TN09 : Stage assistant ingénieur

Conception d’une base de données pour

la mise en place d’un observatoire des agricultures

Cirad, Saint-Denis de la Réunion



Figure  : L'agriculture familiale

Encadrement :

Tutrice entreprise : Sandrine Auzoux

Responsable entreprise : Mialet-Serra Isabelle

Tutrice UTC : Marie-Hélène Abel

Stagiaire : Lucas Le Moine

Lucas.le-moine@etu.utc.fr

Du 01/09/2020 au 12/02/2021

Remerciements

Je souhaite remercier en premier lieu toutes les personnes qui ont rendu ce stage possible.

Tout d’abord, je remercie ma suiveuse de stage à l’UTC, madame ABEL pour l’appui qu’elle a pu être pendant cette période.

Ensuite, ma tutrice de stage, Madame AUZOUX mérite, pour son accueil, son aide et sa clairvoyance, mes plus sincères remerciements.

Je remercie également profondément Madame MIALET-SERRA, coordinatrice du projet de l’observatoire des agricultures de l’océan Indien, qui m’a permis d’effectuer ce stage au sein de ce groupe.

Je voudrai aussi remercier Messieurs Bosc et Bélières qui ont été d’excellents interlocuteurs et qui ont permis un contact avec le domaine métier durant la conception de la base de données.

Aussi, je remercie tous les membres du CIRAD à Saint Denis pour leur accueil et les connaissances qu’ils m’ont transmises de par la diversité de leurs domaines.

Enfin, je tiens à remercier madame LY pour ses réponses claires et rapides malgré la charge de travail qu’elle supporte.

Table des matières

[I. Table des illustrations 5](#_Toc60150566)

[II. Introduction 6](#_Toc60150567)

[III. Présentation du CIRAD 7](#_Toc60150568)

[A. Faits et chiffres 7](#_Toc60150569)

[B. Objectifs actuels 7](#_Toc60150570)

[C. Département 8](#_Toc60150571)

[D. Contexte du projet 8](#_Toc60150572)

[IV. Dans les grandes lignes 10](#_Toc60150573)

[A. Missions 10](#_Toc60150574)

[B. Méthodologie 10](#_Toc60150575)

[C. Outils utilisés 10](#_Toc60150576)

[1. CodeIgniter 10](#_Toc60150577)

[2. Leaflet 10](#_Toc60150578)

[3. Git et GitHub 11](#_Toc60150579)

[4. D3.js 11](#_Toc60150580)

[5. Bootstrap 11](#_Toc60150581)

[6. SQL Power Architect 11](#_Toc60150582)

[7. PostgreSQL 11](#_Toc60150583)

[8. Microsoft Office Access 11](#_Toc60150584)

[9. Gantt Project 12](#_Toc60150585)

[D. Planning 12](#_Toc60150586)

[V. Déroulement du projet 13](#_Toc60150587)

[A. Compréhension des intentions 13](#_Toc60150588)

[B. Bibliographie 13](#_Toc60150589)

[C. Benchmark 13](#_Toc60150590)

[D. Recueil des besoins des utilisateurs 15](#_Toc60150591)

[E. Recueil des données 16](#_Toc60150592)

[F. Réalisation du cahier des charges 16](#_Toc60150593)

[G. Conception de la base de données 17](#_Toc60150594)

[1. Général 17](#_Toc60150595)

[2. Contraintes 17](#_Toc60150596)

[3. Problèmes rencontrés 19](#_Toc60150597)

[H. Intégration de 2 bases de données 19](#_Toc60150598)

[I. Vues 20](#_Toc60150599)

[J. Réalisation du prototype 20](#_Toc60150600)

[1. Base de l’application Web 20](#_Toc60150601)

[2. Data visualisations 22](#_Toc60150602)

[K. Documentation 23](#_Toc60150603)

[VI. Bilan d’expérience 24](#_Toc60150604)

[A. Difficultés 24](#_Toc60150605)

[B. Apports personnels 24](#_Toc60150606)

[VII. Glossaire 25](#_Toc60150607)

# Table des illustrations

[Figure 1 : L'agriculture familiale 1](#_Toc60150608)

[Figure 2 : Carte du CIRAD 7](#_Toc60150609)

[Figure 3 : Logo de la FAO 8](file:///C:\wamp64\www\OA-OI_git\Autres_livrables\Rapport_de_stage.docx#_Toc60150610)

[Figure 4 : Benchmark 14](#_Toc60150611)

[Figure 5 : Partie du diagramme Power Architect 18](#_Toc60150612)

[Figure 6 : Schéma architecture Modèle\_Vue\_Controleur 21](#_Toc60150613)

[Figure 7 : L'observatoire 22](#_Toc60150614)

[Figure 8 : Composition des revenus 22](#_Toc60150615)

[Figure 9 : Dataviz Afrique 23](#_Toc60150616)

[Figure 10 : Histogramme du CIRAD 26](#_Toc60150617)

# Introduction

Le stage d’assistant ingénieur TN09, dans le cadre du parcours universitaire de l’Université Technologique de Compiègne, branche Génie Informatique a été réalisé au sein du Centre de Coopération Internationale en Recherche Agronomique pour le développement (CIRAD). Ce stage de 24 semaines a été effectué en télétravail entre le 1er et le 18 septembre, puis sur le site de la direction régionale du CIRAD à Saint Denis de la Réunion entre le 18 septembre et le 12 février. Il n’a pas été affecté par le reconfinement fin octobre.

Ce rapport de stage contient une brève présentation de l’institut d’accueil et du contexte du projet. Sont ensuite discutés les objectifs et des livrables attendus ainsi que les technologies employées. Enfin, certaines réflexions personnelles seront évoquées à la fin du rapport.

# Présentation du CIRAD

## Faits et chiffres

Le CIRAD est un Établissement Public à caractère industriel et commercial (EPIC) français fondé en 1984. Il est spécialisé dans la recherche agronomique dans les régions aux climats chauds. Ce type d’établissement est créé pour répondre à un besoin qui, pour des raisons concurrentielles, ne peut pas être satisfait par une entreprise privée. En raison de la diversité de ses activités, le CIRAD est à la fois placé sous la tutelle du ministère de l’Enseignement supérieur et de la Recherche et du ministère des Affaires étrangères.

Le CIRAD emploie actuellement 1650 personnes, dont 800 chercheurs, répartis en 3 départements scientifiques et 33 unités de recherche (mixtes ou propres). Le partenariat fait partie de l’identité du CIRAD et les 12 directions régionales (dont celle de la Réunion) permettent une grande coopération. Par ailleurs, le partage des connaissances, la qualité de la recherche, l’ouverture internationale ainsi que l’engagement pour le développement sont les clés de l’action de cet établissement.



