Test document for unit testing

Tx is a Java library to support conversion and analysis of text documents in CRIMSy. Tx itself depends on Apache Tika and some other libraries. These libraries have been upgraded to some newer versions. These upgrades resulted in some issues, which went undetected due to missing unit tests. Early indication of these problems would have saved some time. It is therefore important to provide some good test coverage of the code.