



Results of IndexMed GRAIL Days 2016: How to use standards to build GRaphs and mIne data for environmental research? IndexMeed consortium for data mining in ecology

Romain David, Jean-Pierre Féral, Anne-Sophie Archambeau, Fanny Arnaud, David Auber, Nicolas Bailly, Loup Bernard, Cyrille Blanpain, Vincent Breton, Denis Couvet, et al.

► To cite this version:

Romain David, Jean-Pierre Féral, Anne-Sophie Archambeau, Fanny Arnaud, David Auber, et al.. Results of IndexMed GRAIL Days 2016: How to use standards to build GRaphs and mIne data for environmental research? IndexMeed consortium for data mining in ecology. RDA Ninth Plenary Meeting "Data Infrastructures for Open Science", Apr 2017, Barcelona, Spain. , 2017, 10.13140/RG.2.2.32504.44806 . hal-02295974

HAL Id: hal-02295974

<https://hal.science/hal-02295974>

Submitted on 24 Sep 2019

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.

*Romain David^{*1}, Romain David^{*1}, Jean-Pierre Féral¹, Anne-Sophie Archambeau², Fanny Arnaud³, David Auber⁴, Nicolas Bailly⁵, Loup Bernard⁶, Cyrille Blanpain⁷, Vincent Breton⁸, Denis Couvet⁹, Anna Cohen-Nabeiro¹⁰, Aurélie Delavaud¹⁰, Alrick Dias¹¹, Sophie Gachet¹, Robin Goffaux⁸, Karina Gibert¹¹, Manuel Herrera¹², Dino Ienco¹³, Romain Julliard⁹, Julien Lecubin⁷, Yannick Legre¹⁴, Michelle Leydet¹, Grégoire Lois⁹, Victor Méndez Muñoz¹⁵, Jean-Charles Meunier¹⁶, Isabelle Mougenot¹⁷, Sophie Pamerlon³, Jean-Claude Raynal¹⁸, Geneviève Romier⁸, Dad Roux-Michollet¹⁹, Alison Specht⁹, Christian Surace²⁰, Thierry Taton¹ and the IndexMeed community

Contacts : Romain DAVID, Jean-Pierre FERAL, Dino IENCO, Thierry TATONI romain.david@imbe.fr

Animated by:



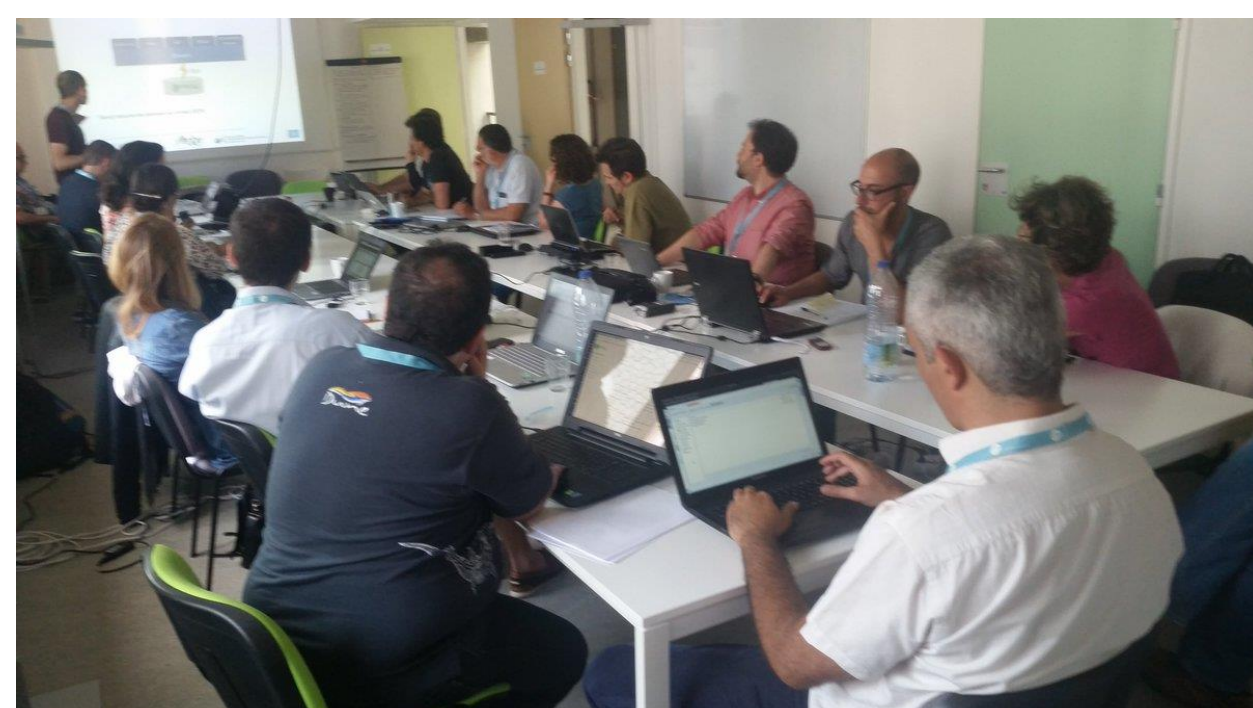
Supported by:



The *Inter operating data* challenge: How to mine heterogeneous non linked environmental data for research and management?

Data produced by biodiversity research projects that evaluate and monitor Good Environmental Status have a high potential for use by stakeholders involved in environmental management. The lack of specific scientific objectives, poor organizational logic, and a characteristically disorganized collection of information leads to a decentralized data distribution, hampering environmental research. In such a heterogeneous system across different organizations and data formats, it is difficult to efficiently harmonize the outputs. There are few tools available to assist.

Standards (including TDWG recommendations) and specific protocols can be applied to interconnect databases. Such semantic approaches greatly increase data interoperability. The aim of this poster is to present the 2016 IndexMed workshop results (<https://indexmed2016.sciencesconf.org>) and recent actions of the consortium (renamed IndexMeed - Indexing for Mining Ecological and Environmental Data): new approaches to investigate complex research questions and support the emergence of new scientific hypotheses.



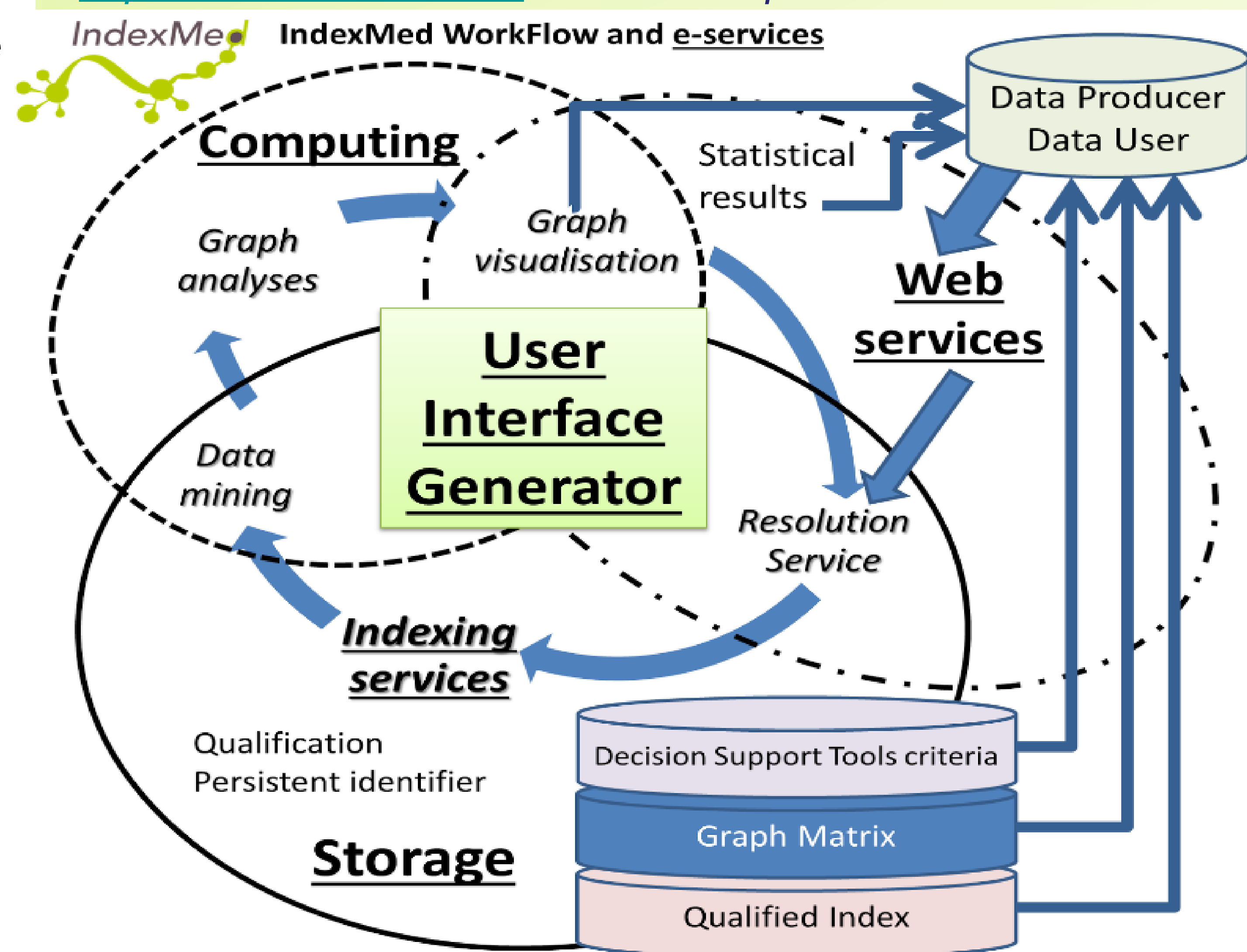
Current developments in data mining based on graphs, the potential for important contributions to environmental research, particularly about strategic decision-making, and new ways of organizing data were also discussed at the workshop.

