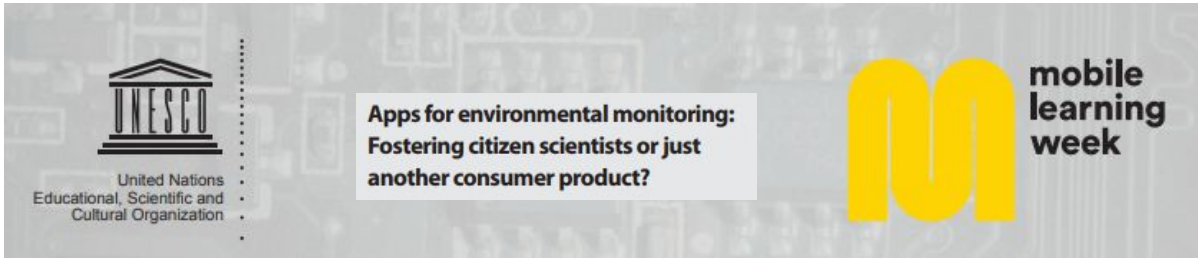
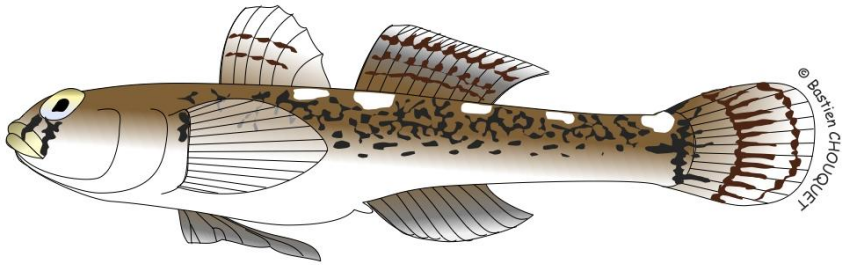
Computer aided identification apps and citizen science

Towards combining human and automatic process

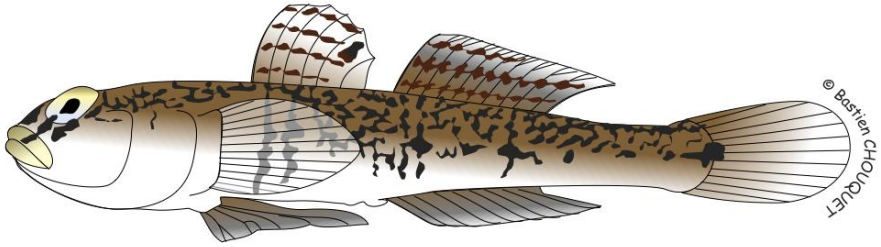
Regine Vignes Lebbe
Institut de Systematique, Evolution, Biodiversite
(Sorbonne Universite, MNHN, CNRS, EPHE, Univ. des Antilles)



Identification

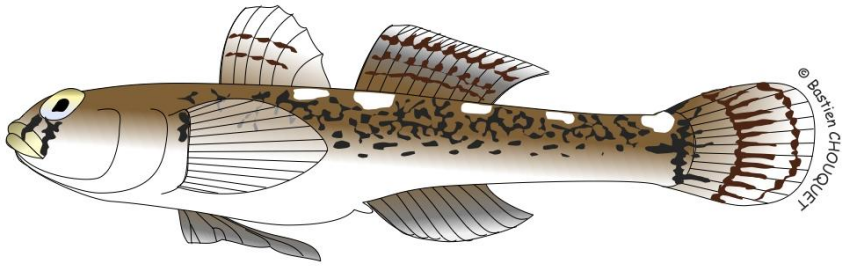


?

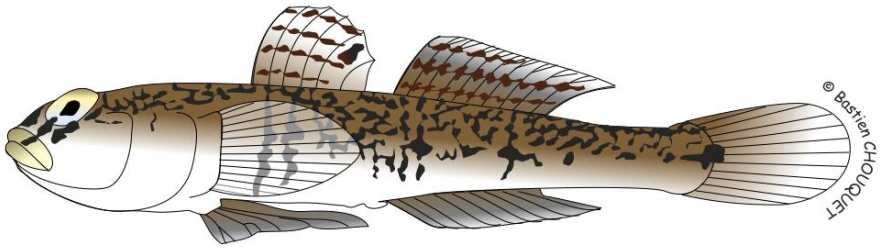


?

Identification



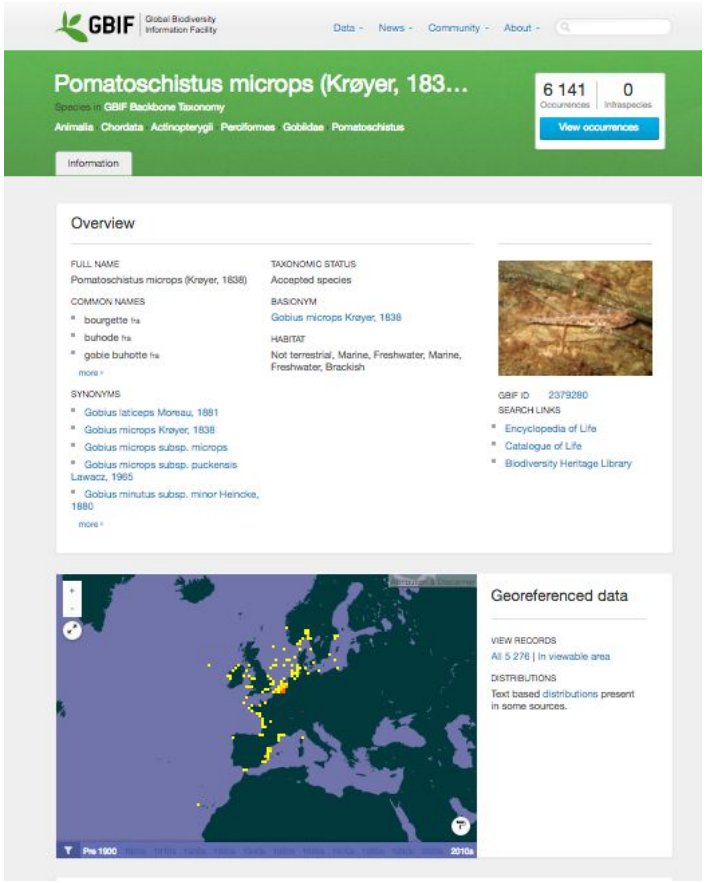
Pomatoschistus microps
(female)



Pomatoschistus minutus
(male)

Taxonomic names are the key access

Fishbase



GBIF

Identification is critical

For scientific activities (scientific names)

For discovering biodiversity (scientific and vernacular names)

For many citizen science project.

