AI Advisory Agent for eTMF Creation and Management

Tara Mezzanotte & Dr. Scott M. Martin
Inventors, Architects, Assigned to Odyssy AI Inc.
January, 30th, 2026

In the pharmaceutical industry in the United States today, documentation for clinical trials is extremely rigorous --- everything needs to be tracked, audited and stored forever. This is typically accomplished through the Trial Master File (TMF), which can be thought of as the Trial's central record. The TMF is a collection of essential documents for the clinical trial, organized to keep everything transparent and compliant. It may include documents such as the protocol, the investigator's brochure, consent forms, case report forms, and safety records, ethics approvals, monitoring reports and other documents as are further described below. The TMF is typically maintained in electronic form where it is required to be kept secure and unalterable.

With the advent of Artificial Intelligence (AI) capabilities, there is both the need and the opportunity to introduce an AI Advisory Agent for eTMF Creation and Management.
The AI Advisory Agent for eTMF Creation and Management, as described in greater detail below, is a regulator-friendly intelligence platform for an AI-powered electronic Trial Master File that will successfully survive audits, make clinical trial operation teams more productive and efficient, and improve patient safety.

The AI Advisory Agent for eTMF Creation and Management is a regulator-friendly intelligent platform for an AI-powered electronic Trial Master File that will successfully pass all audits, make clinical trial operation teams more productive and efficient, and improve patient safety.

The AI Advisory Agent eTMF is a validated, GxP (Good Clinical Practice) compliant system that:

Manages all essential trial documents (ICH-GCP Section 8)
Continually assesses accuracy, completeness, quality, and timeliness
Uses multi-task AI agents to classify, extract, reconcile, and predict risks
Maintains full auditability, trackability, traceability, and explainability

---

Core Design Principles that are Non-Negotiable In Clinical Trials:

Regulatory-first
ICH-GCP E6(R2/R3)
FDA 21 CFR Part 11
EMA Annex 11
The AI is advisory, not authoritative
Human-in-the-loop for final decisions -
Explainability over potential raw data inaccuracy
Unassailable trackable, traceable and audit trail
Integrated validation and verification of model governance

System Definitions:

Let:

Studies: $S = \{s_1, s_2, \ldots\}$
Sites: $I_s = \{i_1, i_2, \ldots\}$ for study s
TMF Artifacts (per DIA TMF RM): $A = \{a_1, a_2, \ldots a_N\}$
Documents: $D = \{d_1, d_2, \ldots\}$

Each document d is represented as:

$$d = (x_d, m_d, t_d)$$

Whereby:

x_d = raw content (PDF, scan, slides, image)
m_d = extracted metadata
t_d = timestamps (creation, effective, uploads)

High-Level Architecture:

User Experience
Clinical Research Associate (CRA) | Clinical Operations | QA | Auditors | AI

↓

AI Services & Computer Intelligence Layer
* Document Classification (OCR/VLM and PBLLM)
* Metadata Extraction (NLP)
* Completeness & Quality Scoring (ML)
* Risk Prediction using Anomaly Detection, Risk Scoring, & Behavior Analytics (Bayesian)
* Audit Readiness (ensures GCP, Regulatory Requirements, SOPS, ALCOA)
* Natural Language Search (PBLLM/RAG)

↓

AI eTMF Core Platform (Validated)
* TMF Reference Model, Drug Information Association (DIA)
* Versioning & Lifecycle Management
* Role-Based Access (RBAC)
* Review & Approval Workflows
* Part 11 eSignatures
* Inspection Readiness Dashboard

↓

Secure Content & Data Layer
* Encrypted Document Store
* Metadata Repository
* Audit Logs, Write Once, Read Many (WORM), immutable, tamper-proof records

* Retention and Archiving

↓

Integration & Interoperability
- Clinical Trial Management System (CTMS)
- Electronic Investigator Site File (eISF/Site Systems)
- Electronic Data Capture (EDC)
- Safety / Pharmacovigilance (Audit and Inspection Lens)
- Regulatory Submission Systems

---

eTMF Intelligent Document Ingestion (IDI):

Auto-classifies docs to Drug Information Association (DIA) TMF zone/artifact (CDISC)
Extracts:
Study ID
Country/Site
Investigator(s)
Effective dates

Flags misfiled or duplicate content
Human confirms and AI learns, if approved (validated and verified feedback loop)

Document Classification (Advisory, Not Authoritative)

Goal: Assign document to DIA artifact with confidence

$$P(a_k \mid d) = f_\theta^{class}(x_d)$$

- f_θ : validated OCR/VLM + PBLLM ensemble
- Output:

Top - k artifact suggestions
Confidence score $c_d \in [0,1]$

Human-in-the-loop constraint

$$\text{FinalArtifact}(d) = \begin{cases} a_k & \text{if human accepts} \\ \emptyset & \text{otherwise} \end{cases}$$

Metadata Extraction as Structured Inference:

For each required metadata field m_j

$$P(m_j = v \mid x_d) \; f_\theta^{NLP}(x_d)$$

Tracked per field:
Extraction confidence, Source text span, Model version
(This supports ALCOA (FDA data integrity) and explainability)

Continuous AI eTMF Monitoring:
AI agent cross-checks:
Protocol version (versioning)
Trial phase(s)
Countries/sites activated
Milestones, First Patient-In, Last Patient-In, Database Lock (FPI, LPI, DBL)

Results/Outputs:
% completeness by:
Study
Country
Site

Predicted missing documents before inspections

Completeness Model (Deterministic + Probabilistic)
Let:

$$R(s,i,t) \subseteq A$$

= required artifacts for study *s*, site *i*, at trial time *t*
(based on protocol, country, phase, milestones)

