

Préserver les données dans un monde contraint

Alexis Arnaud
(GRICAD, UGA)

2024-10-02
R.I.P. Data

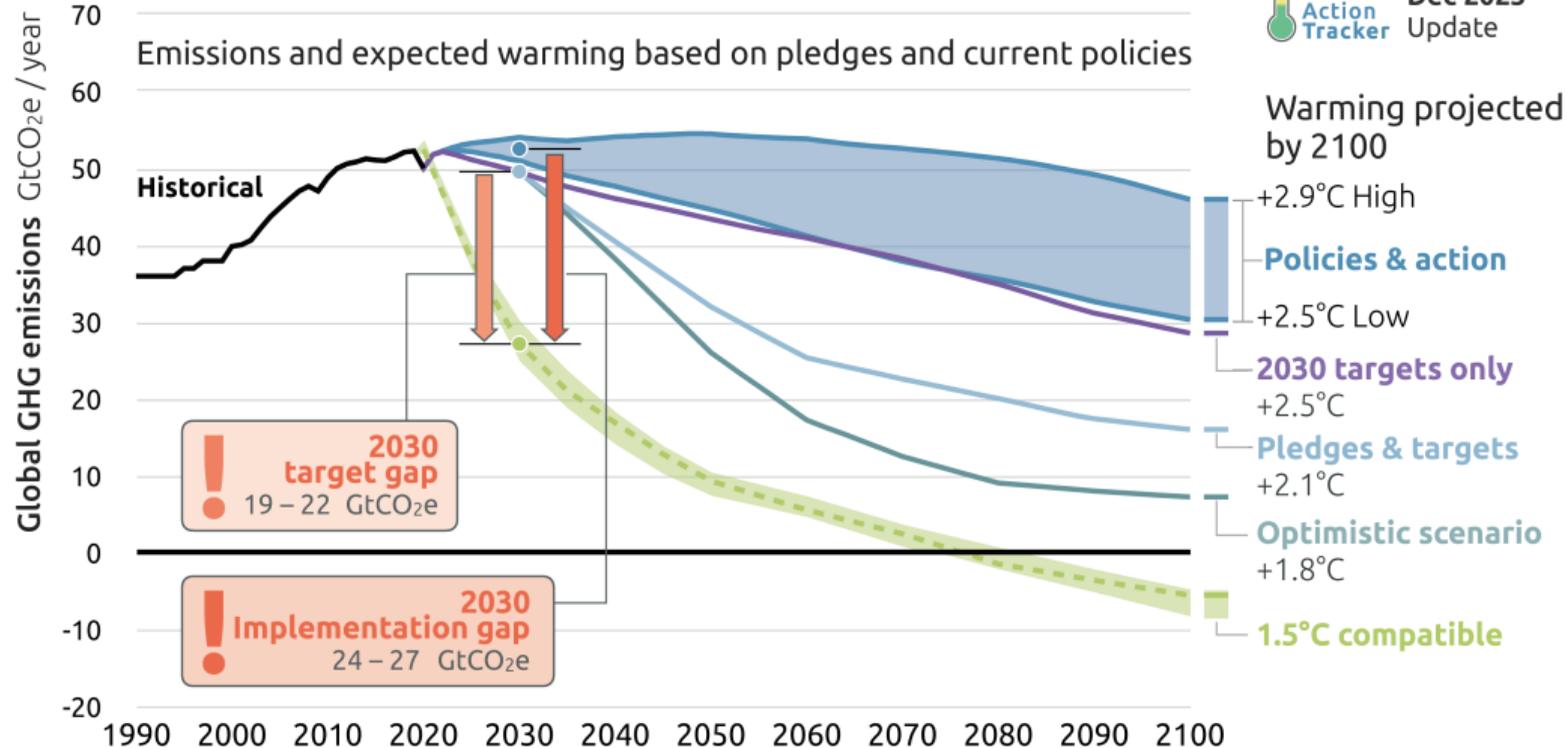


Plan

- 1 Contexte climatique
- 2 Impacts environnementaux du numérique
- 3 Les données (de recherche) dans tout ça ?
- 4 Ressources

