

文化情報学とデータベース A Database for Culture and Information Science

村上 征勝

Masakatsu Murakami

同志社大学 文化情報学部, 京田辺市多々羅都谷 1-3

Doshisha University, 1-3 Tataramiyakotani, Kyotanabe, Kyoto

あらまし：文化情報学という学問領域におけるデータベースの構築とその活用法、および同志社大学文化情報学部のカリキュラムにおけるデータ分析とデータベースに関係した講義を紹介し、文化情報学部でのデータサイエンス教育について概観した。

Summary : This paper discusses how to make and use database for Culture and Information Science. Also we introduce the undergraduate program of study in Culture and Information Science at Doshisha University.

キーワード：人文科学, 文化情報学, データベース, データ分析

Keywords : Humanities, Culture and Information Science, Database, Data Analysis

1. 文化情報学におけるデータベース

“情報”という言葉が付加された新しい学問領域は、一般に、従来の学問より積極的にコンピュータを活用し情報の分析を試みるが、文化情報学もまた例外ではない。特にデータ分析において多種多様の大量のデータの保存を可能とするデータベースの利用は文化研究に新たな道を拓きつつある。

データベースはいろいろな研究に利用できるように構築された汎用性の有るデータベースと、ある個別の問題の解明のために構築された汎用性の無い特殊なデータベースがある。汎用性の有るデータベースは一般に構造がシンプルである。たとえば、青空文庫のような文学作品の原文がデジタル化されたデータベースや、浮世絵の画像を収録したデータベースなどはそうであろう。汎用性のあるデータベースは、そこからすぐ問題解明に必要な情報が得られることは少ない。しかし、たとえば、ある作家の文体の経年変化を調べるために、青空文庫の原文データベースを利用し、デジタル化された原文に形態素解析等を施し、必要な付加情報をつ

けることによって問題解明のための特殊なデータベースを構築することができる。また汎用性の有るデータベースを利用せず、最初から、問題解明に必要な情報を含むデータベースを構築することがある。著者たちが日蓮遺文の真贋論争、源氏物語の著者問題、西鶴作品の著者問題の解明のために構築したデータベースはそのようなデータベースである。

2. 問題解明を目的としたデータベース

文化現象に最初から数量データが付随しているわけではない。したがって、計量的な観点から現象を解明しようとする、まず現象を表わすデータを作成することが必要となる。しかし、文化現象を表すデータの種類は一般に膨大である。たとえば、作家の文体に関する情報としては、文の長さ、単語の長さ、品詞の使用率、単語の使用率、品詞の接続関係等が考えられる。データサイエンスを用いた現象の解明においては、データ、そしてデータベースは問題が解明できるように作成するという考え方が重要となる。どのような種類

の情報、どのようなスケールを用いて測定すべきか、沈黙考が必要となる。

3. 同志社大学文化情報学部のカリキュラム

同志社大学文化情報学部のカリキュラム(図1)には、データベースに関する講義としては「データベースシステム」があるだけである。しかし、必修科目の「文化情報学入門」、「文化情報学演習」、「ジョイント・リサーチ」、「卒業研究」の4科目中、「文化情報学入門」以外の3科目は、問題発見からデータ分析による問題解明までの一連のプロセスを実習する科目である。この必修科目を通じ、小規模であるが問題解明に必要なデータをどのように作成し、データベースをどのように構築し、どのように用いるかを習得させる。

データ分析に関係する選択科目は、データサイエンス科目として以下の科目を設置している。

データサイエンス入門	データサイエンス 入門演習
データサイエンス基礎	データサイエンス 演習
定量的データ分析	定量的データ分析演習
定性的データ分析	定性的データ分析演習
データマイニング	心理データ解析
テキストマイニング	空間データ解析
データハンドリング	データベースシステム
アルゴリズムとデータ構造	
また、文化クラスター科目として以下の科目を設置している。	
時系列解析	決定理論
社会調査入門	社会調査
社会調査演習	文化計量学入門
文献解析	音楽解析
文化解析	行動計量学
数理統計 I	数理統計 II
確率・統計	言語とコーパス

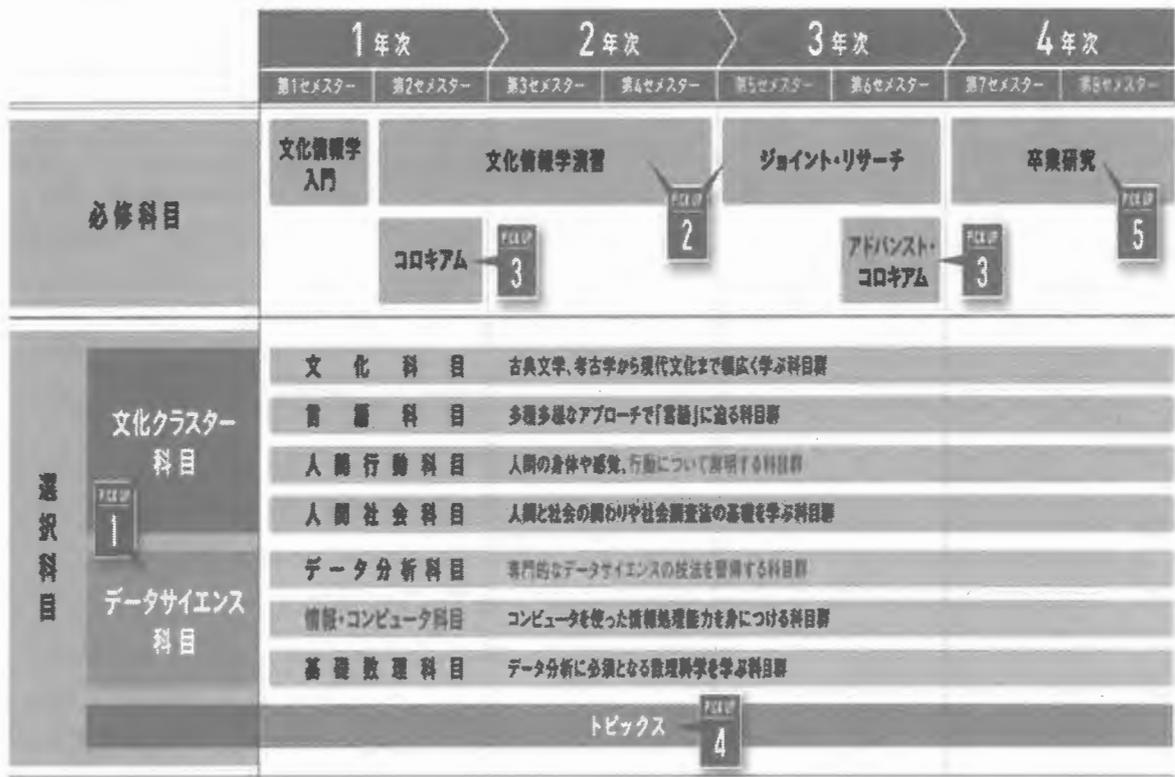


図1. 文化情報学部のカリキュラム