SNAC Ingest Pipeline

EAD MARC21 XML (source of data) → Extraction → **Duplicative (institution-centric) EAC-CPF XML** → Deduplication → **Self-unique (institution-centric) EAC-CPF XML**

EAC-CPF XML (source of data) → **Schematron Validation (structure)** ← Self-unique (institution-centric) EAC-CPF XML

Other Formats Database, JSON → External Processing

Schematron Validation (structure) → **SNAC-valid EAC-CPF (structure, not data) XML**

External Processing

Processing
- match, create vocab terms
- match, create resources
- holding institution + locations

Identity Reconciliation (identity, relations)

Optional Processing Refine to partial I.C. for updating an existing I.C.'s data components

IR / OpenRefine Vocabulary Lookup Resource Create

Partially-valid Constellation (valid structure) JSON → Processing

RestAPI (API boundary)

SNAC EAC-CPF Serializer

SNAC EAC-CPF Parser

SNAC Datastore (Postgres, Elastic Search, Neo4j)

REST Clients SNAC WebUI A.Space

Valid Identity Constellation (structure + data) JSON — view/edit create new → RestAPI

SNAC Server

Legend