

# Collec-Science - version 2.2.3

Stockez et retrouvez vos échantillons !



Éric Quinton

`eric.quinton@irstea.fr`

IRSTEA – UR Écosystèmes aquatiques et changements globaux

*Document distribué sous licence CC-BY*

<https://creativecommons.org/licenses/by/4.0/fr/legalcode>



18 mars 2019

# Sommaire

- 1 Objectifs
  - Stockage
  - Traçabilité
  - Description
  - Droits
- 2 Technique
  - Plate-forme technique
  - Sécurité
  - Terrain
  - Obtenir le logiciel
- 3 Vie du projet
  - Site web
  - Listes de diffusion
  - Versions
  - Pérennité

# Objectifs

# Objectifs

- assurer la traçabilité des échantillons
  - ▶ où est-il stocké ?
  - ▶ qui l'a manipulé ?
  - ▶ d'où vient-il ?
  - ▶ qu'est-il devenu ?



Liste des conteneurs



Liste des échantillons



Entrée rapide d'un échantillon dans le stock



Entrée / sortie par scan multiple de QRCODE



Sortie rapide d'un échantillon du stock

# Où est-il ?

- gestion des containers (ou contenants) sous forme de « poupées russes »
  - ▶ types de contenants configurables
    - ★ sites, bâtiments, pièces, armoires, caisses...
  - ▶ position géographique d'une pièce
    - ★ affichage sur une carte
- étiquetage
  - ▶ qrcode sur les étiquettes
  - ▶ lecture optique pour les entrées/sorties
- toutes les entrées/sorties sont conservées
- possibilité de faire des inventaires
  - ▶ identifier les objets manquants

The screenshot shows a web application interface for container management. The main section is titled 'Détail d'un conteneur' (Container details). It includes a map showing the container's location at '3 Hall technologique'. Below the map, there are sections for 'Informations générales' (General information), 'Événements' (Events), 'Mouvements' (Movements), and 'Conteneurs présents' (Current containers).

**Informations générales**

- UID et référence: 3 Salle technologique
- Type: Invulsière - Pièce
- Statut: État normal
- Latitude: 44.734461460370
- Longitude: -0.68172031164950
- Emplacement: 3 Hall technologique

**Événements**

Date	Type	Reste disponible (échantillons)	Commentaire
Aucune donnée disponible dans le tableau			

**Mouvements**

Date	Sens	Conteneur	Emplacement	Commentaire	Utilisateur
28/01/2017 10:19:45	Entrée	3 Hall technologique	CTL1		eric.quenton

**Conteneurs présents**

UID	Identifiant ou nom	Autres identifiants	Statut	Type	Dernier mouvement	Conditions de stockage

## D'où provient-il ? Qu'est-il devenu ?

- saisie possible des coordonnées de prélèvement
  - ▶ coordonnées GPS
  - ▶ lieux pré-déterminés
    - ★ peuvent être associés à une collection
- un identifiant métier principal : identifier (ou *sample\_name*)
- ajout possible de tous types d'identifiants complémentaires
- exemples :
  - ▶ IGSN : numéro international des carottes prélevées en milieu aquatique
  - ▶ numéro d'inventaire
- possibilité d'indiquer la clé provenant d'une base externe
  - ▶ nom de la base + UID
- recherche multi-critères, sur tous les types d'identifiants

# Échantillon dérivé ou sous-échantillonnage ?

À partir d'un échantillon :

- on extrait un élément
  - ▶ il est identifiable individuellement (étiquetage dédié, p. e.)
  - ▶ il peut être d'un type différent
- exemples : un otolithe de poisson, une section de carotte de sondage
- c'est un **échantillon dérivé**

# Échantillon dérivé ou sous-échantillonnage ?

À partir d'un échantillon :

- on extrait un élément
  - ▶ il est identifiable individuellement (étiquetage dédié, p. e.)
  - ▶ il peut être d'un type différent
- exemples : un otolithe de poisson, une section de carotte de sondage
- c'est un **échantillon dérivé**
- on dispose d'une quantité de matière ou de matériel
  - ▶ on peut prélever ou remettre une parcelle de celui-ci pour analyse
  - ▶ il est impossible d'identifier précisément ce qui est prélevé par rapport au **tout**
- exemples : 5 écailles prélevées sur un poisson,  $10 \text{ cm}^3$  de matière
- c'est du **sous-échantillonnage**
- l'échantillon est le **tout**