Figure : Carte du CIRAD

## Objectifs actuels

Le CIRAD vient de signer avec ses ministères de tutelle son nouveau Contrat d’Objectifs et de Performances pour la période 2019-2023. Le CIRAD s’engage autour de quatre grandes ambitions pour contribuer à l’atteinte des Objectifs du développement durable au Sud, en particulier ceux sur l’éradication de la faim et de la pauvreté (ODD 2, ODD 1).

Pour se donner les moyens d’accompagner les sociétés du Sud dans des trajectoires de développement durable, le CIRAD s’engage ainsi sur quatre grandes ambitions pour la nouvelle période 2019-2023 : la science, le partenariat, la formation et l’impact.

## Département

Le stage s’est déroulé au sein de la Direction Régional Mayotte Réunion (DRRM). Néanmoins, Sandrine AUZOUX, la tutrice entreprise lors du stage, fait partie de l’unité de recherche Aïda (recherche "Agroécologie et intensification durable des cultures annuelles"). Cette unité se positionne sur l’intensification et la durabilité de la production des cultures annuelles en quantité et, lorsque pertinent, en qualité, en milieu tropical particulièrement contraint. Pour cela, ses recherches visent la pleine valorisation des ressources disponibles, en mobilisant les processus écologiques qui régissent leur dynamique au sein des agrosystèmes. Sandrine AUZOUX est spécialiste de la gestion des données et était donc parfaitement compétente pour superviser le stage.

## Contexte du projet

La décennie 2019-2028 a été proclamée comme celle de l’agriculture familiale par les Nations Unies. Ce type d’agriculture représente pas moins de 90% des exploitations agricoles dans le monde qui sont ainsi au nombre de 500 millions. Elles produisent également 80% de la production alimentaire mondiale alors que près des 3 quarts d’entre elles ne dépassent pas un hectare. Pourtant, la majorité de ces agriculteurs sont touchés par la pauvreté et sont vulnérables aux aléas météorologiques et aux fluctuations des prix des marchés.

Par ailleurs, les Nations Unies ont ciblé 17 objectifs du développement durable (ODD) à réaliser d’ici à 2030. Ces objectifs interconnectés ont pour but d’assurer que chaque être humain connaisse l’égalité, la prospérité, la santé, la justice et la paix plutôt que la faim et la pauvreté. Les deux premiers objectifs, et les plus importants, étant éradiquer la faim et la pauvreté, l’agriculture familiale a un rôle immense à jouer pour atteindre ces objectifs. Au total, ce ne sont pas moins de 10 des ODD qui sont liés directement à l’agriculture familiale.

Aujourd’hui, à l’heure où les effets du changement climatique se font ressentir, les exploitants agricoles doivent produire pour nourrir une population toujours plus nombreuse en limitant leur impact sur l’environnement.

Malgré leur importance, autant en terme de nombre qu’en terme de potentiel, l’agriculture familiale n’est que très peu connue et reconnue. Très peu de données existent sur les structures de ces exploitations familiales, les problèmes qu’elles rencontrent ou le travail qu’elles génèrent. C’est dans ce contexte que Pierre-Marie BOSC, ancien chercheur au CIRAD aujourd’hui détaché à la FAO, a lancé le concept de l’observatoire des agricultures du monde (OAM).

Figure : Logo de la FAO

Projet WAW, porté par Pierre-Marie Bosc.

Recensement et études lims variables de sortie

Typologie

Récupérer note de concept

Stratégie d’investissement différenciées…

COI

Collecter données sur le terrain après celle qui manquent

Le stagiaire a ainsi pu bénéficier des conseils de sa tutrice. Il a été en contact réguliers avec Messieurs Bélières et Bosc et Madame Darras durant la conception de la base de données. Les contacts avec Madame Mialet Serra étaient pour leur part limitées aux interactions avec les partenaires. Jean baptiste

# Dans les grandes lignes

## Missions

La mission principale du stage est la conception de la base de données qui regroupera toutes les données sur les exploitations (familiales ou non) de la zone Océan Indien. Dans le même temps, un cahier des charges devait être réalisé concernant la création de l’Observatoire des Agricultures de l’Océan Indien (OA-OI). Cet outil, sous la forme d’un site web, doit permettre aux utilisateurs d’accéder aux données essentielles selon leur profil (décideur politique, producteur, chercheur, …). Enfin, en fonction du temps disponible, un premier prototype de l’observatoire devait être développé.

## Méthodologie

Une méthodologie dite AGILE, où les clients et les réalisateurs du projet se réunissent régulièrement aurait pu être très appréciable. En effet, le projet ayant beaucoup d’inconnues, cela permet d’ajuster précisément les besoins des clients et d’avoir une grande flexibilité. Néanmoins, elle n’était pas employable car ce projet ne disposait pas réellement de client.

Le choix a donc été de réaliser un cahier des charges et de s’y tenir.

## Outils utilisés

De multiples outils ont trouvé leur utilisé au cours du stage. Chacun d’eux à nécessiter une période d’apprentissage plus ou moins profonde comme cela est détaillé dans le section suivante.

### CodeIgniter

CodeIgniter est un Framework PHP. L’utilisation de Framework permet de faciliter et d’accélérer la production de site web. Ce Framework se distingue par une grande légèreté associée donc à une grande rapidité d’apprentissage. Néanmoins, avec son orientation objet, il regroupe toutes les fonctionnalités nécessaires à la programmation Web. Madame Auzoux étant habitué à utiliser ce Framework, ce choix était logique.

Il a été préféré à Symfony, qui est un autre Framework PHP car Symfony est bien plus long à apprendre et Madame Auzoux n’était pas apte à guider le stagiaire dans cette voie.

### Leaflet

Leaflet est une bibliothèque JavaScript libre permettant de faire apparaître des cartes interactives sur des pages html. Elle est notamment utilisée par le projet de cartographie libre et ouverte OpenStreetMap.

### Git et GitHub

Git est un outil de gestion de version et Github est un service d’hébergement de projets git. C’est-à-dire qu’il est possible d’enregistrer sur GitHub les différentes versions d’un code au fur et à mesure du temps. Cela est très pratique pour retrouver une erreur qui s’est glissée dans le code. Git est également très utilisé pour développer des fonctionnalités indépendantes en même temps. Il est donc primordial afin de travailler en équipe. Les commandes init, push, pull, commit, add et checkout sont les plus basiques et celles qui ont le plus servi au cours du développement du prototype.

### D3.js

D3js est aussi une bibliothèque JavaScript. Elle est très utile pour représenter les données avec des visualisations originales et attrayantes. Les possibilités sont très nombreuses. Néanmoins, il faut un temps non négligeable pour en comprendre le fonctionnement.

### Bootstrap

Bootstrap est un Framework CSS qui permet de faciliter et d’accélérer la présentation des pages du site web. Il est reconnu pour l'adaptation extrêmement simplifiée d’un site web en format téléphone (responsive design). Bootstrap est dit orienté smartphone.

