Documentation Technique

Module REDACTion

Analyse et Ajustements intelligents des documents

BLEX (Business Life & Excellence)

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  |  |  |
|  |  |  |

Sommaire

[Architecture du Systeme 1](#_Toc321140622)

[Technologies Utilisées 2](#_Toc321140623)

[Fonctionnalités Techniques 3](#_Toc321140624)

# Architecture du Système

## Writing

Le module **Writing** est conçu pour analyser, fusionner et corriger des documents en utilisant des technologies de NLP et d’IA. Il est structuré en plusieurs composants :

• Frontend : Interface utilisateur permettant l’upload, l’édition et l’affichage des documents.

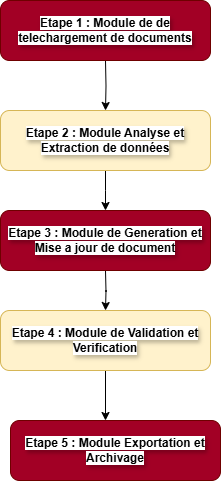
• Backend : Gère la logique métier, les interactions avec la base de données et l’IA.

• Base de données : Stocke les documents, les templates et les historiques des modifications.

• Moteur d’analyse et de fusion : Compare les documents, détecte les différences et propose des corrections.

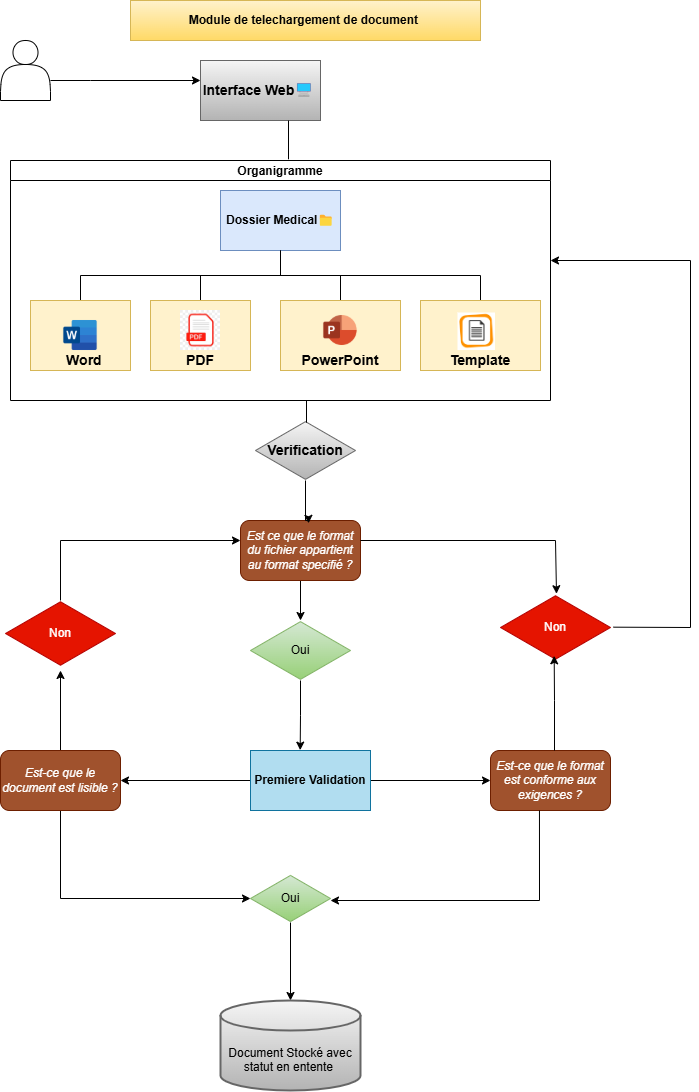
## Schema architecture

Ce schéma donne l’architecture globale de cette application puis on va décortiquer chaque étape



