Data Warehousing
資料倉儲

Introduction to Data Warehousing

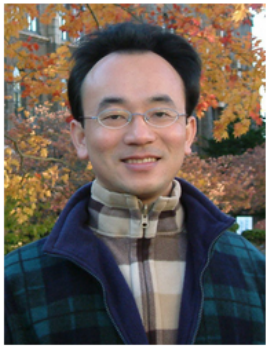
1001DW01
MI4
Tue. 6,7 (13:10-15:00) B427

Min-Yuh Day
戴敏育
Assistant Professor
專任助理教授
Dept. of Information Management, Tamkang University
淡江大學 資訊管理學系

http://mail.im.tku.edu.tw/~myday/
2011-09-06

Min-Yuh Day

mail.im.tku.edu.tw/~myday/



Min-Yuh Day

Assistant Professor
Department of Information Management, Tamkang University

Tel: 886-2-26215656 ext. 2347
Fax: 886-2-26209737
Office: I716, Chueh-sheng Memorial Hall [Office Hours]
Address: No.151, Yingzhuan Rd., Danshui Dist., New Taipei City 25137, Taiwan (R.O.C.)
Email: myday@mail.tku.edu.tw
Web: http://mail.im.tku.edu.tw/~myday/

[Vita | Education | Research | Publications | Teaching | Professional Activities] [中文版]

Vita:

Dr. Min-Yuh Day is an Assistant Professor in the Department of Information Management at Tamkang University, Taiwan. Prior to joining the faculty at TKU in 2011, he was a Postdoctoral Fellow in the Intelligent Agent Systems Lab, Institute of Information Science, Academia Sinica, Taiwan. He received the Ph.D. degree from the Department of Information Management at National Taiwan University, Taiwan. He received his MBA in Management Information System from Tamkang University, Taiwan. His current research interests include Knowledge Management, Electronic Commerce, Information Systems Evaluation, Social Media Service, Question Answering Systems, Data Mining and Text Mining, and Biomedical Informatics. He has published papers in *Information & Management, Decision Support Systems, Integrated Computer-Aided Engineering, ACM Transactions on Asian Language Information Processing*, and a number of international conference proceedings.

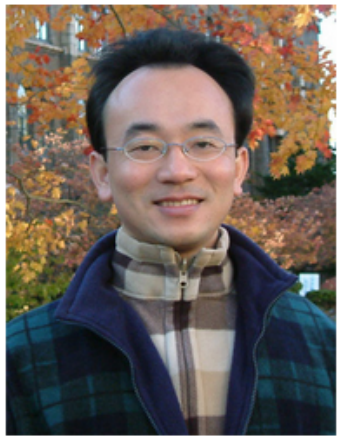
Education:

- Ph.D. Department of Information Management, National Taiwan University, 2001-2010
  Dissertation: A Study of Evaluation Model of User Satisfaction with Social Network Services
  Advisor: Dr. Chorng-Shyong Ong
- M.B.A. Department of Information Management, Tamkang University, 1993-1995
  Thesis: Research of Applying Genetic Algorithms to Fuzzy Forecasting - Focus on Sales Forecasting
  Advisor: Dr. Hung-Chang Lee
- B.B.A. Department of Information Management, Tamkang University, 1989-1993

http://mail.im.tku.edu.tw/~myday/

戴敏育 (Min-Yuh Day)

mail.im.tku.edu.tw/~myday/cindex.htm



戴敏育 博士 (Min-Yuh Day, Ph.D.)

專任助理教授
淡江大學 資訊管理學系

電話： 02-26215656 #2347
傳真：02-26209737
研究室：I716 (覺生綜合大樓) [Office Hours]
地址： 25137 新北市淡水區英專路151號
Email： myday@mail.tku.edu.tw
網址：http://mail.im.tku.edu.tw/~myday/

簡介 (Vita):

戴敏育博士目前是淡江大學資管系專任助理教授。他於2011年加入淡江大學專任教師之前，曾任職於中央研究院資訊科學研究所智慧型代理人系統實驗室博士後 研究員。 他於2010年取得國立台灣大學資訊管理博士學位，他在淡江大學資訊管理學系取得碩士學位。他目前的研究興 趣包括知識管理 (Knowledge Management)、電子商務 (Electronic Commerce)、資訊系統評量 (Information Systems Evaluation)、社會媒體服務 (Social Media Service)、問答系統 (Question Answering Systems)、資料與文字探勘 (Data Mining and Text Mining)、生物醫學資訊 (Biomedical Informatics)。他的學術研究論文已發表在*Information & Management, Decision Support Systems, Integrated Computer-Aided Engineering, ACM Transactions on Asian Language Information Processing*等國際期刊和 許多國際研討會論文集。

教 育 (Education):