### SQL Power Architect

SQL Power Architect est un logiciel permettant de réaliser du Forward Engineering. Cela signifie que l’utilisateur crée graphiquement les tables et les relations de la future base de données. Le logiciel peut alors créer automatiquement le code SQL de la base ou directement implémenter la base de données grâce à des drivers permettant de faire des liens avec les SGBD.

### PostgreSQL

Le SGBDR (système de gestion de base de données relationnelle) utilisé a été PostgreSQL afin de profiter du module POSTGIS qui peut être très utile dans le cadre de données géo référencées. Finalement, aucune donnée géo référencée n’a pu être intégrée au prototype mais le choix de ce SGBD pourra s’avérer payant plus tard. L’interface utilisé était PgAdmin.

### Microsoft Office Access

Access permet de gérer des bases de données relationnelles. Ce logiciel n’a pas été utilisé pour la conception de la base de données car ce n’est pas le plus adapté pour cet usage. Néanmoins, pour intégrer des données, il est extrêmement pratique et c’est donc dans cet objectif qu’il a été utilisé.

### Gantt Project

Gantt Project est un outil de création de diagramme de Gantt. Ces diagrammes sont très utilisés dans le cadre de la gestion de projet afin d’estimer le temps nécessaire à la livraison d’un projet. Cela permet aussi de mettre en évidence le chemin critique.

## Planning

Le stage s’est déroulé en plusieurs phases qui sont regroupées dans le diagramme de Gantt ci-après. Celui-ci a été réalisé avec le logiciel Gantt Project.

# Déroulement du projet

## Compréhension des intentions

Dans un premier temps, le stage s’est concentré sur la compréhension des intentions et des objectifs voulus par le CIRAD et par Monsieur BOSC. Il est apparu à de nombreuses reprises que les différentes parties prenantes du projet n’avaient pas exactement la même manière de voir ce projet. Il a donc été nécessaire de se concerter pour concentrer nos efforts dans la même direction.

Les objectifs se résument à la réalisation d’un outil informatique permettant :

* Au grand public d’accéder aux données sur les exploitations agricoles et en particulier les agricultures familiales,
* Aux techniciens de saisir des données sur le terrain concernant les exploitations étudiées,
* Aux décideurs politiques d’élaborer des politiques différenciées d’investissement, censées être plus efficaces car plus appropriées,
* Aux producteurs d’avoir une évaluation de leur exploitation agricole selon les 5 capitaux (financier, humain, social, naturel et physique),
* Aux chercheurs de publier des études basées sur des données extraites de l’outil, et pouvant contenir des « paroles d’agriculteurs ».
* Et éventuellement à tout chef de projet de stocker ses données et d’avoir un aperçu de l’impact du projet.

## Bibliographie

Une fois les objectifs assimilés, l’objectif a été de rechercher dans les dataverse du CIRAD, de l’IRD et de l’INRA (un Dataverse est une application web à code source ouvert permettant de préserver, partager, citer, rechercher et analyser des données de recherche), des éventuelles données qui pourraient être utiles dans le cadre de l’observatoire. Malgré les quantités importantes de données de recherche, il s’est avéré qu’aucune base de données ou fichier Excel ne contenait des informations utilisables. Cela représentait un point négatif pour l’observatoire qui n’allait pas disposer de beaucoup de données au démarrage. Néanmoins, cela signifiait également que l’observatoire allait pouvoir subvenir, par la suite, à d’éventuels besoins en données.

## Benchmark

Après avoir saisi et incorporé les différentes intentions derrière l’OA-OI, un benchmark a été réalisé sur les différents outils du CIRAD. Habituellement un Benchmark est utilisé pour comprendre les points forts et les points faibles de concurrents sur un marché. Ici, la situation est un peu différente. Il s’agit d’étudier les outils mis en place par le CIRAD et de savoir s’ils pourraient servir dans le cadre de l’observatoire. Leur éventuel succès ou échec est aussi pris en compte.

Tous les outils étudiés ici n’avaient donc pas vocation à être la base du futur observatoire. Néanmoins, il est intéressant de noter les points forts et les points faibles de chaque outil en lien avec nos besoins. L’inventaire présenté ici n’est qu’une ébauche ; un inventaire plus détaillé est fourni en annexe. Ce qu’il faut retenir, c’est qu’aucun outil, à l’heure actuelle, ne répond aux besoins formulés par l’observatoire. En effet, la diversité des échelles voulues, des acteurs impliqués et des objectifs affichés fait de l’observatoire un cas à part.

Par ailleurs, l’hypothèse a été émise de pouvoir réutiliser certains de ces outils en post-traitement si cela apportait une réelle plus-value. Par exemple, si on se rend compte qu’une exploitation agricole a des efforts à faire concernant l’utilisation de produits phytosanitaires, on l’orientera vers Phyto’aide. Ce dernier est un outil qui permet de cerner les différents leviers qu’un agriculteur a, à sa disposition, pour limiter l’impact environnemental des produits qu’ils utilisent.

|  |  |  |
| --- | --- | --- |
| Outils | Avantages | Inconvénients |
| AGREF | -Plateforme  -Saisie données  -Export données  -Application mobile  -Différents profils utilisateurs | -Design  -Sécurité  -Non opérationnel  -Nombreuses variables inutilisées  -Pas de données agrégées  -Pas d’indicateurs  -Pas de gestion filière |
| AEGIS | -Design  -Analyse  -Collecte des données et côté ludique  -Echelle parcelle | -Pas échelle exploitation  -En cours de développement |
| E-watch | -Requête/cartographie  -Données par projet  -Développé pour tablette  -Statistiques par zone/projet | -En refonte  -Peu adapté à divers profils |
| IDEA RUN | -Auto-évaluation rapide  -Co-créée avec acteurs locaux | -Adaptée uniquement Réunion  -Juste une méthode  -Trop de données supplémentaires |
| FAST | -Critères de recherche | -Pas de classement par ordre préférentiel |
| LASER | -Package sous R pour effectuer les calculs poussés |  |
| Phyto’aide | -Différents leviers d’action explicités | -Modèle très simple |
| HiH | -Evolution données avec « vidéo » sur la carte  -Curseur droite/gauche pour comparaison  -Carte et graphique liés |  |

Figure : Benchmark

## Recueil des besoins des utilisateurs

#### Questionnaire

Cette partie du travail a été assez particulière car les acteurs ne sont pas réellement des clients. En effet, les acteurs sont les producteurs, les agences de développement, les ONG, les chercheurs ou encore les décideurs politiques. Ces derniers n’ont pas demandé directement l’observatoire. Il était donc nécessaire de faire des compromis entre les besoins réels des futurs utilisateurs et les intentions des porteurs de projet.

