Cahier des charges

1. Contexte du projet

FashMatch est une application de recommandation de vêtements destinée à offrir aux utilisateurs une expérience d'achat personnalisée.

La première version du système (FashMatch V1) utilisait des données sensibles, fonctionnait sans supervision humaine, et présentait de fortes non-conformités vis-à-vis du RGPD et de l'AI Act. Le traitement de données multimédias (photos, vidéos), l'analyse automatisée du visage et les décisions systématiquement automatisées exposent l'entreprise à des risques juridiques, éthiques et techniques majeurs.

Afin de corriger ces dérives, le projet FashMatch V2 a pour objectif de proposer une architecture entièrement refondue, sécurisée, responsable et conforme aux exigences réglementaires. Cette nouvelle version élimine tout traitement de contenu visuel ou multimédia, repose uniquement sur des données déclaratives et comportementales non sensibles, et intègre des mécanismes structurants de supervision humaine, de journalisation complète, d'explicabilité et de minimisation des données.

La V2 vise ainsi à fournir un moteur de recommandation performant, transparent, auditable, traçable et adaptable, garantissant la protection des utilisateurs tout en respectant les règles européennes en vigueur.

2. Objectifs du projet

Le projet FashMatch V2 a pour finalité de concevoir un système de recommandation conforme, sécurisé et transparent, en rupture totale avec les pratiques problématiques identifiées dans la V1. Les objectifs s'organisent autour de quatre axes structurants : conformité, performance, gouvernance et protection des utilisateurs.

2.1. Objectifs fonctionnels

- Proposer à chaque utilisateur des recommandations pertinentes basées uniquement sur :
    - ses préférences déclarées (style, tailles, couleurs, budget, marques),
    - son comportement de navigation (likes, dislikes, historiques de consultation).
- Fournir une interface claire présentant les recommandations et leurs explications.

- Permettre à l’utilisateur de contrôler à tout moment ses préférences, son profil et ses consentements.

- Offrir un mécanisme de feedback permettant d’améliorer continuellement le moteur de recommandation.

2.2. Objectifs techniques

- Implémenter un moteur de recommandation hybride (collaboratif + contenu non visuel).

- Établir un Feature Store contenant exclusivement des données non sensibles et pseudonymisées.

- Développer un module d’explicabilité basé sur SHAP ou LIME, capable d’expliquer chaque recommandation.

- Garantir une supervision humaine pour toute décision sensible ou automatisée susceptible d’avoir un impact sur l’utilisateur.

- Assurer un stockage sécurisé, pseudonymisé et segmenté des données.

2.3. Objectifs de conformité et sécurité

- Garantir une conformité stricte au RGPD :
  minimisation, pseudonymisation, base légale clairement définie, gestion des droits, conservation limitée.

- Répondre aux obligations AI Act pour les systèmes à risque élevé :
  supervision humaine (HITL), explicabilité, registre IA, documentation technique (model cards, data cards), tests de biais et dérive.

- Mettre en place un système complet de journalisation (logs immuables) pour assurer la traçabilité des décisions du moteur IA.

- Assurer qu' aucune donnée sensible, aucun média et aucune information biométrique ne soient collectés ou traités.

- Sécuriser la plateforme via chiffrement, contrôle d’accès strict (RBAC) et gestion des clés.

2.4. Objectifs organisationnels

- Définir clairement les rôles entre équipe IA, équipe technique, équipe sécurité et DPO.
- Instaurer des procédures d'audit IA et de revue régulière des modèles.
- Structurer un registre RGPD et IA Act centralisé (Console DPO).

3. Périmètre du système

Le périmètre fonctionnel et technique de FashMatch V2 définit précisément ce que le système couvre, ce qu'il exclut, et les limites imposées par les exigences réglementaires et de sécurité. Le périmètre du projet découle directement des risques identifiés dans la V1 et des recommandations du DPO.

3.1. Périmètre inclus (ce que FashMatch V2 fait)

Fonctionnel

- Création et gestion d'un compte utilisateur.
- Vérification de l'âge (modèle conservateur + supervision humaine).
- Gestion du consentement (opt-in, opt-out, retrait).
- Saisie et modification des préférences utilisateur.
- Analyse de l'historique de navigation (likes, dislikes, produits consultés).
- Génération de recommandations personnalisées.
- Affichage d'explications concernant les recommandations.
- Possibilité pour l'utilisateur d'exercer ses droits RGPD (accès, rectification, suppression).
- Feedback utilisateur améliorant le modèle.

Technique / IA

- Pipeline IA basé exclusivement sur des données déclaratives et comportementales non sensibles.
- Vectorisation et extraction de features non sensibles.

- Moteur hybride de recommandation (collaboratif + contenu non visuel).
- Module d'explicabilité par méthode XAI (SHAP/LIME).
- Supervision humaine pour les décisions sensibles.
- Journalisation complète des décisions IA.
- Monitoring du modèle : performance, dérive, biais.