- 博士　國立台灣大學資訊管理研究所 (2001-2010)
  博士論文：社交網路服務使用者滿意度評量模式之研究
- 碩士　淡江大學資訊管理研究所 (1993-1995)
  碩士論文：應用遺傳演算法發展模糊預測之研究－以銷售預測為例
- 學士　淡江大學資訊管理學系 (1989-1993)

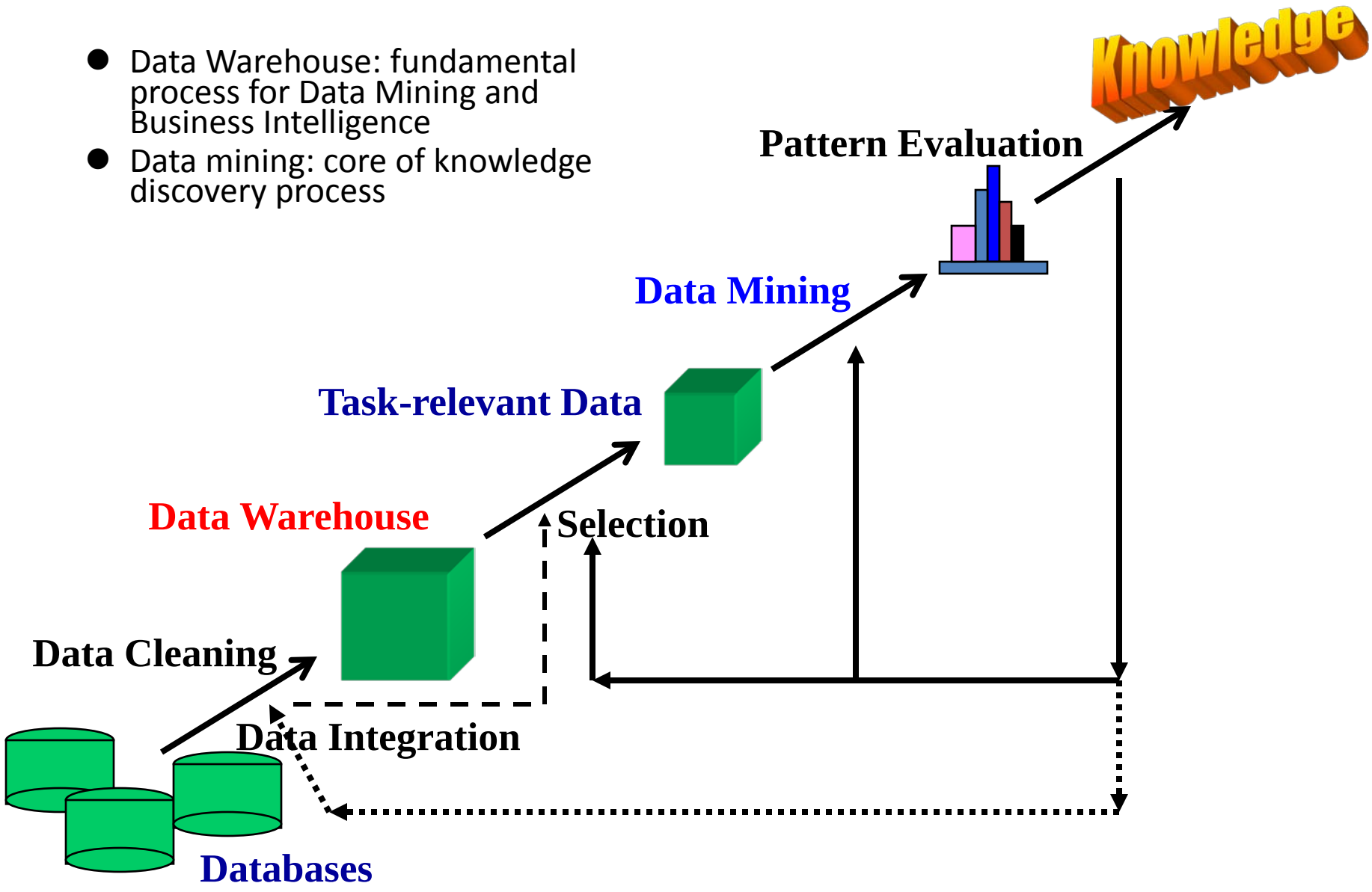
http://mail.im.tku.edu.tw/~myday/

課程資訊

- 課程名稱：資料倉儲 (Data Warehousing)
- 授課教師：戴敏育 (Min-Yuh Day)
- 開課系級：資管四 (MI4)
- 開課資料：選修 單學期 2學分
- 上課時間：週二 6,7 (Tue 13:10-15:00)
- 上課教室：B427