## Description « métier »

- il est possible d'associer des données « métier »
  - ▶ schémas de métadonnées associés aux types d'échantillons
  - ▶ possibilité de recherche dans les métadonnées
  - ▶ stockage au format JSON
- schémas simples
  - ▶ une liste de champs
    - ★ multi-valeurs possibles dans certains cas (listes)
  - ▶ ne remplace pas une application dédiée
- les métadonnées peuvent être imprimées sur l'étiquette

# Gestion des habilitations poussée

- Cinq droits globaux gérés :
  - ▶ **consult** : visualiser les échantillons, les containers, les mouvements
  - ▶ **gestion** : créer des mouvements, des containers, des échantillons
  - ▶ **import** : import de masse
  - ▶ **collection** : import de masse, gestion des collections
  - ▶ **param** : paramétrage global
  - ▶ **admin** : gestion des droits et des utilisateurs

# Les collections

- Notion de **collection** :
  - ▶ ensemble cohérent d'échantillons représentatifs d'une famille de recherche :
    - ★ carothèque sédimentaire
    - ★ poissons capturés selon un échantillonnage normalisé dans un bassin versant
    - ★ peut comprendre des échantillons divers : poissons, otolithes, prélèvements génétiques, etc.
  - ▶ un échantillon est rattaché à une collection
  - ▶ seuls les personnes membres d'une collection peuvent :
    - ★ créer un nouvel échantillon
    - ★ modifier les informations
    - ★ visualiser les métadonnées
- très structurant : nécessite de bien réfléchir aux implications lors de la mise en place de Collec-Science

**Technique**

# Plate-forme technique

- Serveur Linux, Apache, PHP 7, Postfix pour l'envoi des mails (mots de passe perdus)
- Base de données Postgresql 9.5 ou ultérieur
- navigateur web pour accéder à l'application
- imprimante à étiquettes, douchette ou smartphone équipé d'un lecteur laser
  - ▶ pour la lecture optique des échantillons
  - ▶ matériels testés, liste disponible
  - ▶ modèles d'étiquettes proposés et adaptables par les utilisateurs
- possibilité de gérer plusieurs bases de données avec la même application

# Sauvegarde et duplication



- Sauvegarde de la base de données à programmer tous les jours
  - ▶ copie des *dump* de sauvegarde dans un second serveur
- si refus de perte d'un enregistrement :
  - ▶ mise en place d'une réplication vers un second serveur Postgresql nécessaire
  - ▶ procédure décrite

# Sécurité



- traces applicatives dans la base de données
  - ▶ qui fait quoi ?
  - ▶ purge au bout d'un an
- messages d'erreur stockés dans syslog (système de gestion de traces de Linux)
  - ▶ possibilité de moissonnage par des systèmes de gestion de traces
- en cas de mot de passe erroné :
  - ▶ blocage temporaire du mot de passe
  - ▶ envoi d'un mail aux administrateurs
- conformité du code vis à vis de la nomenclature ASVS  
[https://www.owasp.org/index.php/Category:OWASP\\_Application\\_Security\\_Verification\\_Standard\\_Project](https://www.owasp.org/index.php/Category:OWASP_Application_Security_Verification_Standard_Project)
  - ▶ résistance aux attaques dites opportunistes (niveau 1)
- logiciel homologué au titre de la sécurité des systèmes d'information par Irstea

# Saisie sur le terrain



- possibilité d'installer la plate-forme dans un Rasperry ou dans un micro-ordinateur portable (Windows ou autre)
  - ▶ procédure d'installation simplifiée (containers *docker*)
  - ▶ imprimantes portables, pilotées depuis le serveur
  - ▶ connexion des terminaux en wifi

# Saisie sur le terrain



- possibilité d'installer la plate-forme dans un Raspberry ou dans un micro-ordinateur portable (Windows ou autre)
  - ▶ procédure d'installation simplifiée (containers *docker*)
  - ▶ imprimantes portables, pilotées depuis le serveur
  - ▶ connexion des terminaux en wifi
- une fois de retour au bureau :
  - ▶ export des échantillons saisis dans un fichier CSV
  - ▶ import du fichier CSV dans la base centrale
    - ★ conservation de l'identifiant généré sur le terrain
    - ★ étiquette « terrain » utilisable
    - ★ association semi-automatique des tables de référence des bases « terrain » avec celles de la base centrale (types d'échantillon, p.e.)
  - ▶ mécanisme utilisable pour échanger des échantillons entre laboratoires

