



GCOE 集中講義 「数学と自然科学・社会科学 III」
***From Information Retrieval to Web Mining:
a mathematical tour***

多様な人材育成プログラムの一環として、GCOE集中講義「数学と自然科学・社会科学 III」を開講します。

講義名: 数学と自然科学・社会科学 III

この集中講義は2つの講義により行われ、評価はその2つを合わせたもので行なう。
もう一つの講義は追って掲示をする。

講師: Mei Kobayashi 氏 (IBM Research-Tokyo)

日程: 10月25日(月)・26日(火) のそれぞれ 15:00~17:00

場所: 京都大学理学部3号館108号室

タイトル: **From Information Retrieval to Web Mining: a mathematical tour**
Part I: Overview
Part II: Selected Topics

アブストラクト:

When I was a graduate student, Linear Algebra and Matrix Computations was my favorite course, because it was a beautiful bag filled with magical tricks. Many years down the road, while working in industry, I was delighted to find that the theorems, fast algorithms, and methods for minimizing error are not only being put to good use by search engines, but constitute the essential core of their service. Of course, no one knows all of the details about all search engines since so much of the knowledge is proprietary, but some basic algorithms that lie at the heart of some famous search systems have been published.

The first of this two-part lecture will be an overview of methods that led to the development of Web search and mining systems. The review will include the evolution of information retrieval systems (from static to dynamic databases), topic detection and tracking, clustering, and some basic algorithms used in Web search engines that exploit link structures. A selection of interesting, emerging topics will be introduced in the second lecture. Topics will include: On-line Discussion Mining, Blog Mining, Web Community Mining, and Recommendation Systems. Some algorithms developed at IBM Research (Almaden and Tokyo Laboratories) will be presented.

Note: This two-part lecture will be accessible to any mathematics or science major with limited background in Web search and mining algorithms. Knowledge of matrix computations will be helpful. Computer science majors familiar with Web-related technologies may want to skip the first lecture.