Knowledge Discovery (KDD) Process

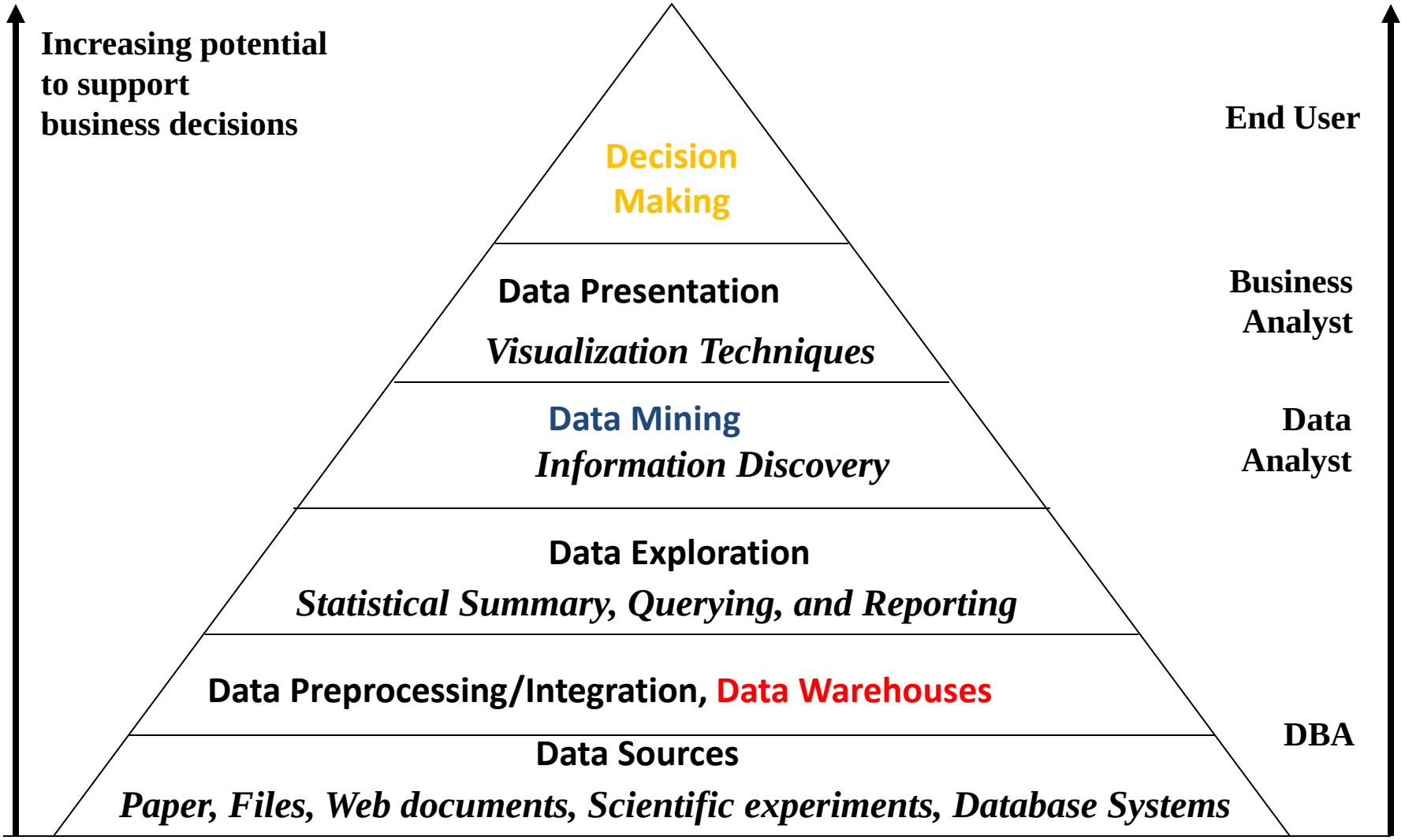
- Data Warehouse: fundamental process for Data Mining and Business Intelligence
- Data mining: core of knowledge discovery process



Source: Han & Kamber (2006)

Data Warehouse

Data Mining and Business Intelligence



Source: Han & Kamber (2006)

課程簡介

- 本課程介紹資料倉儲的基本概念及技術。
- 課程內容包括資料倉儲、OLAP、資料探勘、商業智慧、即時分析處理，資料方塊，關聯分析、分類、分群、社會網路分析、文字探勘、與網頁探勘。

Syllabus

週次	日期	內容（Subject/Topics）
1	100/09/06	Introduction to Data Warehousing
2	100/09/13	Data Warehousing, Data Mining, and Business Intelligence
3	100/09/20	Data Preprocessing: Integration and the ETL process
4	100/09/27	Data Warehouse and OLAP Technology
5	100/10/04	Data Warehouse and OLAP Technology
6	100/10/11	Data Cube Computation and Data Generation
7	100/10/18	Data Cube Computation and Data Generation
8	100/10/25	Project Proposal
9	100/11/01	期中考試週

Syllabus

週次	日期	內容（Subject/Topics）
10	100/11/08	Association Analysis
11	100/11/15	Classification and Prediction
12	100/11/22	Cluster Analysis
13	100/11/29	Sequence Data Mining
14	100/12/06	Social Network Analysis
15	100/12/13	Link Mining
16	100/12/20	Text Mining and Web Mining
17	100/12/27	Project Presentation
18	101/01/03	期末考試週

Course Introduction

- This course introduces the fundamental concepts and technology of data warehousing.
- Topics include data warehousing, data mining, business intelligence, OLAP, data cube, association analysis, classification, cluster analysis, social network analysis, text mining, and web mining.

