HYU:MA – A Model for Library-Supported Projects in Japanese Digital History

Peter Broadwell, Academic Projects Developer, UCLA Digital Library
Broadwell@library.ucla.edu

Tomoko Bialock, Japanese Studies Librarian, UCLA Library
tbialock@library.ucla.edu
Macro-scale Views of the 10\textsuperscript{th}-15\textsuperscript{th}-century Imperial Collections of Waka Poetry

N-gram Search and Visualization Interface for Japanese Poetic Texts on the Bookworm

LDA Similarity Metrics of Corpus on Heat Map

Confusion Matrix Visualization of Original Genre Classifications (vertical axis) Versus naïve Bayes Classifier Classifications (horizontal axis) of Waka Poems from the Imperial Anthologies

Naïve Bayes classifier: 
https://github.com/mimno/naivebayes/blob/master/Readme.md
MeCab: http://taku910.github.io/mecab/
UniDic: http://pj.ninjal.ac.jp/corpus_center/unidic/
Confusion of Spring Poems Classified as Winter Poems on the Mimno’s Naïve Bayes Classifier at Micro-Scale View

Original: 春 | NB: 冬
LDA (Latent Dirichlet Allocation) Topic Modeling for Large Japanese Corpus, Aozora Bunko on the dfr-Browser

Aozora Bunko package prepared for Himawari: http://www2.ninjal.ac.jp/lrc/index.php...

Genji Monogatari prepared for Himawari: http://textdb01.ninjal.ac.jp/LCgenji/himawari_package.html

dfr: https://agoldst.github.io/dfr-browser/
Aggregate Visualization of *Ezukushi Banzuke* on the Coverspace

References