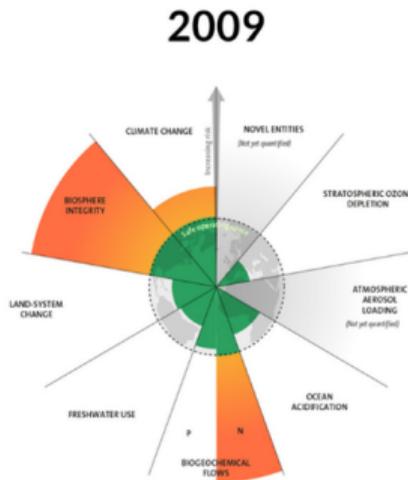
Évolution de la température moyenne mondiale

2100 WARMING PROJECTIONS



Source : Climate Analytics and NewClimate Institute

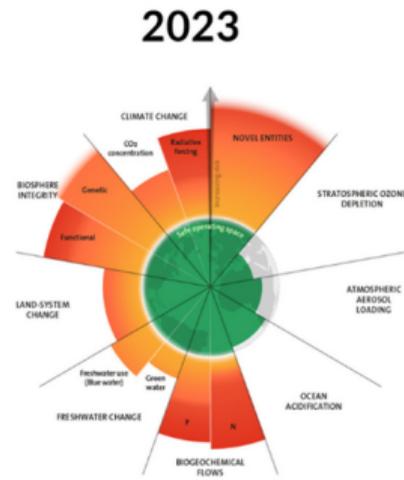
Frontières planétaires



7 boundaries assessed,
3 crossed



7 boundaries assessed,
4 crossed

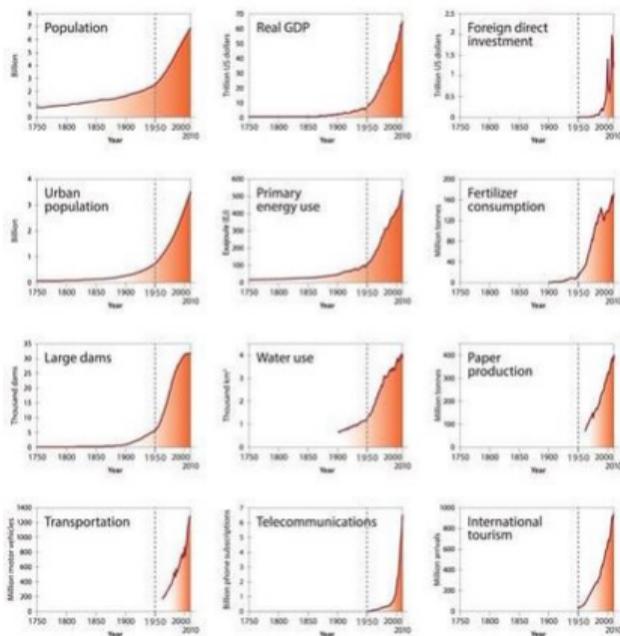


9 boundaries assessed,
6 crossed

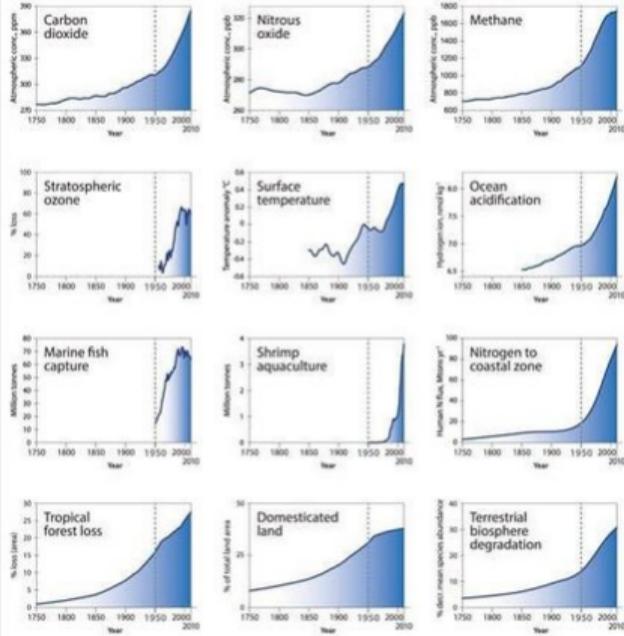
The evolution of the planetary boundaries framework. Credit : Azote for Stockholm Resilience Centre, Stockholm University. Based on Richardson et al. 2023, Steffen et al. 2015, and Rockström et al. 2009