Objective

- Students will be able to understand and apply the fundamental concepts and technology of data warehousing.

教學目標之教學策略與評量方法

- 教學目標
  - 學生將能夠瞭解及應用資料倉儲的基本概念及技術。
- 教學策略
  - 課堂講授、分組討論
- 評量方法
  - 出席率、報告、討論、期中考、期末考

教材課本

- Data Mining: Concepts and Techniques, Second Edition, Jiawei Han and Micheline Kamber, 2006, Elsevier
- 參考書籍
  - 資料探勘：概念與方法，王派洲 譯，2008，滄海
  - SQL Server 2008 R2 資料採礦與商業智慧，謝邦昌、鄭宇庭、蘇志雄，2011，碁峯
  - 資料庫理論與實務SQL Server 2008，施威銘研究室，2010，旗標
  - Web 資料採掘技術經典，孫惠民，2008，松崗

Data Mining: Concepts and Techniques (Second Edition)

http://www.amazon.com/Data-Mining-Concepts-Techniques-Management/dp/1558609016

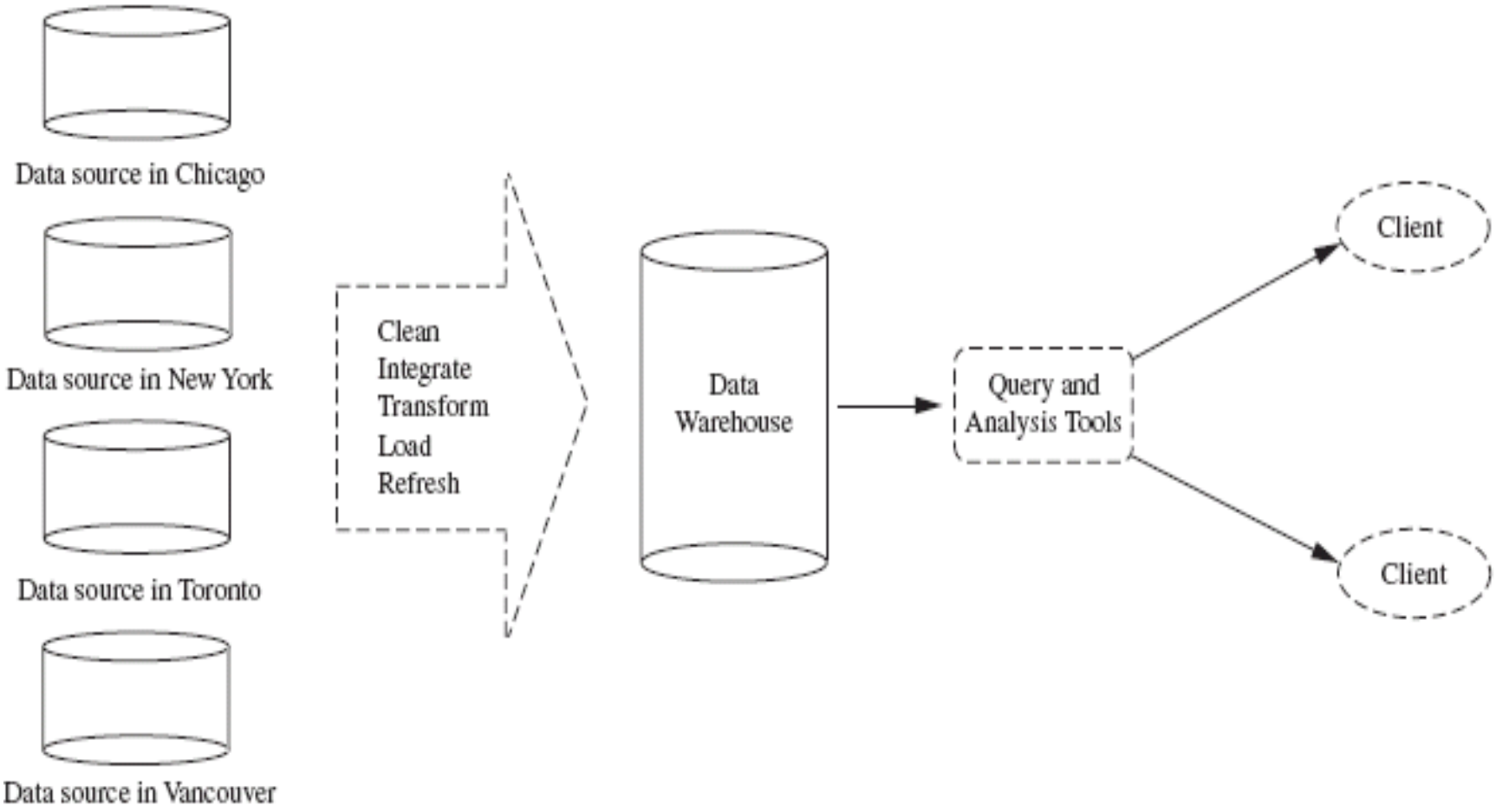
作業與學期成績計算方式

- 批改作業篇數
  - 1篇（Team Term Project）
- 學期成績計算方式
  - ☑期中考成績：30 %
  - ☑期末考成績：30 %
  - ☑作業成績： 20 % (Team Term Project)
  - ☑其他（課堂參與及報告討論表現）： 20 %

Team Term Project

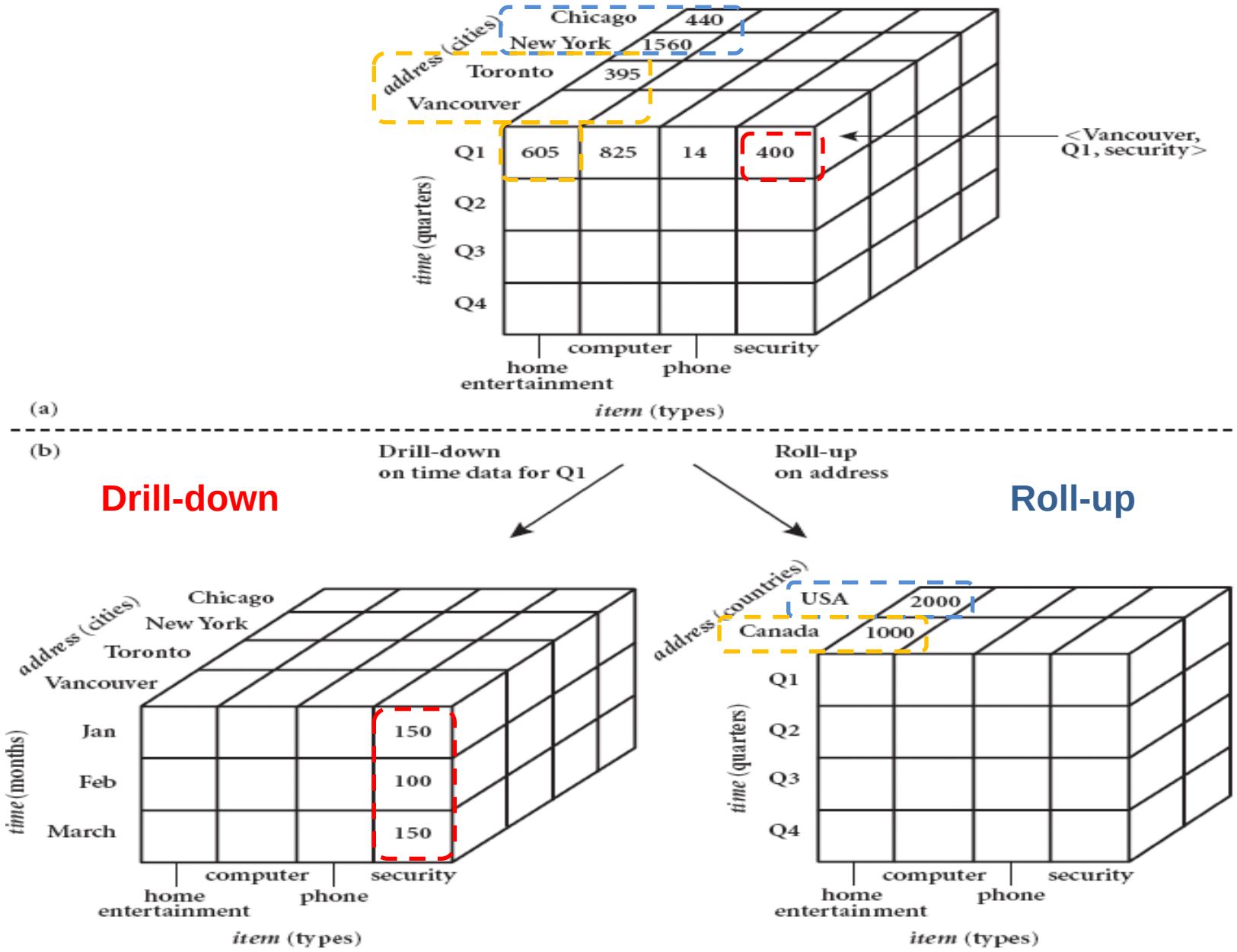
- Term Project Topics
  - Data Warehousing
  - Business Intelligence
  - Data mining
  - Text mining
  - Web mining
  - Social Network Analysis
  - Link Mining
- 3-5 人為一組
  - 分組名單於 2011.09.20 (二) 課程下課時繳交
  - 由班代統一收集協調分組名單

