

AI KDD: Livelink Metadata Quality Improvement Tools - DRAFT

Context

EBRD is seeking to enable processing of authoritative records managed in Livelink using the capabilities of generative Artificial Intelligence (AI).

Data Requirements

While generative AI offers effective techniques for processing unstructured data, on its own it does not provide means for identifying records required to be processed. This ability needs to be provided by more traditional indexing and retrieval methods. This capability is particularly critical in applications supporting significant business decisions that must rely on trustworthy data.

Availability of governed metadata associated with a record and business transaction it represents is a critical enabler of such processing. Metadata associated with a record should be managed within their authoritative source of truth.

Minimum metadata managed in Livelink required to support AI applications include:

- Unique identifier
- Type (e.g. Final Review Memo, Contribution Agreement)
- Date of issue
- Unique identifiers of associated entities (e.g. parties or projects) required for backward compatibility

Additional record metadata or metadata associated with the entity represented by the record (e.g. industry, country, product) may be required to support search, filtering or retention management.

Metadata Generation Tool Requirements

Considering significant volumes of records non-conformant with the above metadata requirements, there is a need for tooling to help

- Data Profiling: discover record metadata and assess their conformance with standards;
- File Comparison: Identify duplicate files
- Unique Id generation: Create human readable unique identifiers as per the standard
- Entity Extraction: Retrieve specific data items from inside the record (party, project, invoice, contract)
- Classification: Classify document type and status based on its content, name and/or location

- Entity Resolution: Resolve the extracted metadata items against master data repositories incl. EBX, Monarch, SAP or Summit
- Metadata Update: Can upload created clean metadata back to Livelink.
- Integration: Integrates with Open Text Content Server 22.1 and higher; integrates with reference data management system, using open standards

Solution Options

ShinyDocs

Custom

Analysis