VIGIENATURE

Un réseau de citoyens qui fait avancer la science

How to help identification ?

Morphological characters (keys)

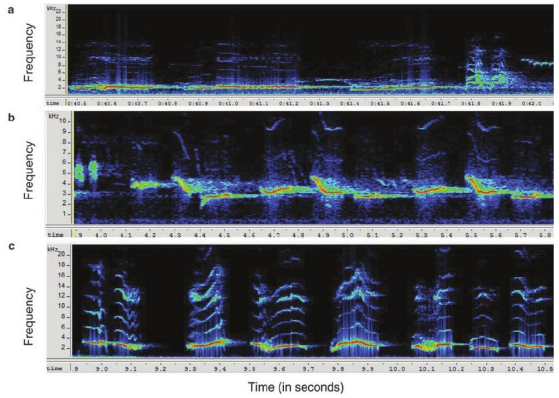
Measures and geometry (morphometry)

Molecular data (barcoding)

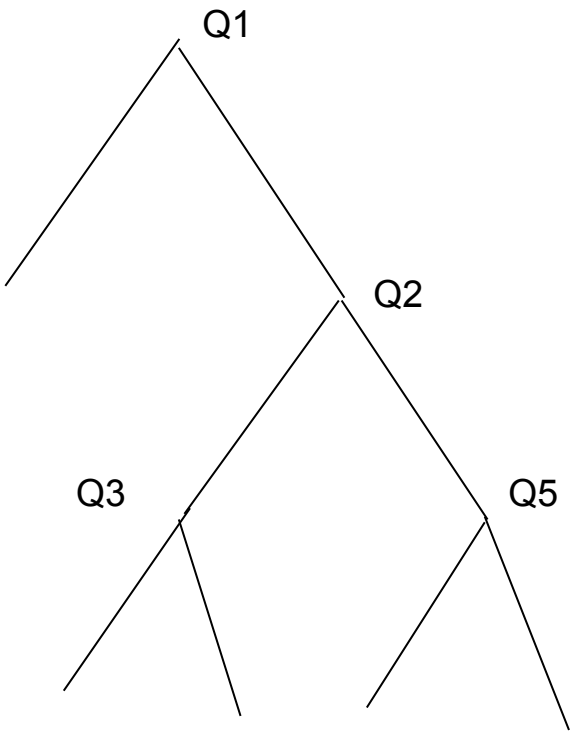
Images (data mining and neural network)

Acoustic

Behavior ...



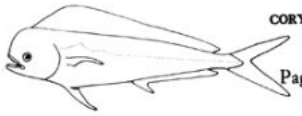
Identification key



Dicho/polytomous key
= Single access key
= Mathematical tree
= Decision tree

Section G: Pelvic Fins Thoracic With Exactly One Spine and Five Soft-rays (continued)

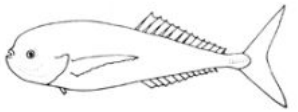
9b Tail forked; lateral line wavy .....



CORYPHAENIDAE Dolphin-fishes Pages 150–151

10a *From 7a page 22: spines in dorsal fin* Dorsal and anal fins with spines only, no soft-rays .....

10b Dorsal fin with both spines and soft-rays 11

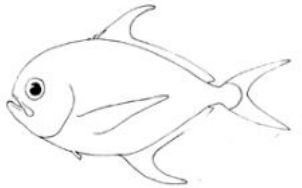


LUVARIDAE Louvar Page 198

11a Dorsal fin with 5 or more spines 13

11b Dorsal fin with only 3 or 4 spines 12

12a Snout blunt; area over eye enlarged; dorsal fin rays highest in anterior portion of fin .....



BRAMIDAE Pomfrets (fanfish) Pages 150–151

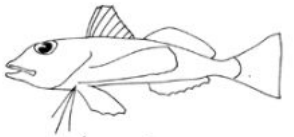
12b Snout elongate; dorsal rays highest in posterior portion of fin .....

CENTROLOPHIDAE Medusafishes Page 198

13a Lowermost rays of pectoral fin not detached from fin 15

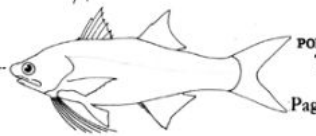
13b Lowermost rays of pectoral detached and thread-like or barbel-like 14

14a Three lowermost pectoral rays detached .....



TRIGLIDAE Searobins Page 113

14b Five to 9 lowermost pectoral rays detached .....

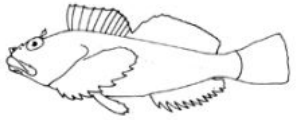


POLYNEMIDAE Threadfins (bobos) Pages 168–169

15a Suborbital stay absent..... 20 (next page)

15b Suborbital stay present..... 16

16a Anal spines absent; body scaleless or only partly scaled .....