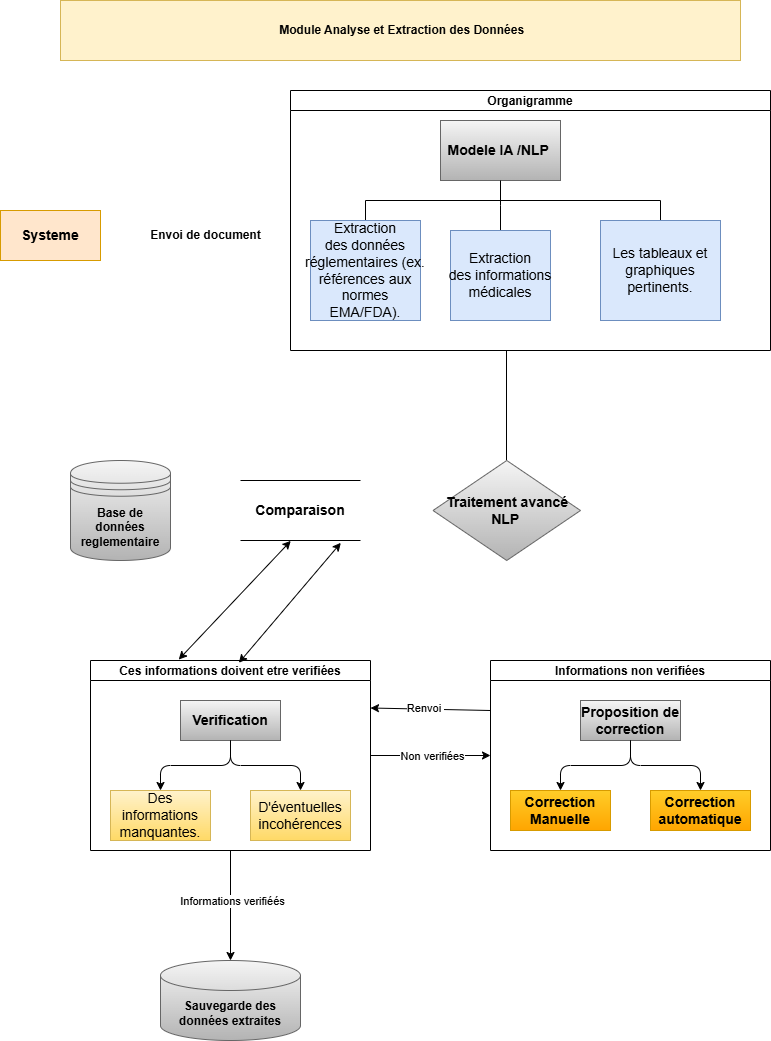
## Module telechargement document

Dans ce module on se familiarise avec l’interface de l’application pour uploader des documents selon le type et les stocker dans la base de données après être passé par une série de vérification et authenticité



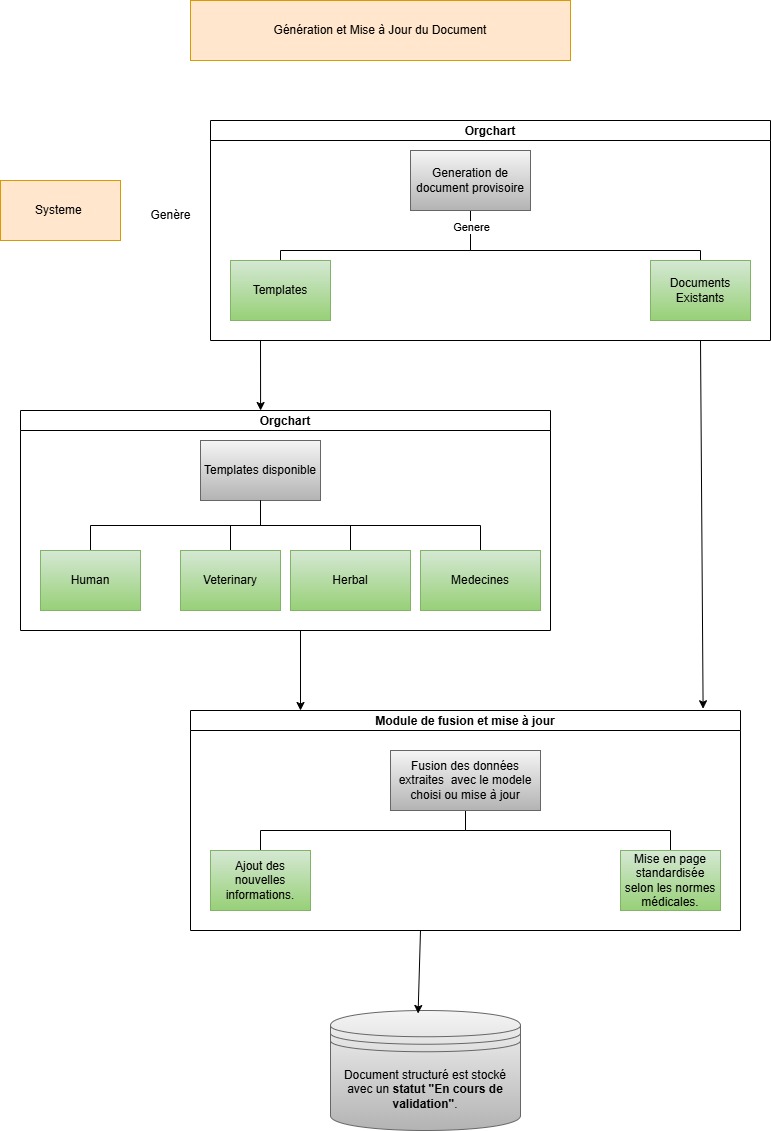
## Module Analyse et Extraction des données