La grande accélération

Socio-economic trends



Earth system trends



Updated Great Acceleration Graphs

Source: Will Steffen et al. "The trajectory of the Anthropocene: The Great Acceleration." The Anthropocene Review, March 2015

Source : [The Great Acceleration \(Steffen et al., 2015b\)](#)

Plan

- 1 Contexte climatique
- 2 **Impacts environnementaux du numérique**
- 3 Les données (de recherche) dans tout ça ?
- 4 Ressources

Le numérique, un secteur impactant

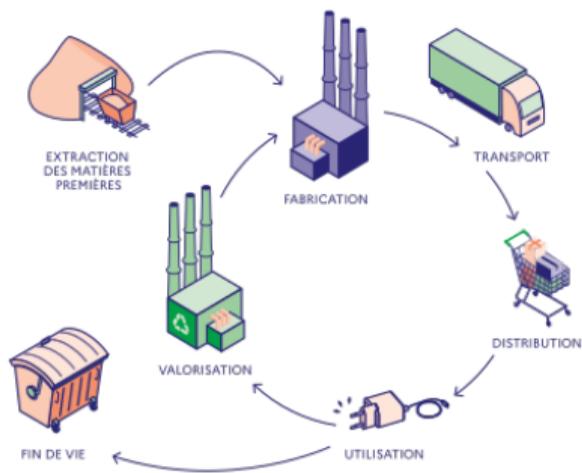
Quelques ordres de grandeur au niveau mondial

- $\approx 3.5\%$ des émissions mondiales de GES ($\approx 2.5\%$ des émissions de la France), et avec une forte croissance
- 67 millions de serveurs
- 1.1 milliard d'équipements réseaux
- 19 milliards d'objets connectés en 2019 : 48 milliards en 2025 selon les estimations

Source : [ADEME, En route vers la sobriété numérique, 2022](#)

Le numérique, un secteur impactant

C'est le matériel qui a le plus d'effets directs



Cycle de vie d'un produit.

Source : [MiNumEco](#), la Mission Interministérielle Numérique Écoresponsable



Agbogbloshie : décharge de déchets d'équipements électriques et électroniques en provenance de pays industrialisés (légalement et illégalement). Détruite en 2021 par le gouvernement. Source : [Wikipedia](#)