Whereby Observed Artifacts:

$$O(s,i,t) \subseteq A$$

Completeness Score:

$$\text{Completeness}(s,i,t) = \frac{|O(s,i,t) \cap R(s,i,t)|}{|R(s,i,t)|}$$

This is:
Deterministic
Transparent
Fully auditable

Quality & Timeliness Scoring:
AI evaluates:
File naming compliance
Signature presence
Date consistency

Redaction errors
Upload lag vs milestone

Simple Example Quality Score:
Site 206 – Quality: 72%

Risk Flags:
Missing PI signature on IB v4
Late IRB approval (12 days post activation)

Each document d has quality dimensions:

Dimension	***Indicator***
Naming	regex / SOP compliance
Signature	presence + validity
Dates	logical consistency
Redaction	ML flag
Timeliness	milestone lag

Let quality vector be:

$$q_d = (q_1, q_2, \ldots, q_n), \qquad q_j \in [0,1]$$

Document Quality Score

$$Q(d) = \sum_{j=1}^{n} w_j q_j$$

Whereby:
w_j are fixed, validated, SOP-approved weights
No black-box weighting

Site-Level Quality

$$Q(s,i) = \frac{1}{|D_{s,i}|} \sum_{d \in Ds_{a,i}} Q(d)$$

Timeliness Modeling:
For document d:

$$\Delta_{td} = t_{\text{upload}} - t_{\text{expected}}$$

Timeliness penalty:

$$T(d) = \exp(-\lambda \Delta_{td})$$

Whereby:

λ is fixed per document class
Late ≠ invalid, just risk-increasing

Inspection Readiness:

eTMF Multi-Task AI agents have a natural language interface (with citations) that can provide accurate responses to example queries -

"Show all essential docs missing for Munich, Germany prior to LPLV"
"Summarize TMF risks for FDA inspection in oncology study #6.4.03"

All results/outputs are traceable, trackable, reference-linked, and exportable

---

Predictive Missing Document Modeling (Bayesian):

Goal: predict probability of future absence, not fabricate data

Let:

$M_a(t)$: indicator if artifact a exists at time t

We estimate:

$P(M_a\ (t+\Delta) = 0 \mid H_{s,i})$

Where $H_{s,i}$ includes:

Historical site behavior
Country risk priors
Trial phase
Past inspection outcomes

Using Bayesian updating technique:

$$P(\text{Missing} \mid H) \propto P(H \mid \text{Missing}) \cdot P(\text{Missing})$$

AI Agent Predictive Risk Intelligence:

Using historical trials -

Predicts:

Sites likely to fail inspections
Countries/cities with chronic document mismanagement
Countries/cities with past poor timeliness Scoring

Inspection Risk Score (Composite, Bounded)

Let:

$$R_{inspect}(s,i) = \alpha(1 - \text{Completeness}) + \beta(1 - Q) + \gamma(1 - T) + \delta P_{\text{missing}}$$

Whereby:

$\alpha, \beta, \gamma, \delta$ are validated constants
Output range is normalized to [0,1]

Risk thresholds trigger alerts, not actions.

Governance, Validation & Compliance:

Multi-Task AI Model Governance -
- Model versioning
- Training data lineage (history of training data)
- Bias, hallucination risk scoring & drift monitoring
- Locked inference models for validated releases

Multi-Task AI Agent Decomposition (No Autonomous Authority)

Each task τ_k is isolated:

$$A = \{\tau_{class}, \tau_{extract}, \tau_{quality}, \tau_{risk}\}$$

Each agent has:
- its own model version;
- locked inference weights post-validation;
- and logs independently.

No multi-task agent can:
- Delete
- Finalize
- Override humans

Human-in-the-Loop Final Control System:

AI suggestions ≠ system of record

Mandatory human acceptance for:
- Classification
- Quality exceptions
- Risk escalation

Human-in-the-Loop as the Final Control System

Human decision $h \in \{0,1\}$

$$\text{SystemState}_{t+1} = \begin{cases} \text{Accepted} & h = 1 \\ \text{Rejected} & h = 0 \end{cases}$$

Human feedback updates future models only after re-validation:

$$\theta_{t+1} \neq \theta_t$$

Full Auditability:

- Every AI action logs:
  - Model version
  - Confidence Score/Risk Score
  - Input features
  - Human override (if needed)

Full Auditability (Mathematically Traceable, Trackable)

- Every AI output logs:

$$\langle \text{ModelID}, \theta, x, f(x), c, h, t \rangle$$

Stored in:
- WORM logs
- Immutable hashes

Guarantees:
- Traceability
- Reproducibility
- Inspector/audit replayability

Security & Data Integrity:

- End-to-end encryption
- Role-based Access Control (RBAC) whereby -
  - Study
  - Role
  - Geography
- Data residency controls (EU vs US)
- WORM audit logs
- Automated retention & destruction rules and requirements

Example Process:

- CRA -
  - Uploads site docs → auto-classified
  - Sees real-time completeness & risk scores
  - Fixes outstanding issues before Quality Assurance (QA) escalates

- Clinical Operation Team Lead -
  - Portfolio-level TMF health dashboard
  - Predictive alerts weeks ahead of inspections

Auditor -

Read-only easy AI-assisted navigation
Instant artifact trackable and traceability
Tagged embedded note taking

Implementation Strategy:

Phase 1 -

Validated eTMF backbone and AI classification & extraction

Phase 2 -

Quality scoring and completeness prediction

Phase 3 -

Inspection and portfolio intelligence

Summary:

- *AI enhances the clinical trial process, not the authority over it*
- *All documents and status thereof is explainable, trackable, and traceable*
- *Human control remains intact*
- *Regulators see consistency and transparency*