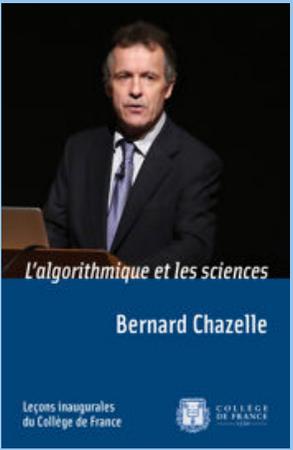
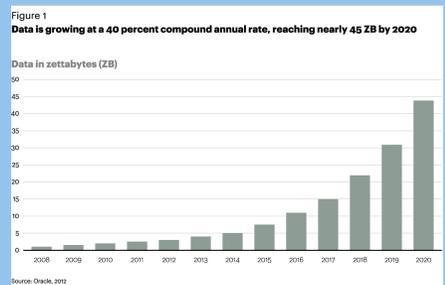
Rebonds

Une révolution numérique en cours

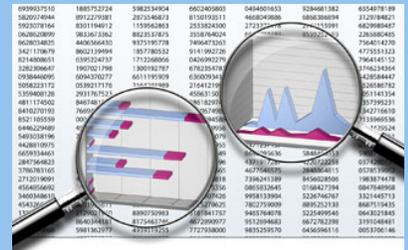
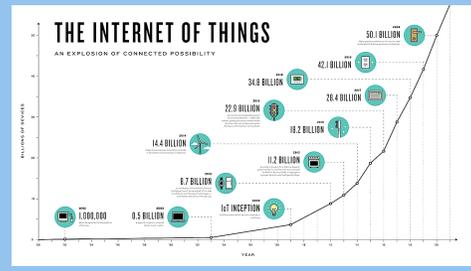
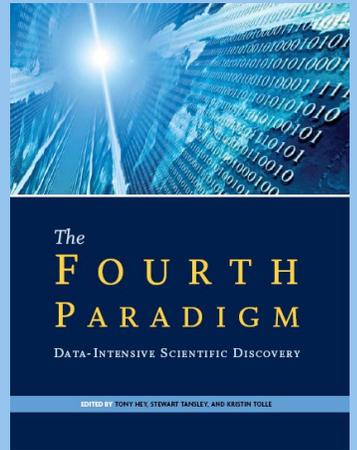
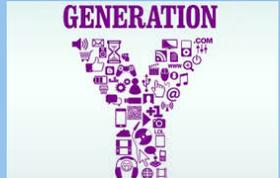