Sécurité & conformité

- Pseudonymisation systématique des données utilisateur.
- Stockage chiffré et cloisonné.
- Registre RGPD + registre IA Act.
- Architecture Privacy by Design / Security by Design.
- Documentation complète IA (model cards, data cards).

3.2. Hors périmètre (exclusions explicites)

Conformément aux risques identifiés dans la V1 et aux recommandations du DPO, les éléments suivants sont formellement exclus du système :

Données multimédias

- Pas d'upload de photos.
- Pas d'upload de vidéos.
- Pas d'upload de sons.
- Pas d'analyse visuelle de l'utilisateur.
- Pas d'extraction de caractéristiques d'images.

Données biométriques ou sensibles

- Pas de reconnaissance faciale.
- Pas de détection d'émotions.

- Pas de données relatives à la santé, religion, origine, orientation, etc.
- Pas de scoring psychologique ou comportemental sensible.

Fonctionnalités intrusives

- Pas de recommandations basées sur des profils psychométriques.
- Pas de push marketing sans consentement explicite.
- Pas de décision automatisée sans supervision humaine.
- Pas d'enregistrement de données qui ne sont pas strictement nécessaires.

Non objectifs

- Pas d'analyse ou collecte via réseaux sociaux.
- Pas d'algorithme d'analyse de l'apparence.
- Pas de prise de décisions avec impact juridique, financier ou social.

3.3. Justification du périmètre

Le périmètre a été défini pour :

- Éliminer tous les risques identifiés dans la V1.
- Se conformer au RGPD (minimisation, absence de données sensibles).
- Se conformer à l'AI Act (supervision, traçabilité, transparence).
- Garantir une architecture contrôlable, stable et industrialisable.
- Limiter le système aux données strictement nécessaires à la recommandation.

4. Parties prenantes

Le développement de FashMatch V2 implique plusieurs acteurs aux rôles complémentaires. La réussite du projet dépend d'une coordination étroite entre les équipes techniques, l'équipe conformité et les décideurs stratégiques.

4.1. Direction et pilotage

Direction exécutive (CEO / CTO)

- Définition de la vision stratégique.
- Validation des décisions techniques et fonctionnelles.
- Arbitrage des choix liés à la conformité et au budget.

Direction produit

- Définit le positionnement fonctionnel de l'application.
- Priorise les fonctionnalités.
- S'assure de la cohérence de l'expérience utilisateur.

4.2. Équipe Data & IA

Data Scientists

- Conception du moteur de recommandation.
- Développement du pipeline IA (vectorisation, modèle hybride).
- Implémentation du module d'explicabilité.
- Tests de performance, robustesse, équité et dérive.

Data Engineers

- Mise en place du Feature Store et des pipelines de données.
- Sécurisation du stockage et des flux de données.
- Intégration du moteur IA dans le système global.

AI Engineers

- Versionnage des modèles (Model Registry).
- Monitoring en production (drift, biais, disponibilité).
- Documentation technique (model cards).

4.3. Équipe backend / infrastructure

Développeurs backend

- Construction de l'API Gateway.
- Mise en place des microservices (profil, consentement, recommandation).
- Gestion des droits, authentification et vérification d'âge.

DevOps / Cloud Engineers

- Déploiement sur infrastructure sécurisée.
- Gestion du chiffrement, des certificats, des backups.
- Surveillance de l'intégrité du système.

4.4. Équipe conformité & gouvernance

DPO (Délégué à la Protection des Données)

- Analyse d'impact RGPD (AIPD) et registre des traitements.
- Validation des bases légales et minimisation des données.
- Supervision du système de logs.
- Vérification des procédures d'exercice des droits.
- Contrôles réguliers et audits de conformité.

Responsable gouvernance IA

- Maintien du registre IA Act.
- Documentation et traçabilité du cycle de vie des modèles.
- Définition des critères de supervision humaine (HITL).
- Réalisation des audits de biais, robustesse et sécurité.

Opérateurs humains / personnel HITL

- Revue et validation des décisions sensibles.
- Gestion des cas bloquants (vérification d'âge, anomalies IA).
- Documentation des interventions humaines.

4.5. Parties prenantes externes

Autorités de régulation (CNIL / UE)

- Consultation en cas de risque élevé.
- Contrôle du registre IA et RGPD.

Utilisateurs finaux

- Fournissent les informations nécessaires au service.
- Bénéficient d'un droit total sur leurs données.
- Interagissent avec l'interface et donnent du feedback.

4.6. Justification du rôle de chaque partie prenante

Chaque partie joue un rôle essentiel pour :

- Garantir la conformité (DPO, responsable IA).
- Maintenir la performance du moteur de recommandation (Data/IA).
- Assurer la sécurité et la fiabilité du système (DevOps / backend).
- Protéger l'utilisateur et assurer sa confiance (pilotage + UX).

5. Fonctionnalités

Les fonctionnalités de FashMatch V2 sont organisées en quatre catégories : fonctionnalités utilisateur, fonctionnalités IA, fonctionnalités de sécurité/conformité et fonctionnalités de supervision.

