

OCR Application

L'application d'analyse de facture du futur

Sommaire

Contexte

Analyse du
Besoin

Organisation
du projet

Schema
Fonctionnel

Maquette figma

Fonctionnement

OCR

OpenCV

PyTesseract
et Tesseract

Spécifications fonctionnelles

Acceuil

Scanner

Connexion
inscription

Historique

Monitoring

Test de l'application

Contexte

Analyse du Besoin et Organisation du projet

Client et besoin:
ESN : Evolution dans la
procédure de pré-
traitement des factures

Organisation :
Analyse
Veille technologiques
développement
backend (api, ocr)
développement front
end

Analyse du Besoin et Organisation du projet

Technologies Envisagées:

OCR:

- TesseractOCR
- Azure Computer Vision
- EasyOCR
- Google cloud Vision OCR

API:

- FastApi
- Flask

Front:

- React
- html
- streamlit

Technologies Envisagées:

OCR:

- TesseractOCR
- OpenCV (preprocessing image)
- Azure Computer Vision

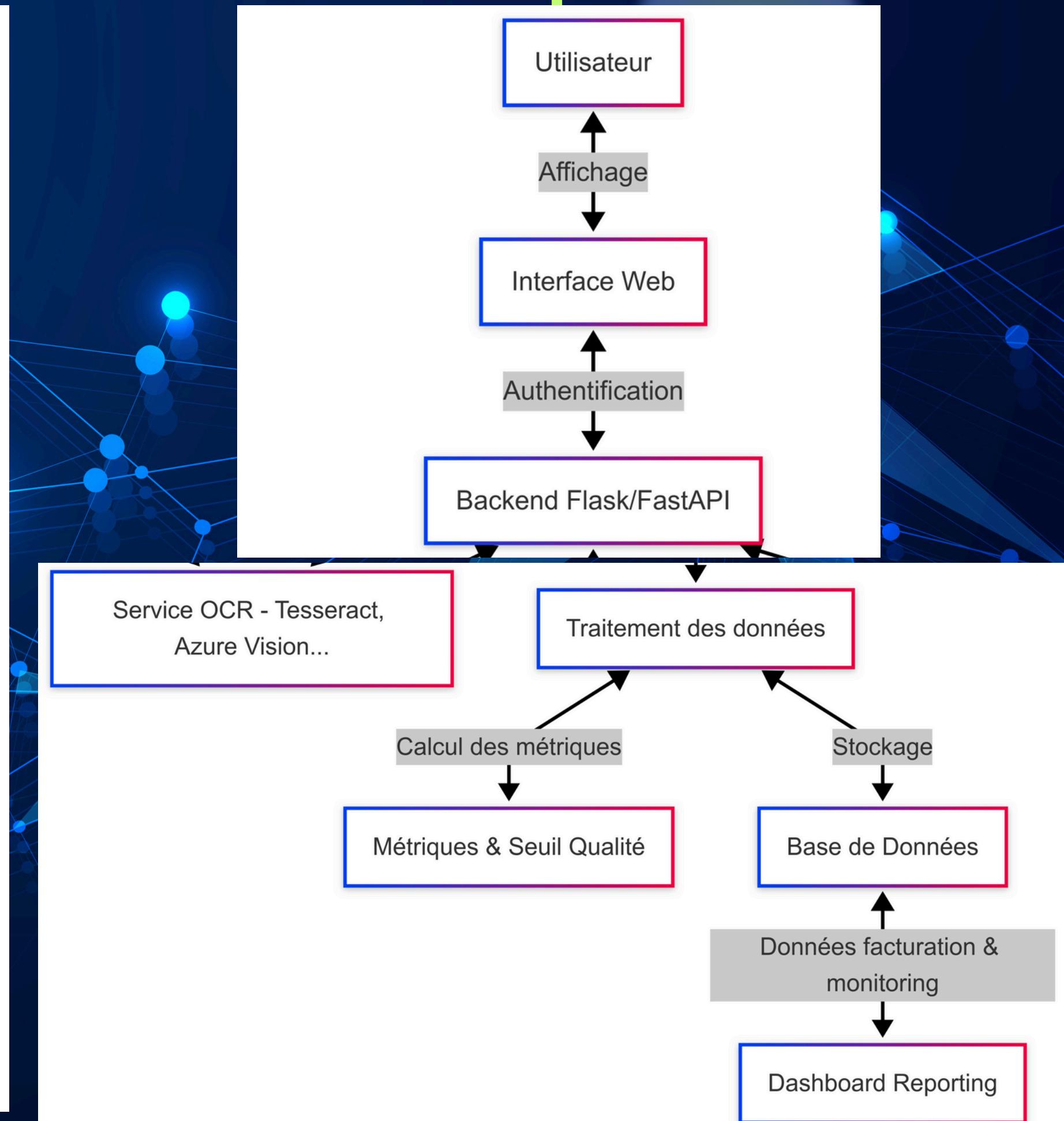
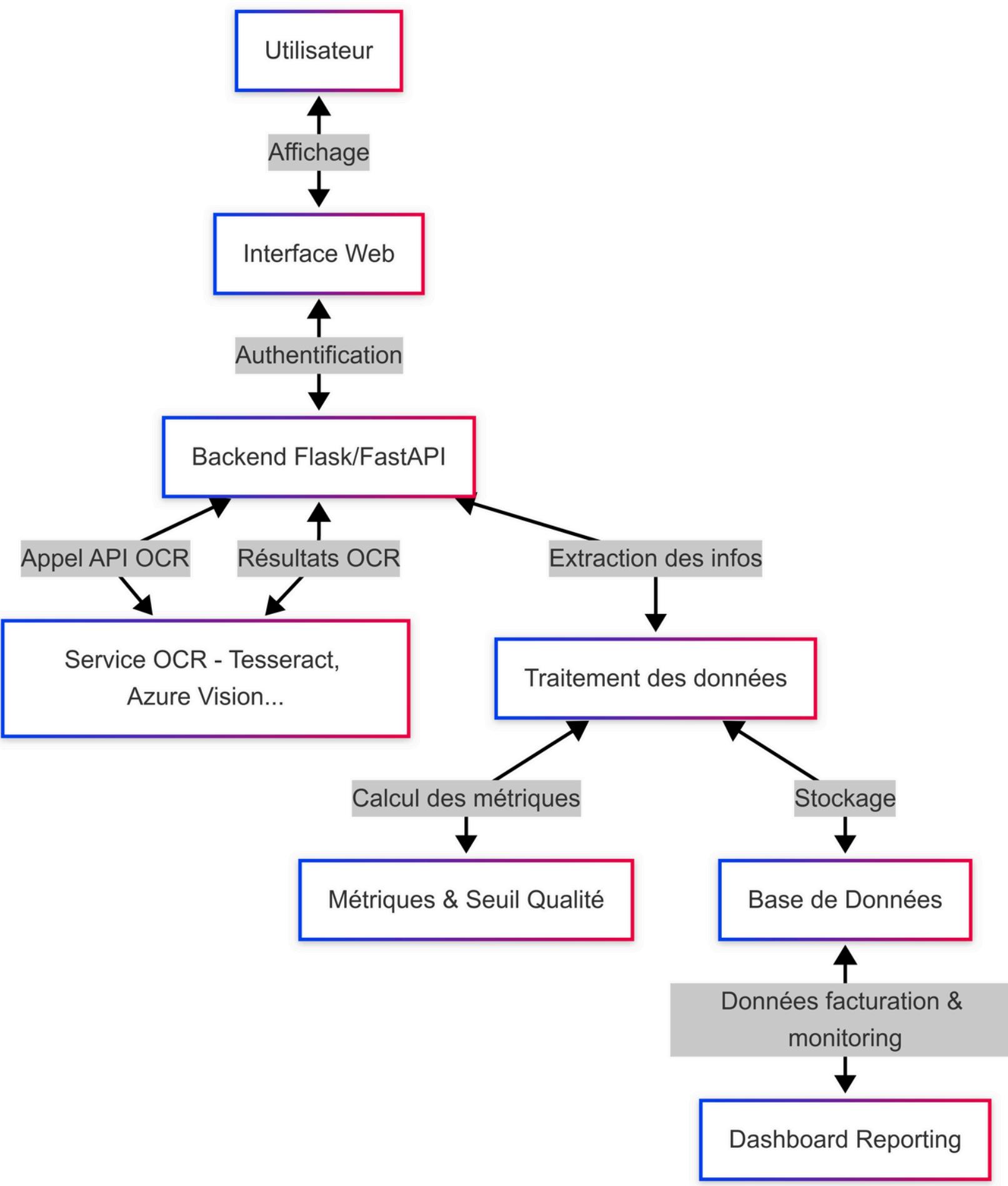
API:

- FastApi

Front:

- Javascript
- html (template Jinja2)

Schema Fonctionnel (complet)