Un questionnaire a donc été réalisé sur Google Form (qui facilite grandement l’exploitation des résultats car ces derniers sont visualisables sur un tableur Google Sheet) afin d’identifier les partenaires possibles, les données dont ils disposent, et d’avoir un aperçu général de leurs besoins potentiels. Ce questionnaire devait ensuite être suivi par des interviews pour spécifier les attentes de chacun. Le cahier des charges devait ensuite être élaboré à partir de ces informations. Ce questionnaire a donc été soumis aux contacts du CIRAD (des agences de développement et des organisations de producteurs principalement). Malheureusement, seuls 4 organismes contactés sur 15 ont répondu.

Conscient que ce travail était d’une importance capitale pour le bon déroulement et la pérennité du projet, le stagiaire a voulu relancer les partenaires qui n’avaient pas répondu et approfondir les contacts avec les répondants, comme cela été prévu. Cependant, malgré plusieurs demandes à Madame Mialet-Serra, celle-ci a jugé que ces préoccupations n’étaient pas de première importance et devraient être remises à plus tard. Les informations pour rédiger le cahier des charges étaient donc maigres.

#### DAAF

Durant le stage, un rendez-vous a été pris avec la DAAF de la Réunion (Direction de l'Alimentation, de l'Agriculture et de la Forêt) qui dispose d’un département statistique et donc de données sur les exploitations agricoles de la Réunion qui pouvaient s’avérer intéressantes. En effet, la DAAF récupère les données des recensements agricoles qui sont des études approfondies des exploitations (hors revenus) effectuées tous les 10 ans. De plus, elle dispose des résultats des études RICA qui sont réalisées annuellement et se concentre sur les résultats économiques des exploitations.

Le but du département de statistiques de la DAAF est de réaliser des études statistiques afin de publier des études sur les exploitations agricoles. Ils réalisent aussi des études sur des demandes précises.

Malgré les demandes, la DAAF n’était pas autorisée à nous transmettre des données car celles-ci étaient confidentielles même après anonymisation. Par ailleurs, le directeur du département statistique, Monsieur Létoublon, a signifié que le CIRAD avait déjà reçu des données de leur part, sans recevoir la contrepartie convenue.

## Recueil des données

Dans le questionnaire cité plus haut, les participants devaient également renseigner les coordonnées des personnes à contacter pour récupérer les données qu’ils collectent. En plus de cela, ils ont pu exprimer le niveau de partage qu’ils souhaitaient pour leurs données. De cette manière, les données d’un projet du réseau SOA ont été récupérées. Néanmoins, elles concernaient directement le projet (qui consistait à aider des jeunes à se lancer dans l’agriculture) et pas les exploitations agricoles dans leur ensemble.

Néanmoins, Monsieur Bélières, un membre du CIRAD ayant travaillé à Madagascar sur les exploitations agricoles a pu fournir deux bases de données au moins partiellement exploitables, à l’échelle de l’exploitation. Elles contenaient des données intéressantes sur un total de plus de 800 exploitations. Ces données n’étaient pas présentes dans le dataverse du CIRAD. Cela démontre la nécessité, pour les chercheurs, de partager leurs travaux et leurs données sur des plateformes communes.

## Réalisation du cahier des charges

Le stagiaire a réalisé un cahier des charges. Ce dernier implique, à la différence d’une méthodologie AGILE, d’avoir une vision d’ensemble dès le démarrage du projet. La rédaction de ce livrable a été le fruit d’une collaboration avec Monsieur Bosc, bien qu’il ne soit pas l’utilisateur final du produit.

Monsieur Bosc avait auparavant regroupé les différentes fonctionnalités qui pouvaient être intéressantes à implémenter dans le cadre de l’observatoire. Ces dernières étaient ///

Concernant la partie charte graphique, le but est de reprendre les principaux éléments de la plateforme Prérad-OI (Plateforme Régionale de recherche Agronomique pour le Développement dans l’Océan Indien) depuis laquelle l’observatoire sera accessible.

En outre, le cahier des charges contient également :

* Une architecture possible du site
* Les modes d’accès à l’observatoire,
* Les matrices des droits des utilisateurs,
* Les spécifications techniques
* Les améliorations à envisager
* Des idées de composition des tableaux de bord

Ce livrable attend des informations complémentaires de la part des utilisateurs pour être finalisé.

## Conception de la base de données

### Général

La conception de la base de données avait un double objectif. En effet, elle devait servir, dans un premier temps, pour l’OA-OI. Dans un deuxième temps, elle devait être réutilisée en tant que « base modèle » pour les futures études concernant les exploitations agricoles réalisées par le CIRAD.

La conception de la base de données a été réalisée suite au travail de Madame Darras. Cette dernière ayant élaboré, en contact avec Messieurs Bosc et Bélières, un inventaire détaillé des variables devant être collectées et stockées dans la base. Son travail a grandement facilité la conception de cette dernière.

Une notion à absolument prendre en compte durant cette conception est la notion de temps. Presque toutes les entités présentes ont des variables qui en dépendent. Afin de concevoir une base de données en 3NF (3ème forme normale), il a donc été nécessaire de séparer certaines entités en deux selon si elles dépendaient du temps ou non. Certains choix comme le sexe ou le nom de famille sont discutables et ne sont pas fixés dans le marbre.

Pour rappel, la 3e forme normale est une manière de structurée la base de données qui assure de ne pas avoir de redondance. Cette dernière pouvant mener à l’invalidité de données dans le cas de modifications.

Finalement, la base de données contient 39 tables dont 31 directement pour la partie collecte des données des exploitations et 8 pour la partie utilisateur de l’interface Web qui reste à modifier. Ce sont ainsi quelques 200 variables qui doivent être renseignées lors d’une collecte. Ces 200 incluent des variables qui ne varieront vraisemblablement pas d’une année sur l’autre (comme les infrastructures ou les terres possédées). Pour chaque table, il sera possible, lors de la collecte de données, de signifier qu’aucun changement n’a été opéré. Cela permettra de remplir les champs de la même manière que la collecte précédente en changeant simplement l’année.

### Contraintes

Les contraintes liées à des relations d’héritage ont été appliquées à la base de données contrairement aux contraintes concernant les listes de valeurs possibles. En effet, ces dernières étant encore en évolution, elles ont simplement été spécifiées dans un document à part.

D’autres contraintes concernant des compositions ont aussi été ajoutées (par exemple, la superficie d’une parcelle est inférieure à celle de la terre sur laquelle elle se trouve).