Toutes les fonctionnalités ont été redéfinies afin d'assurer une conformité stricte et d'éliminer les dérives identifiées dans la V1.

5.1. Fonctionnalités utilisateur

5.1.1. Création et gestion de compte

- Création d'un compte utilisateur avec email et mot de passe.
- Vérification d'âge obligatoire avant l'accès au service.
- Possibilité de modifier les informations du profil.

5.1.2. Gestion des préférences

- Déclaration des préférences personnelles : style, couleurs, tailles, budget, marques.
- Mise à jour à tout moment.
- Interface simple et intuitive.

5.1.3. Historique et interactions

- Likes / dislikes sur les produits.
- Historique de consultation.
- Utilisation du comportement pour affiner l'IA.

5.1.4. Restitution des recommandations

- Suggestions basées sur préférences + comportement.
- Filtrage par marques, prix, catégories.
- Feedback utilisateur.

5.1.5. Explicabilité

- Affichage du "Pourquoi cette recommandation ?"
- Raisons basées sur préférences ou interactions.

5.1.6. Gestion des droits RGPD

- Accès aux données.

- Suppression du compte.
- Export des données.
- Gestion du consentement.

5.2. Fonctionnalités IA

5.2.1. Recommandation hybride

- IA basée uniquement sur données déclaratives + comportementales.
- Filtrage collaboratif + contenu (non visuel).
- Pondération automatique.

5.2.2. Construction de features

- Vectorisation des préférences + interactions.
- Aucune donnée sensible.
- Aucune donnée multimédia.

5.2.3. Explicabilité

- Méthodes SHAP/LIME.
- Décomposition de la prédiction.
- Explications lisibles côté utilisateur.

5.3. Fonctionnalités de sécurité & conformité

5.3.1. Vérification d'âge

- Modèle conservateur. (Un modèle volontairement strict qui préfère refuser plutôt qu'accepter)
- Jamais de refus automatique définitif.
- Supervision humaine obligatoire.

5.3.2. Gestion du consentement

- Enregistrement opt-in / opt-out.
- Blocage en absence de consentement valide.
- Traçabilité complète.

5.3.3. Pseudonymisation

- Identifiant pseudo-aléatoire.
- Séparation identité / données IA.
- Aucun traitement sensible.

5.3.4. Sécurité des données

- Chiffrement TLS + AES-256.
- Environnement UE certifié.
- RBAC.
- Logs immuables WORM.

5.4. Fonctionnalités de supervision & gouvernance IA

5.4.1. Supervision humaine (HITL)

- Validation des décisions sensibles.
- Interface interne.
- Journalisation des interventions.

5.4.2. Registre IA

- Documentation cycle de vie.
- Model Cards & Data Cards.
- Monitoring (biais, dérive, performance).

5.4.3. Console DPO

- Visualisation logs IA.
- Registre RGPD.
- Gestion des demandes utilisateurs.
- Rapports d’audit.

6. Contraintes du système

Les contraintes définissent les limites réglementaires, techniques, organisationnelles et fonctionnelles imposées au développement de FashMatch V2. Elles garantissent la faisabilité, la conformité et la sécurité du système.

6.1. Contraintes réglementaires

6.1.1. RGPD (Règlement Général sur la Protection des Données)

- Minimisation des données.
- Interdiction des données sensibles et biométriques.
- Pseudonymisation systématique.
- Journalisation complète des traitements.
- Droit d’accès, rectification, suppression, export.
- Conservation limitée et documentée.
- Consentement explicite et traçable.
- Horodatage des modifications de consentement.

6.1.2. AI Act (Règlement Européen sur l’IA)

- Supervision humaine obligatoire (HITL).
- Explicabilité imposée.
- Documentation modèle (Model Cards).

- Documentation données (Data Cards).
- Monitoring biais, robustesse, dérive.
- Registre IA obligatoire.
- Journalisation des inférences.
- Validation humaine des décisions à impact.

6.1.3. Droit de la consommation

- Pas de profilage manipulatoire.
- Pas de push marketing sans consentement.
- Pas de discrimination algorithmique.

6.2. Contraintes techniques

6.2.1. Données traitées

- Données non sensibles uniquement.
- Aucune donnée multimédia.
- Aucune donnée biométrique.
- Données autorisées : préférences, interactions, compte.

6.2.2. Architecture

- Architecture modulaire et segmentée.
- Frontend → API → Profil → IA → Logs → DPO.
- Stockage exclusivement UE.
- Feature Store non sensible.

6.2.3. Sécurité

- Chiffrement TLS 1.3.

- Chiffrement AES-256 au repos.
- RBAC strict.
- Journalisation immuable (WORM).
- Rotation des clés.
- Monitoring sécurité.

6.3. Contraintes IA

6.3.1. Supervision humaine (HITL)

- Aucune décision automatisée sans validation humaine.
- Cas concernés : âge, anomalies, risques.
- Interface interne dédiée.
- Journalisation des actions humaines.