In particular, this workshop promoted decisions on how:

- to analyze heterogeneous distributed data spread in different databases,
- to create matches and incorporate some approximations,
- to identify statistical relationships between observed data and the emergence of contextual patterns, and
- to encourage openness and the sharing of data, in order to value data and their utilization.

From last 2016 seminar, IndexMed became **IndexMeed (Indexing for Mining Ecological and Environmental Data)**. The task of the newly created consortium of IndexMeed is to index biodiversity data (and to provide an index of qualified existing open datasets) and make it possible to build graphs to assist in the analysis and development of new ways to mine data and to build new Intelligent Decision Support Systems (IDSS). Architecture of this project was defined at EGI Workshop “design your e-infrastructure” (Amsterdam) (see below)

<http://www.indexmed.eu> and soon <http://www.indexmeed.eu>



The IndexMeed project participants are now exploring the ability of two scientific communities (ecology *sensu lato* and computer sciences) to work together. The uses of data from biodiversity research demonstrate the prototype functionalities and introduce new perspectives to analyze environmental and societal responses including decision-making. Output of the seminar lists scientific questions that can be resolved by the new data mining approaches and proposes new ways to investigate heterogeneous environmental data with graph mining.

Indeed, we are delighted to invite you to the

2017 IndexMeed Seminar "SAGES DAYS"

Sciences and Algorithms around Graphs in Environment and Societies that will take place on **November 15th -17th 2017 in Paris**, at the CNRS headquarters.

Following the successful workshops of 2014, 2015 and 2016, this 4th seminar will promote exchanges between participants, the acquisition of practical methods and will explore new ways to **develop algorithms to mine graphs constructed with heterogeneous environmental data**. Thought as a collaborative space between experts in ecology and biodiversity and experts in the field of ICST, this one has the vocation of developing new research based on the approaches related to the graphs, and in particular to construct a project bearing

- on the **theoretical aspects (ICST) of algorithms for digging graphs** according to scientific questions and
- on the **theoretical aspects of integrating heterogeneous environmental data with a view to construct bio-scenarios** and to assist in environmental decision-making.

The 2016 IndexMeed Seminar welcomes abstracts for original oral and poster contributions until the 1st of October 2017. The online submission process will be opened soon until 1st of November on <https://indexmeed2017.sciencesconf.org>.

Contact : indexmeed2017@sciencconf.org

* DAVID et al 2016. David R., J.-P. Féral, A.-S. Archambeau, N. Bailly, C. Blanpain, V. Breton, A. De Jode, A. Delavaud, A. Dias, S. Gachet, D. Guillemain, J. Lecubin, G. Romier, C. Surace, L. Thierry de Ville d'Avray, C. Arvanitidis, A. Chenuil, M.E. Çınar, D. Koutsoubas, S. Sartoretto, T. Taton ; "IndexMed projects : new tools using the CIGESMED DataBase on Coralligenous for indexing, visualizing and data mining based on graphs". In : Sauvage S, Sánchez-Pérez J.-M., Rizzoli, A.E. (Eds.), *Proceedings of the 8th International Congress on Environmental Modelling and Software, Environmental modelling and software for supporting a sustainable future*, Vol. 3, pp.656-665, Toulouse, France, July 2016

This SAGES seminar is open to all.
Registration is free and open to any researcher. PhD and Master students are welcome.

- The first day will be devoted to the results of the workshops and the exchanges between ecological researchers and researchers in ICST.
- The second day will be dedicated to a reflection in the form of workshops concerning the theoretical locks to be lifted (ICST) concerning environmental data, and the elaboration of a response to a call for projects for a specific action on the graphs of environmental data.
- A third day will bring together small groups of researchers to structure the response to call for projects according to coherent Work Packages.