

# Bibliothèque du Savoir — Version Améliorée

**Statut** : Draft exécutable

**Résumé** : Document synthèse enrichi — table des matières, schémas opérationnels (Mermaid), workflow anti-biais, KPI étendus, timeline ISO, glossaire et encadré "next steps" pour partenaires.

---

## Sommaire

1. Vision
  2. Principes fondateurs
  3. Architecture conceptuelle 3.1 Pipeline global (diagramme) 3.2 Modules détaillés 3.3 Interopérabilité et RAG
  4. Schéma de métadonnées
  5. Protocole de vérification (workflows opérationnels) 5.1 Diagramme : du claim à la publication 5.2 Workflows anti-biais et audits
  6. Badges & échelles de confiance
  7. Gouvernance & décentralisation (diagramme)
  8. Stack technique MVP
  9. Indicateurs de succès (KPI) — étendu
  10. Sécurité & robustesse
  11. Stratégie d'adoption & engagement
  12. Scénarios d'usage
  13. Feuille de route synthétique & jalons
  14. Vers une norme ISO de l'information & de l'éducation 14.1 Timeline ISO (diagramme)
  15. Glossaire
  16. Engagement / Next steps (appel à action)
  17. Annexes (modèles JSON, checklist sécurité, modèle de panel d'experts)
- 

## 1. Vision

(Identique — créer une bibliothèque open-source du savoir vérifié, consensus méthodologique, traçabilité complète, validation IA+humaine, gouvernance multipartite.)

## 2. Principes fondateurs

- Transparence (provenance, horodatage, hash, pipeline visible)
- Validation hybride (IA assiste, humans validate, community contests)
- Open-source (Apache-2.0 pour code, CC-BY-SA pour contenus)
- Neutralité procédurale, Mesurabilité, Éducation intégrée

## 3. Architecture conceptuelle

### 3.1 Pipeline global (diagramme mermaid)

```
flowchart TD
    U[Utilisateurs / Contributeurs] -->|Soumission/Édition| B[Backend Wiki]
    B --> E[Extraction NLP]
    E --> S[Scoring (check-worthiness)]
    S --> R[Recherche de preuves (RAG)]
    R --> A[Analyse sémantique & NLI]
    A --> H[Validation Humaine]
    H --> G[Graphe de connaissances (RDF/PROV)]
    G --> UI[Interfaces / API]
    UI --> F[Feedback / Correction]
    F --> B
```

*Note* : Ce diagramme est prêt à être transformé en SVG haute-résolution pour inclusion dans un pitch deck.

### 3.2 Modules détaillés

(Tableau synthétique — similaire au document d'origine, enrichi d'une colonne "interopérabilité" et "remplaçable/plug-in")

### 3.3 Interopérabilité et RAG

- RAG avec sources ouvertes (Gallica, PubMed, CORE) + bases académiques privées via API
- Politique de sandboxing des LLM externes : toutes les réponses RAG doivent être accompagnées d'un "explainability bundle" (sources + passage exact + confidence score)
- Versioning obligatoire des prompts et des embeddings utilisés pour chaque vérification

## 4. Schéma de métadonnées

(Maintien du JSON fourni, ajouté : `model_card`, `embedding_id`, `verification_pipeline_version`, `confidence_interval`, `language`)

## 5. Protocole de vérification

### 5.1 Pipeline (opérationnel)

1. Soumission (formulaire structuré)
2. Canonicalisation & détection du claim
3. Scoring check-worthiness
4. Recherche automatique (RAG + scraping contrôlé + API académiques)
5. Analyse sémantique (NLI, fact-comparison)

6. Revue communautaire (crowd review, micro-tâches)
7. Panel d'experts si nécessaire
8. Publication & journal immuable (hash / ledger)
9. Appel & Versioning (processus SLA)

## 5.2 Diagramme : du claim à la publication

```

flowchart LR
    Claim[Claim soumis] --> Canon[Canonicalisation]
    Canon --> Scoring
    Scoring --> Fetch[Recherche de preuves]
    Fetch --> Analysis[Analyse sémantique]
    Analysis --> Crowd[Crowd review]
    Crowd --> Expert{Besoin d'expert ?}
    Expert -->|Oui| Panel[Panel d'experts]
    Expert -->|Non| Pub[Publication]
    Panel --> Pub
    Pub --> Ledger[Ledger/Trace immuable]

```

## 5.3 Workflows anti-biais et audits

- **Audit trimestriel** : échantillonnage aléatoire (1% des claims) + vérification indépendante.
- **Bias-tracker** : dashboard mesurant représentation par langue/région/genre/sujet.
- **Model cards & dataset cards** : pour chaque modèle et dataset utilisé, publication d'une fiche explicative (in/out of domain, known failure modes).
- **Contestations** : SLA de contestation (réception, triage <48h, résolution <14j).