Une fois le document stocké dans la base données on passe au module Analyse et Extraction qui permet d’extraire les informations médicales et pertinentes ainsi que les graphiques avant de les faire faire un traitement NLP ensuite une vérification ou une proposition de correction pour les documents non vérifiées



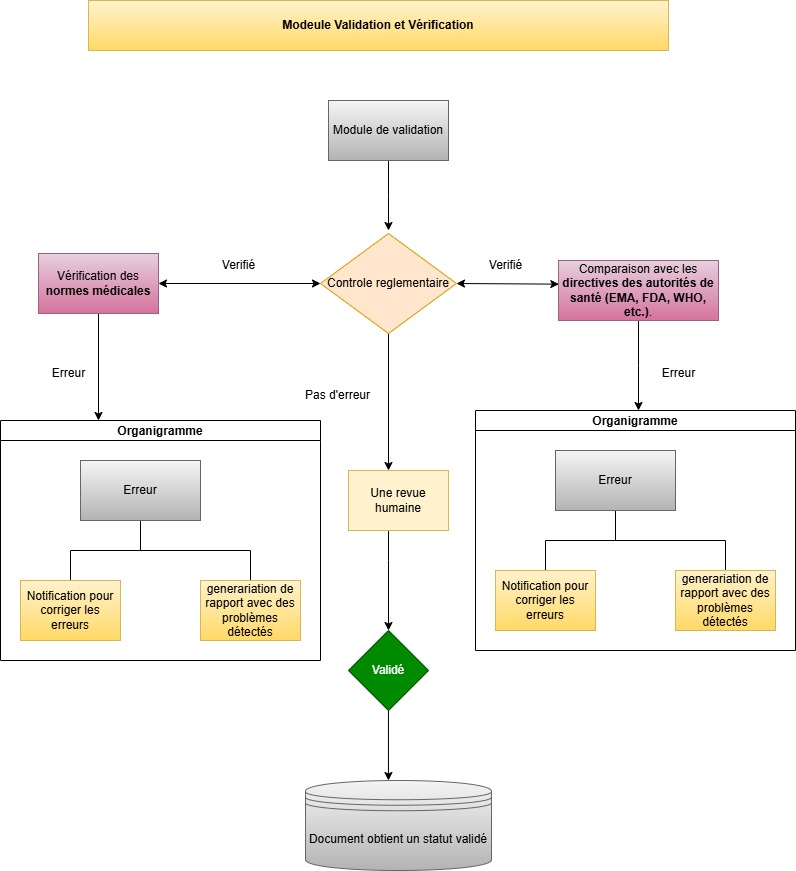
## Module Génération et Mise à jour du document

Dans ce module on va commencer la génération provisoire pour voir comment jauger le document ou le Template une fois téléchargé selon le type qu’on veut. Après ceci on va commencer la fusion et mise a jour du nouveau document en ajoutant les nouvelles modifications ensuite on va stocker ce document dans la base de données avec le statut **‘En cours de validation’**