Impacts directs du numérique

Tout au long du cycle de vie du matériel



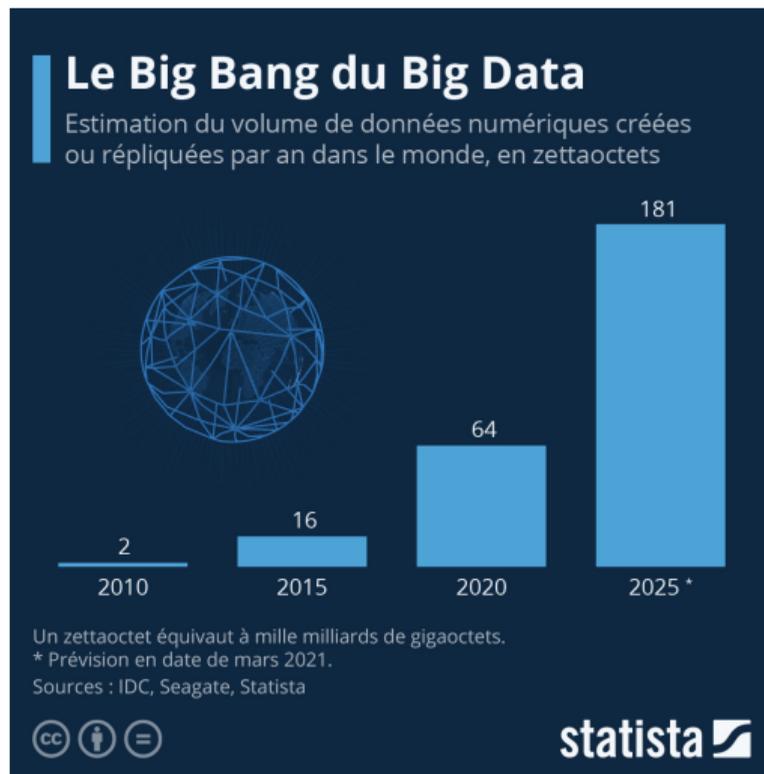
Source : J. Combaz - 2019

Plan

- 1 Contexte climatique
- 2 Impacts environnementaux du numérique
- 3 Les données (de recherche) dans tout ça ?**
- 4 Ressources

Une explosion de données produites et disponibles

- **Un volume de données en croissance exponentielle, à cause :**
 - ▶ des usages : vidéos, IA, santé, objets connectés, etc.
 - ▶ du matériel disponible : fibre optique, 5G, large démocratisation des smartphones



Source : [statista.com](https://www.statista.com)

Une utilisation de données qui devient problématique

Cas de l'IA générative

IA et consommation d'eau

- **Microsoft** : +34% entre 2021 et 2022.
- **Google** : +20% entre 2021 et 2022.

Source : [LeMondelInformatique.fr](https://www.lemondeinformatique.fr)

IA et consommation d'électricité

L'électricité consommée par les centres de données numériques dans le monde devrait doubler d'ici à 2026, principalement en raison de l'essor de l'IA et des cryptomonnaies, selon un [rapport publié le 24 janvier 2024 par l'Agence internationale de l'énergie](#).

Source : [LeMonde.fr](https://www.lemonde.fr)

Intégrer les enjeux environnementaux à la conduite de la recherche Une responsabilité éthique

C'est au monde de la recherche lui-même d'ouvrir en son sein un large débat sur ces questions. Pour le COMETS, il s'agit là d'une **exigence première**, bien avant toute mise en place d'instances ou de critères d'évaluation environnementale des projets de recherche, lesquels sont loin d'être inutiles mais pourraient contribuer à routiniser un questionnement qui exige avant toute chose d'être constitué en réflexion collective approfondie.

Source : COMETS, Comité d'éthique du CNRS

Recommandations a l'adresse de la direction du CNRS et des personnels de recherche

- **Reconnaître que la prise en compte de l'environnement fait partie intégrante de l'éthique de la recherche** ; affirmer à ce titre la responsabilité des acteurs et actrices de la recherche de penser leur activité au regard des enjeux environnementaux ; cette responsabilité concerne non seulement l'empreinte des pratiques de recherche mais plus généralement l'impact environnemental négatif ou positif que le choix de tel ou tel sujet de recherche et de telle ou telle voie pour le traiter peut engendrer pour l'environnement au sens large, à court, moyen ou long terme.
- **Multiplier les espaces de discussion** permettant à l'ensemble des personnels de recherche de débattre des enjeux et contours de cette responsabilité.

