📋 Audit Trail EU AI Act - Guide Complet

Article 11 - Annexe IV | Documentation Technique Obligatoire

🎯 1. DESCRIPTION GÉNÉRALE DU SYSTÈME

📝 1.1 - Objectif prévu, nom du fournisseur, version

Temps audit : 15 minutes | ⚠️ AUCUNE NORME ISO équivalente

Ce que l'auditeur va demander :

> *"Présentez-moi la fiche d'identité complète de votre système IA"*

Documentation OBLIGATOIRE à fournir :

Identité du système IA :

```
Nom du système : CreditScoring_Retail_BNP_v2.1.3
Fournisseur : BNP Paribas Personal Finance
Contact responsable : Marie Dupont (marie.dupont@bnpparibas.com)
Version actuelle : 2.1.3 (production depuis 15/01/2025)
Date de création : Décembre 2024
Dernière mise à jour : 20/01/2025
```

Objectif précis du système :

```
Finalité : Évaluation automatique solvabilité prêts personnels 5K-50K€
Input : Données financières client (revenus, dettes, historique)
Output : Score 0-1000 + recommandation APPROUVER/REJETER/EXAMINER
Impact : Décisions d'accès au crédit (droit fondamental économique)
Utilisateurs finaux : Conseillers bancaires + clients via web
Volume : 200-300 évaluations/jour ouvré
```

Classification EU AI Act automatique :

```
Risque : HIGH RISK
Justification : Annexe III, Section 5(b) - "Systèmes destinés à évaluer
              la solvabilité de personnes physiques"
Conséquences : Soumis à TOUTES les obligations EU AI Act (Articles 8-15)
Délai compliance : Août 2026 (systèmes en production)
```

🚨 **Point critique :** Classification erronée = non-conformité totale ✅ **ConforMind automatise :** Classification instantanée + justification réglementaire

---

🔗 1.2 - Interaction avec hardware/software/autres systèmes IA

Temps audit : 20 minutes | 🏢 Infrastructure générale = ISO 27001 SUFFIT

Ce que l'auditeur va demander (focus IA uniquement) :

> *"Quels systèmes alimentent DIRECTEMENT les décisions de votre IA ?"*

Intégrations CRITIQUES pour l'IA :

```
APIs externes influençant décisions :
├── FICO Credit Score API
│   ├── Impact : 30% du score final
│   ├── Données : Score externe 300-850
│   └── Risque : Biais algorithme FICO non-auditable
├── Banque de France API
│   ├── Impact : Veto si incidents détectés
│   ├── Données : Incidents paiement historiques
│   └── Risque : Données historiquement biaisées
└── CRM interne BNP
    ├── Impact : 20% du score (ancienneté, produits)
    ├── Données : Historique comportemental
    └── Risque : Biais dans données historiques

Autres modèles IA connectés :
├── Modèle détection fraude (preprocessing)
├── Système anti-blanchiment (post-processing)
└── IA scoring comportemental (enrichissement)
```

🏢 Infrastructure réseau/serveurs : → ISO 27001 couvre déjà | L'auditeur validera juste le certificat

🚨 **Point critique :** APIs externes = sources de biais non-contrôlées ✅ **ConforMind détecte :** Cartographie automatique des risques d'intégration

---

⚙️ 1.3 - Versions logiciels/firmware

Temps audit : 15 minutes | 🏢 OS/Infrastructure = ISO 27001 SUFFIT

Ce que l'auditeur va demander (focus ML uniquement) :

