

Modular Data Hub & Data Integration

Reusable data layer for connecting information from different sources, structuring it consistently and making it available for operational, analytical and AI-enabled applications.

What we offer

EBRE AI has developed a configurable Data HUB foundation for organisations that need to bring together information from different projects, systems and formats. Data can be organised through reusable templates, versioned schemas, catalogues and APIs while preserving its source, context and history.

The platform supports data ingestion, validation, structured capture, transformation and access. Once the information is connected and reliable, the same foundation can feed dashboards, indicators, AI components, reporting tools or other applications developed for the specific use case.

Access, permissions and data-handling rules can be configured according to the organisation, users and sensitivity of each dataset.

Where it creates value

Data integration

- Connect multiple internal and external sources
- Structure tabular, document and captured data
- Relate information while preserving source context

Data quality & traceability

- Validation rules and review workflows
- Versioned schemas and historical data context
- Raw, canonical and derived values kept distinguishable

Reusable data foundation

- Configurable schemas, templates and catalogues
- APIs and exports for existing systems and applications
- Extend the same foundation to new datasets and use cases

Working foundations

Configurable Data HUB: Projects, hierarchies, reusable templates, columns, catalogues and structured data management.

Versioned data model: Schema history, pinned template versions, raw and canonical values, validation and traceability.

Data capture and import: CSV import, manual entry, document/photo capture and review workflows with explicit confirmation.

Integration and access: FastAPI-based interfaces, exports, MongoDB persistence and Keycloak-based identity and permissions.

How a pilot can start

1. Choose one data-integration problem or operational use case and identify the minimum sources required.
2. Connect a limited number of sources and define the shared structure, relationships, validation and access rules.
3. Deliver a working data flow and one useful output, such as a dashboard, indicator, export, API or AI-enabled function.
4. Extend the validated foundation to additional sources, teams or applications when it creates value.

What a pilot should prove

Data reliability

Can different sources be connected, validated and interpreted consistently while preserving their origin and history?

Operational usefulness

Can users access or reuse the combined information in a way that improves the target workflow or application?

Extensibility

Can the same data foundation absorb new sources and support additional applications without rebuilding the core?