COTTIDAE Sculpins (cabezon) Page 118

16b (next page)

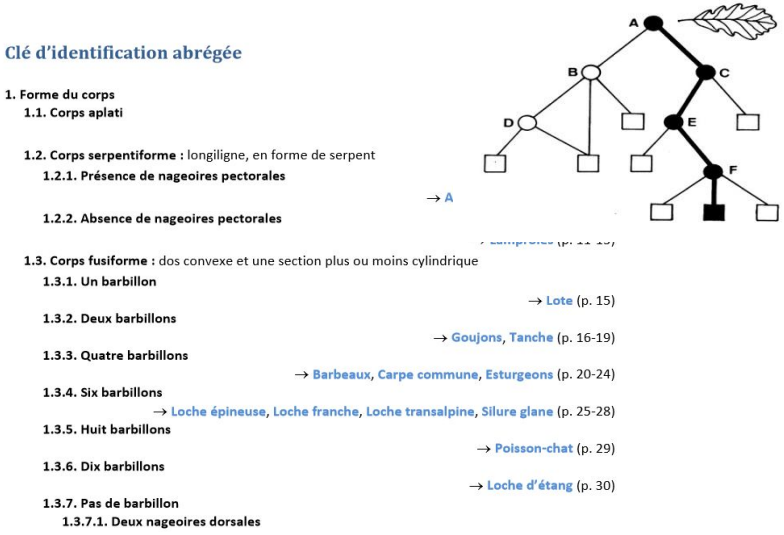
Guide To The Coastal Marine Fishes of California

Biological keys and algorithms

To compute a tree

To explore databases

Single access polytomous key

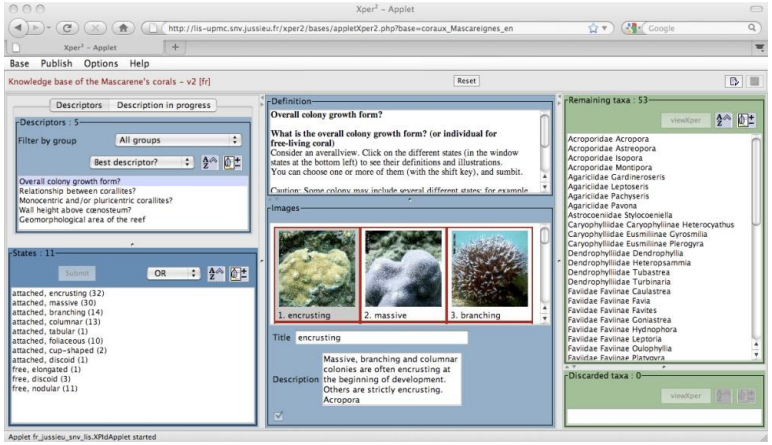


Clé d'identification abrégée

1. Forme du corps
   1.1. Corps aplati
   1.2. Corps serpentiforme : longiligne, en forme de serpent
      1.2.1. Présence de nageoires pectorales → A
      1.2.2. Absence de nageoires pectorales
   1.3. Corps fusiforme : dos convexe et une section plus ou moins cylindrique
      1.3.1. Un barbillon → Lote (p. 15)
      1.3.2. Deux barbillons → Goujons, Tanche (p. 16-19)
      1.3.3. Quatre barbillons → Barbeaux, Carpe commune, Esturgeons (p. 20-24)
      1.3.4. Six barbillons → Loche épineuse, Loche franche, Loche transalpine, Silure glane (p. 25-28)
      1.3.5. Huit barbillons → Poisson-chat (p. 29)
      1.3.6. Dix barbillons → Loche d'étang (p. 30)
      1.3.7. Pas de barbillon
         1.3.7.1. Deux nageoires dorsales

- Single access
- Answer must be « yes » / « no » at each question
- No fuzzy or vague answer

Interactive or **free access** Key



- Free access
- Possibility of doubting,
- Additional helpful information (text, images, links to external resources)
- Guidance to choose characters

