



# REDCap :

## Du terrain à la conformité, de la recherche clinique à l'écologie ?

Atelier “Carnets de terrain électroniques”

29 mars 2018

Montpellier

Thierry Valéro, Bruno Granouillac, et al.  
(IRD/MIVEGEC, IRD/TransVIHMI)



# *Contexte*



# Atelier 2016

« Un premier sondage fait apparaître qu'il y a une demande sur :

- **les outils Cybertracker/GeoODK**
- **la mise en place d'un système d'information complet (de la collecte de la donnée à leur mise à disposition dans une Base de Données). »**

D'après :

*Compte rendu de la Réunion 'Acquisition de données dans le cadre de suivis de populations' (Carnets de terrain électronique)*

Mercredi 13 Avril (12h00) - Jeudi 14 Avril (15h00) 2016, Sète, Station Marine de l'Université de Montpellier

Organisée avec le soutien de l'OSU OREME, de la Zone Atelier Antarctique (ZATA), et du CEFE (Plateformes 'Systèmes d'Information en Ecologie' et 'Programmes à Long Terme')





# Séminaire IRD / Plateau bioinfo 2017

## MIVEGEC & TransVIH avec le plateau bioinformatique. Retour d'expérience des systèmes de collecte de donnée sur le terrain (technologies ODK).

Ce séminaire a pour objectif de présenter les techniques informatiques de collecte de données sur le terrain, et les retours d'expériences des équipes de MIVEGEC, de TransVIHMI et d'autres unités proches de l'IRD et de partenaires.

Ce séminaire s'adresse principalement aux agents de l'unité MIVEGEC, des unités proches comme celles du plateau bio-informatique ou des départements SAS et ECOBIO, qui envisagent de mettre en oeuvre ces technologies.

- Introduction à la technologie **Open Data Kit (ODK)** et démonstrations pratiques par Aurélien Cheylan. (CEFE.)
- Retour d'expérience avec **Cybertracker**, par Emeline Sabourin (MIVEGEC).
- Retour d'expérience avec le prestataire **SurveyCTO**, par Sonia Fortin (Nutripass).
- Retour d'expérience avec le prestataire **ONA.io**, par Sébastien Pion (TransVIHMI).
- Retour d'expérience avec le système **Kobo Toolbox**, par David Chavernac et al. (CIRAD ASTRE).
- Retour d'expérience avec **Cybertracker**, par Marie-Claude Quido (CEFE).
- Retour d'expérience avec **REDCap** (ou ODK ?), par Frédéric Lardeux (MIVEGEC)



# REDCap

**REDCap est une application e-CFR** (electronic Case Report Form) disponible via un navigateur Web ou une application smartphone/tablette avec possibilité de collecte hors ligne. Il permet la gestion des bases de données et formulaires nécessaires à la réalisation de projet de recherche clinique.

**La gestion des données est un besoin prioritaire de l'unité MIVEGEC** et dépasse largement la collecte de données. La technologie ODK (Open Data Kit) reste incontournable mais atteint vite ses limites pour la gestion de la donnée et la « conformité ». D'un autre côté, des logiciels comme OpenClinica mettent trop l'accent sur des fonctionnalités spécifiques à la recherche clinique.

**Le choix de REDCap est un compromis** : le logiciel a été conçu pour la recherche médicale, mais il est utilisé au-delà de ce champ disciplinaire car il propose la traçabilité des données et un format standard pour exporter une étude pour archivage (format CDISC ODM).



# Financement : CEMEB



- Instrument CEMEB: **FINANCEMENT DE SEJOUR EN IMMERSION POUR FORMATION SCIENTIFIQUE OU TECHNIQUE –Techniciens, ingénieurs, doctorants, post-doctorants**
- **Objectif : Outils et gestion de données (REDCap) dans une institution.**  
Acquérir une expérience pratique dans la gestion des données et la mise en œuvre du système REDCap dans un cadre opérationnel (étude dans le domaine de la santé), en environnement tropical, avec des contraintes de « conformité », en suivant la donnée (« data flow ») et sa gestion par les différents intervenants lors d'une étude clinique, en relation avec le plan de gestion des données.
- **En relation avec :**
  - groupe « collection » du CEMEB ;
  - atelier Carnets de terrain électroniques à Montpellier (CEMEB, OSU OREME), 28 et 29 mars 2018.