Typical framework of a data warehouse



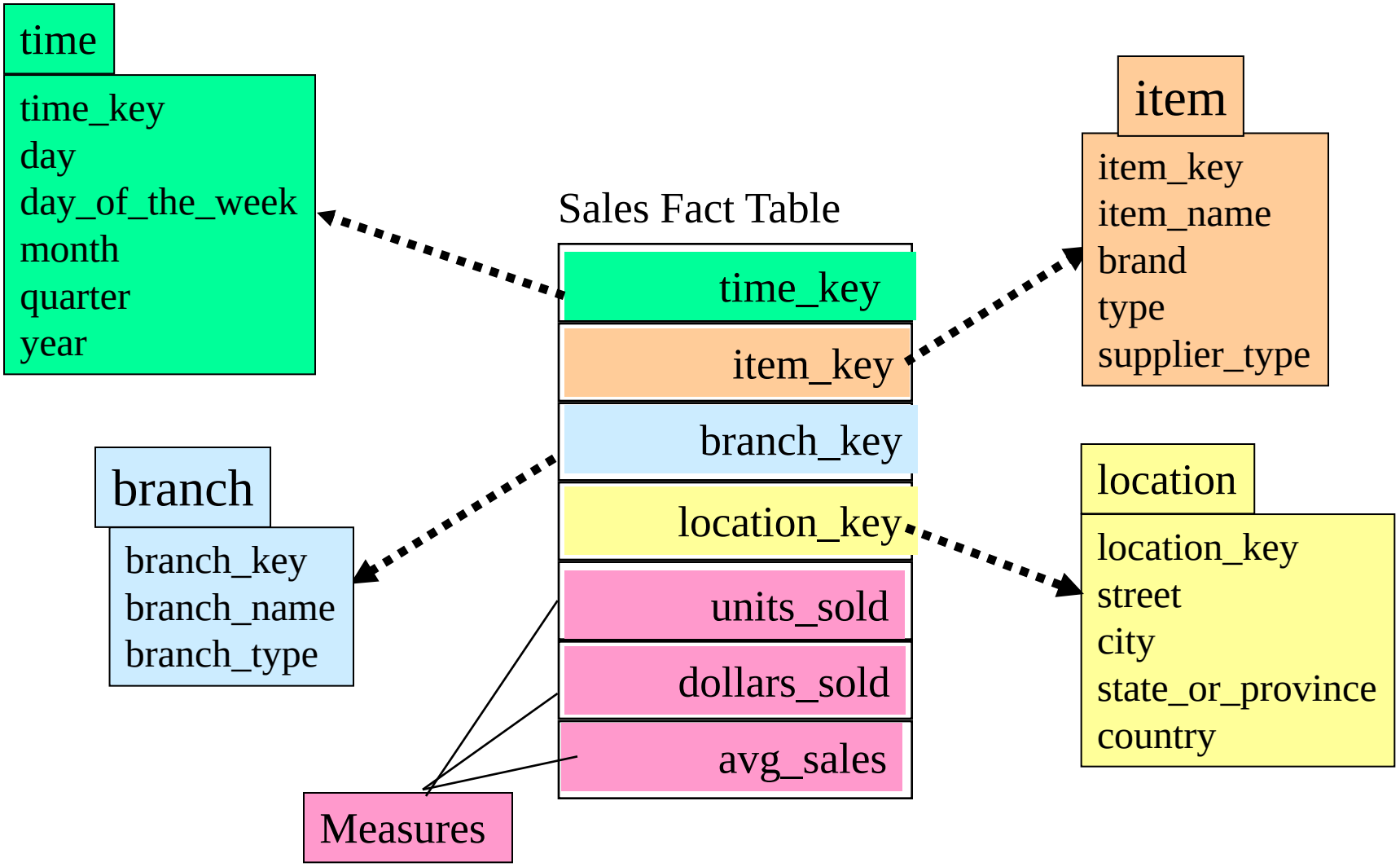
Source: Han & Kamber (2006)