6.3.2. Explicabilité

- Explication obligatoire pour chaque recommandation.
- Méthodes SHAP / LIME.
- Explication compréhensible par un non-expert.
- Aucun modèle "black box" non justifiable.

6.3.3. Monitoring IA

- Détection de dérive.
- Audit trimestriel des biais.
- Suivi performance (précision, couverture, diversité).
- Versionnage modèles (Model Registry).

6.4. Contraintes organisationnelles

6.4.1. Gouvernance

- Rôles distincts : IA, DPO, opérateurs HITL.
- Séparation des responsabilités.
- Processus d’audit régulier.

6.4.2. Documentation

- Registre RGPD.
- Registre IA Act.
- AIPD.
- Model Cards & Data Cards.
- Procédures HITL.
- Politique de conservation.

6.4.3. Formation

- Formation obligatoire des opérateurs humains.
- Formation IA aux contraintes AI Act.
- Sensibilisation aux biais et dérives IA.

6.5. Contraintes de performance

Temps et disponibilité

- Recommandation en < 500 ms.
- Disponibilité ≥ 99.5%.

Scalabilité

- Scalabilité horizontale.
- Résilience aux pics de trafic.

Robustesse

- Gestion des données incomplètes.
- Tolérance aux anomalies.

7. Architecture du système

7.1. Vue d'ensemble de l'architecture

L'architecture de FashMatch V2 repose sur une segmentation stricte des modules afin d'assurer :

- la conformité RGPD et AI Act,
- la sécurité des données,
- la traçabilité des traitements,
- la supervision humaine,
- et la performance du moteur IA.

L'architecture suit une approche **Privacy by Design** et **Security by Design**, où chaque composant est isolé, journalisé et contrôlé.

7.2. Composants principaux

7.2.1. Frontend

- Interface utilisateur Web/Mobile.
- Collecte des préférences.
- Interaction : likes, dislikes, navigation.
- Restitution des recommandations et explications.

7.2.2. API Gateway

- Point d'entrée unique du système.
- Authentification (OAuth2 + MFA).
- Contrôles d'accès.
- Rate limiting.

- Filtrage des requêtes.

7.2.3. Module Consentement & Âge

- Gestion des consentements utilisateur.
- Vérification d'âge avec modèle conservateur.
- Supervision humaine systématique en cas de doute.
- Horodatage des validations.

7.2.4. Module Profil & Données utilisateur

- Stockage pseudonymisé du profil.
- Gestion des préférences.
- Historique d'interactions.
- Séparation stricte identité / données IA.

7.3. Data Layer

7.3.1. Base Utilisateur pseudonymisée

- Stockage des données de profil non sensibles.
- Données isolées par identifiant pseudonymisé.
- Chiffrement au repos (AES-256).

7.3.2. Feature Store non sensible

- Vecteurs anonymisés des préférences et comportements.
- Strictement aucune donnée multimédia.
- Utilisé par le moteur IA de recommandation.

7.3.3. Catalogue Produits

- Base de produits avec caractéristiques textuelles.

- Métadonnées nécessaires au scoring IA.

7.3.4. Logs immuables (WORM)

- Journalisation complète des traitements IA.
- Logs de consentement, âge, supervision, explication.
- Base non modifiable (Write Once Read Many).

7.4. Pipeline IA

7.4.1. Vectorisation

- Transformation des préférences et interactions en vecteurs exploitables.
- Normalisation et agrégation.

7.4.2. Moteur de Recommandation Hybride

- Filtrage collaboratif (utilisateurs similaires).
- Filtrage basé contenu (caractéristiques produits).
- Pondération automatique selon le contexte utilisateur.

7.4.3. Module d’explicabilité

- Explication des recommandations via SHAP/LIME.
- Génération de phrases explicatives compréhensibles.
- Journalisation des explications.

7.5. Supervision et conformité

7.5.1. Supervision humaine (HITL)

- Intervention humaine obligatoire sur décisions sensibles.
- Interface interne pour les opérateurs.
- Journalisation complète de toutes les actions.

7.5.2. Console DPO

- Accès aux logs IA.
- Gestion du registre RGPD.
- Gestion du registre IA Act.
- Suivi des demandes d'exercice des droits.

7.6. Flux global (par étapes)

Étape 1 — Utilisateur

L'utilisateur interagit avec l'application et renseigne ses préférences.

Étape 2 — Consentement & âge

Le système vérifie les consentements + âge (avec supervision humaine en cas de doute).

Étape 3 — Profil & données

Les préférences et interactions sont stockées de manière pseudonymisée.

Étape 4 — Feature Store

Construction de vecteurs non sensibles à partir des données du profil.

Étape 5 — IA

Le moteur IA hybride génère des recommandations + explications.

Étape 6 — Supervision

Les décisions sensibles sont évaluées par un opérateur humain (HITL).

Étape 7 — Restitution

L'utilisateur reçoit ses recommandations accompagnées d'explications.

