1. **Introduction**
2. **Etude préalable**
3. Analyse de l’existant.
4. Spécification fonctionnelle du DSS.
5. Stratégie de développement :

3.1. Modèle de développement adopté. (En spiral)

3.2. Dispositif de coordination entre les acteurs. (Réunions && Outils de travail)

1. Plan de développement :

3.1. Découpage WBS.

3.2. Analyse des risques.

3.3. Planification et ordonnancement des tâches. (Diagramme de Gantt)

1. **Etude détaillée globale**
2. Description technique de la solution proposée.
   1. Architecture logicielle. (Pattern MVC)
   2. Description de la dynamique de fonctionnement. (Communication et schéma d’exécution)
3. Production de maquettes IHM.
4. **Interrogation des données**
5. Analyse du système d’extraction et d’annotation sémantique : @Web.

1.1. Présentation du système.

1.2. Principe de fonctionnement.

1.3 Ontologie relative à la bioraffinerie.

1. Structuration de données.

2.1 Les données annotées.

2.2 Les données de la base de données PostgreSQL.

1. Développement du module d’interrogation.

3.1. Technologies et services employés.

3.2. Diagramme de contexte.

3.2. Description du mécanisme de fonctionnement.

1. **Introduction**

Le principal obstacle qui freine l'exploitation des données scientifiques est leur format textuel et leurs structures hétérogènes. Dans ce contexte, l'utilisation des méthodes de l'ingénierie des connaissances (KE) semble pertinente pour structurer l'information expérimentale et de l'exprimer dans un vocabulaire normalisé. Cette structuration peut être faite en utilisant une ontologie spécifique pour modéliser les données expérimentales d'intérêt.

C’est dans cette même perspective qu’interviennent les systèmes d’aide à la décision **(Decision Support System - DSS)** pour aider les chercheurs impliqués dans la conception de la bioraffinerie à prendre des décisions rationnelles fondées sur des données et des connaissances exprimées par les experts du domaine dans la littérature scientifique.

Notre projet s’inscrit dans le cadre de la conception et la mise en œuvre du pipeline de traitement de données, qui prend en ***entrée*** un ensemble de documents scientifiques extraites de ressources bibliographiques sur le Web et génère en ***sortie*** un classement des procédés de bioraffinage basée sur une évaluation de l'impact environnemental.

1. **Etude Préalable**
2. **Analyse de l’existant :**

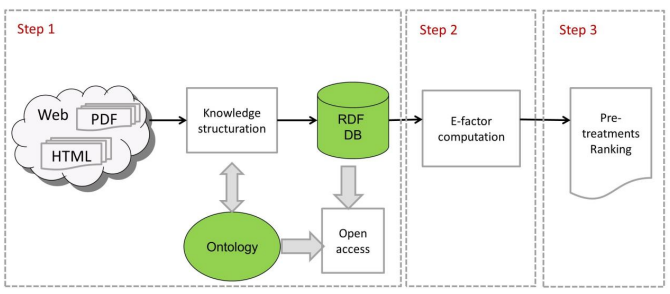
Au jour d’aujourd’hui, il n’existe pas de **DSS** qui mettent en œuvre un pipeline complet tel que celui présenté sur la figure-1 afin d'évaluer les impacts environnementaux des procédés de transformation de la biomasse.

Ce **DSS**, qui sera décrit plus en détail ultérieurement, est un pipeline composé de trois étapes:

* **Etape 01 :** L'annotation guidé par une ontologie de données expérimentales publiées dans des documents scientifiques ;
* **Etape 02 :** L'extraction de données annotées et calcul d’indicateurs E-Facteur ;
* **Etape 03 :** Visualisation du rendement des procédés et des indicateurs E-facteur dans des cartes graphiques.

La première étape du pipeline, à savoir la structuration des données expérimentales, est implémentée en utilisant l'outil sémantique d’annotation de données @Web (Annotated Tables of the Web).

Le présent projet vise à l’implémentation de la seconde et la troisième étape de ce pipeline.



