

Date Normalization

Date normalization

| | |
|-------------------|--|
| Name: | David Loy |
| Version: | January 31, 2006 |
| Status: | Test |
| Specification: | datenorm_documentation.doc |
| Download: | datenorm.jar |
| More information: | Curation home page |

This date normalization package was developed as part of the [American West](#) project.

Date Normalization Utility

This Java utility takes non-machine readable Common Era dates as input and outputs machine-readable dates in order to enhance digital collections to support date range queries. The output is compliant with the [TEMPER \(Temporal Enumerated Ranges\) specification](#).

Note that this code is unsupported by CDL; use at your own risk.

Description

[datenorm.jar](#) contains all of the necessary Java source and JAR files to produce a working date normalizer.

The date normalization was developed to create a standardized Common Era (CE) date range in the form year1.year2.year3-year4.year5-year6, year1 and year2 are individual years while year3-year4 and year5-year6 are year ranges.

| Input | Outputs |
|----------|---|
| [1880's] | date.normalize: 1880-1889 date.temper: 1880-1889 date.decade: 1880 date.token: 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 |

For a detailed description of the algorithm, inputs and outputs, consult the August 2005 draft of the specification.

[datenorm_documentation.doc](#)

Requirements

datenorm.jar requires Java 1.4 and Apache Ant in order to run.

Installation

1. Create a directory.
2. Download [datenorm.jar](#) to the new directory.
3. Unpack the JAR by issuing the command: `jar -xvf datenorm.jar`

Running the Utility

Within the newly expanded ant directory, execute the following command: `ant run-date`

As packaged, `resources/FrameworkLocal.properties` is set to accept `data/example-date.txt` (a file that only has dates) as input. Changing `DateTest.type=datetoDateTest.type=text` will tell the Ant script to expect input at `data/example-text.txt` (a file of text with dates inside).

Output is directed to `SYSDOUT`.