

# PROPOSITION DE CAHIER DES CHARGES DATA INSIGHT PLATFORME

## 1. Introduction

### 1.1 Contexte général

Dans un contexte économique mondial en constante mutation, les institutions financières, les investisseurs, les gouvernements et les cabinets d'étude ont besoin d'outils performants pour analyser, comparer et anticiper les dynamiques économiques. Or, l'accès à des données multisources (BCEAO, Banque Mondiale, FMI, etc.) reste souvent fragmenté, peu interprétable ou non valorisé par des outils modernes d'analyse.

### 1.2 Objectif du projet

Le projet **DIP (Data Insight Plateforme)** vise à concevoir une application web innovante, intégrant des données économiques multisources pour :

- Explorer et visualiser dynamiquement les données,
- Automatiser les analyses macroéconomiques,
- Générer des scores de risque pays,
- Fournir des prédictions économiques,
- Accompagner la prise de décision à travers l'intelligence artificielle.

Cette application se positionnera comme un outil d'aide à la décision basé sur les données, au service des institutions publiques, des décideurs économiques et des experts analystes.

### 1.3 Enjeux stratégiques

- Centraliser les données économiques fiables et actualisées dans une plateforme unique.
- Offrir une lecture augmentée des dynamiques économiques nationales et régionales.

- Faciliter l'accès à l'information via des interfaces intuitives, des modèles prédictifs et des analyses générées par IA.
- Permettre aux utilisateurs d'effectuer leurs propres analyses via des requêtes libres (*self-service analytics*).

## 2. Périmètre fonctionnel

Cette section présente les principales fonctionnalités de la plateforme DIP, structurées autour de ses objectifs stratégiques : analyse de données économiques, modélisation prédictive, intelligence artificielle générative, notation des pays, et autonomie utilisateur.

Chaque fonctionnalité est décrite avec ses objectifs, cas d'usage, éléments techniques clés et interactions éventuelles avec d'autres modules.

### A- Fonctions principales

- Visualisation de données
- Machine Learning supervisé (classification des pays selon les risques / développement/Education /santé/ économique..... )
- Machine Learning non supervisé (clustering des marchés)
- Création d'un indice composite (adoption tech & inclusion)
- Analyse prédictive (comparaison avec les projections du FMI)
- Intelligence Artificielle générative pour l'interprétation
- Self-service analytics (requêtes en langage naturel)
- Intégration des projets ETL et GeoInsights

### B- Description des fonctionnalités

#### 2.b.1 Visualisation de données économiques

**Objectif** : Offrir une lecture graphique claire et interactive des indicateurs économiques par pays, région ou période.

**Description** :

- Tableaux de bord interactifs (PIB, inflation, taux de change, balance commerciale, etc.).
- Filtres dynamiques par pays, années, catégories économiques.
- Visualisation géospatiale via cartes (intégration GeoInsights).
- Export en formats PDF, Excel, PNG.

**Sources :** BCEAO, Banque Mondiale (WDI), FMI (WEO), OWID.

## **2. b.2 Machine Learning supervisé – Classification des pays selon le risque**

**Objectif :** Classer les pays selon leur niveau de risque dans des domaines comme économie , le développement , l'éducation etc..... à partir de modèles entraînés.

### **Description :**

- Modèles ML supervisés (Random Forest, XGBoost, etc.).
- Entraînement sur données historiques (indicateurs macro, notation Bloomfield).
- Affichage des scores par pays avec explication des variables importantes (SHAP/LIME).
- Mises à jour automatiques selon l'évolution des données.

## **2. b.3 Machine Learning non supervisé - Segmentation des marchés**

**Objectif :** Identifier des groupes homogènes de marchés ou pays présentant des similarités économiques.

### **Description :**

- Algorithmes de clustering (K-means, DBSCAN...).
- Utilisation pour recommandations stratégiques (groupes cibles).
- Visualisation des clusters sur carte.