# Intérêt pour CEMEB ?

Si des unités de la communauté CEMEB ont déjà une expérience avec la collecte de données sur le terrain ou avec des chaînes d'acquisition institutionnelles, ou avec la bio-informatique, aucune des unités de la communauté n'a d'expérience avec ces **techniques et méthodes issues de la santé**.

Les bailleurs, les éditeurs de revue, et les tutelles imposent de plus en plus de contraintes réglementaires (« conformité », qualité, traçabilité, archivage, mise en œuvre du protocole de Nagoya) sur la gestion des données.

L'expérience acquise par MIVEGEC avec REDCap serait utile aux autres unités du LabEx qui coordonnent des activités de recherche sur la biodiversité au sens large, et manipulent des bases de données recensant divers matériels biologiques provenant du monde entier, ou qui recueillent des données de questionnaires sur le terrain, soit directement (application mobile) ou a posteriori (questionnaire saisi avec une double entrée par ex.) .



# *Déroulement de la mission*





# Thématiques abordées

**Etudes cliniques** (avec REDCap)

**Gestion opérationnelle de l'informatique** et virtualisation.

**Technologies** : Docker, git, intégration continue (git et Quay), monitoring (Portainer), illustré avec REDCap

**Logiciel biobanque** (développement local)

**REDCap** (informatique)



# Staff de l'informatique (IT) ...en cours de recrutement

IT Manager	
Title and background	Task
<b>IT Officer,</b> <u>BSc, 1 année d'expérience,</u> Windows, Active Directory, etc.	<ul style="list-style-type: none"> <li>• Support users locally and remotely</li> <li>• Install and configure software</li> <li>• Run network cables &amp; make connectors</li> <li>• Troubleshoot (PC, Mac, Printers, Scanners, etc.)</li> </ul>
<b>System and Network admin</b> <u>BSc IT, 3 années d'expérience</u> Windows 7 & 10 et Server Virtualization (VMWare) Linux (Ubuntu, etc.), Cisco switches, etc.	<ul style="list-style-type: none"> <li>• Install and configure software</li> <li>• Troubleshoot PC, Switches, Servers, etc.</li> <li>• Maintain accounts Active Directory, MS Exchange</li> <li>• Implement and troubleshoot network issues</li> </ul>
<b>Web Developer / DB Designer</b> <u>BSc IT or higher</u> Experience in web applications Web, DB, HTML & PHP Linux, Windows, Docker, Git. WordPress, Dryupal, Alfresco, REDCap	<ul style="list-style-type: none"> <li>• <u>In collaboration</u>, analyze and formalize the needs</li> <li>• Propose optimized solutions</li> <li>• Develop, document, maintain and support</li> <li>• Follow the security best practices</li> <li>• <u>Advise and assist the other potential actors involved</u></li> <li>• Follow the trends of technologies</li> </ul>



# Périmètre et projets du service IT

L'informatique (IT) dépend du Directeur Administratif et Financier.

Fonctions couvertes :

- Infrastructure IT et réseaux.
- Administration RH,
- Compatibilité et finance,
- Dossier Patient Informatisé (DPI).
- Serveurs Web,
- Informatique BioMed (équipements de laboratoire).
- Biobanque,
- Serveur de gestion de données (REDCap),
- Autres systèmes de support à la recherche.

Difficulté à suivre / coordonner les projets de système d'information en appui à la recherche

- Avant le recrutement du **Web Developer / DB Designer**,
- Qui aura aussi des compétences en **système d'information scientifique**  
**« de gestions »**



# *Questionnement ?*



*“Although IT and data management are services provided to support research, and not ends in themselves, they can cause considerable problems to trial units and even invalidate results if not carried out properly.” [ECRIN-Raising-Standard]*

*“Les études cliniques génèrent une quantité croissante de données, mais il n’y a pas de standard de qualité suffisamment pratique, réaliste et en accès libre pour leur traitement par les structures académiques.” [ECRIN-Thérapie]*

Quelles sont les particularités et point d’attention pour la collecte et la gestion de données avec REDCap à MIVEGEC ? Y compris en environnement tropical.

Le contexte tropical, parfois qualifié de “*Resource Poor Environment*”, comporte des difficultés spécifiques bien connues comme l’accès à Internet, le réseau électrique, les ressources humaines, fournisseurs, etc.

Ceci étant dit, les développements techniques récents (mini NAS, conteneurs, REDCap) font baisser le ticket d’entrée aux outils.

Quel référentiel pour la collecte et la gestion de données, qui puisse servir pour MIVEGEC ? Et d’autres ?