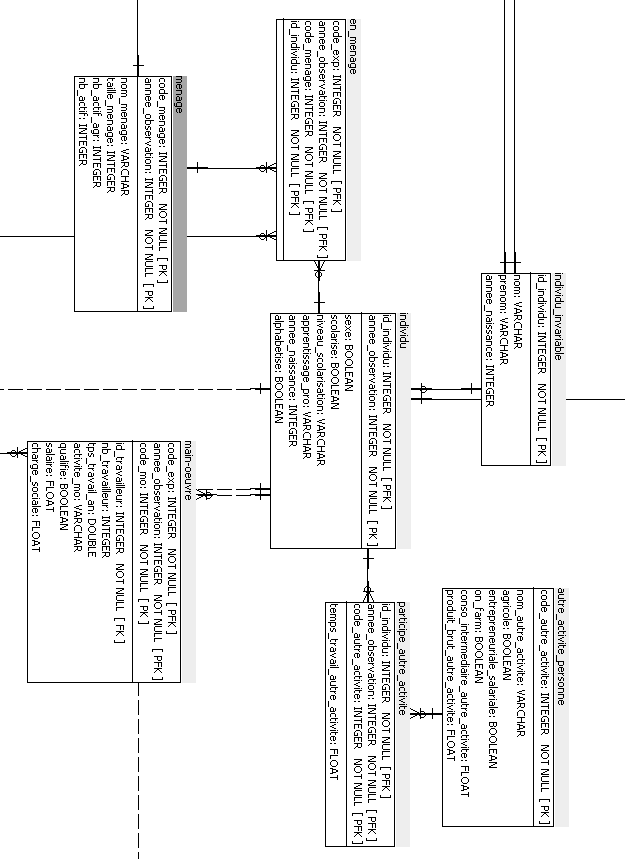


Figure : Partie du diagramme Power Architect

### Problèmes rencontrés

Afin que la collecte de données sur le terrain soit réalisable dans un temps raisonnable, il a été nécessaire de réduire cette liste de variables. Cela a occasionné de nombreux changements, supprimant parfois des tables entières. Ainsi, cette partie a pris un temps considérable. Plusieurs versions de la base de données ont été conservées selon les quantités d’informations nécessaires ou présentes.

Un autre aspect qui a été nécessité une certaine réflexion est celui des unités à utiliser pour les quantités. Que ce soit pour les consommations intermédiaires ou les quantités issues de la production, il faut pouvoir mesurer les quantités pour les comparer. Il a donc été décidé d’ajouter une variable unité et une variable « équivalent\_kg ». De plus, une table « conversion », regroupant les conversions des masses en kg, a été créée. Grâce à l’utilisation d’un trigger, lors de l’insertion d’une donnée, la colonne « équivalent\_kg » est remplie automatiquement.

Le problème est que les agriculteurs mesurent parfois leurs quantités en charrettes, en sceaux ou en flacons et ne connaissent pas nécessairement la masse que cela représente. Pour ces cas-là, il sera nécessaire d’estimer le volume de ces unités afin de trouver la masse grâce aux masses volumiques des produits. Il s’agira bien évidemment d’une estimation.

Par ailleurs, les données potentiellement récoltées provenant de diverses sources, il fallait faire preuve de flexibilité. Par exemple, certaines exploitations ou certaines enquêtes enregistrent l’utilisation de consommation intermédiaires à l’échelle de l’exploitation alors que d’autres le font à l’échelle de la parcelle. Certaines informations paraissent ainsi redondantes alors qu’il s’agit d’une réelle volonté de pouvoir récupérer des données plus précises si elles sont disponibles et plus générales dans le cas contraire.

D’autre part, étant donnés les pays dans lesquels l’observatoire a vocation à s’implanter, certaines situations particulières doivent être envisagées. En effet, en France, un ménage peut posséder et/ou vivre sur une ou plusieurs exploitations agricoles. Dans les pays de la COI (Commission de l’Océan Indien), il est possible d’avoir plusieurs ménages sur une seule exploitation. Ce type de situations compliquent la base de données. Les deux entités « ménage » et « exploitation » étant séparées, il est parfois difficile de savoir si une variable doit être liée à l’une ou à l’autre de ces entités.

Enfin, les mots employés par les Messieurs Bélières et Bosc n’étaient pas toujours connus de tous. Le monde agricole possède un vocabulaire spécifique et c’est dans ce type de cas que la communication facilite le travail.

## Intégration de 2 bases de données

Après la conception de la base de données, le but était de récupérer les données disponibles et de les intégrer à notre base. Cela permettait plusieurs choses :

* De confronter notre base de données « théorique » à des données collectées sur le terrain. De cette manière, de multiples modifications ont été effectuées, notamment sur l’aspect flexibilité évoqué plus haut. Par ailleurs, certaines variables ont été ajoutées et d’autres retirées, en discussion avec Messieurs Bosc et Bélières.
* De « remplir » la base de données afin de réaliser

Cette partie du stage, bien qu’essentielle est extrêmement fastidieuse. En effet, les bases de données récupérées ne possèdent ni les mêmes tables, ni les mêmes clés primaires, ni les mêmes noms de variables (qui sont parfois difficiles à comprendre lorsqu’aucune documentation n’existe). Parfois, il était même nécessaire de remonter aux questionnaires utilisées lors de la collecte de données pour comprendre ce que les noms des variables impliquaient.

Cela a d’ailleurs permis de souligner l’importance de la clarté des questionnaires d’enquête. Un questionnaire demandait d’abord les informations du chef d’exploitation, puis, des membres du ménage. Certains enregistrements faisaient ainsi apparaître le chef d’exploitation dans les membres de ménage et pas d’autres. Afin de ne pas corrompre les données, la décision a été prise de ne pas comptabiliser les chefs d’exploitation dans les membres du ménage pour de ne pas avoir de doublons.

Les bases de données témoignent parfois d’un manque de connaissance en gestion des données. En effet, sur les deux bases de données, l’une était constituée d’une seule table et des dizaines de colonnes alors que l’autre ne respectait pas certaines contraintes de clés étrangères. Il est par exemple arrivé d’avoir des engrais utilisés sur un champ qui n’avait jamais été enregistré. Avant d’insérer les données, il a donc été nécessaire de réaliser un long nettoyage. De nombreuses données ont ainsi été perdues.

## Vues

## Réalisation du prototype

### Base de l’application Web

Développer création application

L’application Web est basée sur une architecture MVC (Modèle-Vue-Contrôleur). Il s’agit d’un design pattern permettant de séparer le code pour une plus grande lisibilité. Pour chaque page html, il y a donc au minimum un fichier pour la vue, un fichier pour le modèle et un fichier pour le contrôleur. Chaque partie a donc une fonction bien précise.

Le Modèle :

* Le Modèle regroupe toutes les fonctions liées aux interactions directes à la base de données.
* L’utilisation de CodeIgniter, comme de beaucoup de Framework, facilite sensiblement ces interactions. Dans CodeIgniter, le modèle est en effet un objet qui hérite d’une classe AbstractController présente dans le Framework dès la création du projet. Cette classe implémente déjà de nombreuses fonctions (comme ->find(), ->where(), ->, …) qui rendent la création de requêtes SQL bien plus rapide.