## **2. b.4 Indice composite de l'adoption technologique et inclusion financière**

**Objectif :** Construire un indicateur unique mesurant l'intégration des technologies financières par les populations.

**Description :**

- Pondération de variables issues de la BCEAO (taux d'accès aux services digitaux, nombre de comptes mobile money, etc.).
- Normalisation, scoring, affichage des tendances.
- Comparaison entre pays.

## **2. b.5 Notation souveraine des pays**

**Objectif :** Fournir une notation synthétique de la solvabilité ou du risque économique d'un pays, inspirée des pratiques de Bloomfield Investment.

**Description :**

- Méthodologie transparente (notation AAA à C).
- Intégration de critères objectifs (dette/PIB, croissance, etc.).
- Historique des notes et alertes en cas de dégradation.

## **2. b.6 Analyse prédictive et projection économique**

**Objectif :** Prédire l'évolution d'indicateurs clés (PIB, inflation...) et comparer avec les projections officielles.

**Description :**

- Régression, séries temporelles (ARIMA, Prophet...).
- Affichage comparatif : prévision DIP vs. prévision FMI.
- Courbes interactives, scénarios alternatifs.

## **2. b.7 Intelligence Artificielle générative (interprétation automatique)**

**Objectif :** Générer automatiquement des commentaires intelligents sur les résultats présentés.

**Description :**

- Intégration d'un agent conversationnel IA (LLM).
- Résumé des tendances, anomalies, recommandations.
- Multilingue (FR/EN), basé sur les données visibles à l'écran.

## **2. b.8 Self-service Analytics Requêtes libres utilisateur**

**Objectif :** Permettre aux utilisateurs de poser des questions en langage naturel et d'obtenir des résultats analytiques.

**Description :**

- Interface de requête textuelle
- Résultats en tableau, graphique ou réponse textuelle.
- Historique des requêtes, suggestions automatiques.

## **2. b.9 Intégration des modules ETL et GeoInsights**

**Objectif :** Automatiser l'extraction, le traitement et la structuration des données, ainsi que leur représentation géographique.

**Description :**

- Pipeline ETL (extraction des données BCEAO, WEO...).
- Nettoyage, normalisation et structuration des données.
- Géovisualisation (indicateurs sur carte interactive, calques par région ou pays).

## **3. Public cible et utilisateurs**

**Note :** À ce stade du projet, la définition des cibles (utilisateurs finaux, profils métiers, etc.) a été mise en attente à la demande de la hiérarchie. Cette section pourra être complétée ou précisée ultérieurement en fonction de l'évolution des besoins ou des orientations stratégiques.

### **3.2 Bénéfices pour les utilisateurs**

- Accès centralisé à des données multisources fiables.
- Visualisation rapide des dynamiques économiques.
- Outils d'aide à la décision avancés intégrant IA et ML.
- Possibilité de poser des questions en langage naturel et d'explorer les données de manière autonome.

## 4. Contraintes et exigences techniques

La réussite de l'application DIP repose sur la mise en œuvre d'une architecture technique robuste, scalable et sécurisée, en cohérence avec les standards actuels du développement web, de la data science et des systèmes d'analyse intelligents.

**Note importante :**  
Les technologies, outils et frameworks mentionnés dans cette section sont proposés à titre indicatif sur la base des recherches et benchmarks réalisés dans le cadre de la préparation de ce cahier des charges. Le choix définitif des technologies sera affiné et validé en concertation avec l'équipe technique, en fonction des compétences internes, des contraintes d'architecture, et des objectifs de performance et de sécurité.

### 4.1 Contraintes fonctionnelles

- **Accessibilité Web** : L'application doit être accessible via un navigateur moderne (Chrome, Firefox, Edge).
- **Responsive design** : L'interface doit s'adapter aux différents supports (PC, tablette, smartphone).
- **Multilingue** : Interface disponible au minimum en français et en anglais.
- **Temps de réponse rapide** : Moins de 3 secondes pour afficher les résultats d'une requête ou d'un tableau de bord.