Nos activités sont de plus plus en plus “régulées” et doivent démontrer leur conformité. Mais les critères de conformité (CNIL, RGDP, ISO, GxP, DMP, etc.) ne sont pas toujours connus à l’avance. Et ces critères ne sont pas suffisant pour piloter des activités de systèmes d’information.



# *Quel référentiel ?*

## *Réponse : ECRIN!*



# Réponse : ECRIN !

Référentiel ECRIN pour la conformité  
aux bonnes pratiques de gestion des données  
des essais cliniques multinationaux

Collette Croy<sup>1,2</sup>, Anne Dubouché<sup>3</sup>, Céline Cyprien<sup>4</sup>, Amélie Le Geyec<sup>5</sup>, Rachida Egan<sup>6</sup>, Isabelle Vignat<sup>7</sup>,  
Fabrice Bachelot<sup>8</sup>, Anne Lévesque<sup>9</sup>, Xavier Lapeere<sup>10</sup>, Céline Clavel<sup>11</sup>, Françoise Gaudelot<sup>12</sup>,  
Alexandra Elie<sup>13</sup>, Fabrice Akwa<sup>14</sup>, Hugues Chenuan<sup>15</sup> et Christine Bangue<sup>16</sup>  
et le Comité de Pilotage du groupe de travail d'élaboration de procédures de recueil des DC.

1. INSERM U1153, Centre de Recherche en Santé, Université de Bordeaux, France  
2. Centre de Recherche en Santé, Université de Bordeaux, France  
3. INSERM U1153, Centre de Recherche en Santé, Université de Bordeaux, France  
4. INSERM U1153, Centre de Recherche en Santé, Université de Bordeaux, France  
5. INSERM U1153, Centre de Recherche en Santé, Université de Bordeaux, France  
6. INSERM U1153, Centre de Recherche en Santé, Université de Bordeaux, France  
7. INSERM U1153, Centre de Recherche en Santé, Université de Bordeaux, France  
8. INSERM U1153, Centre de Recherche en Santé, Université de Bordeaux, France  
9. INSERM U1153, Centre de Recherche en Santé, Université de Bordeaux, France  
10. INSERM U1153, Centre de Recherche en Santé, Université de Bordeaux, France  
11. INSERM U1153, Centre de Recherche en Santé, Université de Bordeaux, France  
12. INSERM U1153, Centre de Recherche en Santé, Université de Bordeaux, France  
13. INSERM U1153, Centre de Recherche en Santé, Université de Bordeaux, France  
14. INSERM U1153, Centre de Recherche en Santé, Université de Bordeaux, France  
15. INSERM U1153, Centre de Recherche en Santé, Université de Bordeaux, France  
16. INSERM U1153, Centre de Recherche en Santé, Université de Bordeaux, France