# Saisie sur le terrain



- possibilité d'installer la plate-forme dans un Raspberry ou dans un micro-ordinateur portable (Windows ou autre)
  - ▶ procédure d'installation simplifiée (containers *docker*)
  - ▶ imprimantes portables, pilotées depuis le serveur
  - ▶ connexion des terminaux en wifi
- une fois de retour au bureau :
  - ▶ export des échantillons saisis dans un fichier CSV
  - ▶ import du fichier CSV dans la base centrale
    - ★ conservation de l'identifiant généré sur le terrain
    - ★ étiquette « terrain » utilisable
    - ★ association semi-automatique des tables de référence des bases « terrain » avec celles de la base centrale (types d'échantillon, p.e.)
  - ▶ mécanisme utilisable pour échanger des échantillons entre laboratoires
- mise au point réalisée par
  - ▶ Julien Ancelin (INRA - Saint-Laurent de la Prée)
  - ▶ Christine Plumejeaud (Lienss - Université de La Rochelle)

# Licence

- Logiciel distribué sous licence Affero (AGPL)
- dépôt du logiciel auprès de l'Agence de Protection des Programmes par Irstea en octobre 2016
- code disponible dans <https://github.com/Irstea/collec>

## **Vie du projet**

# Site web

<https://www.collec-science.org>



Collec-Science

Stockez et retrouvez vos échantillons !

[Introduction](#)[Installation](#)[Tutoriels vidéo](#)[FAQ](#)[Communauté](#)[Report a bug](#)[Contact Us](#)

- nouveautés
- installation
- utilisation
- tutoriels vidéo
- FAQ
- vie du projet...

# Liste de diffusion

- Deux listes de diffusion

- ▶ <https://groupes.renater.fr/sympa/info/collec-users>
  - ★ poster des messages : ouvert à tous
  - ★ archives privées
  - ★ abonnements contrôlés
- ▶ <https://groupes.renater.fr/sympa/info/collec-dev>
  - ★ réservée aux développeurs

# Versions

- 6 versions majeures depuis octobre 2016
- 240 tickets ouverts sur la période
- base de données :
  - ▶ 31 tables dans le schéma des données
  - ▶ 10 tables dans le schéma de gestion des comptes et des droits
- version 2.2.3 :
  - ▶ 8700 lignes de code PHP pour l'application *stricto-sensu*
  - ▶ 9400 lignes pour le *framework* (le moteur gérant les droits, l'accès aux données, les modules, etc.)
  - ▶ 7400 lignes HTML et Javascript (interface graphique) pour l'application, et 1000 pour le *framework*
- qualité du code :
  - ▶ analyse réalisée avec SonarCloud :  
[https://sonarcloud.io/dashboard?id=lrstea\\_collec](https://sonarcloud.io/dashboard?id=lrstea_collec)

# Évolutions envisagées

- exportation d'un contenant et des échantillons présents, pour importation dans une autre instance (envoi d'une boîte avec des tubes à un autre laboratoire, p. e.)
- gestion des prêts d'échantillon
- mise en place de services web permettant de moissonnage d'une collection :
  - ▶ récupération de la liste des échantillons et des contacts associés
  - ▶ avec gestion des droits (ou collection définie comme publique)
- fonctionnalité de moissonnage :
  - ▶ soit implémentée dans Collec-Science
  - ▶ soit implémentée dans une application dédiée
  - ▶ choix à faire...

# Pérennisation

- le logiciel est aujourd'hui considéré comme stable
  - ▶ les fonctionnalités majeures sont disponibles
  - ▶ mais des évolutions sont toujours envisagées
- Une équipe « projet » réduite
  - ▶ une implication variable dans la durée (autres projets à mener)
- un seul développeur en continu
  - ▶ des renforts ponctuels (Chrono-Environnement et Lienss en 2017 et 2018)
  - ▶ des besoins particuliers nécessitant des compétences pointues
- un besoin d'appui :
  - ▶ formation des gestionnaires
  - ▶ appui méthodologique
  - ▶ compléter le site web (guides, exemples, vidéos, etc.)
- forte implication des Zones-Ateliers

**Comment pérenniser sur le long terme ?**