Laboratory of Informatics and Systematics

Produce and publish your keys

Online collaborative system to create your own knowledge base and keys

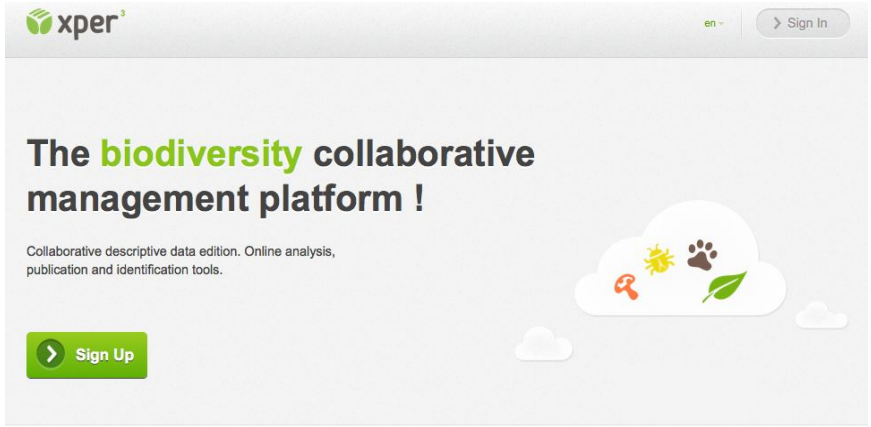
Open since october 2013

> 2500 accounts

> 3000 knowledge bases

xper³

http://www.xper3.fr/

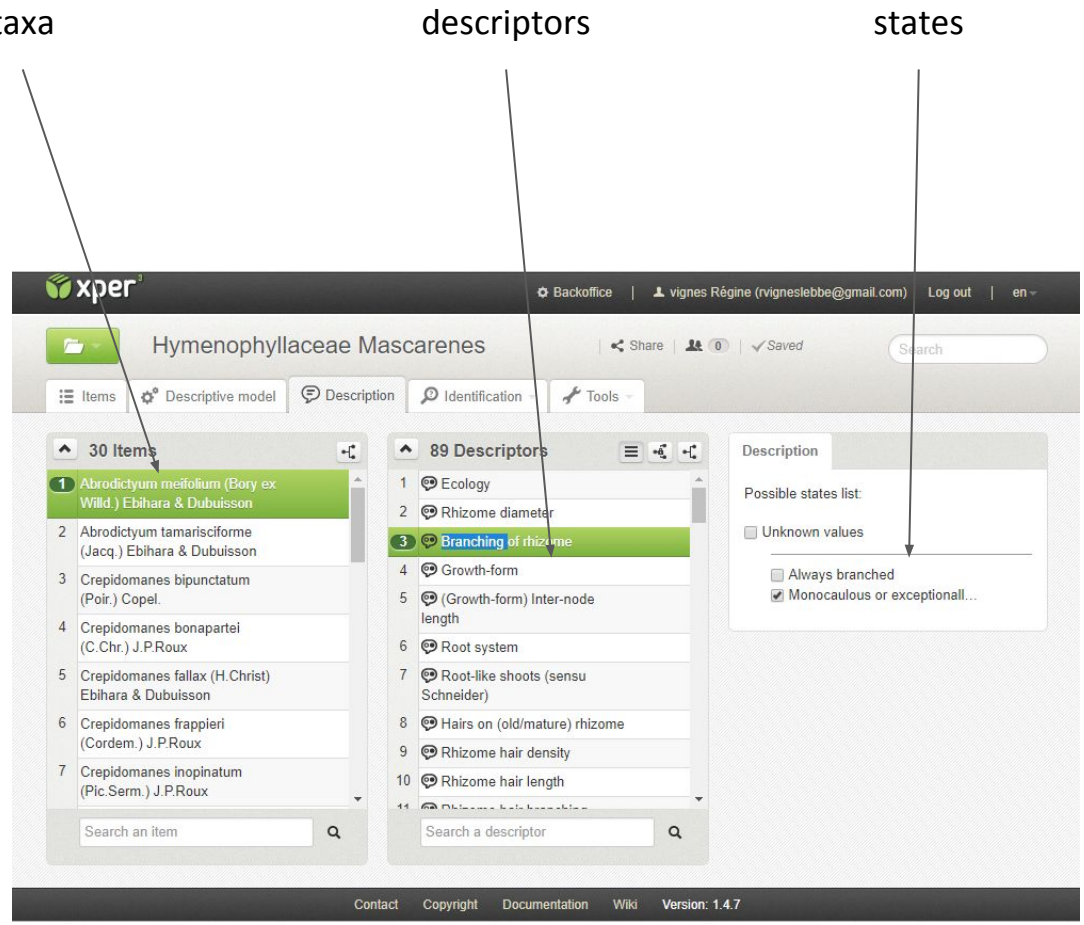


All taxa are described using a **common terminology**.

Taxa (items), descriptors and states are **documented** by texts and images.

The content is **accessible** for all if the author(s) publish the content and/or identification key online.

Here a knowledge base describing Hymenophyllaceae of Mascarenes.



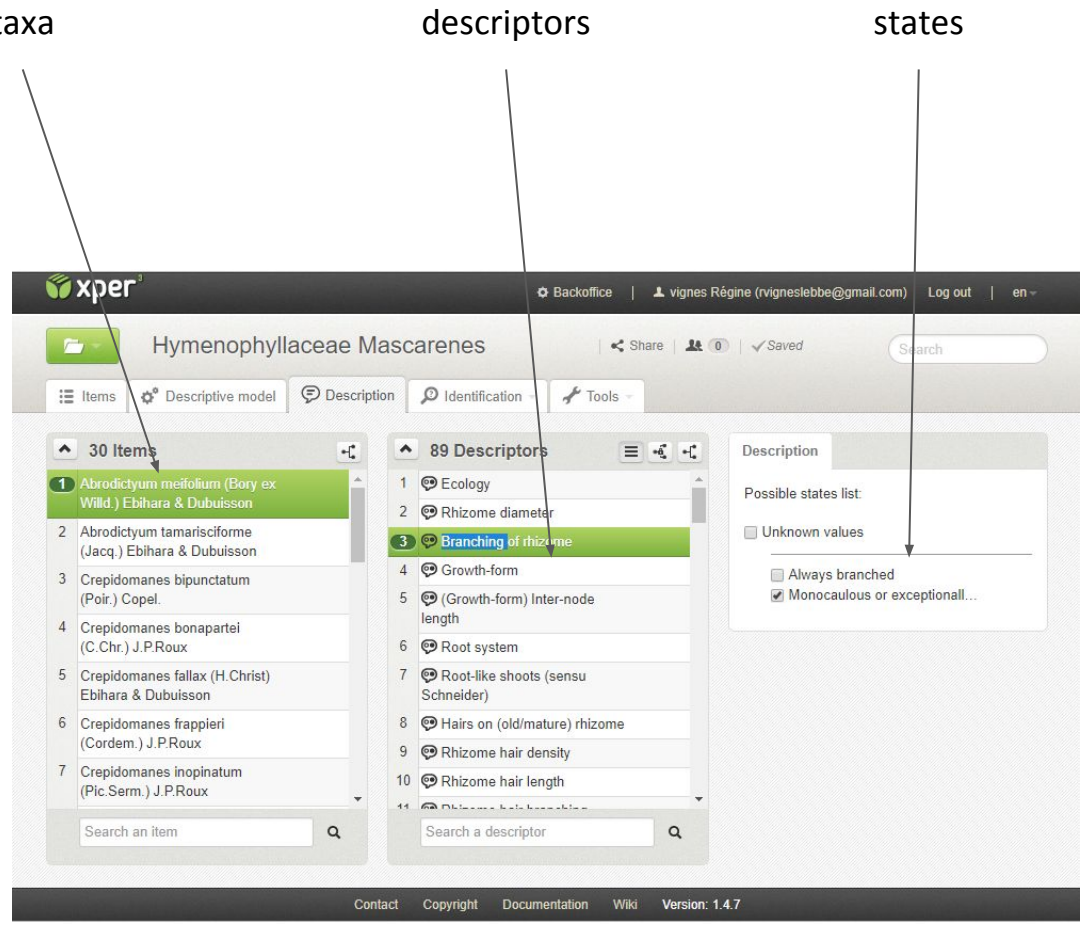
xper 3

All taxa are described using a **common terminology**.

Taxa (items), descriptors and states are **documented** by texts and images.

The content is **accessible** for all if the author(s) publish the content and/or identification key online.

Here a knowledge base describing Hymenophyllaceae of Mascarenes.