Révolution technologique
 Le super-calcul à la portée de tous

Révolution scientifique
 Le retour de l'intelligence artificielle



Révolution culturelle
 La Data au cœur de l'innovation



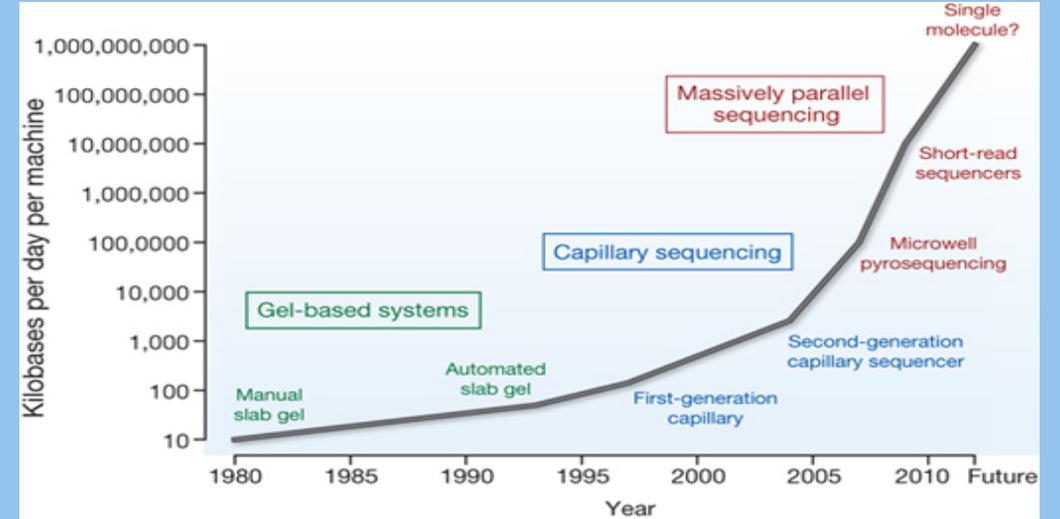
(big) Data mining

Progrès des technologies de production

Nouvelles technologies de séquençage pour les données génomiques