18 mars 2016

## Quel référentiel pour la collecte et la gestion de données, pour MIVEGEC et al.

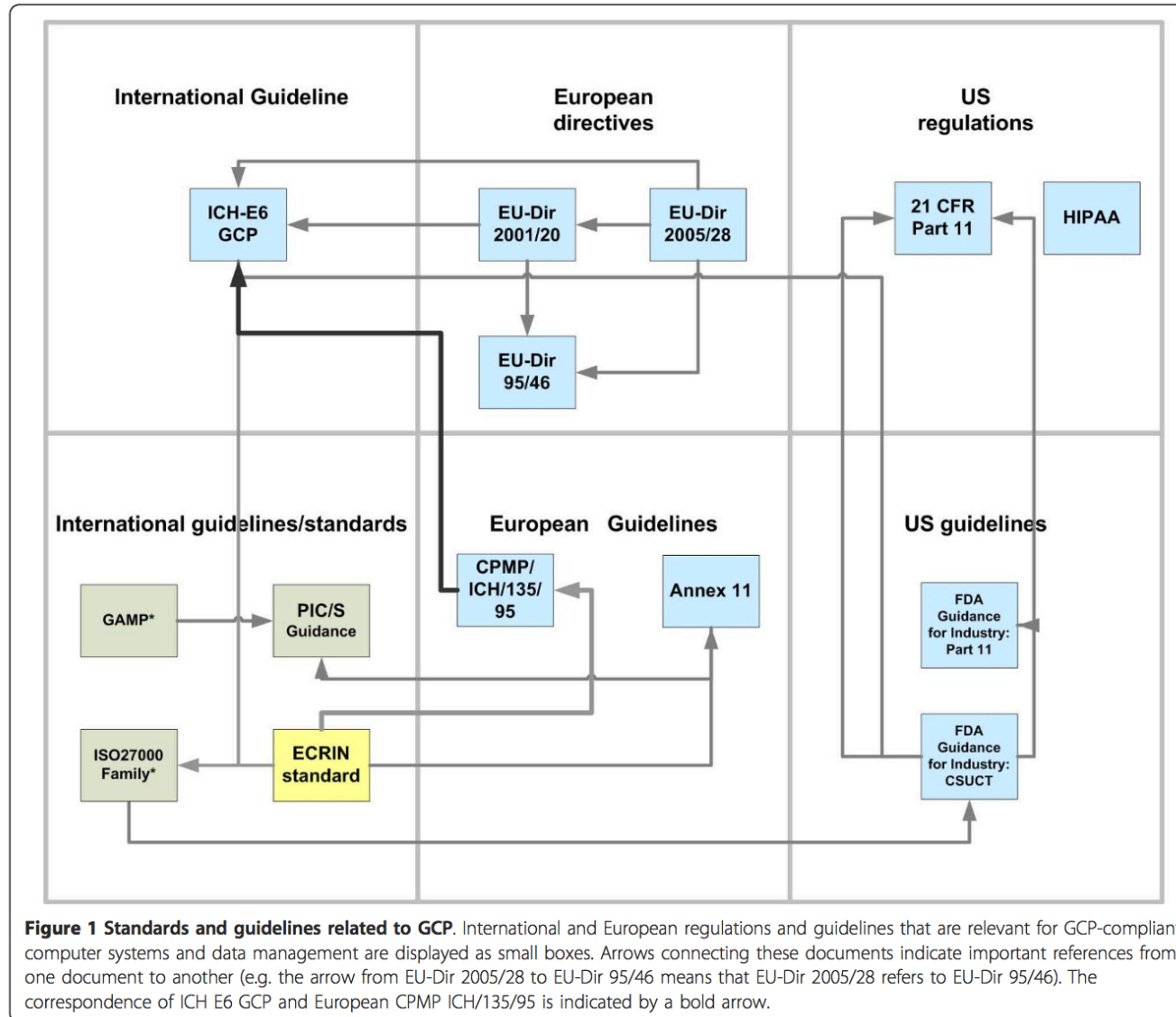
Référentiel ECRIN pour la conformité aux bonnes pratiques de gestion des données des essais cliniques multinationaux



Résumé : Contexte. Les études cliniques génèrent une quantité croissante de données, mais il n'y a pas de standard de qualité suffisamment pratique, réaliste et en accès libre pour leur traitement par les structures académiques. ECRIN (European Clinical Research Infrastructures Network) a publié un référentiel pour la certification de centres de gestion de données. Nous publions une version française du référentiel. Méthodes. Un groupe d'experts a produit ce référentiel par consensus, révisé après son utilisation lors d'audits pilotes pour la certification ECRIN. Résultats. Le référentiel comporte 21 listes de 5 à 10 exigences, divisées en trois groupes : technologies de l'information, gestion de données (data management [DM]) et « généralités ». Conclusions. Ce référentiel offre une description claire des bonnes pratiques en informatique et en DM. Initialement créé pour servir de référentiel de certification des centres de gestion de données ECRIN, il peut être utile à de nombreuses structures académiques développant ce type d'activités.