Litterature, collections, expertise

bottleneck

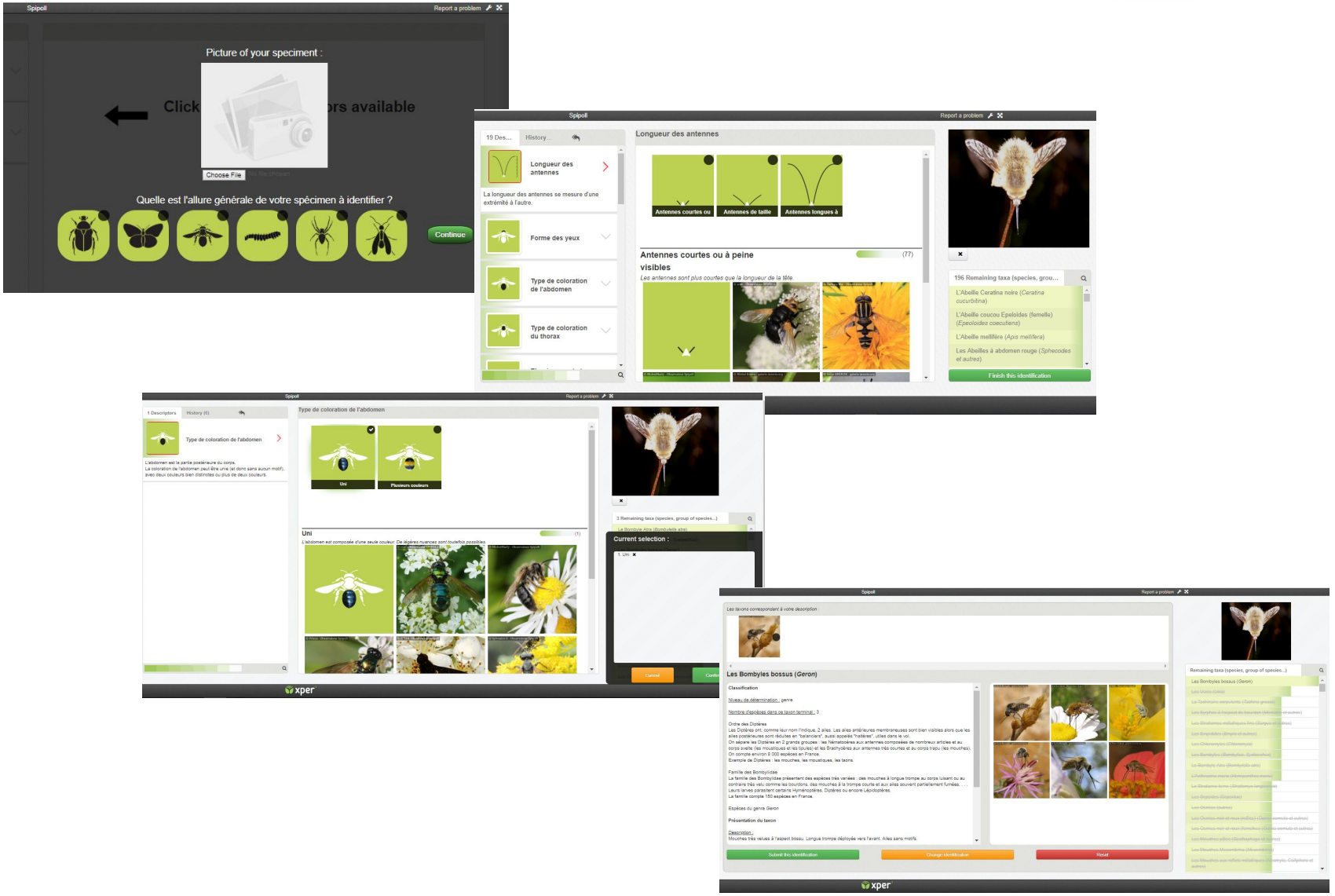
Knowledge bases, Structured data

Spipoll

Example of identification service for a citizen project about flower vsitors

http://spipoll.org

Spipoll: key demo online



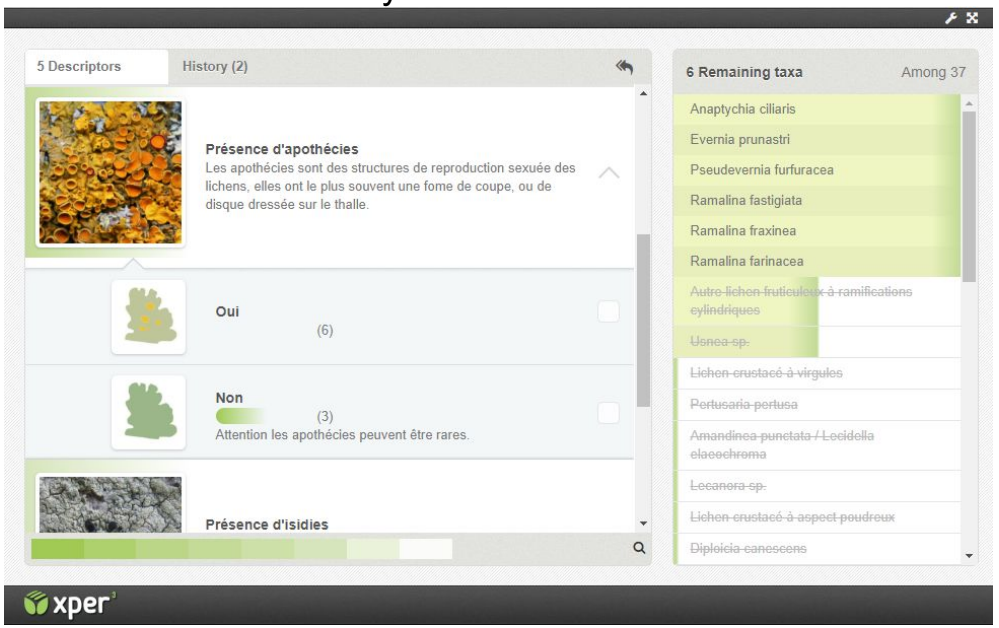
Lichen GO!

