

e-learning 用『古事類苑』DB 構築

計画二年目成果報告

[進捗状況]

- (1) 第1年目において全 51 巻(約 69,000 頁)の本文ページ画像のデジタル化(Jpeg 形式 300dpi)および五十音索引(約 64,000 項目)のテキスト化を完了。
- (2) 第2年目前半の作業として8月までに五十音索引の総見出語(約 64,300 語)に「よみがな」を付すことを完了。
- (3) 第2年目後半の作業として、「五十音索引」の見出語(約 64,300 語)および国文学研究資料館より提供のあった「総目録索引」の項目語(約 40,700 語)を「本文ページ画像」とリンクすると同時に、見出語また項目語にリンクされた本文ページの中に、該当語が最初に出現する箇所を色網掛表示する作業を進行中である。本作業は3月中旬に完了予定である。

[計画の見通し]

- (1) 来年度計画3年目には『古事類苑』DBの基盤構築を完了し、それに多機能な検索システムを施して e-learning 用『古事類苑』検索ソフトを完成できる見通しである。

[準備データ]

- 本文ページの画像は Jpeg 形式(各頁 300dpi)として入力済み。総頁数は約 67,100 頁。データ容量は 109.5GB。
- 「五十音索引」は Excel データ(通番・見出語・よみがな・部名・巻数・頁)として入力済み。総見出語は約 64,300 語。データ容量は 14,594KB。
- 「総目録索引」は Excel データ(通番・項目語・部名・巻数・頁)として入力済み。総項目語は約 40,700 語。データ容量は 3,867KB。