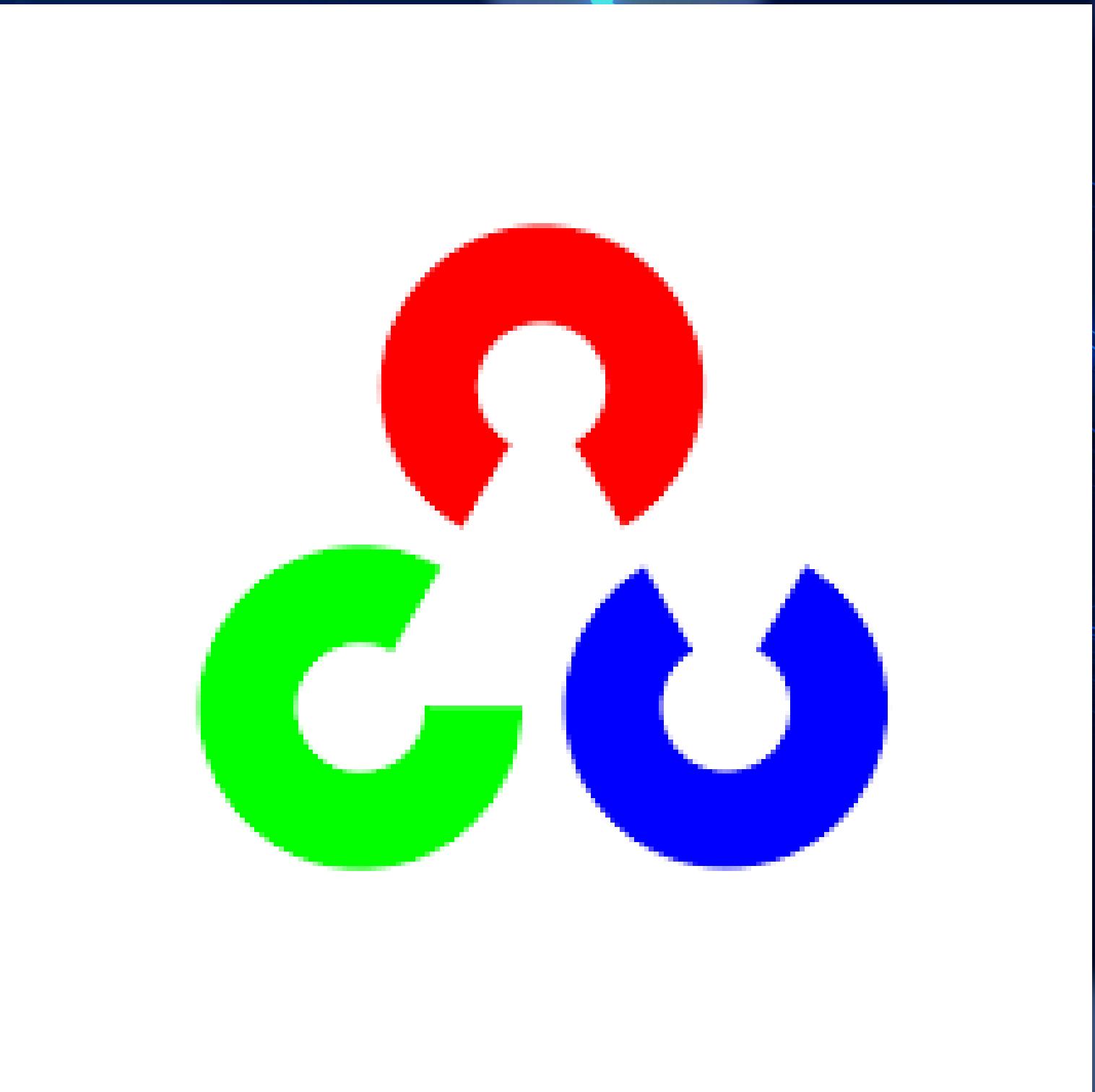


Fonctionnement OCR

Fonctionnement OCR : Open CV

Preprocessing de l'image:

- Redimensionnement de l'image
- Application d'un masque sur l'image
- Convertir en niveau de gris
- Définir le seuil



Fonctionnement OCR: Tesseract et PyTesseract

Traitement OCR:

- Choix de la configuration Optimal
- Création d'un dictionnaire pour stocker les data extraites

Traitement Texte extrait:

- Normalisation des données
- Vérifications des erreurs courantes
- Extractions des informations importantes dans le texte (regex)

Specification fonctionnelles

Fonctionnalité Application : Accueil

- Inscription/Connexion
- Page d'accueil
- Scanner OCR :
- Historique des factures
- Dashboard

The screenshot shows the homepage of the OCR Project application. At the top, there is a dark header bar with the text "OCR Project" and navigation links for "Accueil", "Scanner", "Historique", and "Dashboard". On the far right of the header are "Connexion" and "Inscription" buttons. The main content area has a white background. At the top center, it says "Bienvenue sur OCR Project" and "Votre solution intelligente pour la reconnaissance de texte". Below this are three buttons: "Connexion" and "Inscription" (both blue), and a larger button labeled "Commencer". To the right of these buttons are three sections: "Scanner de Documents" (with a description and a "Commencer" button), "Historique" (with a description and a "Voir l'historique" button), and "Export Facile" (with a description and a "En savoir plus" button). The background of the entire page features a subtle network of blue lines and dots.