PARTICITAE

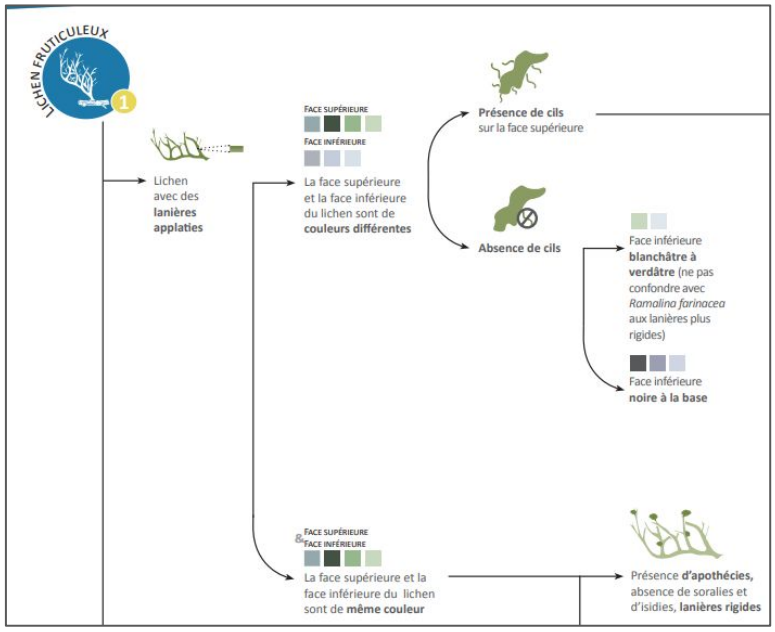
LICHENS-GO!

ÉVALUEZ LA QUALITÉ DE L'AIR EN OBSERVANT LES LICHENS

Online free access key



Printed single access key



Mobile access in the field ...

online

xper 3

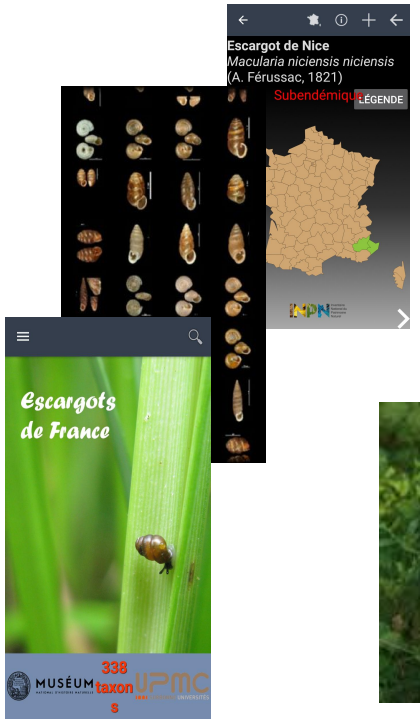
Common identifier

XML exchange format

preprocessing

SDD file to SQLite DB

Online or offline

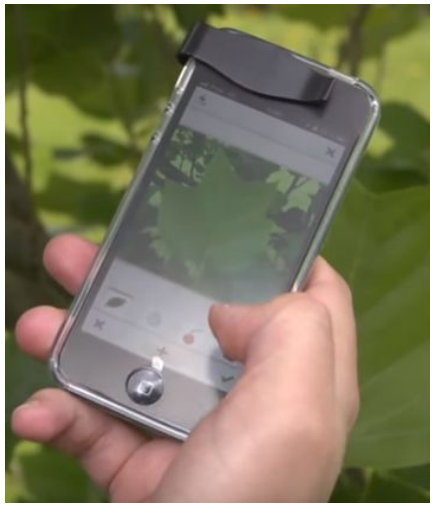


- Mollusca
- Carnivora
- Snakes
- Owls
- ...