La Vue :

* La vue est la partie du code qui produit le document html.
* Dans CodeIgniter, il est possible de passer directement des variables du contrôleur à la vue au moment de l’appel à la fonction render(« view\_name », $data). Chaque variable passée dans l’ensemble « $data » peut être utilisé entre des balises php dans la vue appelée.

Le Contrôleur :

* Le contrôleur est quant à lui chargé de gérer les requêtes http qui lui sont envoyées. Par la suite, il appelle les modèles et les bonnes vues en fonction des informations dont il dispose.
* Dans CodeIgniter, c’est dans ce fichier que l’on instancie les modèles et que l’on renvoie les vues (grâce à la fonction render(« view\_name »).

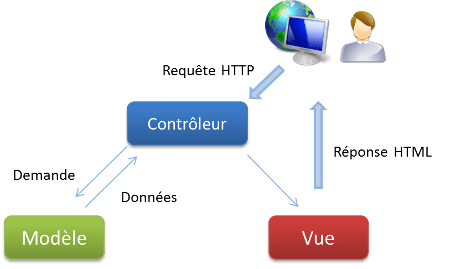


Figure : Schéma architecture Modèle\_Vue\_Controleur

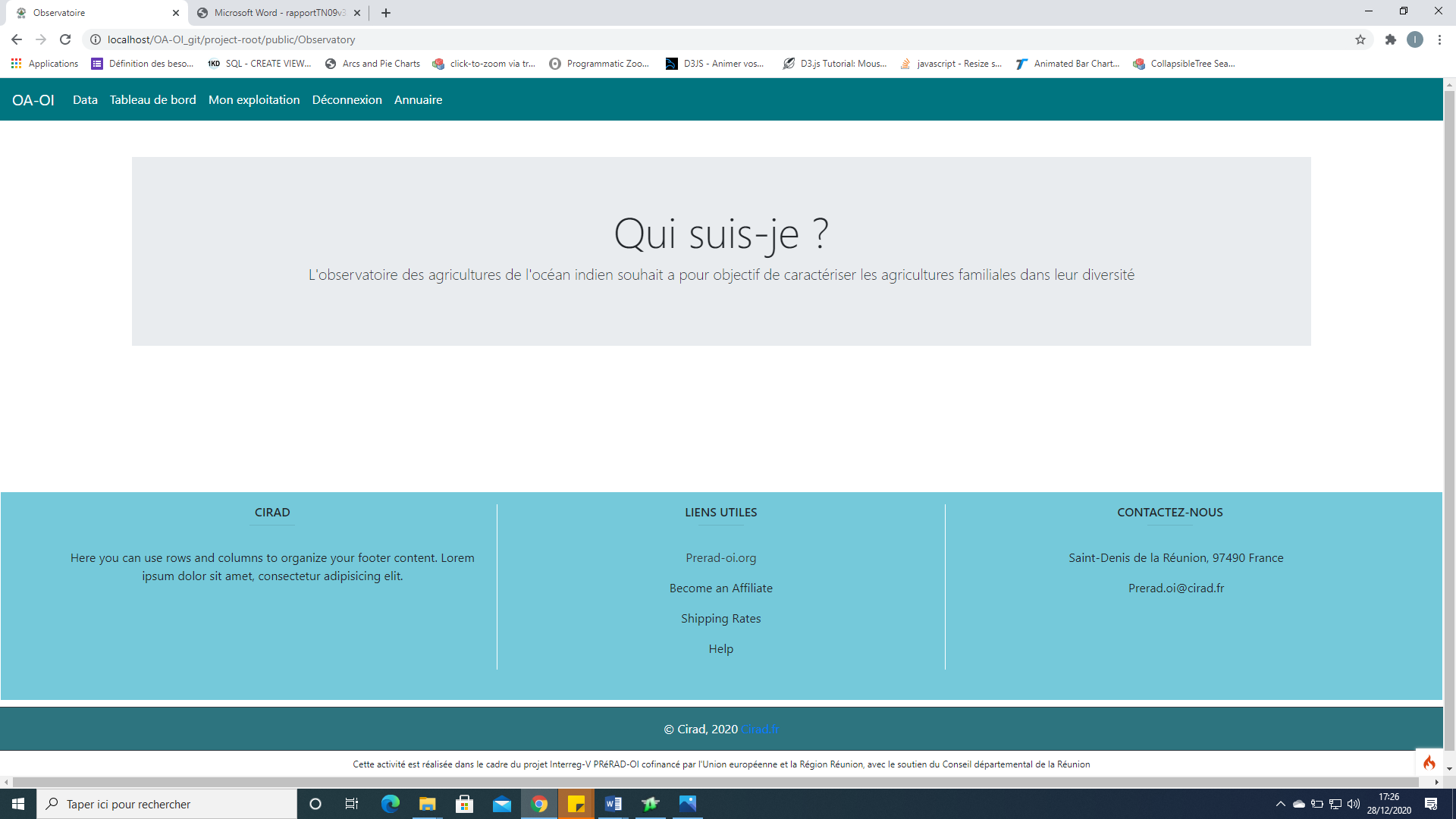


Figure : L'observatoire

### Data visualisations

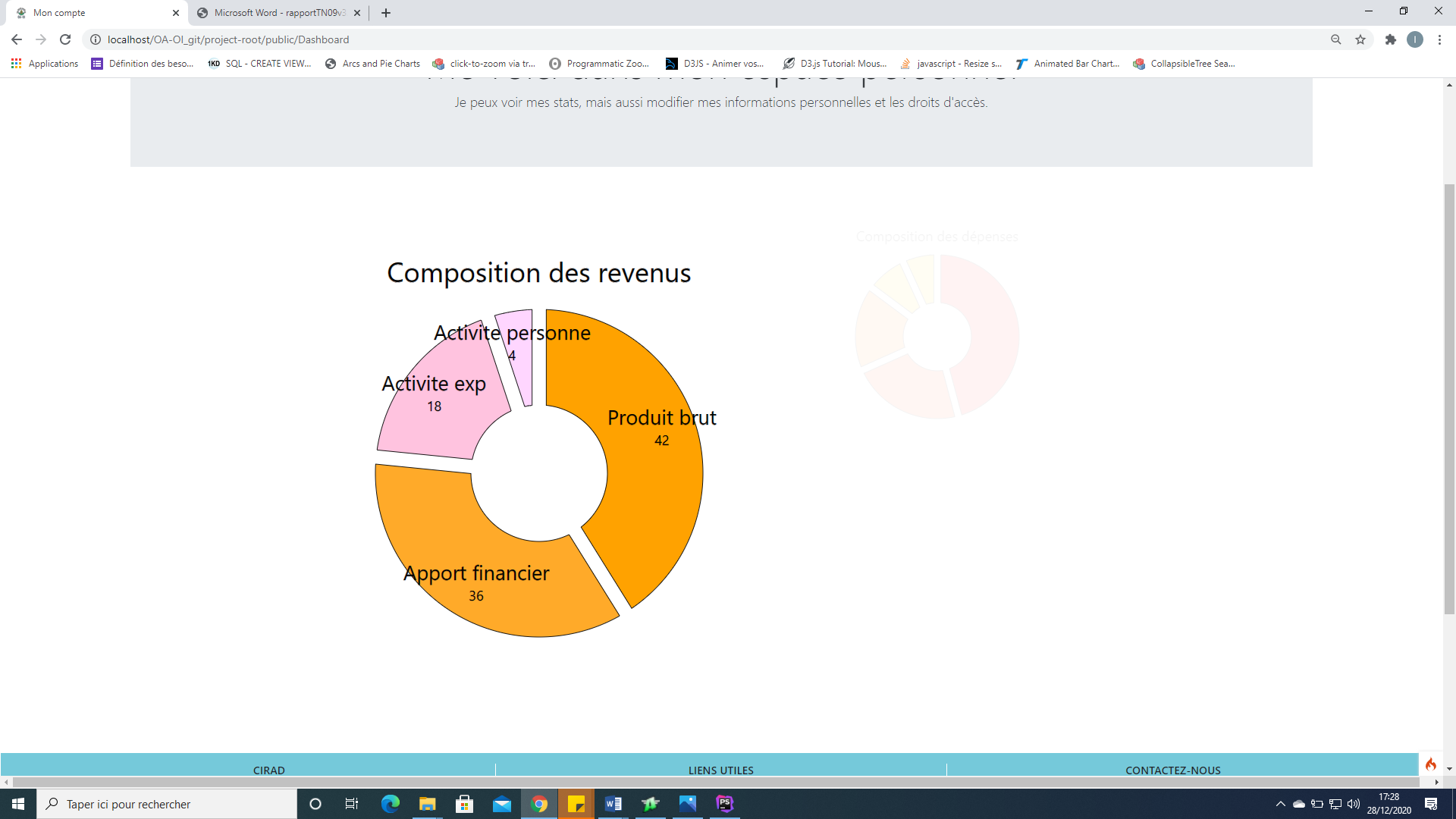


