日本語学会秋季大会2014.10.19
札幌・北海道大学学術交流会館

平安時代漢字字書のリレーションシップ

池田 証壽[†]　李 媛[†]　申 雄哲[†]　賈 智[‡]　斎木 正直[†]　（北海道大学大学院文学研究科[†]　日本学術振興会外国人特別研究員[‡]）

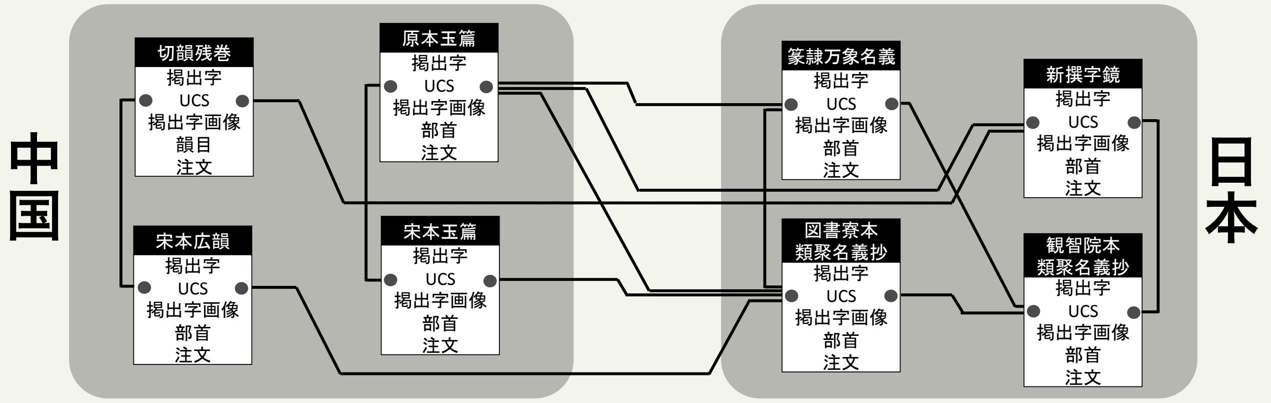
目的と方法

平安時代漢字字書総合データベース（略称HDIC）

平安時代の漢字字書に見える掲出字・注文を電子テキスト化し検索するシステム

中国側漢字字書とのリレーションシップによる平安時代漢字字書の解読の効率化

UCSコードによるリレーションシップ設定


中国
切韻残巻
掲出字
UCS
掲出字画像
韻目
注文
宋本広韻
掲出字
UCS
掲出字画像
部首
注文
原本玉篇
掲出字
UCS
掲出字画像
部首
注文
宋本玉篇
掲出字
UCS
掲出字画像
部首
注文
篆隷万象名義
掲出字
UCS
掲出字画像
部首
注文
図書寮本類聚名義抄
掲出字
UCS
掲出字画像
部首
注文
新撰字鏡
掲出字
UCS
掲出字画像
部首
注文
観智院本類聚名義抄
掲出字
UCS
掲出字画像
部首
注文
日本

リレーションシップ利用の実際

篆隷万象名義と新撰字鏡

篆隷万象名義
約16,000字
⟷
新撰字鏡
約21,000字

掲出字の不対応
- 頊 ←→ 項
- 一 二 三 五 六 七 八 廿 卅 百 千
- 男 釋 勇 川 音 私 季 歴 孤 乏 災

篆隷万象名義と図書寮本類聚名義抄

符号化のずれ
- 入力方針の違い ⇒ 個別処理が必要

図書寮本類聚名義抄		篆隷万象名義
康熙字典体に変換		原形のまま
渴（U+6E34）	≠	渇(U+6E07)
溫(U+6EAB)		温(U+6E29)

掲出字の不採録
- 「認」⇒ 原本玉篇・篆隷万象名義に項目無
  説文解字も無

HNG　CHISE

文字処理

高山寺本篆隷万象名義

Unicode

CJK統合漢字
IDS（漢字構成記述文字列）
IDC（漢字構成記述文字）
⿰ ⿱ ⿲ ⿳ ⿴ ⿵ ⿶ ⿷ ⿸ ⿹ ⿺ ⿻

IDS処理

𨞶 ⿰⿱喜心否　cf.⿰憙否

本文解読

瀛涯敦煌韻輯　フランス国立図書館Gallica

掲出字データベース

画像データ：原本字形の細部を把握
テキストデータ：文字同定⇒ リレーションシップ

本文解読と補正

切韻残巻の原本調査
既刊翻刻の補正

許諾

享和本新撰字鏡

所蔵者

高山寺
宮内庁書陵部図書課図書寮文庫
天理大学附属天理図書館

組織

平安時代漢字字書総合データベース編纂委員会

付記：本発表は平成25年度～27年度日本学術振興会科学研究費補助金基盤研究（C）「大規模漢字集合フォント対応のデータベース構築による平安時代漢字字書研究」（課題番号25370506、代表：池田証壽）、および平成25年度～27年度日本学術振興会特別研究員奨励費「日本古辞書における唐代の字体規範の受容に関する実証的研究」（課題番号2503302、賈智）成果の一部である。