# Le référentiel ECRIN et les autres

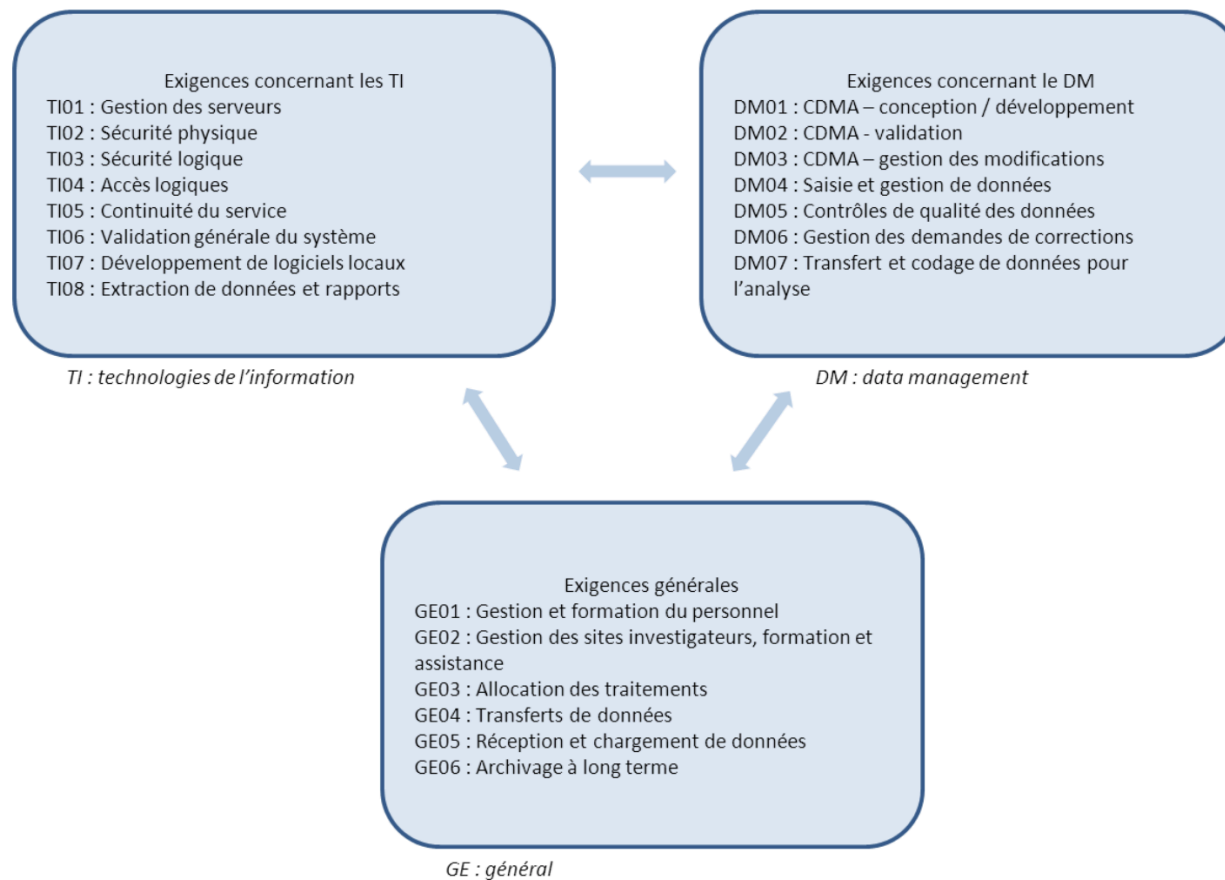


D'après Ohmann et al. *Trials* 2011, 12:85  
<http://www.trialsjournal.com/content/12/1/85>





# Le référentiel ECRIN IT, DM et GF



**Fig. 1.** Liste des exigences par thème. Avec l'aimable autorisation de Trials.<sup>[7]</sup>

D'après CORNU, Catherine, DONCHE, Anne, COFFRE, Carine, et al. *Référentiel ECRIN pour la conformité aux bonnes pratiques de gestion des données des essais cliniques multinationaux*. *Thérapie*, 2015.