> *"Quelles sont les versions exactes de votre stack IA ?"*

Stack ML SPÉCIFIQUE obligatoire :

```
Environnement IA :
├── Python : 3.11.5
├── Frameworks ML :
│   ├── XGBoost : 1.7.6
│   ├── scikit-learn : 1.3.2
│   ├── pandas : 2.0.3
│   └── numpy : 1.24.3
├── Sérialisation :
│   ├── joblib : 1.3.2 (sauvegarde modèle)
│   └── pickle : intégré Python
└── Explicabilité :
    ├── SHAP : 0.42.1
    └── matplotlib : 3.7.2

Versioning modèle :
├── Git SHA : a1b2c3d4e5f6...
├── Model hash : 7d865e959b2466918c...
├── Training date : 2024-12-15 14:30:00 UTC
└── Dataset version : credit_data_2024Q4_v1.2

Sécurité dépendances :
├── CVE scan : Quotidien automatique
├── Vulnérabilités : 0 critiques détectées
├── EOL components : XGBoost EOL dans 18 mois
└── Update plan : Migration vers LightGBM Q3 2025
```

🏢 OS/Infrastructure (Linux, Docker, Kubernetes) : → ISO 27001 couvre déjà | Simple validation certificat

🚨 **Point critique :** Vulnérabilités CVE dans ML libs = faille sécurité IA ✅ **ConforMind automatise :** Scan CVE continu + recommandations upgrade

---

🏢 1.4 - Formes de commercialisation

Temps audit : 5 minutes | 99% clients = Usage interne UNIQUEMENT

Question simplifiée pour la plupart des clients :

> *"Ce système IA est-il utilisé uniquement en interne ?"*

Usage interne typique :

```
Distribution : Aucune (usage interne BNP uniquement)
Utilisateurs : Conseillers bancaires BNP (450 agences)
Accès : Intranet sécurisé + authentification SSO
Commercialisation : Aucune vente/licence à des tiers
```

Si OUI → Section validée en 2 minutes Si NON → Documentation additionnelle complexe (rare)

💼 Commercial externe nécessiterait :

- Contrats Provider/Deployer selon Article 26
- Documentation utilisateur Article 13
- Support technique Article 16
- (Concerne <1% des clients ConforMind)

🔒 1.5 - Hardware requis

Temps audit : 5 minutes | 🏢 ISO 27001 + SLA bancaires SUFFISENT

Questions simplifiées (validation certificats) :

> *"Infrastructure certifiée ISO 27001 ? Specs suffisantes pour charge IA ?"*

Validation certificats existants :

```
Certifications infrastructure :
├── ISO 27001 : Valide jusqu'à 2026 ✅
├── PCI DSS : Conforme banking ✅
├── SOC 2 Type II : Audité annuellement ✅
└── HDS : Si données santé ✅

Dimensionnement IA :
├── CPU : 32 cores (dimensionné 1000 req/jour)
├── RAM : 128GB (modèle + cache)
├── Stockage : 500GB SSD NVMe
└── Réseau : 1Gbps redondant

SLA existants :
├── Disponibilité : 99.9% (heures ouvrées)
├── Performance : <2s par prédiction
└── Recovery : RTO=4h, RPO=1h
```

🏢 Infrastructure complète déjà auditée ISO : → L'auditeur EU AI Act accepte ces certificats

🚨 **Point critique :** Sous-dimensionnement = dégradation supervision humaine ✅ **ConforMind valide :** Capacité infrastructure vs charge IA prévue

👥 1.6 - Interface utilisateur pour le déployeur

Temps audit : 15 minutes | ⚠️ SPÉCIFIQUE EU AI Act - Aucune norme ISO

Ce que l'auditeur va demander :

