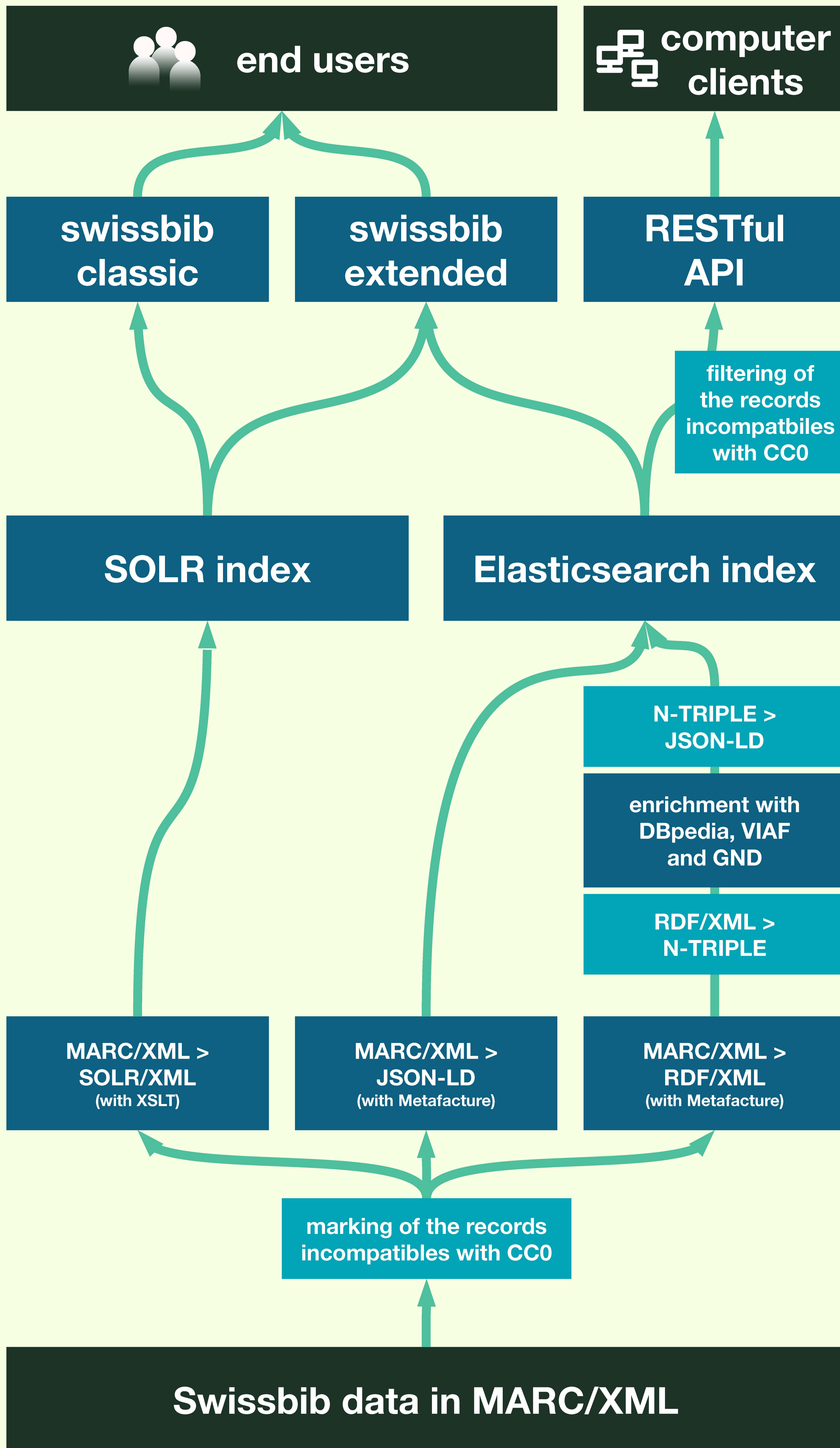


data flows in linked.swissbib.ch

from MARC to Linked Data

the project

The project linked.swissbib.ch aims to integrate the metacatalog swissbib into the semantic web, with the creation of a Linked Data service. This implies the transformation of the metadata from the main swiss library networks into a semantic format based on the RDF model, its enrichment and its publication. The data processing operations have to be fully automated to allow daily updates.



Besides the classic interface, the infrastructure provides an improved interface for the end user (e.g. a searcher), and a data service for other applications.

Massive data reuse is possible through the RESTful API. The swissbib data comes from various providers with various terms of use. To avoid any problem of attribution for the reuser, only data licensed under CC0 is accessible via the API.

The basic index of swissbib is SOLR. The Linked Data index uses Elasticsearch, based on the JSON-LD format. The swissbib extended service for end users needs access to both indices for the data presentation layer.

The enrichment operations compare the swissbib data with data from other reference databases, such as VIAF or DBpedia. To achieve this, the data must be transformed into the N-Triple format, which is also RDF-based.

The main transformation operations happen at this stage. To convert data into the XML format of SOLR, classic XSLT stylesheets are used. The creation of RDF, be it the RDF/XML or the JSON-LD serialization, is realized using "Metafacture".

Swissbib processes the bibliographic metadata in MARC format from some fifteen swiss library networks, and deduplicates it.