OCR Project Accueil Scanner Historique Dashboard

Connexion Incription

Bienvenue sur OCR Project

Votre solution intelligente pour la reconnaissance de texte

Connexion Incription

Scanner de Documents

Numérissez vos documents en quelques clics et convertissez-les en texte modifiable.

Commencer

Historique

Consultez et gérez l'historique de vos documents numérisés.

Voir l'historique

Export Facile

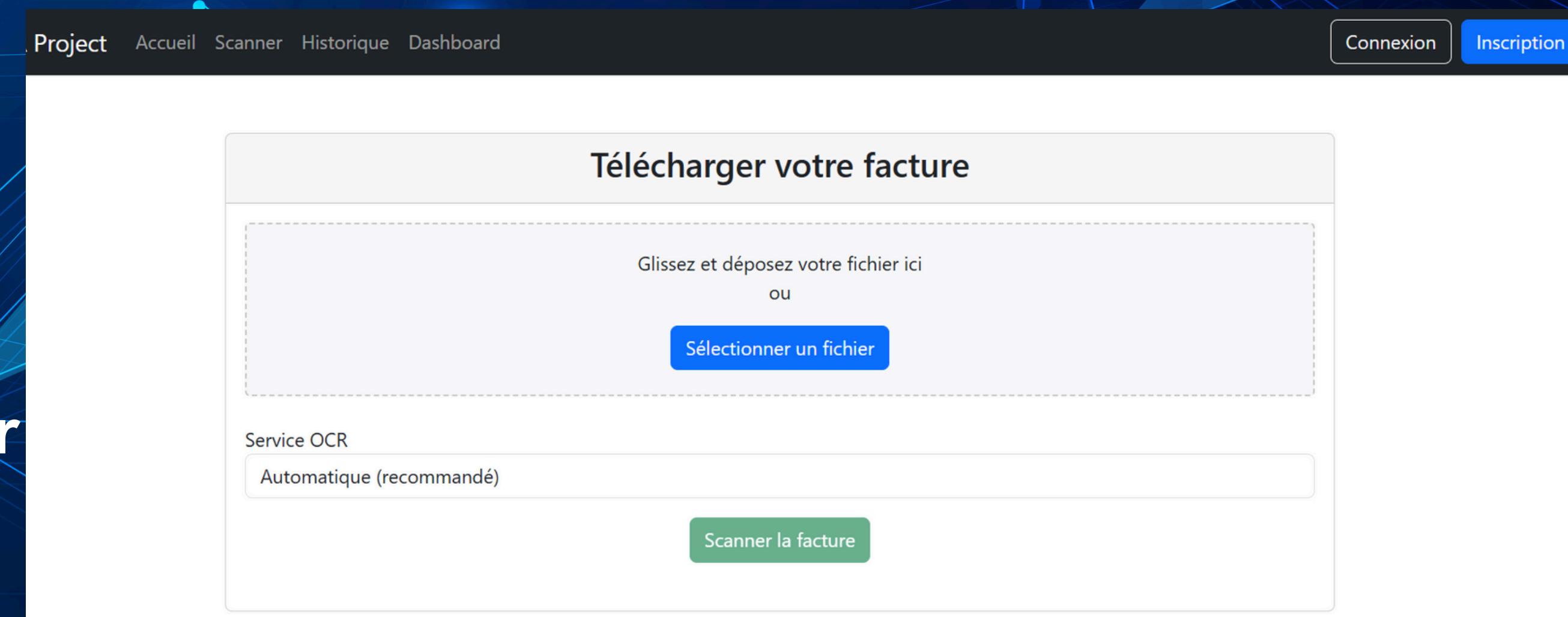
Exportez vos documents numérisés dans différents formats.

En savoir plus

Fonctionnalité Application : Scanner

Scanner OCR:

- Peux scanner une ou plusieurs factures
- Retire au choix des factures du scans
- Permettre un choix automatisé ou personnel pour l'ocr
- Modification si nécessaire des données extraites



Fonctionnalité Application : Historique(pas connecter)

Historique des factures:

- Message permettant de dire à l'utilisateur de se connecter

Historique des documents

Veuillez vous connecter pour voir votre historique de factures.

Nouveau document

Test de l'application

Merci de m'avoir écouté

Avez vous des questions ?