

# Galaxy-Ecology

Une plateforme d'analyse de données accessible

---



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

Muséum National d'Histoire Naturelle  
Centre d'Ecologie et des Sciences de la CONservation  
PATRImoine NATurel

14 mai 2019

# Galaxy-Ecology : Une plateforme d'analyse de données accessible

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

L'analyse des données issues des observatoires de sciences participatives nécessite l'emploi d'outils statistiques pointus qui requièrent des compétences en programmation. Ces dernières peuvent représenter un obstacle à l'apprentissage ou à l'utilisation d'outils par certains acteurs de la communautés des sciences participatives.

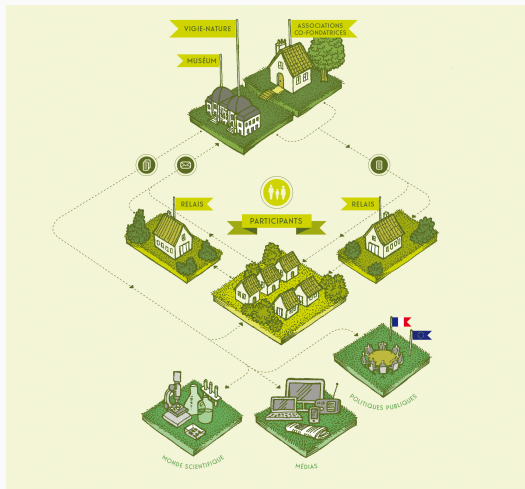
La plateforme galaxy-E permet de développer des outils avec une interface graphique, gérant les dépendances et la reproductibilité des analyses. Elle facilite l'accès à ces outils pour les acteurs des sciences participatives qui analysent les données et produisent des indicateurs.

Nous vous présenterons les derniers développement des outils de la plateforme et sur les projets.

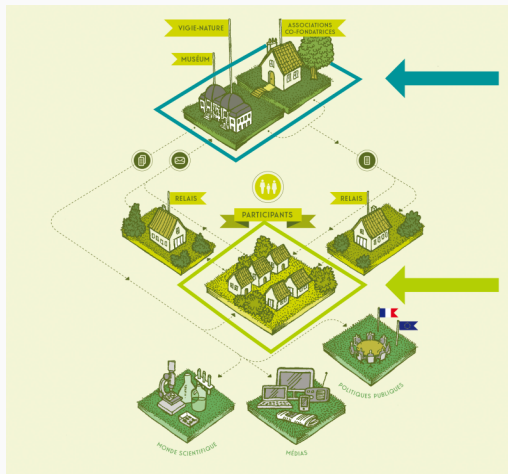
# Vigie nature



# Organisation du réseau d'acteurs



# 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.

# Demandes de la communautés et freins à l'implémentation

- Associations et structures partenaires
- Large communauté
- Compétences hétérogène en analyse et gestion de données
- Indicateurs régionaux
  - Communication auprès des participants, des médias et des instances politiques
- Accès aux données
- Installation et utilisation des outils
  - Dépendances (e.g. versions des packages)
  - Lignes de commande (utilisation de R)

- 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 écologiques
    - Worldclim
  - Davantage de données bientôt
- Bibliothèques de données
  - Créées par les utilisateurs
- Possibilité de faire un pont entre les disciplines

# 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"))
week_day <- as.numeric(strftime(date.serie, format = "%u"))
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$SPECIES[1], SITE =
YEAR = sp_data$YEAR[1], MONTH = month, WEEK = week, DAY = d
COUNT = NA)

count_index <- match(paste(sp_data$SITE, sp_data$DAYNO, sep =
sep = "-"))
all_day_site$count[count_index] <- sp_data$COUNT
site_count_length <- aggregate(sp_data$COUNT, by = list(sp_data
site_count_length$X <- as.character(site_count_length$GR
site_countno <- utils::stack(site_count_length$X)
all_day_site$countno <- NA
all_day_site$countno[count_index] <- site_countno$values # add

# Add zero to close observation season two weeks before and af
first_obs <- min(all_day_site$DAYNO[!is.na(all_day_site$COUNT)]
last_obs <- max(all_day_site$DAYNO[!is.na(all_day_site$COUNT)])

closing_season <- c((first_obs - 1):(first_obs - 7), (last_obs

# If closing season is before day 1 or day 365, simply set the
if (min(closing_season) < 0)
closing_season[1:5] <- c(1:5)
if (max(closing_season) > nday)

```

[illegible]



- Calcul d'indicateurs
- RegionalGAM
- Tendances STOC points d'écoute
- Traitement de données
- Détection et identification de chauves-souris à partir de d'enregistrements.

- Intégration d'outils shiny
- Wallace
- Autres outils interactifs
- Visualisation de données

- Données
- Outils
- Workflow
- Association d'outils
- Résultats
- Rapports

- Réalisé pour RegionalGAM

- 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
- Proposer une plateforme de formation