**Figure-1 : Pipeline proposé par l’approche.**

1. **Spécification fonctionnelle du DSS :**

Le pipeline, tel que présenté sur la Figure-1, doit pouvoir garantir les fonctionnalités suivantes :

1. Collecter, intégrer et structurer les données hétérogènes disponibles dans la littérature scientifique sur les procédés de transformation de la biomasse ;
2. Définir ***un modèle ontologique simple et générique*** de l'information pertinente qui doit être identifié et annoté dans les journaux scientifiques. Le modèle doit être à la fois simple pour garantir la facilité de sa mise à jour par des experts de bio-raffinage (qui ne sont pas des informaticiens), et générique pour être réutilisable afin de capitaliser d'autres types de données gérées par les chercheurs ;
3. Permettre un accès libre aux données et les unités de mesure associée. Cela peut être fait grâce aux ***permalinks*** (URLs utilisés pour référencer des informations de manière permanente ou pour une période temporaire) et un système d'interrogation gérant les conversions d'unité pour faciliter la réutilisation de données ;
4. Evaluer la fiabilité des données (sources) et en tenir compte dans l'évaluation de l'impact environnemental ;
5. Gérer l’imprécision des données sachant que les données expérimentales associées à la biomasse et au procédé de transformation de celle-ci sont sujettes à l'incertitude ;
6. Tenir compte de la variabilité biologique associée aux procédés de transformation de la biomasse et de la propagation de l'incertitude ultérieure au cours du calcul des indicateurs d'impact environnemental ;
7. Calculer les indicateurs de facteur environnemental (E-Factors);
8. Visualiser le classement des processus de la biomasse en fonction du rendement des processus et des E-Factors.

Dans le cadre du présent projet, il nous est demandé de mettre en œuvre les spécifications fonctionnelles 6-7-8 du **DSS**.

1. **Stratégie de développement :**
   1. **Modèle de développement adopté :**

Pour mener à bien notre projet, nous avons opté pour un découpage temporel standard englobant :

* Etude préalable ;
* Etude détaillée & technique ;
* Réalisation ;
* Mise en œuvre.

Contraints par le temps, nous avons décidé de suivre un modèle de développement favorisant l’organisation du développement que sa documentation. Ainsi, nous avons puisé certains principes des méthodes agiles afin de dresser notre mode opératoire, qui consiste principalement en ce qui suit :

* Un responsable fonctionnel définit et ordonne la production des composants de l'application.
* Le projet est structuré en incréments sur des intervalles précis ;
* Une réunion initiale organise chaque incrément en définissant les tâches à réaliser ;
* L’équipe pilote la qualité et la performance du projet de manière consensuelle ;
* Le maintien d’un contact permanant entre les membres de l’équipe afin de situer l’avancement du projet, mettre en évidence les éventuels problèmes rencontrés et la manière de les résoudre.
* Un point de situation est dressé par l’ensemble de l’équipe lors d’une réunion de travail organisée hebdomadairement.
* Le responsable fonctionnel valide le travail de l'équipe et ajuste les besoins entre chaque incrément.
* Des réunions de moyennes fréquences sont organisées avec le maître d’ouvrage afin de valider conjointement le travail réalisé et prendre en considération les éventuels changements qu’il souhaite introduire.
  1. **Dispositif de coordination entre les acteurs :**

Afin de favoriser le travail collaboratif et l’interaction permanente entre les membres de l’équipe, une première solution ***organisationnelle*** consistait, selon la mesure du possible, en la tenue de rencontres journalières entre ses acteurs, en sus d’une réunion de travail hebdomadaire tel que nous l’avons introduit précédemment.

Une solution ***technique*** vient consolider ce dispositif à travers un ensemble d’outils que nous avons manipulés lors des différentes phases du projet :

1. **Cacoo (**[**http://cacoo.com/**](http://cacoo.com/)**) :**

C’est un site web de travail collaboratif très intuitif permettant de créer rapidement des croquis ou des diagrammes complexes.

1. **Framindmap (**[**https://framindmap.org/**](https://framindmap.org/)**):**

C’est un site web de travail collaboratif permettant la création de cartes mentales dans un objectif de prise de notes et de structuration des idées.

