

corpora (English and Russian)

–TXM
import modules (TXT, XML)

–raw and
XML text preparation

–proprietary
format conversion

–textual
units engineering (splitting, merging)

–metadata
editing (CSV, XPath)

–tagging
(part-of speech, lemma...)

Advanced TXM

–full-text
search engine CQL patterns

–statistical
analysis: factorial analysis (clustering), classification

–multi-facet
and parallel corpora

Advanced data sources engineering

–XML-TEI,
XML-TXM

–XSLT2

–Oxygen

–Groovy

References:

TXM Project: <http://textometrie.ens-lyon.fr/?lang=en>

TXM platform development site:

<http://sourceforge.net/projects/txm>

TXM demo portail:

<http://txm.risc.cnrs.fr/demo/?locale=en>