

Galaxy-Ecology



Galaxy
PROJECT



VIGIENATURE

Une plateforme d'analyse de données accessible

Simon Bénateau, Benjamin Yguel, Alan Amossé, Yvan Le Bras

Muséum National d'Histoire Naturelle

UMR Centre d'Ecologie et des Sciences de la COnservation

UMS PATRImoine NATurel

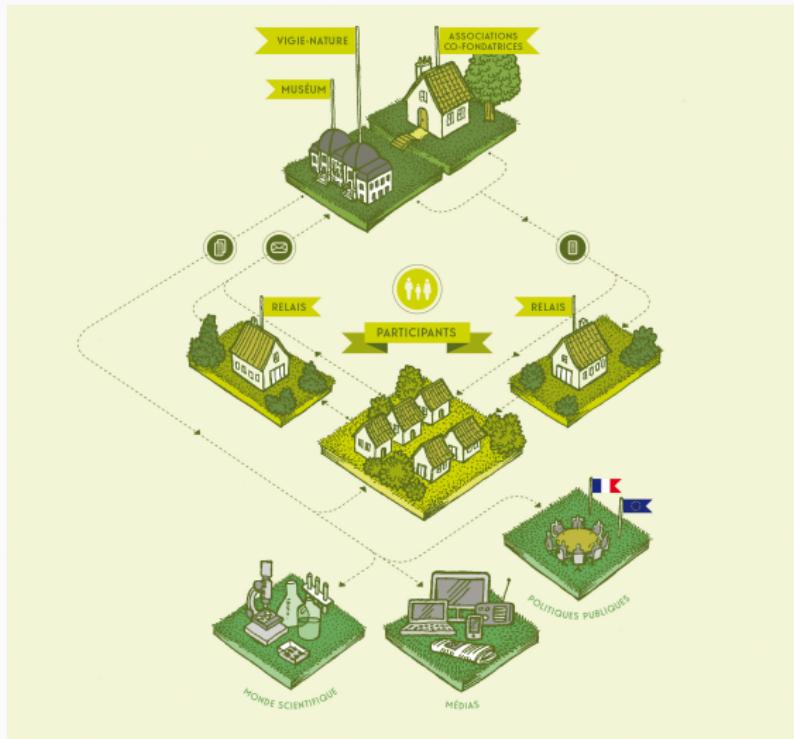
14 mai 2019

Ajouter Logo GAPARS / H2020 # Vigie nature



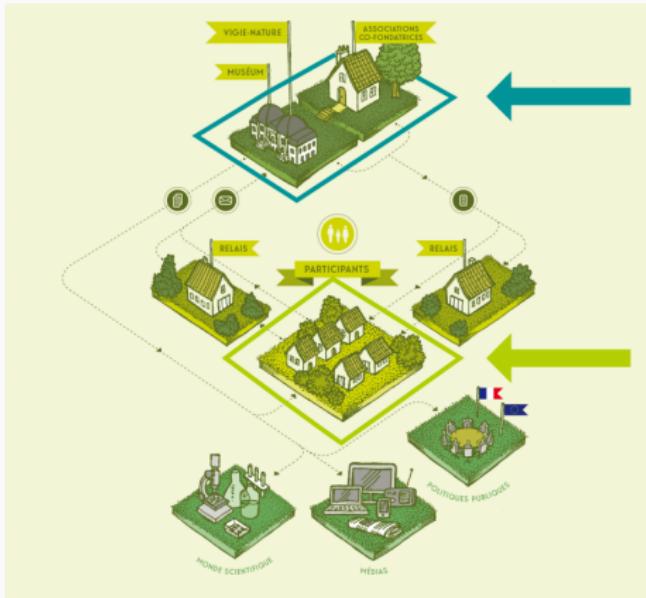
- bulle -> clarifier -> regarder slide Karine

Organisation du réseau d'acteurs



- clair sur les boîtes

Destination des plateformes d'analyse de données



Galaxy-Ecology

Plateforme pour les chercheurs et les acteurs des associations et institutions partenaires.

Galaxy-Bricks

Plateforme collaborative pour les participant.e.s aux observatoires.