1. **Google Drive (**[**https://www.google.com/drive/**](https://www.google.com/drive/)**):**

C’est un service de stockage et de partage de fichiers [dans le cloud](https://fr.wikipedia.org/wiki/Cloud_computing) lancé par la société [Google](https://fr.wikipedia.org/wiki/Google), que nous avons utilisées principalement pour le partage de notre documentation et la préservation d’une copie distante de notre travail.

1. **GitHub *(*https://github.com/):**

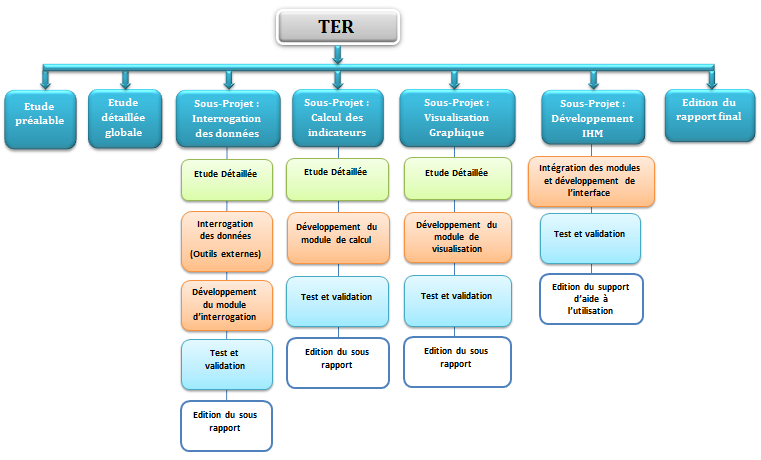
C’est un service web d'[hébergement](https://fr.wikipedia.org/wiki/H%C3%A9bergeur_web) et de gestion de développement de logiciels, utilisant le [logiciel de gestion de versions](https://fr.wikipedia.org/wiki/Logiciel_de_gestion_de_versions) [Git](https://fr.wikipedia.org/wiki/Git). Cet outil nous a été introduit par un intervenant du groupe CGI dans le cadre de l’UE : Gestion des projets.

1. **TeamViewer (Version 11) :**

C’est un [logiciel](https://fr.wikipedia.org/wiki/Logiciel) [propriétaire](https://fr.wikipedia.org/wiki/Logiciel_propri%C3%A9taire) de [bureau à distance](https://fr.wikipedia.org/wiki/Bureau_%C3%A0_distance), disposant de fonctions de [téléassistance](https://fr.wikipedia.org/wiki/T%C3%A9l%C3%A9assistance_informatique), de réunions en ligne, de [conférences Web](https://fr.wikipedia.org/wiki/Conf%C3%A9rence_en_ligne) et de [transfert de fichiers](https://fr.wikipedia.org/wiki/Transfert_de_fichier).

1. **Plan de développement :**
   1. **Découpage en WBS :**

A l’issue des entretiens menés avec le maître d’ouvrage, nous avons pu dresser un découpage hiérarchique en livrables spécifiques des travaux à exécuter, résultant en la structure WBS (Work Breakdown Structure), présentée sur la figure-2, qui définit et traduit visuellement les unités gérables et appréhendables par tous les membres de l'équipe. Chaque niveau du WBS permettant d'approfondir le détail des tâches.