### 4.2 Contraintes techniques

- **Langages / Frameworks**:
  - Back-end : Python (Flask, FastAPI) ou Node.js
  - Front-end : React.js avec Librairie de data viz (Plotly, D3.js, Recharts)
  - Base de données : PostgreSQL MongoDB
  - Modèles ML : Scikit-learn, XGBoost, Prophet, TensorFlow
  - Intégration IA : API OpenAI / LangChain / GPT
- **ETL** : Airflow, Talend ou pipeline personnalisé en Python
- **Géo-visualisation** : Leaflet, Kepler.gl, Mapbox ou intégration GeoInsights

- **Architecture cloud** : Déploiement envisagé sur AWS, Azure

#### 4.3 Sécurité & confidentialité

- Authentification sécurisée (OAuth2, double authentification)
- Chiffrement des données sensibles (au repos et en transit)
- Journalisation des accès et opérations critiques
- Gestion des rôles et permissions par profil utilisateur
- Sauvegardes régulières, plan de reprise d'activité

#### 4.4 Contraintes réglementaires

- Conformité RGPD ou législation locale (protection des données)
- Transparence des modèles algorithmiques
- Traçabilité des sources de données (origines, fréquence de mise à jour)

### 5. Architecture technique de la solution

La plateforme DIP s'appuie sur une architecture modulaire, sécurisée et évolutive permettant l'intégration fluide de différentes sources de données, le traitement intelligent des informations, et la mise à disposition d'interfaces adaptées aux différents profils utilisateurs.

#### 5.1 Architecture générale (vue logique)

[Sources de données]

|

▼

[Module ETL & Data Cleaning]

|

▼

[Base de données centrale]

|

|

|

└─► [Modèles de Machine Learning & Scoring]

|

└─► [Moteur IA générative (requêtes / résumés)]

|

▼

[Back-end API & Services métiers]

|



[Front-end utilisateur]

└─ Tableau de bord interactif

└─ Module Self-Service Analytics

└─ Carte interactive (GeoInsights)

└─ Interface IA conversationnelle

## 5.2 Description des modules

### a) Sources de données

- Données économiques brutes issues de la BCEAO, Banque Mondiale (WDI, DataBank), FMI (WEO), UNU-WIDER, OWID, etc.
- Formats divers : CSV, API REST, fichiers Excel, requêtes SQL.

### b) Module ETL (Extract, Transform, Load)

- Automatisation de l'extraction et du nettoyage des données.
- Intégration des projets (ETL & GeoInsights).
- Harmonisation des indicateurs, gestion des fréquences de mise à jour.

### c) Base de données centrale

- Stockage structuré et historisé des données économiques nettoyées.
- Système relationnel ou hybride selon les besoins.

### d) Modèles de Machine Learning

- Entraînement de modèles supervisés et non supervisés (classification, clustering).
- Stockage et appel des modèles via API.

### e) Moteur IA générative

- Interface avec un LLM pour générer des analyses automatiques ou répondre aux requêtes utilisateurs.
- Intégration possible via OpenAI, LangChain ou équivalents.

### f) API et services métiers



- Interface de communication entre les modules internes.
- Normalisation des accès aux données, sécurisation des flux.

#### **g) Interface front-end**

- Dashboards interactifs, cartes dynamiques, filtres par pays/période.
- Requête en langage naturel (self-service).
- Module de notation des pays et indicateurs composites.

### **6. Design UX/UI et expérience utilisateur**

#### **6.1 Principes généraux**

- **Simplicité d'usage** : L'interface doit rester intuitive pour des profils non techniques (décideurs, consultants, étudiants, etc.).
- **Cohérence visuelle** : Utilisation de la charte graphique de METRICS-DECISION.
- **Lisibilité** : Priorité à la clarté des textes, des chiffres et des graphiques.
- **Feedback utilisateur** : Affichage d'alertes claires (chargement, erreurs, résultats).
- **Responsiveness** : Affichage optimal sur PC, tablette et smartphone.