7.7. Conformité de l'architecture

Principes appliqués

- Minimisation des données.
- Aucune donnée sensible ni visuelle.

- Pseudonymisation systématique.
- Séparation des services.
- Journalisation immuable.
- Supervision humaine obligatoire.
- Documentation IA complète.
- Transparence par design.

Schéma simplifié

```
[Utilisateur]
        ↓
[Application Web / App Mobile]
        ↓
[API Gateway – OAuth2 / MFA]
        ↓
[Module Vérification d'âge + Consentement]
        ↓
[Profil + Données utilisateur pseudonymisées]
        ↓
[Feature Store (données non sensibles)]
        ↓
[Moteur IA hybride de recommandation]
        ↓
[Module d'explicabilité + Supervision humaine (HITL)]
        ↓
[Logs immuables + Console DPO]
        ↓
[Résultats affichés à l'utilisateur]
```

8. Pipeline IA détaillé

8.1. Vue générale du pipeline IA

Le pipeline IA de FashMatch V2 est conçu pour respecter la minimisation des données, la transparence, la supervision humaine et la conformité RGPD/AI Act.
Il repose exclusivement sur des données **non sensibles**, issues du profil utilisateur et de son comportement.

Le pipeline se compose de 6 grandes étapes :

1. Collecte des données déclaratives et comportementales
2. Vectorisation
3. Construction des features
4. Recommandation hybride
5. Explicabilité
6. Supervision humaine

8.2. Étape 1 — Collecte des données

8.2.1. Données déclaratives

- Préférences utilisateur : style, couleurs, tailles, budget, marques.
- Conditions spécifiques : type de vêtements recherchés, saisonnalité.

8.2.2. Données comportementales

- Produits consultés.
- Likes / dislikes.
- Temps passé sur certains types de produits.
- Ordre d'interaction (séquence comportementale).

8.2.3. Données autorisées

- Uniquement données non sensibles.
- **Aucune donnée multimédia ou biométrique.**
- Données pseudonymisées dès l'ingestion.

8.3. Étape 2 — Vectorisation

8.3.1. Normalisation des données

- Transformation des préférences en vecteurs numériques.

- Mise à l'échelle des valeurs (normalisation, encodage).

8.3.2. Encodage catégoriel

- Encodage One-Hot ou embeddings pour les catégories :
    - style, marques, couleurs, catégories de produits.

8.3.3. Transformation comportementale

- Séquences → vecteurs continus.
- Poids spécifiques attribués aux dernières interactions (recent weighting).

8.4. Étape 3 — Construction des features

8.4.1. Agrégation

- Fusion données déclaratives + comportementales.
- Création d'un vecteur final utilisateur.

8.4.2. Feature engineering

- Similarité avec autres profils.
- Similarité avec produits aimés.
- Extraction de caractéristiques du catalogue produit.

8.4.3. Stockage

- Enregistrement dans le Feature Store.
- Strictement aucun identifiant direct → pseudonymisation obligatoire.

8.5. Étape 4 — Moteur IA de recommandation

8.5.1. Filtrage collaboratif

- Recherche d'utilisateurs avec comportements similaires.

- Analyse des patterns de goût.

8.5.2. Filtrage basé contenu

- Rapprochement des caractéristiques des produits aimés.
- Matching avec les attributs produits (texte, tags).

8.5.3. Scoring et ranking

- Score final = combinaison pondérée :
    - score collaboratif,
    - score contenu,
    - score comportemental récent.
- Classement final par pertinence.

8.6. Étape 5 — Explicabilité

8.6.1. Génération explications

- Méthodes XAI : SHAP, LIME.
- Attribution des contributions :
    - préférence couleur,
    - budget,
    - marques aimées,
    - produits similaires consultés.

8.6.2. Restitution

- Phrase explicative claire :
  “Nous vous recommandons ce produit car il correspond à votre style X et vous avez aimé des produits similaires.”

8.6.3. Journalisation

- Chaque explication est enregistrée dans les logs immuables.

8.7. Étape 6 — Supervision humaine

8.7.1. Validation des décisions sensibles

- Recommandations impactantes ou anomalies IA → revue humaine.

8.7.2. Interface opérateur

- Interface interne permettant de valider ou rejeter une décision IA.

8.7.3. Traçabilité

- Chaque intervention humaine est :
    - horodatée,
    - enregistrée dans les logs,
    - accessible au DPO.

9. Sécurité des données

9.1. Principes généraux de sécurité

La sécurité des données constitue un pilier fondamental de FashMatch V2.
Le système doit assurer la confidentialité, l'intégrité et la disponibilité des données à chaque étape du cycle de vie, conformément :

- au RGPD,
- à l'AI Act,
- aux recommandations du DPO,
- et aux normes de sécurité européennes (ISO 27001).

Tous les composants de l'architecture sont conçus selon les approches **Privacy by Design** et **Security by Design**.

9.2. Chiffrement

9.2.1. Chiffrement en transit

- Utilisation obligatoire du protocole **TLS 1.3**.
- Interdiction des versions obsolètes (TLS < 1.2).
- Protection contre les attaques MITM.