**Figure-2 : Découpage en WBS**

* 1. **Analyse des risques :**

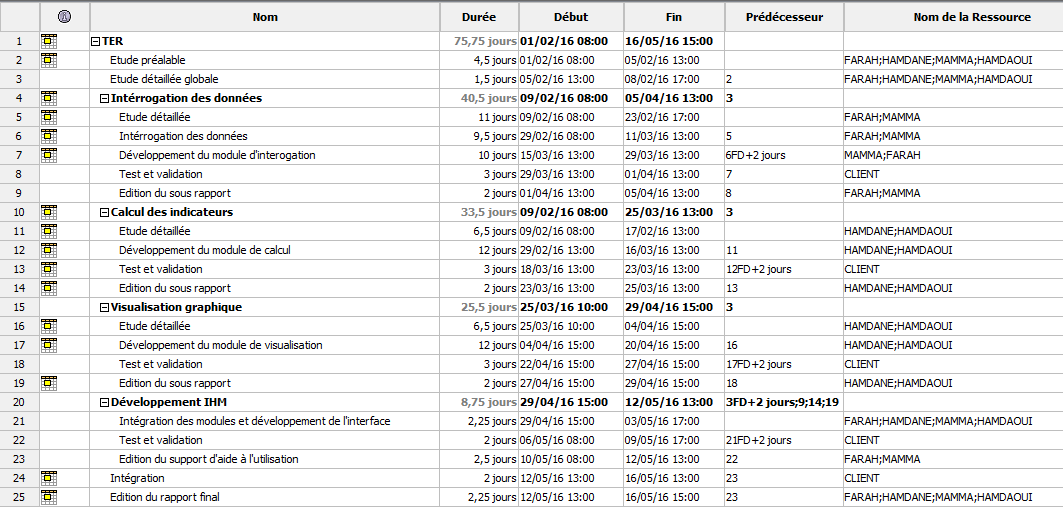
Il est primordial à ce stade de procéder à une analyse des facteurs de risque qui peuvent survenir au cours du projet. Cette analyse permet non seulement d’adapter la planification temporelle des tâches, mais aussi de prévoir des solutions de secours en cas de survenance de certains types de risques.

Le profil de risque de notre projet est présenté sur le tableau qui suit :

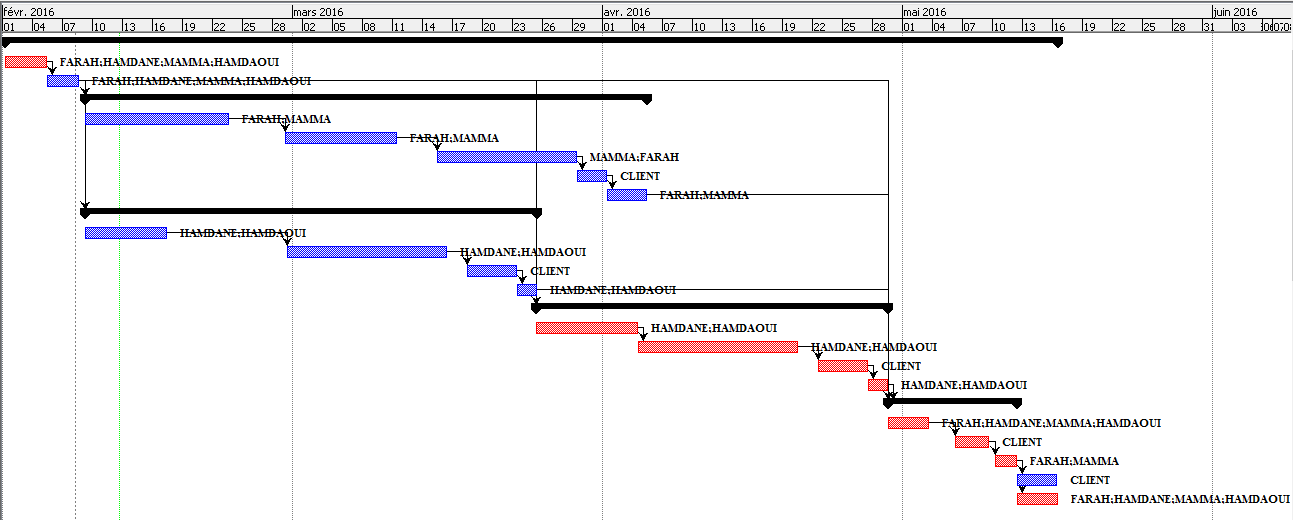
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Nature du risque | Degré de risque pour le projet | | | | | |
| 0 | 1 | 2 | 3 | 4 | 5 |
| Taille du projet/Temps de réalisation |  |  | X |  |  |  |
| Difficulté technique |  |  |  | X |  |  |
| Degré d’intégration |  | X |  |  |  |  |
| Configuration organisationnelle |  |  | X |  |  |  |
| Changement |  | X |  |  |  |  |
| Instabilité de l’équipe du projet |  |  | X |  |  |  |

* 1. **Planification et ordonnancement des tâches :**

Par projection de notre analyse des risques sur la structure en WBS retenue, il ressort la planification et l’ordonnancement des tâches présentées respectivement sur la figure-3 et la figure-4.

