



HAL
open science

Des réponses originales à l'hétérogénéité des données écologiques en Méditerranée dans le cadre du consortium IndexMed

Romain David, Jean-Pierre Féral, Anne-Sophie Archambeau, Loup Bernard,
Cyrille Blanpain, Alrick Dias, S. Gachet, Gibert Karina, Julien Lecubin,
Michelle Leydet, et al.

► To cite this version:

Romain David, Jean-Pierre Féral, Anne-Sophie Archambeau, Loup Bernard, Cyrille Blanpain, et al..
Des réponses originales à l'hétérogénéité des données écologiques en Méditerranée dans le cadre du
consortium IndexMed. Colloque LITEAU, Observation et recherche en appui aux politiques du littoral
et de la mer, Jan 2016, Brest, France. 10.13140/RG.2.1.2925.9280 . hal-04653485

HAL Id: hal-04653485

<https://hal.science/hal-04653485>

Submitted on 18 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Des réponses originales à l'hétérogénéité des données écologiques en Méditerranée dans le cadre du consortium IndexMed



R. DAVID, J.-P. FERAL, A.S. ARCHAMBEAU, L. BERNARD, C. BLANPAIN, A. DIAS, S. GACHET, K. GILBERT, J. LECUBIN, M. LEYDET, G. ROMIER, C. SURACE et la communauté IndexMed

En savoir plus : <http://www.indexmed.eu>



doi: 10.13140/RG.2.1.2925.9280

Introduction

La production de données dans le domaine littoral et marin concernant la **biodiversité et les systèmes socio-écologiques (SSE)** est à la fois coûteuse et peu automatisée. Les suivis de longues séries temporelles et/ou à larges emprises spatiales sont difficiles à mener, dès lors qu'il faut recourir à plusieurs observateurs, et la robustesse et la reproductibilité de l'observation est plus difficile à obtenir.

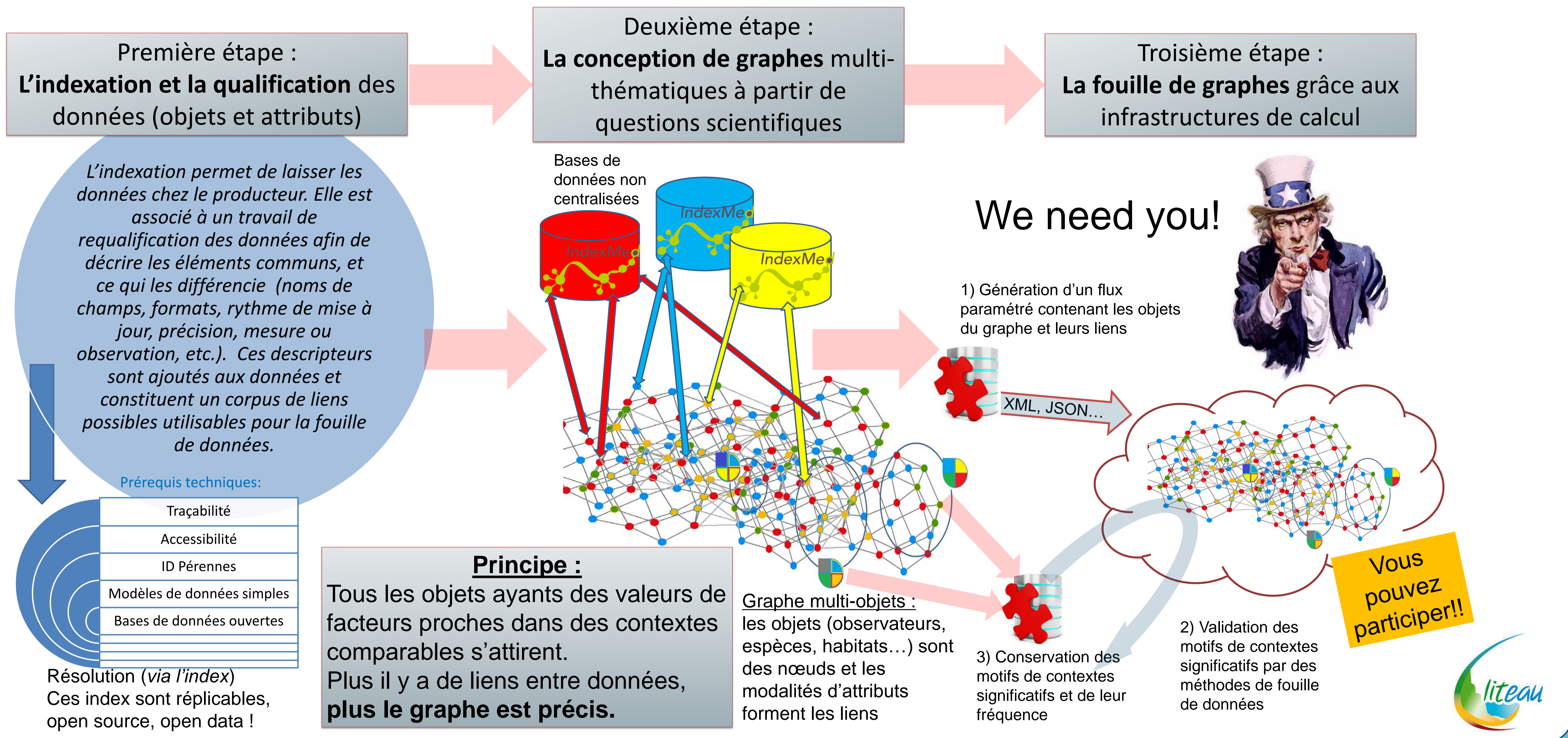
Dans un cadre de production de données multi-sources, l'équivalence de systèmes d'observations et l'inter-calibration d'observateurs deviennent cruciales. Des approches intégratives pluri- voire transdisciplinaires deviennent nécessaires, dans l'étude de systèmes où la production de données dans chaque discipline est **discontinue, peu précise et mal répartie**. Pourtant, toutes les variables (biotiques, abiotiques, pressions anthropiques et naturelles, services rendus et ressentis, image sociétale) de ces systèmes interagissent dans le temps et à chaque échelle spatiale.