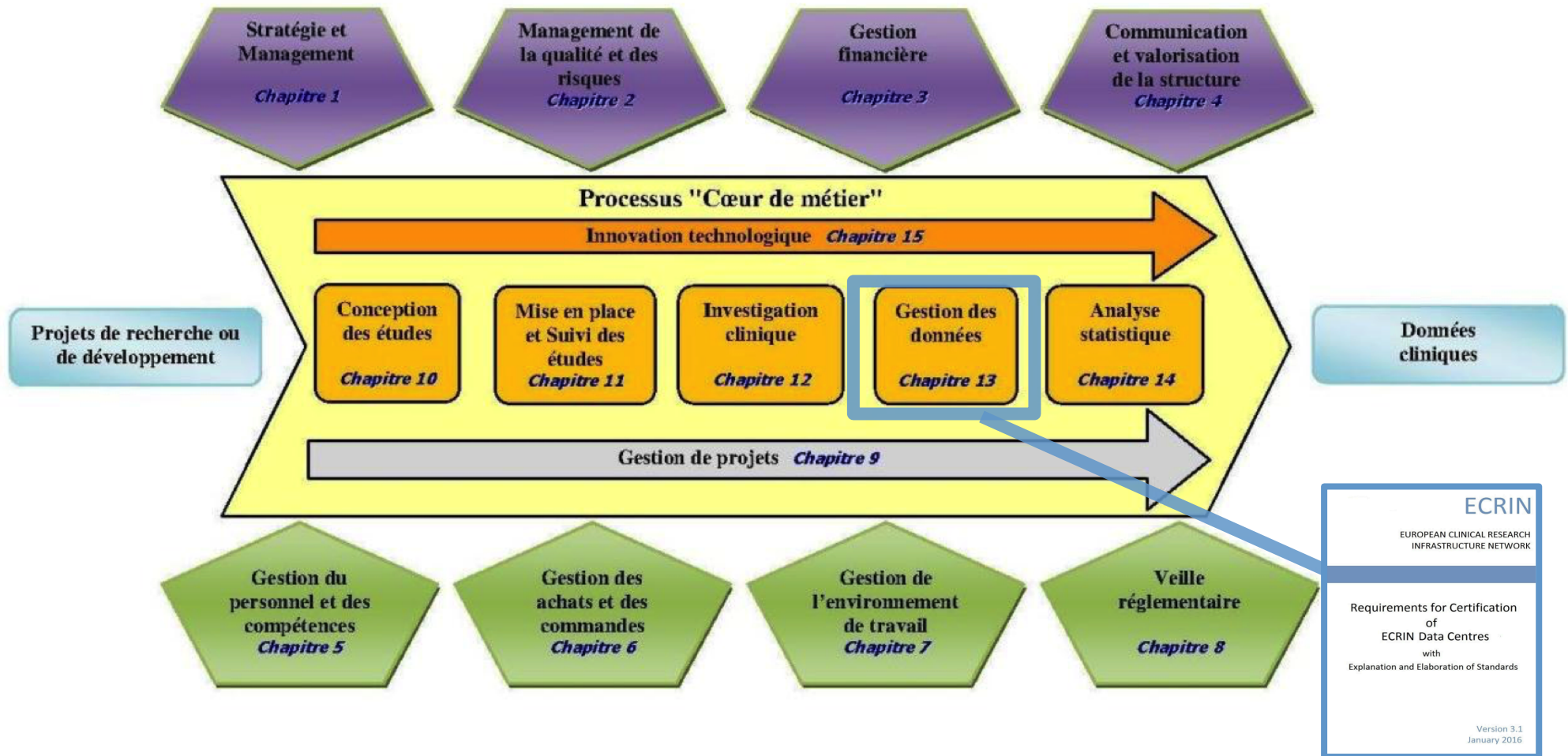
<https://www.journal-therapie.org/articles/therapie/abs/first/therapie150042/therapie150042.html>

Citant Lauritsen JM. *Implementation of good clinical practice software. Internal review of principles and requirements*, University of Southern Denmark, Institute of Public Health, Denmark, 2007 Obtainable from [jlauritsen@health.sdu.dk](mailto:jlauritsen@health.sdu.dk)



