

1 Search Index and Baseline Encoding

2 Search Index and Baseline Encoding

Since performing queries directly on data stored in a potentially distributed and replicated grid environment is neither quick nor straightforward, all relevant information is gathered in a search index consisting of an XML database (metadata, aggregation content, XML-encoded object content) and an RDF triple store. The latter one is fed with relations such as `isDerivedFrom`, `isAlternativeFormatOf`, `hasSchema`, `aggregates`, or `hasAdaptor`. The latter one needs some explanation: To enable structured search and processing capabilities across XML data in the TextGrid Repository, TextGrid developed the so-called [Baseline Encoding](#), a text typespecific encoding which is based on the [TEI P5 standard](#). The transformation of project specific XML into baseline-XML is performed by a so-called Adaptor (i.e. an XSLT stylesheet) with every write and update operations on an object with XML content – provided the `hasAdaptor` relation is set. The baseline instance of an object is kept only in the search index, not in the grid.