> *"Votre interface permet-elle une supervision humaine effective ?"*

Interface OBLIGATOIRE Article 14 :

```
Supervision humaine :
┌─────────────────────────────────────────────────────┐
│ 🤖 RECOMMANDATION IA                                │
│ Score: 847/1000 | Décision: APPROUVER | Conf: 92%   │
├─────────────────────────────────────────────────────┤
│ 🔍 EXPLICATION                                      │
│ Facteurs favorables:                                │
│ • Revenus élevés (+150 pts)                         │
│ • Bon historique crédit (+120 pts)                  │
│ Facteurs défavorables:                              │
│ • Ratio endettement (-30 pts)                       │
├─────────────────────────────────────────────────────┤
│ 🚨 ALERTES                                          │
│ ⚠️ Confiance modèle: 67% (validation recommandée)   │
│ ⚠️ Score FICO indisponible (données partielles)     │
├─────────────────────────────────────────────────────┤
│ 👤 DÉCISION HUMAINE                                 │
│ [✅ APPROUVER] [❌ REJETER] [⚠️ EXAMINER]           │
│ Justification si override: [_____________]          │
│ [📞 ESCALADER AU MANAGER]                           │
└─────────────────────────────────────────────────────┘

Audit trail automatique :
├── 15/01/2025 14:23:15 - Conseiller MARTIN
│   └── Action: Saisie demande (25K€, 60 mois)
├── 15/01/2025 14:24:02 - Système IA v2.1.3
│   └── Action: Score 847, Recommandation APPROUVER
├── 15/01/2025 14:25:30 - Conseiller MARTIN
│   └── Action: Validation humaine APPROUVÉE
│   └── Justification: "Profil excellent, revenus stables"
└── [📄 Export PDF audit] [📊 Rapport détaillé]
```

🚨 **Point critique :** Interface = cœur de la conformité EU AI Act ✅ **ConforMind valide :** Checklist Article 14 + recommandations UX

📊 TEMPS TOTAL ÉTAPE 1 OPTIMISÉE

Avant optimisation : 4-5 heures

- ❌ Re-audit infrastructure complète
- ❌ Redondance avec certifications ISO
- ❌ Focus technique vs conformité IA

Après optimisation : 1h15

- ✅ **Identité IA** : 15 min (spécifique EU AI Act)
- ✅ **Intégrations critiques** : 20 min (focus APIs IA)
- ✅ **Versions ML** : 15 min (stack IA uniquement)
- ✅ **Usage interne** : 5 min (99% des cas)
- ✅ **Validation ISO** : 5 min (certificats existants)
- ✅ **Interface supervision** : 15 min (cœur conformité)

🚀 **Gain de temps : 70% | Focus sur vraie valeur ajoutée IA**

🎯 OUTILS CONFORMIND ÉTAPE 1

🔧 Développement prioritaire MVP :

1.1 Classificateur automatique EU AI Act

- Input : Description objectif + secteur activité + impact
- Processing : Rules engine Annexe III + machine learning
- Output : Classification HIGH/LIMITED/MINIMAL + justification légale
- ROI : 15 min → 2 min | Élimination erreurs classification

1.6 Validator interface supervision

- Input : Screenshots interface + workflow description
- Processing : Computer vision + checklist Article 14
- Output : Score conformité + recommandations UX précises
- ROI : 30 min → 5 min | Détection gaps supervision automatique

1.3 CVE Scanner ML dépendances

- Input : requirements.txt ou code scanning
- Processing : NVD database + ML-specific vulnerabilities
- Output : Rapport sécurité + recommandations upgrade

- ROI : 2h → 5 min | Veille sécurité automatisée

📋 Validation ISO simplifiée :

- **Checker certificats** : "ISO 27001 en cours ?" → ✅/❌
- **Validator specs** : "Infrastructure dimensionnée ?" → ✅/❌

💼 POSITIONNEMENT COMMERCIAL

Message ConforMind :

"Nous ne ré-auditons pas votre infrastructure (ISO 27001 suffit). Nous nous concentrons sur les NOUVELLES obligations EU AI Act : classification automatique, supervision humaine, gouvernance IA."

Différenciation vs consultants :

- **Autres** : Re-auditent infrastructure + IA (redondance)
- **ConforMind** : Focus exclusif spécificités EU AI Act
- **Résultat** : 70% temps économisé + même conformité

ROI client immédiat :

- **Classification manuelle** : 2h expert → 2 min automatique
- **Interface validation** : 1 jour UX audit → 5 min scan
- **CVE monitoring** : Hebdomadaire manuel → Quotidien auto

🚀 ÉTAPE SUIVANTE

Une fois Étape 1 validée → **Étape 2 : Développement et Architecture**

Documentation générée par ConforMind-AI | Conformité EU AI Act optimisée