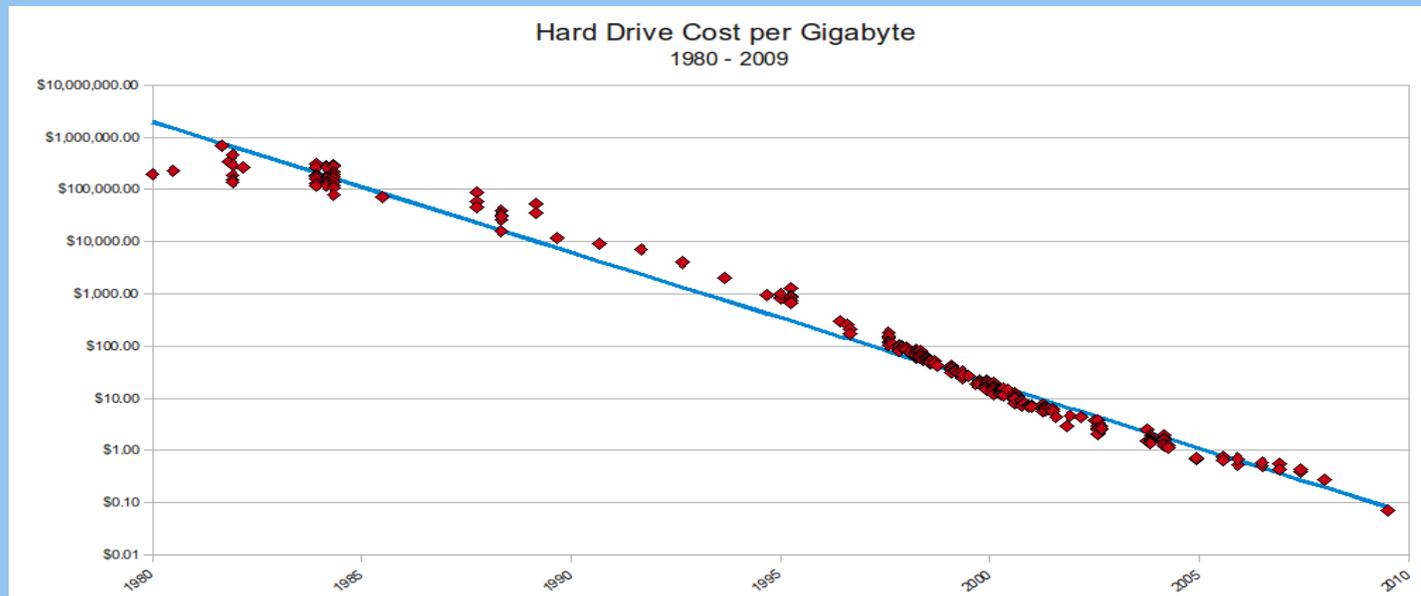
Imagerie HD

Automatisation, robotisation

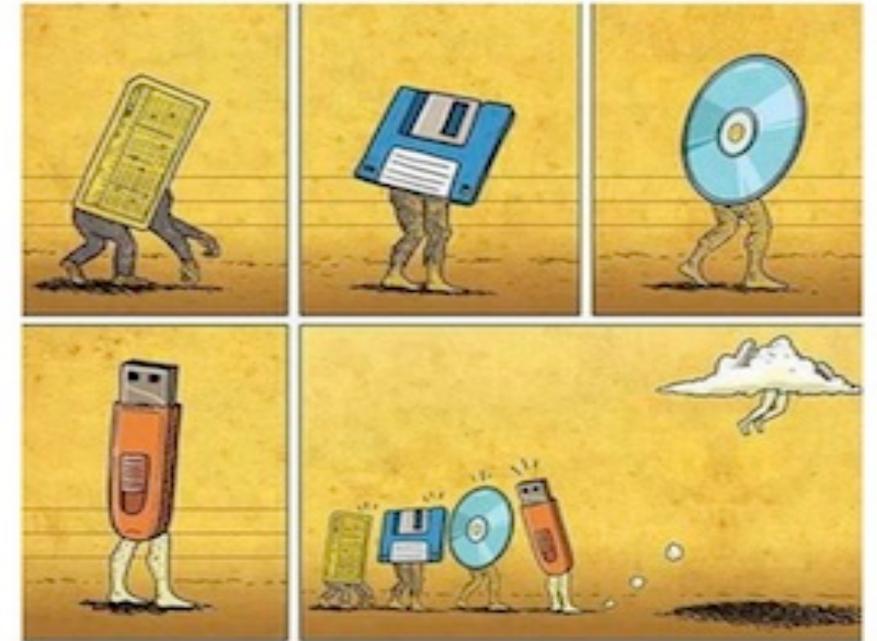


Progrès des technologies de stockage

De la mémoire de moins en moins chère et de plus en plus disponible

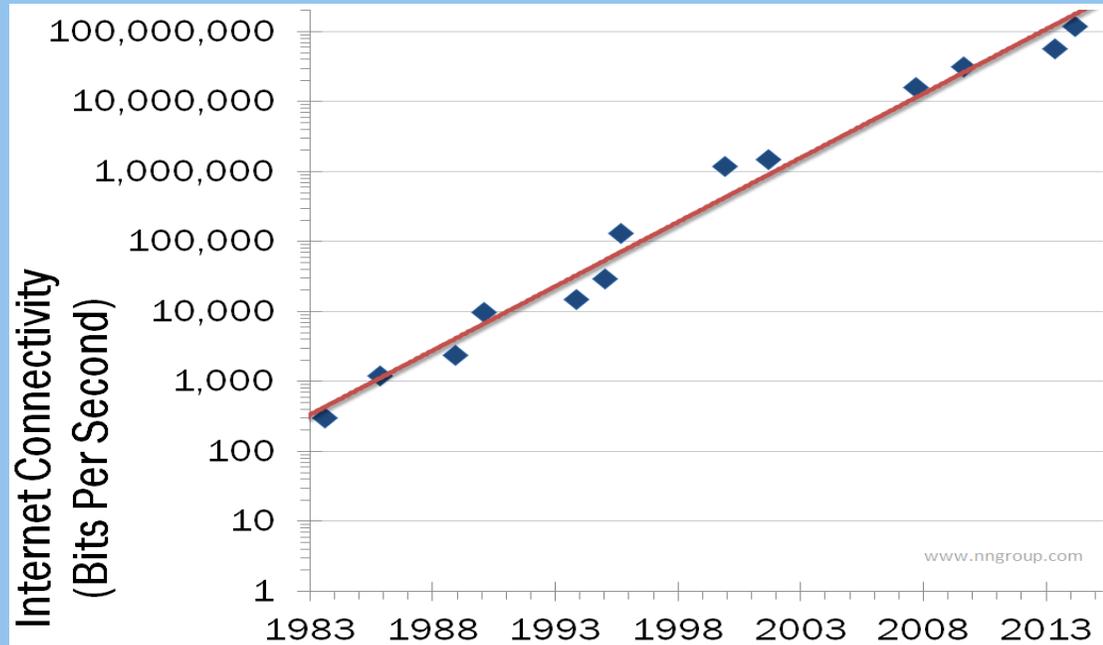


The Evolution of Data Storage