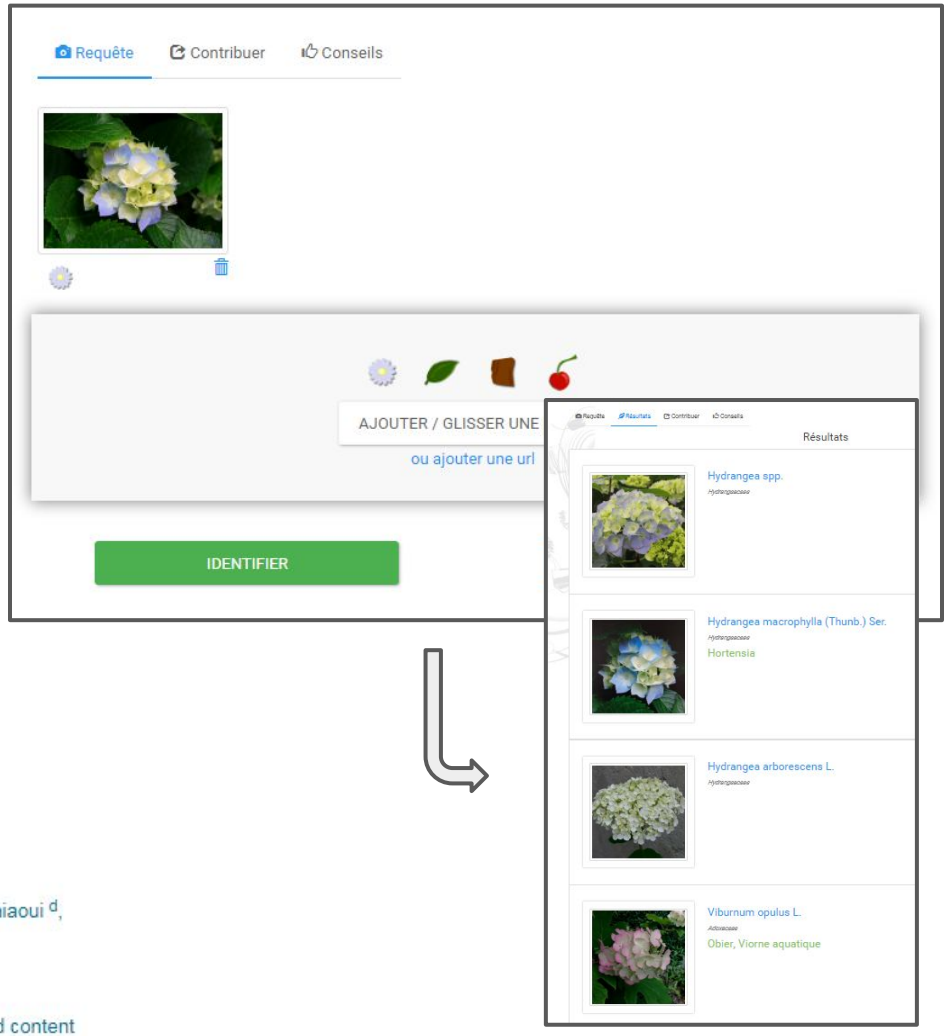
BiodiversiClés

Automatic method

Pl@ntNet



Pl@ntNet Version Web



https://hal.inria.fr/hal-01182775/document

Interactive plant identification based on social image data

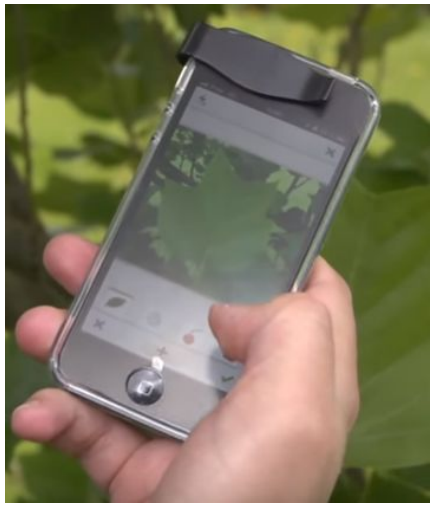
Alexis Joly [a], Hervé Goëau [a], Pierre Bonnet [b], Vera Bakić [a], Julien Barbe [c], Souheil Selmi [a], Itheri Yahiaoui [d], Jennifer Carré [e], Elise Mouysset [e], Jean-François Molino [f], Nozha Boujemaa [g], Daniel Barthélémy [h]

Show more

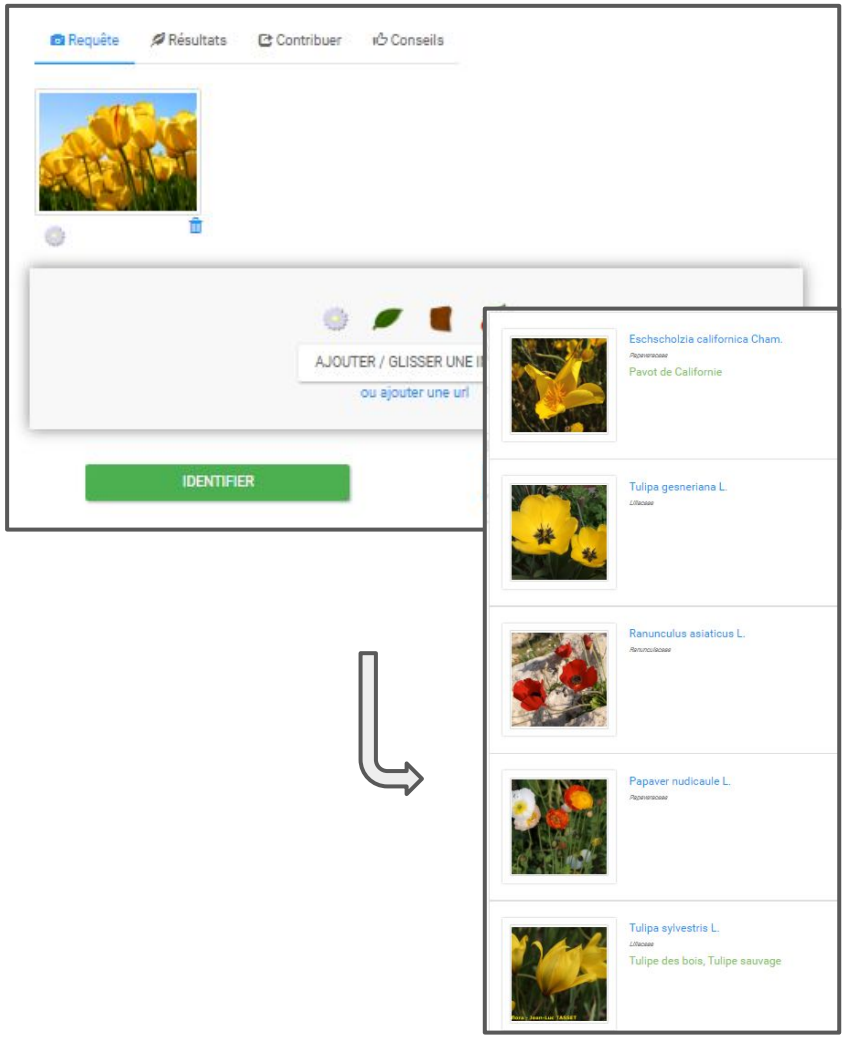
https://doi.org/10.1016/j.ecoinf.2013.07.006