9.2.2. Chiffrement au repos

- Chiffrement des bases utilisateur et du Feature Store via **AES-256**.
- Rotation régulière des clés de chiffrement.
- Gestion centralisée des clés dans un système HSM (Hardware Security Module).

9.3. Contrôle d'accès

9.3.1. RBAC (Role-Based Access Control)

- Droits strictement séparés entre :
    - équipe IA,
    - équipe backend,
    - opérateurs HITL,
    - DPO.
- Accès minimum (principe du **least privilege**).
- Politique de non-accès aux données brutes pour les développeurs.

9.3.2. MFA (authentification multi-facteurs)

- Obligatoire pour tous les comptes administrateurs.
- MFA activé pour les opérateurs HITL.

9.3.3. Journaux d'accès

- Journalisation systématique de toute ouverture de session.

- Détection automatique des accès anormaux.

9.4. Pseudonymisation

9.4.1. Identifiant pseudo-aléatoire

- Chaque utilisateur reçoit un identifiant UUID.
- L'identité réelle (email, nom) est isolée dans une base séparée.

9.4.2. Séparation des environnements

- Base d'identifiants → *stockage isolé et chiffré*.
- Données IA → *stockées dans le Feature Store* sans identifiant direct.

9.4.3. Objectif

- Réduire les risques de réidentification.
- Respecter strictement la minimisation du RGPD.

9.5. Journalisation et traçabilité

9.5.1. Logs immuables (WORM)

- Tous les logs IA, consentements, vérifications d'âge et actions HITL sont enregistrés dans un stockage **Write Once Read Many**.
- Impossibilité de modifier ou supprimer les logs.

9.5.2. Traçabilité IA complète

Pour chaque inférence, le système enregistre :

- version du modèle,
- données d'entrée pseudonymisées,
- score généré,
- explication SHAP/LIME,

- actions humaines associées.

9.5.3. Accès DPO

- Le DPO possède un accès dédié pour consulter les logs.
- Export possible pour audits externes ou AIPD.

9.6. Protection contre les attaques

9.6.1. Anti-intrusion

- Système IDS/IPS (Intrusion Detection & Prevention).
- Détection comportementale des anomalies réseau.

9.6.2. Protection API

- Rate limiting pour éviter le spam.
- Protection contre injection, CSRF, XSS.

9.6.3. Surveillance continue

- Monitoring 24/7.
- Alertes automatique en cas :
    - d’accès illégitimes,
    - d’activité suspecte,
    - de dérive IA anormale.

9.7. Hébergement et conformité

9.7.1. Hébergement 100% européen

- Données stockées dans un cloud certifié :
    - ISO 27001

- ISO 27017 / 27018
- HDS si applicable

9.7.2. Isolation des environnements

- Environnements dev / test / prod strictement séparés.
- Données réelles interdites dans les environnements de test.

9.8. Sauvegardes et continuité

9.8.1. Backups

- Sauvegardes automatiques quotidiennes.
- Chiffrement systématique.
- Tests de restauration réguliers.

9.8.2. Plan de continuité

- Redondance des bases de données.
- Plan de reprise après sinistre documenté (PRA).
- Engagement de disponibilité : **99.5%**.

10. Conformité RGPD & AI Act

10.1. Principes généraux de conformité

FashMatch V2 est conçu pour être **pleinement conforme** aux réglementations européennes en matière de protection des données (RGPD) et d'intelligence artificielle (AI Act).
L'architecture suit les principes :

- Privacy by Design
- Security by Design
- Minimisation des données
- Pseudonymisation systématique

- Traçabilité et transparence
- Supervision humaine obligatoire

Le système exclut **toute donnée sensible, biométrique ou multimédia**.

10.2. Conformité RGPD

10.2.1. Base légale

La base légale retenue pour le traitement est :

- **Consentement explicite** pour l'utilisation des données de préférences et interactions.
- **Intérêt légitime** exclu car trop risqué pour un système de recommandation personnalisé.
- Les données sont collectées uniquement pour fournir les recommandations.

10.2.2. Minimisation des données

- Uniquement données strictement nécessaires :
  préférences, interactions, email, âge.
- Aucun traitement de données sensibles ou biométriques.
- Aucun média utilisateur collecté.

10.2.3. Pseudonymisation

- Toute donnée personnelle est liée à un identifiant pseudonymisé (UUID).
- Séparation stricte :
    - base identité (email)
    - base IA (vecteurs anonymisés)

10.2.4. Gestion des droits

- Droit d'accès (export JSON/CSV).
- Droit de rectification (modification profil).

- Droit à l'effacement (suppression complète).
- Droit à la portabilité.
- Droit d'opposition (arrêt de la recommandation IA).
- Portail de gestion des droits.

10.2.5. Registre RGPD

Le DPO maintient un registre comprenant :

- finalité du traitement,
- catégories de données,
- durée de conservation,
- destinataires,
- mesures de sécurité,
- journaux IA,
- documentation AIPD.

