

京都大学人文科学研究所共同研究実績・活動報告書

(3年計画の 2年度目)

1. 研究課題

(和文) 情報処理技術は漢字文献からどのような情報を抽出できるか

(英文) What information can be extracted from Kanbun texts with computational methods?

2. 研究代表者

(氏名) 山崎直樹

3. 研究期間

平成22年 7月 から 平成 25年 3月 まで

4. 研究目的 (400字程度)

本課題は、現代の情報処理技術が東アジア古代社会の遺産として残された漢字文献からどのような情報を抽出できるかという問題に対し、複数の角度からその可能性を探り、人文情報学の基礎を築くことを目的とする。ここでいう「人文情報学」とは、文献を機械処理することにより、i) 人手では不可能な大量のデータを扱い、ii) 人手による処理では帰納できない類の情報を抽出し、iii) 得られた情報を機械可読なかつ再加工可能な形式で蓄積する学問分野を指す。

本課題では、以下に例として挙げる処理が可能であるかを試みる予定である。a) 白文の古典中国語のテキストを解析して句読点を施す、b) 墓誌銘や地方誌などの定型的なテキストから研究者が望む情報を抽出する、c) 返り点の施された漢文は構文情報をアノテーションしたコーパスであると考え、構文情報ごとコーパス化をする、d) 経典等に付された注釈は元のテキストに付されたメタデータであると考え、それを機械可読な形式で関連づける、e) 定型的な韻文などで常に関連づけられる語句は、ある種のネットワークを成していると考え、それを体系化する、などである。

5. 本年度の研究実施状況 (400字程度)

7月8日に研究会を開催し、先立つ2月18日に開催したシンポジウム『文字と非文字のアーカイブズ/モデルを使った文献研究』で扱った話題のうち、「モデルを使った文献研究」で取りあげられた「ネットワーク分析を用いた文献解析」の可能性を議論した。この議論を9月30日の研究会でさらに突き詰め、11月19日に公開セミナー『ネットワーク科学は工具箱』を人文研で開催した。講師とテーマは以下の通り。

- ・藤原義久 (兵庫県立大学) 「ネットワーク解析の道具を理解しよう」
- ・湯田聡夫 (CREV) 「大規模社会ネットワーク分析の事例と展望」

また、上記公開セミナーに先立つ11月4日の研究会では、メタデータとその処理に関する議論をおこない、『情報の構造とメタデータ』と題するシンポジウムを2012年2月24日開催で決定した。報告者と発表テーマは以下の通り。

- ・安岡孝一（京都大）「マンガにおける異本研究」
- ・C. Wittern（京都大）「TEIテキスト・モデルの今昔」
- ・白須裕之「漢字文献における電子的翻刻の課題——或いは翻刻者の使命」
- ・大向一輝（国立情報学研究所）「CiNiiのメタデータ・デザイン」

このシンポジウムに先立つ2月10日の研究会では、シンポジウムの運営に対する議論を行い、メタデータに関する考察をさらに深めるため、人文情報学研究所所長の永崎研宣氏をコメンテーター兼パネルディスカッションの司会として招聘することを決定した。

6. 研究成果の概要（400字程度）

公開セミナー『ネットワーク科学は道具箱』では、ネットワーク構造の物理モデルに関する初歩的な知識から高度な内容までをまず学び、その後、大規模ネットワークに関する研究の応用例までを学んだ。参加者は、人文科学におけるネットワーク科学の応用について、一定の見通しを得ることができた。

シンポジウム『情報の構造とメタデータ』では、本プロジェクトがその目的に掲げている「人文情報学の基礎を築く」に関して、昨年のシンポジウムに引き続き、「テキストとはどのような構造をしているのか、それはどのようにモデル化できるか」「人文学で扱う情報とはどのような構造をもつか、それはどのように扱うべきか」の諸点について議論を行い、人文情報学の現状と課題について理解を深めることができた。

7. 共同研究会に関連した公表実績（出版、公開シンポジウム、学会分科会、電子媒体など）

公開セミナー『ネットワーク科学は道具箱』および公開シンポジウム『情報の構造とメタデータ』をおこなった。また、これらの資料および当日のUSTREAM録画を、以下のURLで公開している。

<http://kanji.zinbun.kyoto-u.ac.jp/~ymzknk/kanzi/>

8. 本年度の共同利用・共同研究の参加状況

区分	所属機関数	参加人数	延べ人数
学内	1	4	20
国立大学	1	1	3
公立大学	0	0	0
私立大学	4	5	13
大学共同利用機関法人	0	0	0
民間・独立行政法人等	6	6	6
外国の研究機関	0	0	0

(うち大学院生)	(0)	(0)	(0)
計	12	17	42

なお、公開シンポジウム等でのUSTREAM (Web)参加人数はのべ100人を超えているのだが、上記の区分が不明であることから人数に含めていない。

※当該年度の共同利用・共同研究参加者の所属機関数、参加人数、延べ人数を区分に応じて記入して下さい。

※「学内」の所属機関数は「学部数」等を記入して下さい。

※参加人数及び延べ人数の算出方法は、以下の例に基づき算出して下さい。

(例)

- ・ 1つの共同利用・共同研究課題で2人を共同研究員として3日間受け入れた(参加した場合) : 参加人数2人、延べ人数6人

9. 本年度 共同利用・共同研究を活用して発表された論文数

論文数	
上記のうち国際学術誌に掲載された論文数	

※研究者がファーストオーサーであること。学内の紀要等に掲載されたものを除く

なお、高いインパクトファクターを持つ雑誌等に掲載された論文がある場合、その雑誌、掲載論文、そのうち主な論文の詳細等

掲載雑誌名等	論文名	発表者氏名