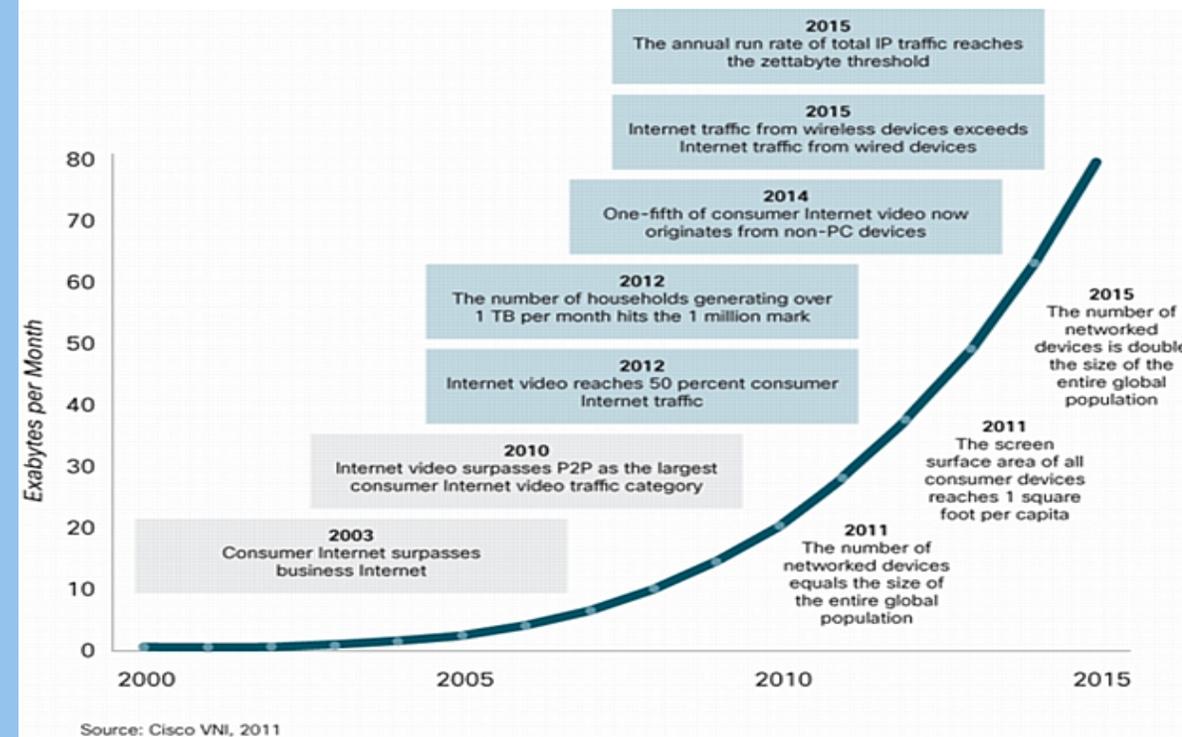
Progrès des technologies du net

Des vitesses de connexion +50% / an



Un trafic en pleine explosion

Figure 1. Five Traffic Milestones and Three Traffic Generator Milestones by 2015



Science 2.0 : une science en transition

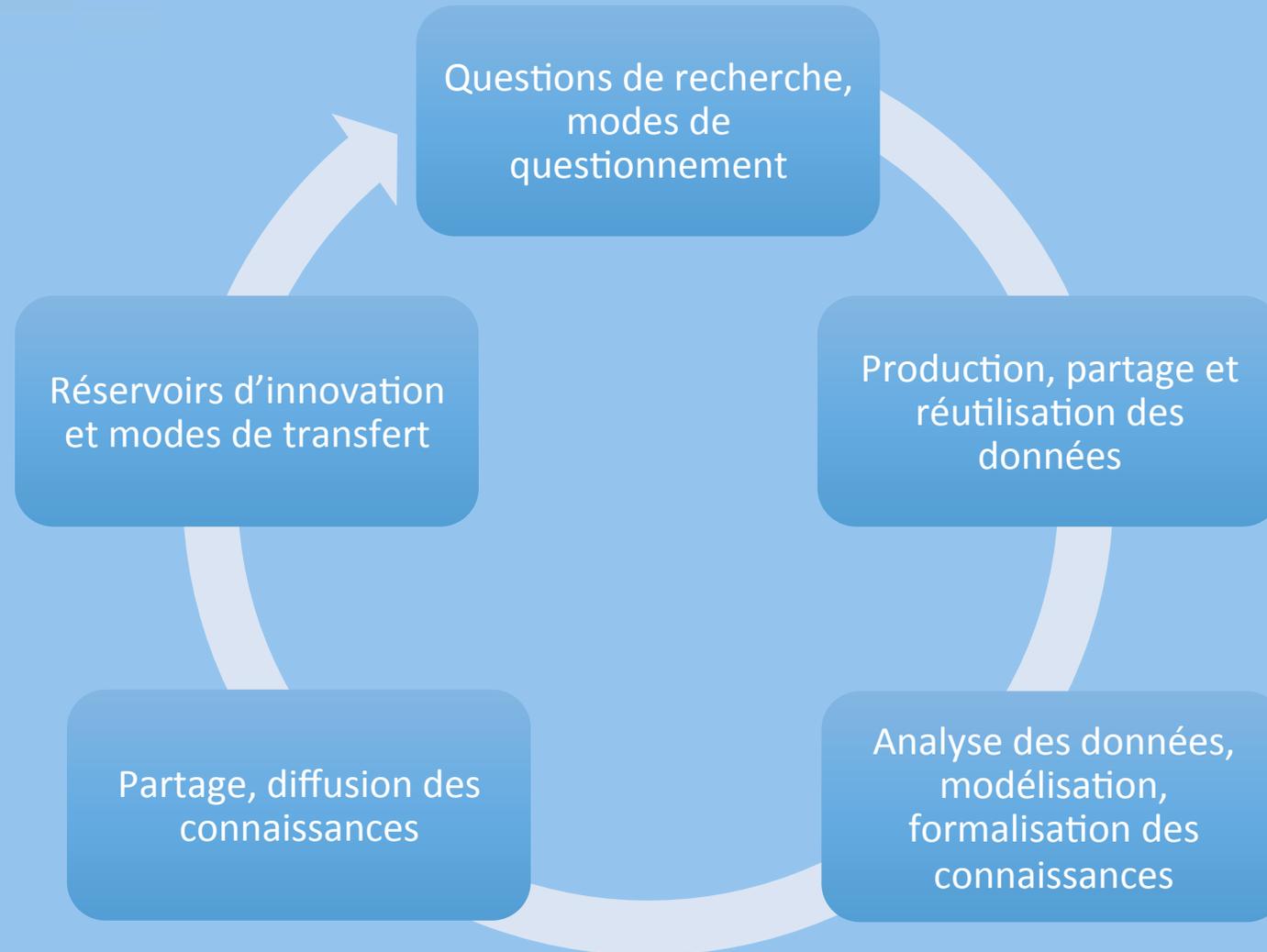
(EC online consultation, 2015)