10.2.6. AIPD (Analyse d'Impact RGPD)

Obligatoire car :

- système d'IA personnalisé,
- risques potentiels d'erreurs de recommandation,
- traitement algorithmique.
  L'AIPD inclut une évaluation des risques résiduels et mesures d'atténuation.

10.3. Conformité AI Act

10.3.1. Classification

FashMatch V2 est classé **Système d'IA à risque élevé** car :

- impact sur les décisions personnalisées,

- nécessite supervision humaine,
- nécessite traçabilité complète.

10.3.2. Obligations de l'AI Act remplies

Le système intègre :

- Supervision humaine obligatoire (HITL).
- Explicabilité : module SHAP/LIME.
- Journalisation complète des inférences IA.
- Modèle documenté (Model Card).
- Données documentées (Data Card).
- Gestion du cycle de vie du modèle (Model Registry).
- Tests réguliers : biais, robustesse, dérive.
- Registre IA obligatoirement tenu par DPO ou responsable IA.

10.3.3. Documentation AI Act

Documentation inclut :

- Objectifs du système IA.
- Données utilisées.
- Méthodologie d'apprentissage.
- Tests d'équité.
- Architecture IA.
- Procédures de monitoring.
- Procédures HITL.
- Historique de versions.

10.3.4. Transparence

L’utilisateur est informé que :

- une IA génère les recommandations,
- aucune donnée sensible n’est utilisée,
- il peut désactiver l’IA à tout moment,
- il peut obtenir une explication lisible.

10.3.5. Gouvernance du cycle de vie IA

- Responsable IA = garant du respect AI Act.
- DPO = supervision conformité RGPD.
- Opérateurs HITL = validation décisions sensibles.
- Journalisation = base de l’auditabilité.

10.4. Durée de conservation

- Consentements : 5 ans.
- Données de profil : durée d’utilisation + 2 ans d’inactivité.
- Logs IA : durée imposée par AI Act (5 à 7 ans).
- Données pseudonymisées : supprimées en fin de vie du compte.

10.5. Exclusions pour conformité

Le système n’inclut **pas** :

- reconnaissance faciale,
- analyse d’images,
- traitement biométrique,
- analyse comportementale sensible,
- scoring des utilisateurs,

- décisions automatisées sans intervention humaine.

10.6. Engagements de transparence

- L'utilisateur peut consulter sa fiche "Comment ma recommandation est calculée ?".
- L'application affiche :
  "Cette recommandation a été générée par une IA supervisée."
- Les explications sont accessibles pour chaque produit proposé.

11. Spécifications techniques principales

11.1. Gestion du compte utilisateur

11.1.1. Création de compte

Élément	Description
Inscription	Email + mot de passe
Vérification email	Code de confirmation obligatoire
Acceptation	CGU + Politique de confidentialité
Sécurisation	Stockage pseudonymisé + chiffrement

11.1.2. Vérification de l'âge

Élément	Description
Déclaration d'âge	Demande obligatoire lors de l'inscription
Modèle IA	Modèle conservateur qui demande une validation humaine en cas de doute
Décision finale	**Jamais automatisée**, toujours validée par un humain en cas d'anomalie
Journalisation	Décisions + interventions HITL horodatées

11.1.3. Connexion / Déconnexion

Élément	Description
Authentification	Email + mot de passe

Sécurité	MFA activable
Sessions	Gestion sécurisée + expiration automatique

11.2. Gestion des préférences utilisateur

Fonctionnalité	Description
Saisie initiale	Style, couleurs, tailles, budget, marques
Modification	Possible à tout moment
Historisation	Changements horodatés
Données sensibles	**Aucune collectée**
Données multimédia	**Aucune collectée**

11.3. Gestion des interactions utilisateur

Interaction	Description
Like / Dislike	Influence la recommandation
Consultation produit	Historique non intrusif
Pondération	Importance accrue des interactions récentes
Traçabilité	Toutes les interactions sont horodatées

11.4. Recommandations personnalisées

Fonctionnalité	Description
Génération	Liste de produits classés par pertinence
Taille de liste	10–20 suggestions
Scoring	Modèle hybride (collaboratif + contenu)
Mise à jour	Après chaque interaction

Filtres utilisateur

Type de filtre	Exemples

Catégorie	T-shirt, pantalon
Marque	Nike, Zara
Budget	Min / Max
Couleur	Bleu, noir

11.5. Explicabilité des recommandations

Élément	Description
Méthodes	SHAP, LIME
Explications	Fondées sur les préférences et interactions
Exemple	“Produit recommandé car vous aimez la marque X”
Journalisation	Explications enregistrées dans les logs immuables

11.6. Gestion du consentement

Élément	Description
Opt-in / Opt-out	Gestion directe par l’utilisateur
Conditions	Consentement requis avant IA
Historisation	Toutes les modifications horodatées
Blocage	L’IA ne fonctionne pas sans consentement valide