Figure : Composition des revenus

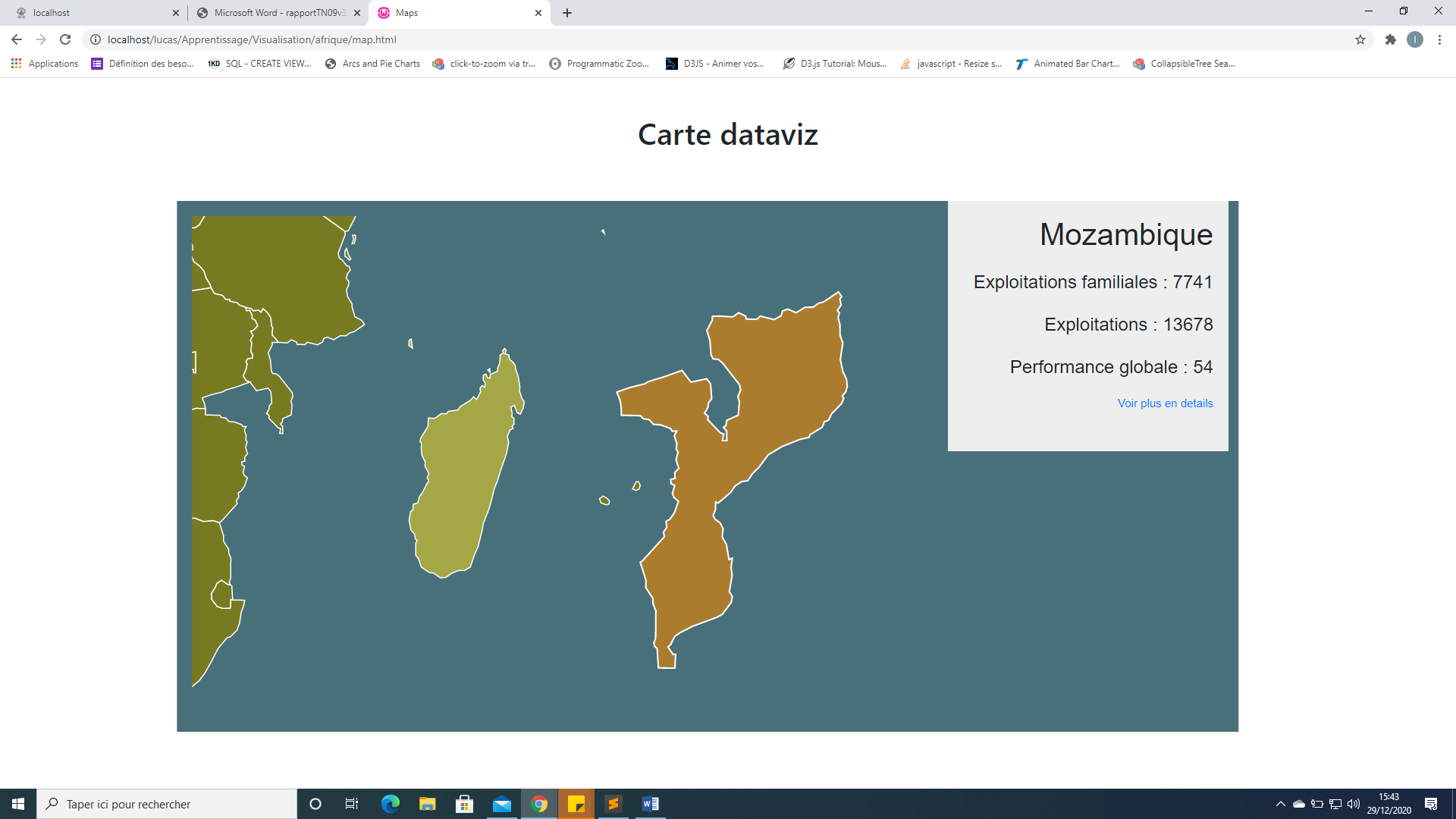


Figure : Dataviz Afrique

## Documentation

Le stage dure simplement 24 semaines et n’est donc pas suffisamment long pour réaliser le prototype complet de l’observatoire. Il est donc primordial de rédiger une documentation fournie et détaillée sur les différents travaux menés.

# Bilan d’expérience

## Difficultés

Ce stage de 24 semaines m’a permis de me confronter à de nombreuses difficultés. Cela m’a permis de déceler les obstacles auxquels un ingénieur doit faire face au cours de l’exercice de son métier.

Coopération pas toujours évidente avec des acteurs dont les connaissances sont très diverses : agronome et informaticien possèdent des vocabulaires très différents. Par ailleurs, tous les acteurs impliqués dans le projet n’étaient pas présents sur place et la communication n’a pas toujours été parfaite.