Multidimensional data cube for data warehousing



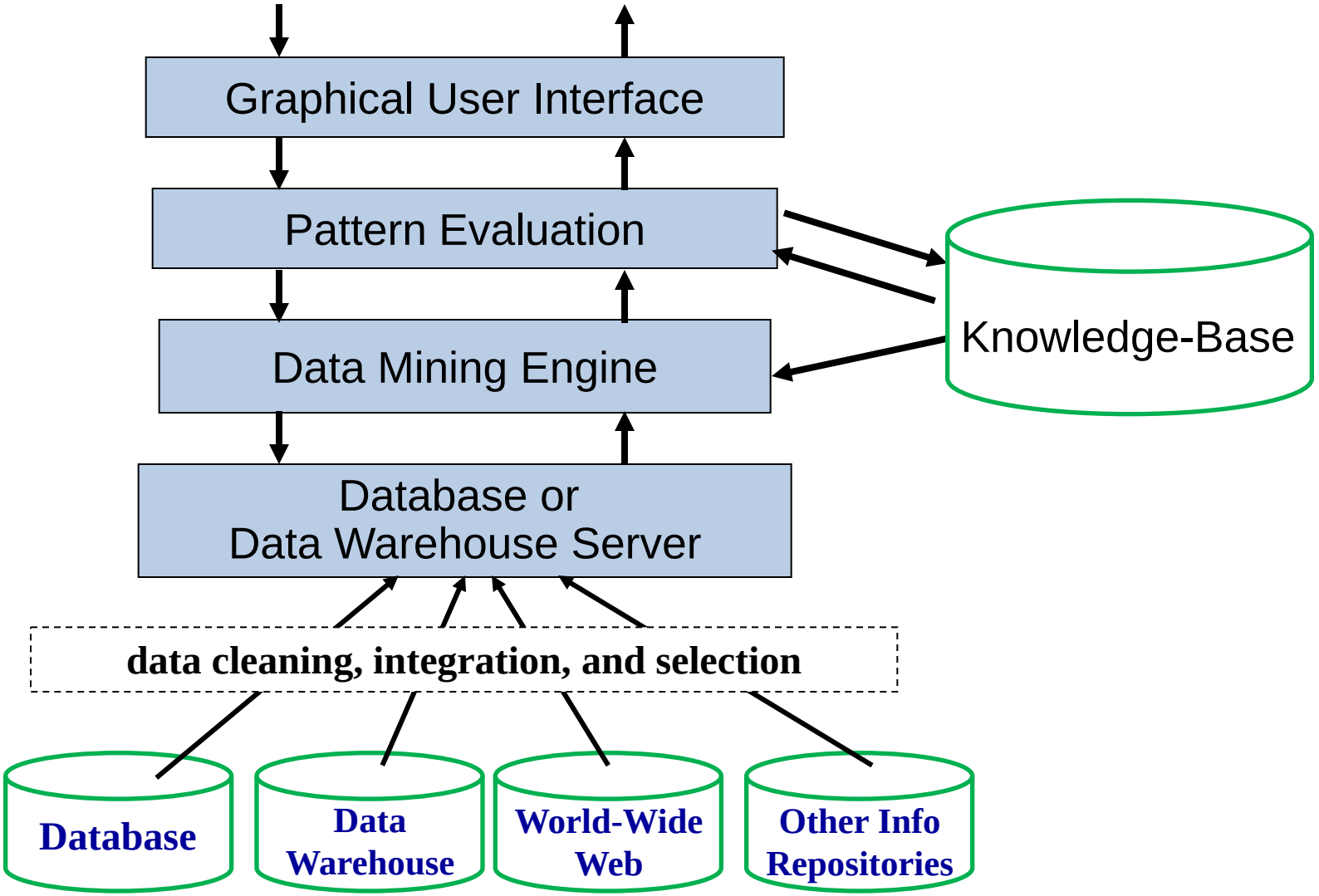
Source: Han & Kamber (2006)