Automatic method

Pl@ntNet



Pl@ntNet Version Web



Bsed on neural networks

Large training sets

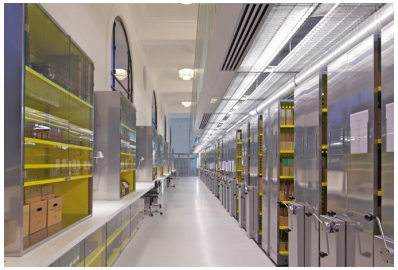
Take care to select the appropriate DB

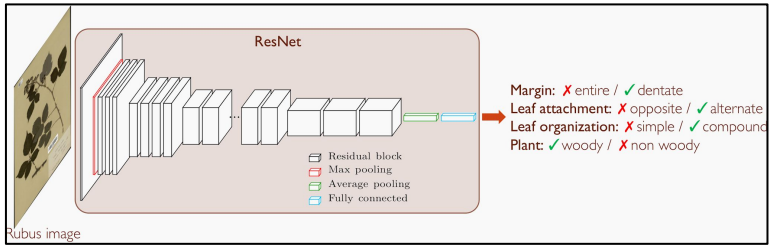
Look at several results

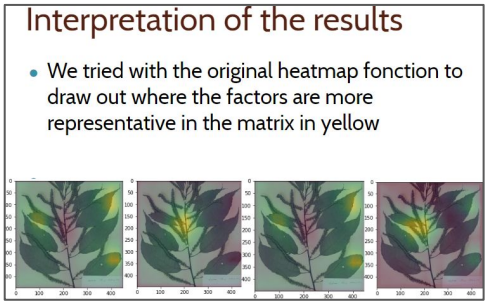
Deep learning to recognize character states

MNHN herbarium, Paris

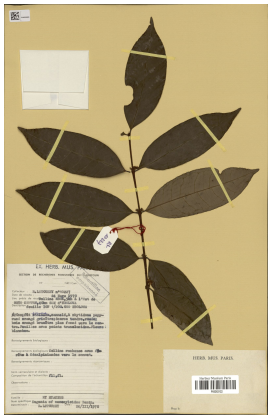
10 millions specimens





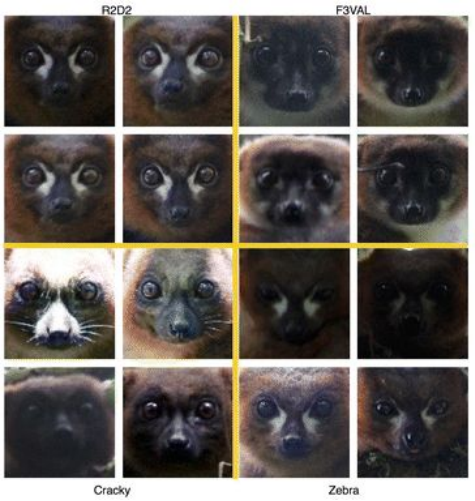


Simple leaf
Opposite phyllotaxy
Smooth border



(Stage Yue Zhu,
Ecole Polytechnique, 2017)
Encadrement du LIP6, SU.
Présenté à TDWG 2017.

Deep learning & biodiversity



Many applications specially on images and sounds at species level, and individual level

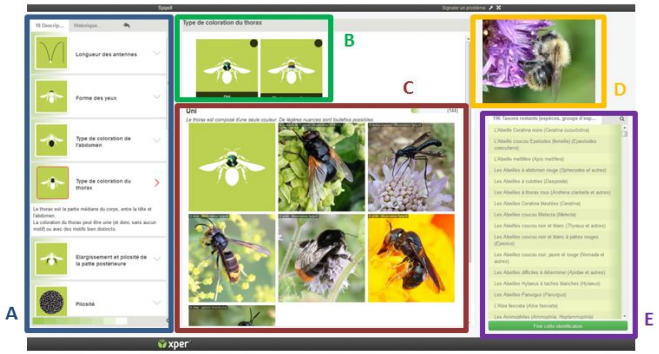
Deep learning with neural network:

- necessity to have a very **large set of images for training** the network
- the identification method gives a result but **no explanation**

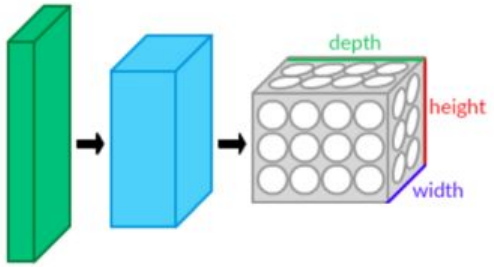
Interesting in monitoring survey, epidemiological survey, camera trap, satellite images …

But few interest at the moment for learning observation of nature And how to recognize.

Human versus automatic process



- Learn terminology of the domain
- Analytic process
- Active observation by human
- Explanation of the identification, why it is not another species



- No verbal description
- Synthetic, global approach
- Automatic,
- Statistical method
- No explanation

How to combine?

Combining human and automatic process

For identification:

Automatic **filter** to restrict identification key to possible taxa (GPS, season, picture recognition)

Backoffice automatic process to check user answer on character states, to propose **advise**, to pop up alert if there is a high probability of misobservation of a character, or miidentification of the species

To create Databases for identification:

Bottleneck of the human production of structured databases

Applying **text mining** to speed up the data harvesting

Teaching applications and knowledge transfer to the public

❖ **Teaching unit (Sorbonne University) : Biodiversity and paleobiodiversity (3ECTS)**
Leaves identification & co-existence approach

Saint Bauzile, France, Miocene

Identification

xper³

Quercus sp.

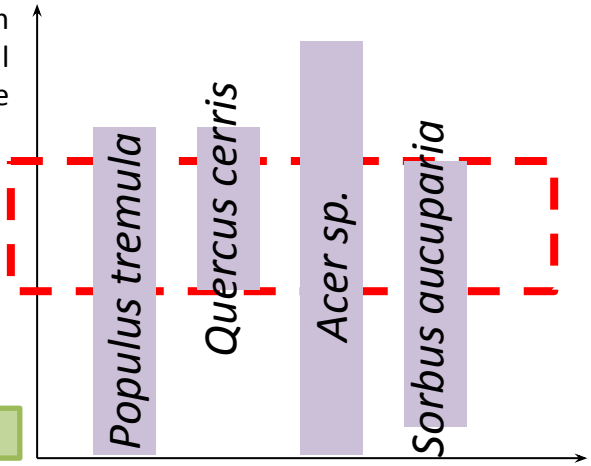
Populus sp.

Sorbus sp.

Acer sp.

Search for :
Nearest living relatives (NLR)
Climatic requirements of NLR

http://palaeoflora.de/



co-existence interval

Climate reconstruction for MAT, MAP, CMMT ...

Thank you for your attention