Comment construire un système d'indexation?

La donnée en écologie et en sciences humaines est extrêmement hétérogène. Pour pouvoir accéder aux différents types et format de données, le système doit être distribué, c'est à dire que les données restent chez le producteur, mais des flux normés sont déployés et peuvent être interrogés par tous les acteurs du réseau, du niveau local au niveau international, grâce à un index répertoriant et **qualifiant chaque enregistrement selon des typologies partagées**.

Le consortium IndexMed a pour objectif d'identifier puis de lever les **verrous scientifiques liés à la qualité des données et leur hétérogénéité**. L'utilisation de l'indexation en vue de construire des graphes permet de les considérer, malgré leur disparité, à un niveau équivalent, et améliore l'aide à la décision en utilisant des méthodes émergentes de fouille de données (clustering collaboratif, machine-learning, fouille de graphes, représentation de connaissances).

Comment interpréter des données hétérogènes et non liées?



La recherche transdisciplinaire est elle efficace?

La révolution du Big Data dans les disciplines « SHS et écologie » tarde, alors qu'elle est considéré par de nombreux chercheurs et industriels comme la plus prometteuse des pistes de progrès et de découvertes.

Les systèmes d'information reliant des objets grâce à des qualifications (des liens) sont omniprésents, le premier étant l'homme, le consommateur. L'entreprise « Google » a initié cette nouvelle science appelée « fouille de données » à partir de pages internet et de liens. Très vite, les systèmes mercantiles ont compris leur puissance analytique (les assurances pour gérer les risques, la française des jeux pour trouver les tricheurs, les banques pour les investissements, les « traders » pour augmenter leur marge, les publicitaires et les réseaux sociaux pour augmenter leurs impacts...).

L'urgence environnementale demande aujourd'hui de répondre grâce à un système **partagé et connecté** à des questions locales et globales à la fois, dépassant les questions scientifiques du type

« **cette dégradation est-elle liée à cette pression ?** »,

mais répondant plutôt à cette question :

« **Comment améliorer/préserver l'état écologique d'un milieu de la manière la plus efficace ?** »

Pourquoi l'indexation?

L'intégration des bases de données en une seule grande base centralisée a été tentée par de nombreux programmes. Cela a toujours conduit à un échec.

Pourtant, nous devons rendre nos données en écologie, économie et sciences sociales interopérables, connectables, comparables, et arriver à les analyser sans en placer une au-dessus des autres.

L'indexation est une alternative à la centralisation, elle permet des approches globales en écologie. **L'écologie globale** considère les systèmes écologiques dans leur complexité en termes de composition, structure et interactions. Elle identifie les paramètres à même de poser les bases d'une gestion durable des ressources et services qu'ils fournissent, de mieux appréhender et anticiper les risques et leurs conséquences et de participer à l'amélioration de la qualité de vie des sociétés.