## 6. Badges & échelles de confiance

(Conserver tableau, ajouter : métadonnées du badge, durée de validité, conditions de retrait, preuve minimale requise pour chaque palier)

## 7. Gouvernance & décentralisation

- Fondation non-profit (hosts assets, legal entity)
- Conseil multipartite (chercheurs, ONG, représentants pays à revenu faible, industrie, utilisateurs)
- Comité scientifique
- Conseil d'éthique
- Processus open review & transparence des votes

### Diagramme gouvernance

```

flowchart TB
    Foundation[Fondation]
    Foundation --> Board[Conseil Multipartite]
    Foundation --> Science[Comité Scientifique]

```

```
Foundation --> Ethics[Conseil d'Ethique]
Foundation --> Ops[Equipe Ops]
Board -->|Nomme| Science
Board -->|Nomme| Ethics
Ops -->|Opère| Platform[Plateforme]
Community[Communauté] -->|Feedback| Board
Community -->|Contribue| Ops
```

## 8. Stack Technique MVP

(Comme fourni, plus : infra infra costs estimate (exemple chiffré), modules remplaçables, policy for managed services vs open-source self-host)

## 9. Indicateurs de Succès (KPI) — étendu

Ajouts proposés : - **Engagement & qualité** : % de claims contestés, temps moyen de résolution des contestations, NPS contributeurs - **Diversité** : % contributions hors top-5 pays, répartition par genre (si fournie), nombre de langues actives - **Qualité IA** : hallucination rate (per 10k inferences), token leak incidents - **Éducation** : % d'écoles/universités ayant intégré le module - Rappels temporels : KPI trimestriels et objectifs 12/24 mois

## 10. Sécurité & Robustesse

- Logging complet & immuable pour audits
- Tests adversariaux automatisés pour modèles (every deploy)
- Protection DDoS, WAF, RBAC, encryption at rest & transit
- Privacy-by-design & conformité GDPR/DSA

## 11. Stratégie d'adoption & engagement

- Programmes pilotes (universités, ONG) — template agreements
- Bourses / subventions pour contributeurs en pays à faible revenu
- Certification pour réviseurs et enseignants
- Pack "starter data" pour institutions (datasets, badges, API keys)

## 12. Scénarios d'usage

(Éducation, Journalisme, Recherche, Gouvernance — cas d'usage détaillés + templates d'intégration technique)

## 13. Feuille de route synthétique & jalons

(Phases et livrables inchangés, mais ajouté : dépendances critiques, ressources estimées FTE, budget indicatif pour MVP)

## 14. Vers une norme ISO de l'information & de l'éducation

### 14.1 Objectifs

- Définition des niveaux (Fait, Analyse, Opinion, Spéculation)
- Protocoles de validation hybrides
- Métadonnées standardisées
- Auditabilité et versioning

### 14.2 Timeline ISO (diagramme mermaid Gantt simplifié)

```
gantt
    title Timeline proposition ISO
    dateFormat YYYY-MM-DD
    section Phase
    Proposition :done, des1, 2025-11-01, 2026-03-31
    Pilote :active, des2, 2026-04-01, 2026-12-31
    Consultation internationale : des3, 2027-01-01, 2027-06-30
    Adoption initiale : des4, 2027-07-01, 2028-06-30
```

*Note* : dates indicatives — à vérifier / aligner avec les instances ISO et comités nationaux.

## 15. Glossaire

- Claim / Affirmation
- Canonicalisation
- RAG: Retrieval-Augmented Generation
- PROV-O, RDF, LLM, NLI
- Model Card, Dataset Card

## 16. Engagement / Next steps (appel à action)

- Invitation à rejoindre la liste de pilotage
- Proposition de réunion d'alignement 0 (1 page, 30 min) — préparation du pitch
- Offerings pour financeurs : MVP budget, risques & mitigation
- Checklist contributions : datasets, reviewers, infra

## 17. Annexes

- Modèle JSON (schema) — extrait
- Checklist sécurité & privacy
- Template : contrat pilote université

---

**Fin du document**

*Ce document est conçu pour être exporté vers PDF ou transformé en page web (DocSite). Les diagrammes Mermaid peuvent être convertis en SVG pour usage dans présentations.*