- Bien mais à voir pour clarifier

Axes de développement et freins à l'implémentation

- Besoins principaux
 - Accès aux données
 - Accès outils d'analyse
 - Manipulation de fichiers
 - Analyses statistique
 - Visualisation)
 - Indicateurs régionaux (pour les gestionnaires)
 - Communication auprès des participants, des médias et des instances politiques
- Freins principaux
 - Formation au méthodes
 - Formation à l'utilisation des outils
 - Installation et utilisation des outils en autonomie
 - Dépendances (e.g. versions des packages)
 - Lignes de commande (utilisation de R)

Accès aux données

- Outils dédiés
 - Importation des données des observatoires
 - Importation depuis des bases de données naturalistes
 - gbif, bison, inat, ebird, antweb, ala, idigbio, obis, ecoengine, vertnet
 - Importation de données environnementales, covariables
 - Worldclim
 - Davantage de données bientôt
- Bibliothèques de données
- Import direct des données
 - URL
 - Fichiers locaux

Outils

```
date.serie <- as.POSIXlt(seq(as.Date(origin.d), length = nday,
dayno <- as.numeric(julian(date.serie, origin = as.Date(origin,
month <- as.numeric(strftime(date.serie, format = "%m"))
week <- as.numeric(strftime(date.serie, format = "%W"))
day <- as.numeric(strftime(date.serie, format = "%d"))
day <- as.numeric(strftime(date.serie, format = "%d"))
site_list <- sp_data[!duplicated(sp_data$SITE), c("SITE")]
all_day_site <- data.frame(SPECIES = sp_data$SPECIES[,], SITE =
YEAR = sp_data$YEAR[,], MONTH = month, WEEK = week, DAY = day,
COUNT = n0)
count_index <- match(paste(sp_data$SITE, sp_data$DAYNO, sep =
" ", sep = " "), all_day_site$COUNT)
all_day_site$COUNT[count_index] <- sp_data$COUNT
site_count.length <- aggregate(sp.data$COUNT, by = list(sp.data
names(site_count.length$) <- as.character(site_count.length$))
site_count <- utils::stack(site_count.length$)
all_day_site$COUNTNO <- NA
all_day_site$COUNTNO[count_index] <- site_count$noValues # add
# Add zero to close observation season two weeks before and after
first_obs <- mta(all_day_site$DAYNO[(ls.na(all_day_site$COUNT))]
last_obs <- max(all_day_site$DAYNO[(ls.na(all_day_site$COUNT))])
closing_season <- c((first_obs - 11):(first_obs - 7), (last_obs
# If closing season is before day 1 or day 365, simply set the
if (min(closing_season) < 1)
closing_season[1:1] <- 6(1:1)
if (max(closing_season) > nday)
```



- Flécher input code vers outil voir dev galaxy - ligne de commande R pour l'exemple

Outils

- Manipulation de fichiers
 - tri / filtre / jointure de table
- SIG
 - GDAL
 - traitement de fichiers raster / vecteur
- Analyse de données
 - Détection et identification de chauves-souris à partir d'enregistrements
 - Calcul d'indicateurs
 - RegionalGAM (Suivi de l'abondance de papillon)
 - Tendances STOC-EPS (Suivi temporel des oiseaux communs - Point d'écoute)

Outils

- Visualisation interactive
 - Librairie D3.JS (Histogramme, nuage de points...)
 - openlayers (Visualisation SIG)
- Autres outils interactifs
 - Jupyter notebook
 - Rstudio
- Intégration d'outils Shiny
 - Wallace (Modélisation de distribution d'espèces)

Possibilité de partage

- Données
- Outils
- Workflow
- Résultats
- Rapports



Galaxy Training!

Contributors View on GitHub Help ▾

Welcome to Galaxy Training!

Collection of tutorials developed and maintained by the worldwide Galaxy community

| Topic | Tutorials |
|---------------------------------|-----------|
| Introduction to Galaxy Analyses | 7 |
| Assembly | 4 |
| ChIP-Seq data analysis | 3 |
| Epigenetics | 3 |
| Genome Annotation | 3 |
| Metabolomics | 3 |
| Metagenomics | 3 |
| Proteomics | 12 |
| Sequence analysis | 5 |

| Topic | Tutorials |
|-----------------------------|-----------|
| Data Manipulation | 6 |
| User Interface and Features | 3 |

Galaxy for Developers and Admins

| Topic | Tutorials |
|------------------------------|-----------|
| Galaxy Server administration | 29 |
| Development in Galaxy | 13 |

How to contribute?

First off, thanks for taking the time to contribute!

You can report mistakes or errors, create more contents, etc. Whatever is your

Perspectives

- Gain de temps pour la production des indicateurs
- Créer des ponts entre les développeurs et les utilisateurs des méthodes
- Améliorer le partage des données et des méthodes d'analyse
- Gérer certains aspects des soucis de reproductibilité
- Faciliter le travail “interdisciplinaire”

Merci