# ECRIN dans un cadre ISO 9001

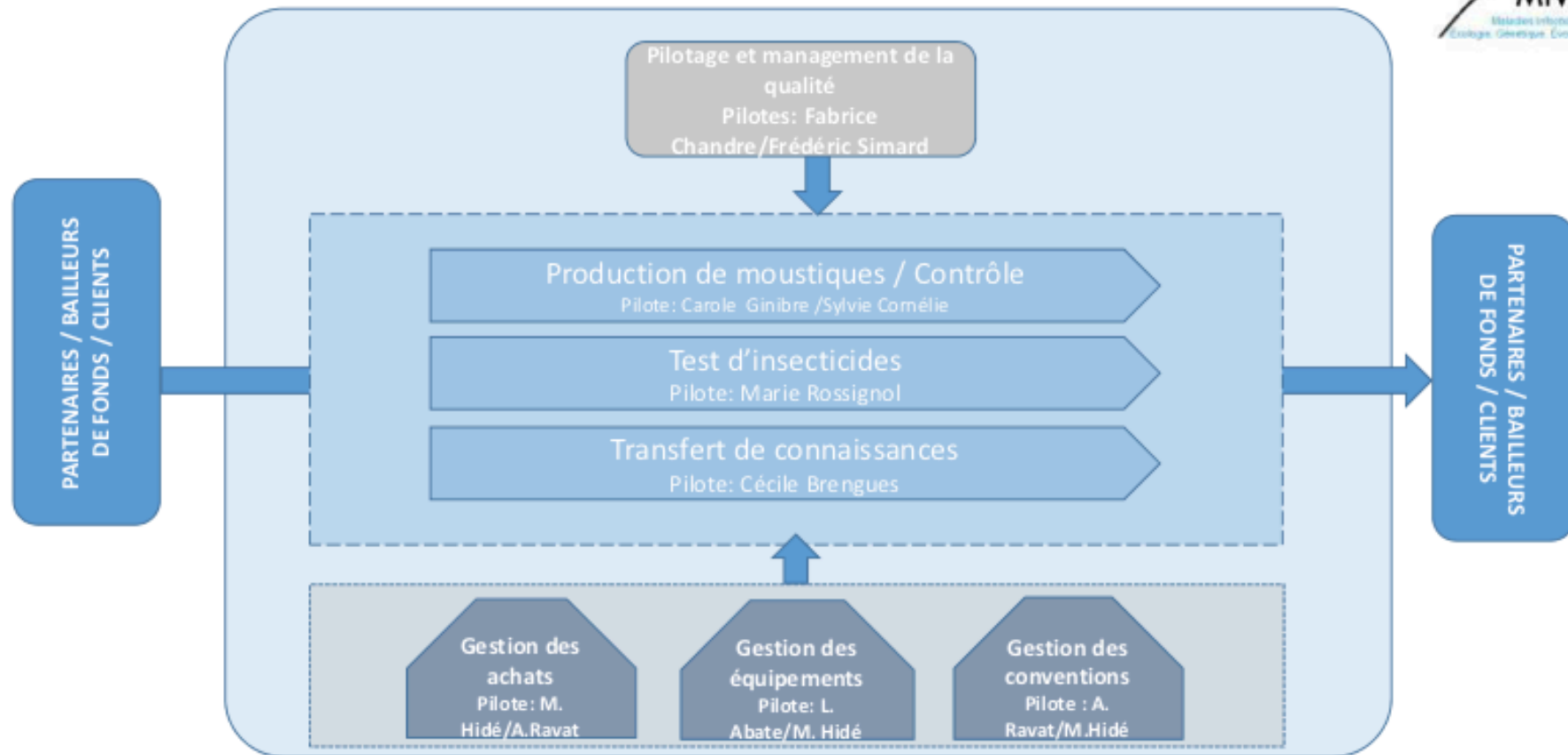
Manuel de bonnes pratiques professionnelles des CIC - Version n° 2 (2017)





# Comparaison : un autre périmètre ?

## CARTOGRAPHIE DES PROCESSUS de l'équipe De stratégie de lutte anti-vectorielle dans l'UMR MiVEGEC



DATE : 30/05/2016

CARTOGRAPHIE  
REALISE PAR : MR/CT/TZ

1-1



# ECRIN : Vocabulaire et Concepts

## Organisation du glossaire ECRIN

<ul style="list-style-type: none"> <li>• Système d'assurance qualité (SAQ)</li> </ul>	<ul style="list-style-type: none"> <li>• Validation</li> </ul>
<p><b>Outils techniques et supports matériels</b></p> <ul style="list-style-type: none"> <li>• Database management system (DBMS)</li> <li>• Clinical data management system (CDMS) comme REDCap, OpenClinica</li> <li>• Clinical data management application (CDMA)</li> <li>• Electronic data capture (EDC)</li> <li>• CRF (<i>case report form</i>) : pCRF , eCRF, iCRF, aCRF</li> </ul>	<p><b>Données</b></p> <ul style="list-style-type: none"> <li>• Données agrégées</li> <li>• Données anonymisées</li> <li>• Données pseudo-anonymisées</li> <li>• Données clinique</li> <li>• Données identifiantes des patients</li> </ul>
<p><b>Documents</b></p> <ul style="list-style-type: none"> <li>• Directives de travail,</li> <li>• Documents contrôlés</li> <li>• Instructions,</li> <li>• SOP, Procédures,</li> <li>• Manuel d'assurance qualité MAQ</li> </ul>	<p><b>Organisations</b></p> <ul style="list-style-type: none"> <li>• Organisation mère</li> <li>• Organisation hôte</li> <li>• Organisation hôte des IT</li> <li>• Centre</li> <li>• Site(s)</li> </ul>

Mais les exigences de gestion de données [DM] du référentiel impliquent un « Business Process » / « Data Management » spécifique aux essais cliniques (dont l'allocation des traitements, l'insu, le gel de la base, le consentement, etc.



# Des avantages... mais des limites

## Des avantages...

- **Les technologies** (Docker, Synology, REDCap) permettraient d'être facilement conformes, après quelques améliorations rapides et faciles (« low hanging fruit »).
- Le référentiel pourrait aider à établir **les rôles et responsabilité** des parties prenantes, y compris en facilitant la rédaction d'une convention d'hébergement ou autre SLA.
- Le référentiel pourrait aider à établir un diagnostic (**auto évaluation**) et à déterminer les actions, et à compléter un DMP (Data Management Plan).