Source : COMETS, Comité d'éthique du CNRS

La science ouverte comme réflexion sur la gestion des données

Diffusion sans entrave des produits de la recherche

● Enjeux :

- ▶ qualité de la science
- ▶ collaboration
- ▶ confiance des citoyens
- ▶ enjeux économiques

● Principes FAIR :

- ▶ diffusion des données
- ▶ réutilisation des données



Source : UNESCO

Proposition d'actions pendant la recherche

Intégrer les impacts environnementaux et sociétaux dans les outils de la science ouverte

- **Importance de réfléchir et estimer les impacts :**

- ▶ On ne peut pas réduire ou éviter ce que l'on ne connaît pas.
- ▶ Réfléchir aux finalités : produire ou non des données ?
- ▶ Réfléchir aux besoins en données : peut-on faire sans produire de nouvelles données ?
- ▶ Réfléchir aux besoins en ressources de calcul : peut-on faire avec les ressources actuelles ?
- ▶ Comparer plusieurs approches, principe de l'analyse en cycle de vie (ACV).

- **Enrichir le plan de gestion des données :**

- ▶ Décrire le coût environnemental des différentes étapes du cycle de vie des données.
- ▶ Garder comme objectif la réutilisation des données (et des codes).
- ▶ Identifier les données nécessitant d'être diffusées : diffuser ou simplement signaler la production de données ?
- ▶ Identifier les durées et les supports de conservation des données.
- ▶ Lier les produits de recherche entre eux.

Proposition d'actions pendant la recherche

Intégrer les impacts environnementaux et sociétaux dans les outils de la science ouverte

- **Discuter collectivement :**

- ▶ Identifier les données qui, par rapport à la crise climatique, sont importantes à produire et comment faire de la science avec ce qui est déjà produit.
- ▶ Réfléchir à de nouveaux moyens de diffusion des métadonnées et des données : diffusion d'extraits de données, stockage à froid, etc.
- ▶ Changer les méthodes d'évaluation pour intégrer les impacts environnementaux aux éléments de science ouverte.

Des éléments à ramener chez soi

- **(Prendre le temps de) Réfléchir :**

- ▶ Accepter le caractère complexe du problème.
- ▶ Intégrer une pensée systémique.
- ▶ Mesurer et questionner les besoins et les finalités.

- **Ralentir :**

- ▶ Nécessaire pour réfléchir.
- ▶ Nécessaire pour garantir la réutilisation des données.
- ▶ Lien avec la slow science.

Merci de votre attention !

Des questions ?

Plan

- 1 Contexte climatique
- 2 Impacts environnementaux du numérique
- 3 Les données (de recherche) dans tout ça ?
- 4 Ressources

Impact environnemental du numérique

- **GDS EcoInfo** : Réduire les impacts environnementaux et sociétaux négatifs des technologies du numérique.
- **Labos 1point5** : Réduire l'empreinte de nos activités de recherche sur l'environnement.
- **Ademe** : Numérique : quel impact environnemental ?

Quelques articles

- Berthoud, Françoise, Bruno Bzeznik, Nicolas Gibelin, Myriam Laurens, Cyrille Bonamy, Maxence Morel et Xavier Schwindenhammer. 2020. "Estimation de l'empreinte carbone d'une heure.coeur de calcul". Research Report. UGA - Université Grenoble Alpes ; CNRS ; INP Grenoble ; INRIA. <hal-02549565>
- Guillaume Charret, Alexis Arnaud, Françoise Berthoud, Bruno Bzeznik, Anthony Defize, et al.. Estimation de l'empreinte carbone du stockage de données. [Rapport de recherche] CNRS - GRICAD. 2020. hal-03573790
- Marion Ficher, Françoise Berthoud, Anne-Laure Ligozat, Patrick Sigonneau, Badis Tebbani, et al.. Évaluation de l'empreinte carbone de la transmission d'un gigaoctet de données sur le réseau RENATER. Centre de Recherche Interdisciplinaire ; GIP RENATER. 2021. hal-04197870
- Didier Mallarino, Sylvie Le Bras, Cyrille Bonamy. Les impacts environnementaux et sociétaux des données : un défi pour l'avenir. Congrès JRES : Les Journées Réseaux de l'Enseignement et de la Recherche, RENATER, May 2022, Marseille, France. hal-03702208