****

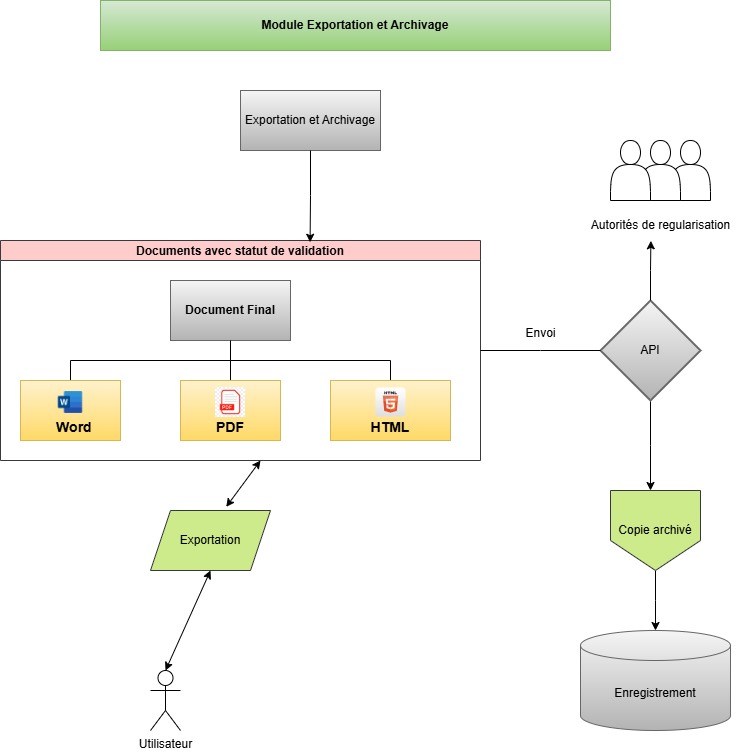
## Module Vérification et Validation

Dans ce module on va faire la vérification du document qui était **‘En cours de validation’** pour voir s’il respecte les normes médicales et une comparaison avec les directives des autorités de santé genre un contrôle règlementaire. S’il y a une erreur on sera notifié pour corriger les erreurs et aussi générer des rapports pour voir les erreurs détectées sinon on valide le document et on change son statut vers **‘Validé’**



## Module Exportation et Archivage

Une fois la vérification et la validation du document complété on passe à l’exportation et l’archivage de ce dernier. On va juste exporter le document sous le format souhaité et l’enregistrer.



# Technologies Utilisées

|  |  |
| --- | --- |
| Composant | Technologies |
| Langage | *Python* |
| Backend | *Django* |
| Stockage | *SQL Lite* |
| Analyse et Fusion | *spaCy, PyMuPDF* |

**Python :** Python est un langage de programmation high-level, orienté objet et dynamique. Il est conçu pour être facile à apprendre et à utiliser, même pour les développeurs non expérimentés. Python est souvent utilisé en tant que langage de scripting, dans l'analyse des données, le développement web, la Machine Learning et l'intelligence artificielle.

**Django :** Django est un Framework web Python qui facilite la création de sites web robustes et maintenables. Il fournit une structure de base pour les applications web, y compris des modèles, des vues, des contrôles et des modèles de données.

**NLP (Traitement Automatique du Langage Naturel) ;** Le traitement automatique du langage naturel (NLP) est un domaine de recherche et de développement qui vise à créer des systèmes capables de comprendre, d'analyser et de traiter les langages naturels comme le français, l'anglais, etc.

**spaCy :** spaCy est une bibliothèque Python pour le traitement automatique du langage naturel (NLP) qui fournit des outils de reconnaissance d'entités nommées (NER), de détection de relations et de tokenisation.

**PyMuPDF :** PyMuPDF est une bibliothèque Python qui fournit des outils pour la lecture et l'analyse de documents PDF.

**SQL Lite :** SQL Lite est une base de données relationnelle open-source qui fournit des outils pour la création, la modification et la suppression de bases de données.

# Contact Information

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Lassana Hamady Sow**  Data Scientist Intern  **Tel** +22238619739  lasssdb4@gmail.com |  | **Yaya Barro**  Data Scientist Intern  **Tel** +2250769079684  yayabarro0184@gmail.com |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

# Company Information

**BLEX (Business Life & Excellence)**