A systemic change in the modus operandi of doing research and organising science, driven by the globalisation of the scientific community, as well as the increasing societal demand to address the Grand Challenges of our times



Vers une science ouverte

Numérique et mutation des pratiques de recherche



Questions de recherche, modes de questionnement

Le numérique facteur d'**enrichissement** de la démarche scientifique du fait de l'explosion des données qui décrivent le monde

Par sa capacité accrue à traiter ce déluge de données, à appréhender, modéliser, simuler la complexité de nos objets, il ouvre dès maintenant à de nouvelles questions de recherche qu'on ne pouvait jusqu'alors aborder, selon des démarches **plus systémiques** et moins analytiques, mais également **plus inductives**

➤ Brain project, traces numériques, ...

Il facilite les interactions entre la recherche et la société civile (sciences participatives, citoyennes, crowd funding, ...), et favorise l'émergence de **nouveaux cadres d'innovation**

➤ Living labs, ...

Production, partage et réutilisation des données

Un accroissement quasi exponentielle de la production et de l'échange de données

➤ Phénotypage, capteurs embarqués, connectés, images satellites



Une dynamique résolue de **partage des données** : ouverture et réutilisation. Les progrès réalisés dans le recensement, le partage, l'annotation et l'intégration des données favorisent en retour la production de données nouvelles, plus originales et spécifiques

➤ Cf. récentes lois du numérique et politique des Instituts pour l'ouverture

Outre un changement de culture pour les chercheurs habitués à traiter les données qu'ils produisent, cette transition impose **d'investir dans les outils et compétences nécessaires à une meilleure gestion de notre patrimoine numérique.**

➤ Vers un portail des données de l'Inra

Analyse des données, modélisation, formalisation des connaissances

Données massives, flux de données, données hétérogènes, données non planifiées, ...

- Une modification radicale des données à analyser, même si petit big data en agriculture

Machine learning, statistique en grande dimension

- Des méthodes de plus en plus riches et complexes

Modélisation intégrative, simulation intensive

- Exploiter toutes les possibilités offertes aujourd'hui pour simuler des agroécosystèmes complexes, vers plus de couplage données – modèles

Réification de la connaissance, un nouvel objet à modéliser

- Des outils d'extraction et de formalisation de + en + nombreux

Partage, diffusion des connaissances

Le numérique révolutionne également les modalités de partage et de diffusion des connaissances (données liées, publications interactives, codes ouverts, livres de connaissance, moocs, serious games...)

- **Moocs Agroécologie ...**

Standardisation des processus et des objets de recherche manipulés

- **Vers des ontologies partagées**

Nouveaux outils de recherche bibliographique, de fouille de documents

- **TDM, Cf plateforme Alvis, ...**

Incitation globale à l'ouverture dans le domaine de la publication (open access, open peer reviewing, ..)