****

**Figure-3 : Planification des tâches**



**Figure-4 : Ordonnancement des tâches (Diagramme de Gantt)**

1. **Etude détaillée globale**
2. **Description technique de la solution proposée :**
   1. **Architecture logicielle :**

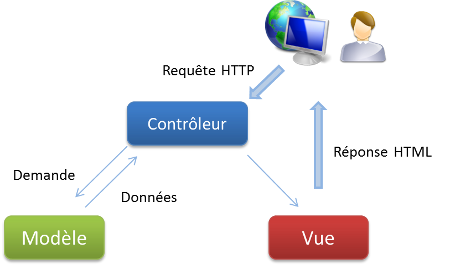
Notre projet consiste à développer un site web traduisant les étapes restantes du pipeline décrit précédemment sur la figure-1.

Ce site web doit comprendre, tel que présenté sur la structure en WBS, trois principaux composants logiciels :

1. **Premier composant logiciel :** garantit l’interrogation des données (Données annotées en format RDF & Données structurés et stockées sur une base de données PostgreSQL) ;
2. **Deuxième composant logiciel :** garantit le calcul des indicateurs E-Factor ;
3. **Troisième composant logiciel :** garantit la visualisation du rendement des procédés et des indicateurs E-facteur dans des cartes graphiques.

Les deux premiers composants logiciels correspondent à une couche de ***traitement des données***, tandis que le dernier composant logiciel correspond à une couche de ***présentation des données***.

Sur la base de cette architecture, nous avons retenu le Modèle MVC (Model – Vue – Contrôleur) pour le développement de notre site web. En effet, ce modèle est destiné à répondre aux besoins des applications interactives à travers une séparation entre le traitement des données et leur présentation (Interface Graphique).



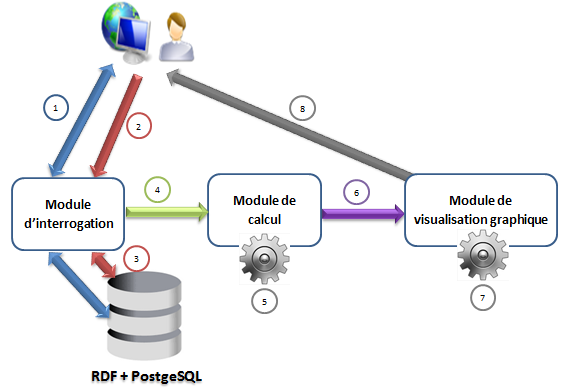
**Figure-5 : Architecture du modèle MVC.**

* 1. **Description de la dynamique de fonctionnement :**

Notre site web vise à proposer à ses utilisateurs, sur la base d’un paramétrage qu’ils auront à définir, la visualisation d’un ensemble de synthèses sur des cartes graphiques,

Ce scénario repose sur un enchainement de traitements interdépendants que nous pouvons décrire comme suite :

1. Interrogation des données pour l’initialisation des paramétrages proposés à l’utilisateur sur l’interface graphique;
2. Récupération du jeu de paramétrage introduit par l’utilisateur et validé sur l’interface graphique ;
3. Interrogation des données pour récupérer les données expérimentales afférentes au paramétrage de l’utilisateur ;
4. Mise en forme des données expérimentales pour amorcer le calcul des indicateurs E-Facteur.
5. Calcul des indicateurs E-Facteur avec prise en compte du paramétrage de l’utilisateur ;
6. Mise en forme des résultats de calcul pour préparer la phase de la visualisation graphique ;
7. Edition des cartes graphiques.
8. Visualisation des cartes graphiques.



**Figure-6 : Dynamique de fonctionnement du site web.**

1. **Production de maquettes IHM :**

Cette phase consiste en la mise en place d’un ensemble de maquettes IHM afin de valider avec le maitre d’ouvrage la dynamique de fonctionnement du site web, son contenu et son ergonomie conformément aux besoins attendus.