11.7. Gestion des droits RGPD

Droit	Fonctionnalité offerte
Accès	Export des données utilisateur
Rectification	Modification du profil
Suppression	Effacement complet du compte
Portabilité	Export JSON/CSV
Opposition	Désactivation du moteur IA

11.8. Interface utilisateur

Critère	Description
Design	Interface claire, moderne, responsive
Navigation	Gestion profil, recommandations, feedback
Accessibilité	Conformité WCAG 2.1 AA
Explication IA	Accessible depuis chaque suggestion

11.9. Console administrateur / DPO

Vue opérateurs HITL

Élément	Description
Décisions sensibles	Doivent être validées ou rejetées
Interface	Tableau des décisions IA douteuses
Historique	Toutes interventions horodatées

Vue DPO

Élément	Description
Logs IA	Consultation des logs immuables
Registre RGPD	Accès complet
Registre IA Act	Accessible et exportable
Auditabilité	Export pour AIPD ou contrôles CNIL

12. Spécifications techniques détaillées

12.1. Architecture technique

Composant	Description
Frontend	Application Web/Mobile, appels API sécurisés
API Gateway	OAuth2, MFA, rate limiting, filtrage
Microservices	Profil, consentement, recommandation, supervision

Data Layer	Base utilisateur pseudonymisée, Feature Store, logs immuables
IA	Vectorisation, moteur hybride, explicabilité
Supervision	Interface HITL + Console DPO

12.2. Technologies recommandées

Domaine	Technologies possibles
Frontend	React / Next.js / Flutter
Backend	Node.js / Python (FastAPI)
IA	Python (TensorFlow / PyTorch / LightFM)
Base de données	PostgreSQL, Redis, Elasticsearch
Infrastructure	Docker, Kubernetes, Cloud UE (Scaleway / OVH / Azure EU)
Sécurité	Vault, IAM RBAC, TLS 1.3

12.3. Exigences techniques

- Réponse IA < **500 ms**.
- Logs immuables (WORM).
- Stockage 100% UE.
- Feature Store non sensible.
- Model Registry obligatoire (MLflow ou équivalent).
- Scalabilité horizontale.

12.4. Intégration IA

Étape	Description
Préparation données	Normalisation, vectorisation
Entraînement	Modèle hybride
Déploiement	API IA dédiée

Monitoring	Dérive, biais, performance
Explicabilité	SHAP/LIME intégré

12.5. Sécurité technique

- Chiffrement AES-256 et TLS 1.3.
- RBAC + MFA pour administrateurs.
- Audit logs accessibles au DPO.
- Isolation des environnements (dev/test/prod).

13. Livrables et critères d'acceptation

13.1. Livrables attendus

13.1.1. Livrables fonctionnels

Livrable	Description
Application Web/App	Interface utilisateur complète
Module de recommandation IA	Modèle hybride, vectorisation, scoring
Module d'explicabilité	SHAP/LIME intégré et opérationnel
Module consentement + âge	Complet avec supervision humaine
Console DPO	Accès aux logs + registre RGPD/IA
Interface HITL	Revue humaine des décisions sensibles

13.1.2. Livrables techniques

Livrable	Description
API Gateway	OAuth2, MFA, rate-limit
Microservices	Profil, recommandation, consentement
Base pseudonymisée	Stockage sécurisé en UE
Feature Store	Vecteurs non sensibles

Model Registry	Versionnage du modèle IA
Pipelines IA	Entraînement, déploiement, monitoring

13.1.3. Livrables de documentation

Document	Description
Cahier des charges V2	Version finale complète
AIPD	Analyse d'impact RGPD
Registre RGPD	Conformité données
Registre IA Act	Conformité IA (risque élevé)
Model Card	Documentation modèle IA
Data Card	Documentation données utilisées
Politique HITL	Procédures supervision humaine
Politique sécurité	Mesures techniques + organisationnelles

13.2. Critères d'acceptation

13.2.1. Fonctionnels

Critère	Condition d'acceptation
Recommandations IA	Pertinentes, rapides (<500ms)
Explicabilité	Affichée pour chaque recommandation
Consentement	Impossible d'utiliser l'IA sans opt-in
Droits RGPD	Tous les droits fonctionnels accessibles
Vérification d'âge	Jamais automatisée à 100%, toujours HITL en cas de doute

13.2.2. Techniques

Critère	Condition
Stockage UE	100% des données hébergées en UE
Pseudonymisation	Aucune donnée personnelle dans le Feature Store

Chiffrement	TLS 1.3 + AES-256
Journalisation	Logs immuables (WORM)
Performance	API < 300–500 ms

13.2.3. Conformité

Critère	Condition
RGPD	AIPD validée, registre complet
AI Act	Registre à jour, documentation modèle
Transparence	Mention IA visible + explications
Supervision humaine	Présente et opérationnelle

13.3. Livraison finale

- Démonstration du système complet.
- Validation DPO (RGPD + IA Act).
- Validation technique (performance, sécurité).
- Remise des sources + documentation.
- Passage en production sous supervision.