- **Les revues Inra incitées à passer dans l'open (ASD)**

Importance croissante du web social (blogs, wiki, réseaux sociaux, ...)

- **ResearchGate, Hypothese.org**

Réservoirs d'innovation et modes de transfert

Le numérique comme vecteur d'innovation incontournable dans les domaines de l'agriculture, de l'environnement ou de l'alimentation (agriculture numérique, e-food, etc.)

➤ cf. rapport AI 2025, vers un portail des données agricoles, IC agriculture numérique ...

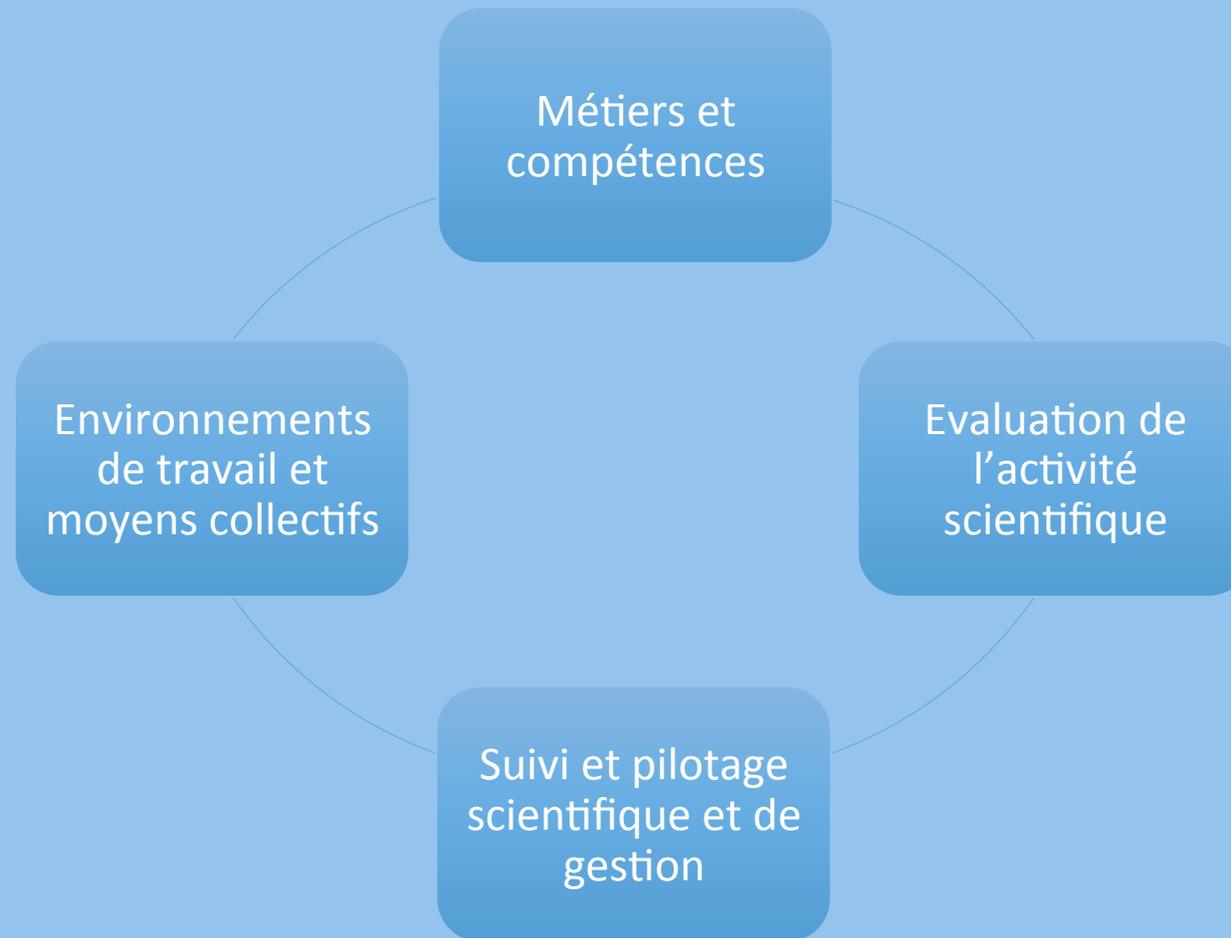
De nouveaux acteurs de la e-économie

➤ Intel à Saclay, IBM, Amazon, et pleins de startups ...