## Des inconvénients

- **La « validation »** ([IT06] General System Validation) **semble hors de portée**, car il semble difficile de les interpréter. Faut-il appliquer les GAMP (Good automated manufacturing practice) ? Faut-il une méthode SDLC (« software development life cycle ») comme dans le GCDMP (« Good Clinical Data Management Practices ») ou les GLP (Good Laboratory Practices) ?
- **Les exigences de gestion de données [DM] sont peu adaptées à nos activités**, et très liées aux... **essais cliniques** (étapes de gestion de données avec « cliquets », gel de la base, etc..). Ce n'est pas adapté à des démarches plus « agiles » de recherche (Vectobol ?) ou moins contraintes (REACT ?).
- ... mais le GCDMP (Good Clinical Data Management Practices) n'est pas mieux !



# *Opportunité ? Les technologies !*





# VM Ware ? Docker ? REDCap ?

## What is Docker ?

La technologie "Open Source" de virtualisation légère à base de conteneurs ("Docker") est aujourd'hui incontournable.

Après avoir été déployée par le secteur privé [1], cette technologie a été évalidée par la recherche académique [2].

### Portabilité

Cette technologie peut être mise en oeuvre dans des armoires (racks), dans des mini serveurs (NAS en outremer), sur les postes de travail, pour le calcul parallèle (HPC), ou pour des nano serveurs de terrain.

### Références :

[1]Forrester Consulting et Dell EMC, Intel, And Red Hat, « Containers: Real Adoption and Use Cases In 2017 | Dell », mars 2017.

[2]A. Silver, « Software simplified », Nature, vol. 546, n° 7656, p. 173-174, 2017.

doi:10.1038/546173a



# Docker à MIVEGEC (avec REDCap)

La technologie Docker (aussi intégrée par Synology) est un vrai facilitateur (« enabler ») pour déployer et gérer des applications dans nos environnements.

- Pour tester, développer, évaluer ...  
...des serveurs ou des applications (composées de serveur)
- Plus légèrement (« agile »), selon de nombreux critères, que ...  
... au moins plus facile à utiliser / installer que ...
- Dans nos environnements favoris : postes de travail et ordinateurs portables, machines virtuelles, mini nas, etc.
- **Plus de sécurité et de qualité et de traçabilité** avec l'intégration continue : gestion des versions, traçabilité, analyse des images, qualité des réseaux, etc.
- Solution technique (temporaire) pour l'intégration continue.
  - Dev : docker-compose, git, gitLab CE, Quay.io
  - Ops : Quay.io, docker-compose, Portainer, WeaveScope, Cockpit





# REDCap (et Docker)

## Docker aideraient...

Docker permettrait de disposer d'**environnements de référence**, du développement à la production en passant par l'évaluation.  
Docker faciliterait la gestion et le déploiement d'applications comme REDCap (ou Galaxy ?).

## Transfert de technologie

- Copie du projet git REDCap de l'IRD,
- Installation d'un serveur git sur une VM,
- Installation et utilisation de SourceTree (git) et gitflow,
- Formation (très) rapide à Docker et docker-compose,
- Création d'un compte gratuit sur Quay.io
- Reconstruction (build) de tous les éléments,
- Envoi et analyse des images sur Quay.io
- Installation d'un porte-conteneur dans une VM Ubuntu (installation de docker, docker-compose, cockpit et cockpit-docker, et portainer).
- Téléchargement des images et personnalisation de la configuration,
- Lancement de l'application REDCap avec docker-compose dans le porte conteneur.



# REDCap, pour la gestion ?

Quand REDCap est installé dans une organisation...

... c'est un sérieux concurrent à Google Forms et Sheets, MS Excel et Access, avec des « projets » (bases de données RECap), pour des applications « secondaires » ou de « conformité ».

## Exemple : QHS (Qualité Hygiène et Sécurité) à MIVEGEC

### Transfert de technologie

- Démonstration des projets QHS (suivi des compétences, gestion des équipements, gestion des améliorations, formulaire de satisfaction), avec les supports de cours.
- Conseil et suivi pour l'élaboration de bases sur le serveur REDCap (pour le suivi des stagiaires)
- Description des fonctionnalités avancées (survey, formulaires répétés, droits d'accès),



# REDCap and Data Management

## Team Approach to Research Design





# Question ? Answer ?

42 !