Case Study: Obsevia

Automated data extraction and compliance workflows for the chemical industry

Julian M. Kleber | Founder | obsevia.com

julian.kleber@sail.black | Last updated: July 18, 2026

Outcome snapshot

- Built three connected products: document management, AI document validation, and a chatbot that answers EU regulation questions from linked compliance knowledge.
- Components run as independently scalable services so intake, processing, and chat can grow without blocking each other.
- Internal pilot models project **49–88% faster turnaround** and **43–87% lower process costs**.
- Target architecture designed for auditable operation, with projected data accuracy up to **99%** and ROI in **4–12 months**.

CONTEXT

Chemical compliance teams still spend significant effort on manual document review and repetitive data transfer across fragmented systems. This process becomes increasingly difficult when document volumes grow, document formats vary by vendor and region, and regulatory rules are updated frequently.

Obsevia was created to turn this high-friction process into a structured and scalable workflow: ingest documents, extract relevant safety and regulatory data, run deterministic checks, and produce outputs teams can trust and audit.

PROBLEM DEFINITION

In the status quo workflow, teams often face:

- high manual effort for Safety Data Sheet (SDS) handling and data transcription,
- inconsistent structure across incoming documents,
- error risks from outdated or misassociated safety information,
- long processing times that delay product and compliance operations.

APPROACH

Obsevia combines proven automation methods with workflow design focused on reliability:

- event-driven full-stack architecture across ingestion, processing, validation, and user-facing applications,
- document management workflows for intake, versioning, metadata handling, and retrieval,
- AI-driven document validation to check completeness, consistency, and quality before downstream decisions,
- Graph RAG retrieval and reasoning flows to support a chatbot focused on EU regulatory questions,

- OCR and layout-aware extraction from unstructured compliance documents,
- rule-based applicability and threshold checks where decision logic is explicit,
- workflow automation for repetitive interactions with legacy environments and operational systems.

SDS processing is the flagship use case. The platform focuses on practical, streamlined extraction and validation flows to reduce complexity and error propagation while keeping outputs traceable for human review.

SYSTEM OUTLINE

- **Document Management System**: Import heterogeneous compliance documents, manage metadata, and maintain traceable document lifecycle states.
- **Extraction Layer**: OCR + layout-aware parsing + classification to produce structured fields.
- **Document Validation AI**: Combine model-based validation and deterministic checks for completeness, consistency, and document-to-product associations.
- **Graph RAG Chatbot**: Build graph-structured regulatory context and retrieval flows for conversational guidance on EU regulation topics.
- **Event Layer**: Event-driven orchestration for task routing, retries, exception handling, and cross-service communication.
- **Governance Layer**: Audit-friendly outputs and explicit decision traceability.

IMPLEMENTATION STACK

- **Backend**: FastAPI services for APIs, orchestration endpoints, and integration logic.
- **Frontend**: Next.js applications for operational dashboards and user-facing workflows.
- **Database**: Supabase (Postgres) for structured data, metadata, and application state.
- **Message Broker**: Kafka used across applications for asynchronous event-driven communication.
- **Deployment**: Containerized services with automated CD pipelines for repeatable, reliable releases.

QUANTITATIVE VALUE (EVIDENCE-BASED)

Based on internal model calculations and pilot projects (2025), with assumptions aligned to documented SDS automation architectures:

- **Time savings**: 49–88% faster turnaround for compliance document analysis.
- **Cost reduction**: 43–87% lower operational process costs.
- **Quality target**: more than 50% fewer manual errors and up to 99% system data accuracy.
- **Payback**: typical amortization window of 4 to 12 months (depending on volume and environment).

WHAT THIS CASE STUDY DEMONSTRATES

- Strong fit between modern automation and high-volume compliance operations.
- Practical path from research-backed methods to production process impact.
- High leverage from combining document management, validation AI, and Graph RAG in one integrated platform.
- Value of event-driven full-stack architecture (FastAPI, Next.js, Supabase, Kafka) for reliability, extensibility, and operational transparency.