Innovations technologiques ou dans le traitement de l'information

> SAS, services en ligne, OAD

Numérique et mutation du travail et de l'organisation collective



Métiers et compétences

Montée en compétence nécessaire pour l'accès et le partage, l'intégration et l'analyse des données, la modélisation et la simulation intensive, l'extraction et la formalisation des connaissances

➤ GPEC et formations

Modification des modes de travail individuels ou collectifs : télétravail, nomadisme, visioconférences, travail collaboratif, environnement partagé, ...

➤ cf. Loi sur le télétravail

Evaluation de l'activité scientifique

Nouveaux produits et nouveaux modes de production



nouveaux critères d'évaluation et indicateurs associés

- altmetrics.com
- Vers des mesures d'impacts personnalisées

Dématérialisation du processus d'évaluation. Un accès direct aux productions formatées

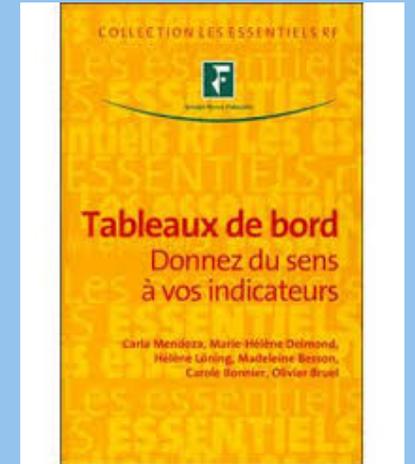
- Quelle place pour les productions informelles ?



Suivi et pilotage scientifique et de gestion

Le numérique permet et encourage un suivi « temps-réel » des activités scientifiques et de gestion

- Mise à disposition à tous les niveaux de management d'instruments de pilotage à base d'informations fiables



Environnements de travail et moyens collectifs

Montée en puissance des infrastructures collectives



- E-infrastructures nationales ou européennes pour le stockage et le calcul
- Des plateformes aux e-infrastructures scientifiques
- Intégration des SI des infrastructures de recherche, couplages aux e-infra

Propriété Intellectuelle du numérique, PI des connaissances

- Importance de la PI dans les développements informatiques (logiciels, bases de données)
- Propriétaire = employeur du développeur. A distinguer de l'auteur.
- Importance du choix des licences d'utilisation.
- Mouvement du logiciel libre. Dynamique forte vers l'utilisation de licences libres (mais elles ne sont pas toutes équivalentes).
- Contradiction à affronter : ouverture / valorisation de nos productions

Qu'en est-il au niveau des « connaissances » ?

Limites juridiques à l'exploration de données (TDM)

« Traitement automatisé de contenus numériques qui peuvent inclure des textes, données, sons, images ou autres éléments ou une combinaison de ceux-ci afin de découvrir de nouvelles connaissances ou des idées » Rapport UE 03/2014

L'exploration peut viser de nombreux objets de droit (texte, chiffres, cartes, graphiques, images, sons, etc.)

2 réglementations s'appliquent principalement : celle sur le droit d'auteur, celle sur le droit des producteurs de bases de données. Limitations avec :

- Il faut prendre possession des contenus et les stocker pour pouvoir les explorer
- La restitution finale du travail d'exploration peut reproduire en partie le contenu d'origine

Nouvelle loi sur le numérique autorise en partie le TDM

Portail des données agricoles

Le Rapport « Agriculture-Innovation 2025 » souligne l'enjeu important que représente l'exploitation des données numériques pour le monde agricole.

Mis en œuvre d'un portail de données pour l'innovation en agriculture

L'objectif de ce portail est de mettre à disposition de l'écosystème agricole un ensemble important de données et d'informations, issues de diverses sources : données publiques, données des entreprises, données des agriculteurs ...

En exploitant les données et informations accessibles depuis ce portail, des acteurs économiques innovants, notamment des startups, pourront développer de nouveaux cas d'usage et des services à destination du monde agricole, dans des démarches d'innovation ouverte. La mise en place du portail pourra ainsi permettre aux agriculteurs de conserver la maîtrise de leurs données et constituera une opportunité pour faire émerger des solutions innovantes pour tous les acteurs de la filière.