

AI Document Processing & Validation Rules

This document defines how Maturion processes, validates, and ingests documents into its reasoning engine.

Ingestion Pipeline

1. File Upload (manual or API)
2. Metadata Extraction (type, domain, owner, etc.)
3. Chunking (semantic slicing)
4. AI validation (structure, clarity, corruption)
5. Knowledge linking (tagged to reasoning pathways)

Validation Checks

- Reject corrupted files (e.g., malformed .docx, binary junk)
- Warn on vague/compound logic
- Support badly formatted uploads from clients
- Pre-process non-standard formats for AI readability

Integrity Rules

- No ingestion without metadata
- Chunk failures logged for owner review
- Backoffice overrides available for superusers only