### **7. Livrables attendus**

#### **7.1 Livrables fonctionnels**

- **Plateforme web opérationnelle** incluant l'ensemble des modules spécifiés :
  - Dashboards et visualisation des données économiques
  - Moteurs de Machine Learning (classification, clustering)
  - Générateur d'indice composite
  - Module de notation pays
  - Interface de requête IA (chatbot analytique)
  - Carte interactive avec données géoéconomiques (GeoInsights)

## 7.2 Livrables techniques

- **Base de données structurée et documentée** (schémas, relations, dictionnaire des données)
- **Pipelines ETL** avec documentation (scripts d'extraction, transformation, chargement)
- **Modèles de Machine Learning déployés** avec :
  - Jeux de données d'entraînement et de test
  - Métriques de performance
  - Explicabilité (SHAP, importance des variables)
- **API back-end** documentée

## 7.3 Documentation et supports

- **Documentation utilisateur** (guide d'utilisation de la plateforme)
- **Documentation technique** (infrastructure, code, dépendances, déploiement)
- **Manuel de déploiement et d'administration**
- **Journal des versions** (historique des évolutions)

## 8. Planning prévisionnel

Le planning prévisionnel permet de faire la suivie des grandes étapes de mise en œuvre du projet DIP.

Phase	Objectifs clés	Durée estimée
Phase 0 : Cadrage	Finalisation du cahier des charges, validation des objectifs, constitution de l'équipe projet	-----
Phase 1 : Conception	Architecture technique, maquettes UI/UX, choix technologiques définitifs, design des modèles de données	-----
Phase 2 : Développement	Développement des modules back-end, front-end, ETL, ML, IA	-----

Phase	Objectifs clés	Durée estimée
<b>Phase 3 : Intégration</b>	Intégration des modules entre eux, tests fonctionnels, sécurité, performance	-----
<b>Phase 4 : Validation</b>	Recette utilisateur, ajustements finaux, validation des livrables	-----
<b>Phase 5 : Déploiement</b>	Mise en ligne, configuration des accès, documentation utilisateur	-----
<b>Phase 6 : Suivi initial</b>	Monitoring, support, corrections post-déploiement, premières itérations	-----

### Remarques :

- Le présent planning constitue une proposition initiale. Il pourra être ajusté en fonction :
  - Des ressources humaines et techniques disponibles,
  - Du périmètre fonctionnel retenu pour la version MVP,
  - Et surtout des décisions collectives prises avec l'ensemble des parties prenantes lors des ateliers de cadrage.
  -

## 9. Annexes

### 9.1 Fichier consolidé (source de données interne)

- **Contenu :**
  - Paysage bancaire de l'UEMOA
  - Analyse macroéconomique de l'UEMOA
  - Notation des pays (Bloomfield Investment)

### 9.2 Sources officielles de données économiques

- BCEAO : <https://edenpub.bceao.int>

- FMI – World Economic Outlook (WEO)
- Banque Mondiale – DataBank : <https://databank.worldbank.org/source/world-development-indicators#>
- Data360 : <https://data360.worldbank.org/en/economy/AGO>
- UNU-WIDER : <https://www4.wider.unu.edu/country-sheet>
- OWID : Our World In Data (fintech, inclusion)
- Goldhub – Production aurifère : <https://www.gold.org/goldhub/data/gold-production-by-country>
- Kpler – Données logistiques & matières premières : <https://www.kpler.com>

### 9.3 Benchmarks analysés

- Coface – Fiches Risques Pays : <https://www.coface.com/.../fiches-risques-pays>
- World Bank Data360
- UNU-WIDER Country Sheets
- BCEAO Edenpub
- Kpler – Visualisation de flux