Aussi, l’apprentissage constant de nouvelles technologies n’a pas été une difficulté en soi mais cela représente un temps de travail considérable. Ce travail est évidemment plus important en période de stage mais les technologies évoluent si vite qu’il faut savoir consacrer une partie de son temps de travail à la veille technologique.

Par ailleurs, un des plus gros problèmes pour la réalisation de l’OA-OI est le manque de données existantes concernant les variables désirées. En effet, afin de conceptualiser la base de données, il serait préférable d’avoir une idée des données qui sont déjà collectées et celles qui sont facilement récupérables. Sans ces informations, il est à prévoir que les variables présentes dans le MCD puissent changer car elles ne correspondent pas à une réalité du terrain.

D’autre part, la synchronisation des différents stagiaires, VSC et autres personnes travaillant sur le projet n’est pas une mince affaire. Madame Mialet-Serra n’étant pas particulièrement habituée à traiter avec des projets qui contiennent de l’informatique, certaines étapes n’étaient pas parfaitement coordonnées. Sur certains aspects, le stage aurait pu être effectué quelques mois plus tard. Le choix des variables utilisées dans le cadre de l’observatoire a été effectué assez tardivement et ne concernaient que l’exploitation agricole. Initialement, elles devaient aussi inclure des données à l’échelle de la filière. De plus, les objectifs précis de l’observatoire se sont aussi préciser au fur et à mesure du stage. Ainsi, le stage s’est apparenté pendant la première partie à faire de l’assistance à maîtrise d’ouvrage selon Sandrine Auzoux. Concilier des objectifs ambitieux, des données manquantes et des faisabilités techniques s’est révélé être très formateur.

Manque de réponse

Etre dans le flou

Mauvais ordre des choses

## Apports personnels

Smart gender

# Glossaire

3NF : 3e Forme Normale

AIDA : Unité de recherche « Agroécologie et intensification durable des cultures annuelles »

CIRAD : Centre de Coopération international de recherche agronomique pour le développement

COI : Commission de l’Océan Indien

DAAF : Direction de l'Alimentation, de l'Agriculture et de la Forêt

Dataverse : Dataverse est une application web à code source ouvert permettant de préserver, partager, citer, rechercher et analyser des données de recherche.

DRRM : Direction Régional Mayotte Réunion

FAO : Organisation des nations unies pour l’alimentation et l’agriculture (Food and Agriculture Organisation)

MCD : Modèle Conceptuel de Données

ODD : Objectifs du développement durable

OA-OI : Observatoire des Agricultures de l’Océan Indien

OAM : Observatoire des agricultures du monde (WAW : World Agriculture Watch)

Prérad-OI : Plateforme Régionale de recherche Agronomique pour le Développement dans l’Océan Indien

SGBDR : Système de Gestion de Base de Données Relationnelle

Annexe

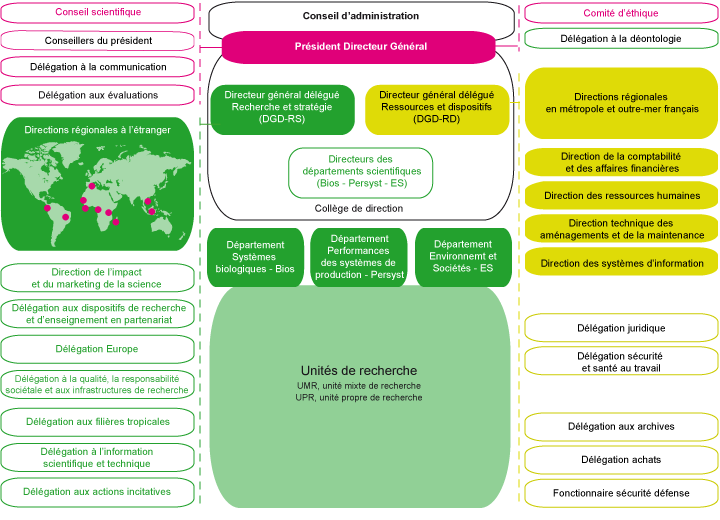


Figure : Histogramme du CIRAD

Idées de choses à développer dans le rapport de stage :

Le plus simple mais qui rassemble tout, tient compte de ce qui existe mais avec des objectifs nouveaux.

Données =pb

Récupération de données :

Confidentialité des données

Compliqué de transférer des données entre deux bases de données différentes

On se rend compte de l’impact qu’ont les questionnaires et leur clarté. Exemple du chef d’exploitation qui apparaît ou non dans la population de l’exploitation. Choix de garder les données même si risque doublons. Le but ici est de commencer à réaliser un prototype et non d’avoir un produit fini. De plus, un stagiaire sera chargé de récupérer toutes les données existantes.

Nettoyage à effectuer

Données qui proviennent d’agronomes dont les bases de données ne sont pas propres avec des relations non existantes. Lien à rajouter, certains enregistrements ne peuvent pas être gardées car ne remplissent pas les conditions de clés primaires ou étrangères.

Problème de matériel, 2 changements de postes informatiques

PGD grosse galère pour récupérer des données individuelles à la daaf. Ne peuvent donner que des données agrégées ce qui est bien mais celles-ci ne sont pas complètes.

Toujours ce problème des utilisateurs finaux.

Echantillonnage pour récupérer données mais comment faire dans le cadre d’une OP. Récupérer les données pour leur donner des indicateurs ?

Module de Q supplémentaires ?

Partie politique entre le cirad et la daaf avec des contreparties des antécédents, on est loin d’avoir un niveau de partage optimal alors que nous travaillons tous pour le bien des agriculteurs.

Travail de longue haleine = récupération de données en 2022-23 mais sûrement retardée encore

Difficile jeu avec plein d’acteurs avec données différentes

Réunion riche en données mais pas les autres.

Baladé entre les technologies, les envies des uns et des autres, les ordres contradictoires, … projet vraiment instable et précaire. On ne sait pas sur qui on va pouvoir compter ni qui va compter sur nous. Pas établi, besoin de cette proof of concept, vraiment ?

Manque de motivation car je ne vois pas le marché derrière.

Désaccord avec le chef de projet (entretien)

Bilan de ce qui ne va pas selon moi :

EN CONTACT AVEC LES PARTENAIRES BORDEL !!!

DESACCORD SUR LA TECHNOLOGIE donc apprentissage dans le vide, perte de temps

UTILISE comme un informaticien et non comme un ingénieur

Ce que je dois faire en tant qu’ingénieur pour pallier à ces problèmes :

Communiquer avec Isabelle pour

Trigger sous requête

Importance femme dans CDC

Faire des requêtes pour montrer le manque de remplissage de la base de données

Importance CodeIgniter vs Symfony combat

Manque de structure

COI