Example of Star Schema



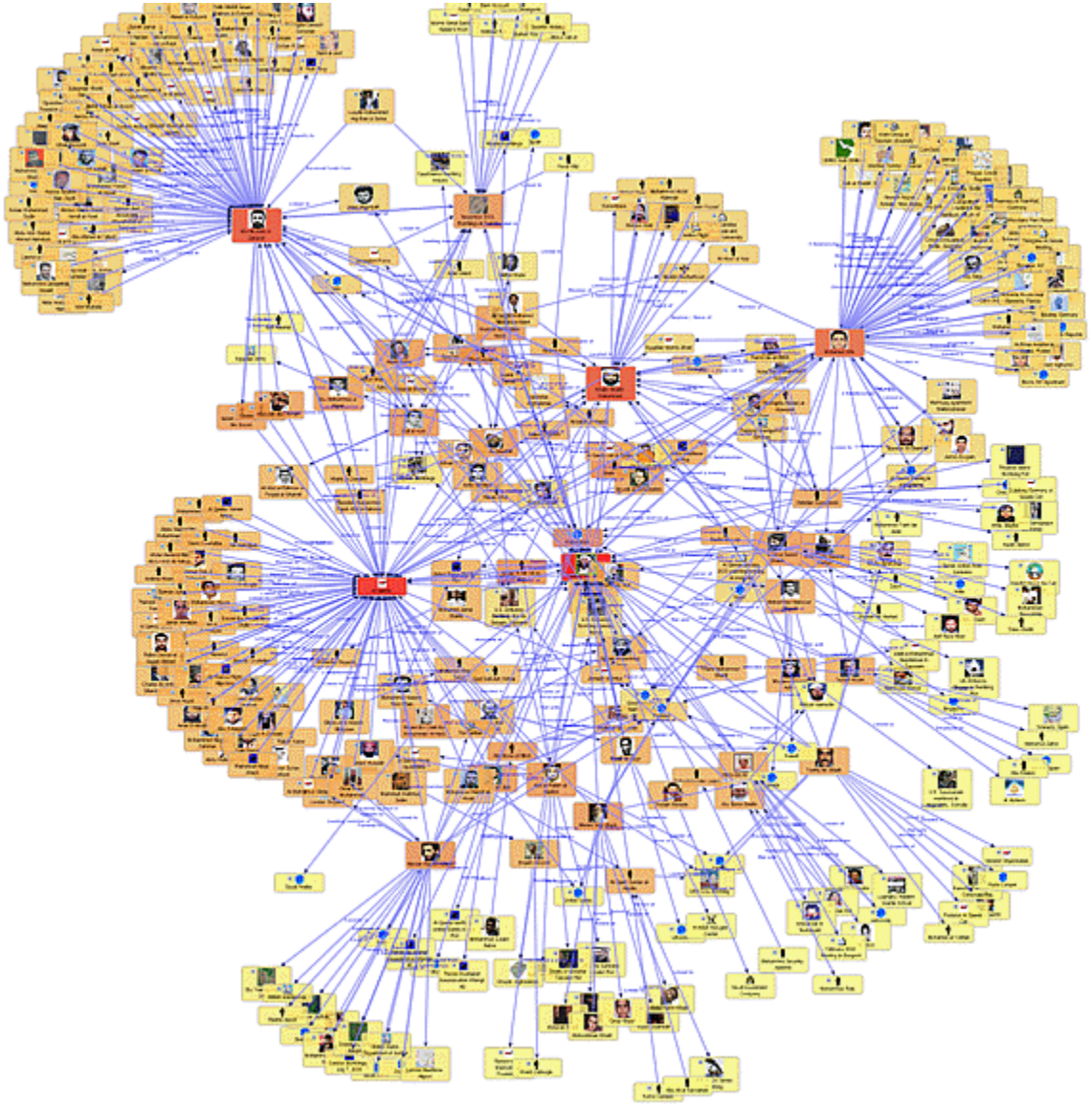
Source: Han & Kamber (2006)

Architecture of a typical data mining system



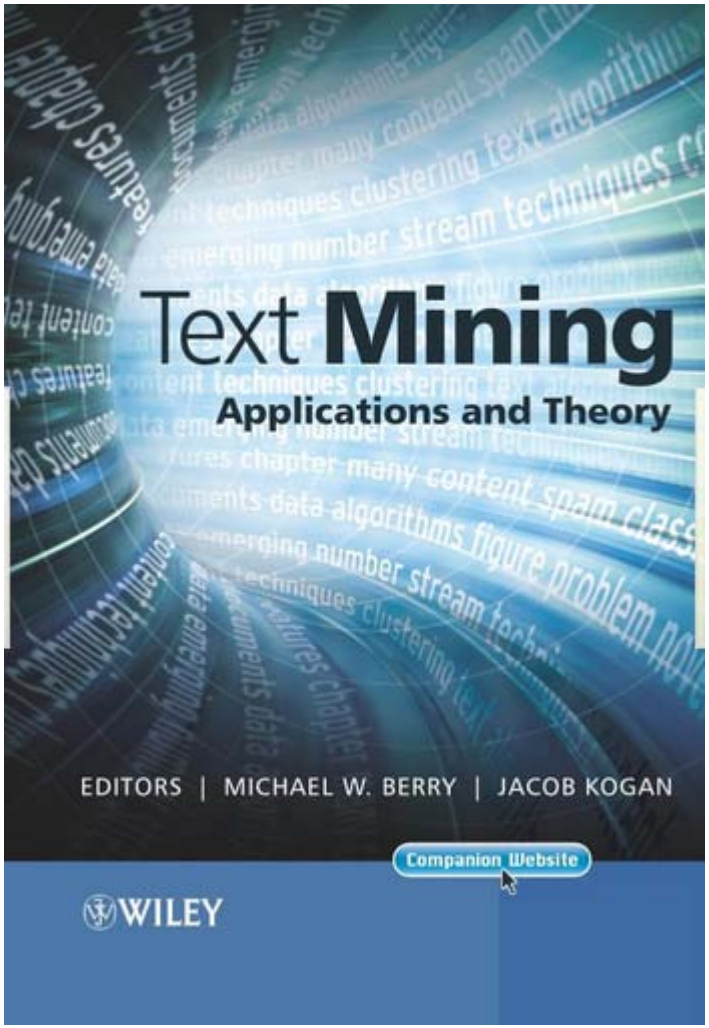
Source: Han & Kamber (2006)

Social Network Analysis



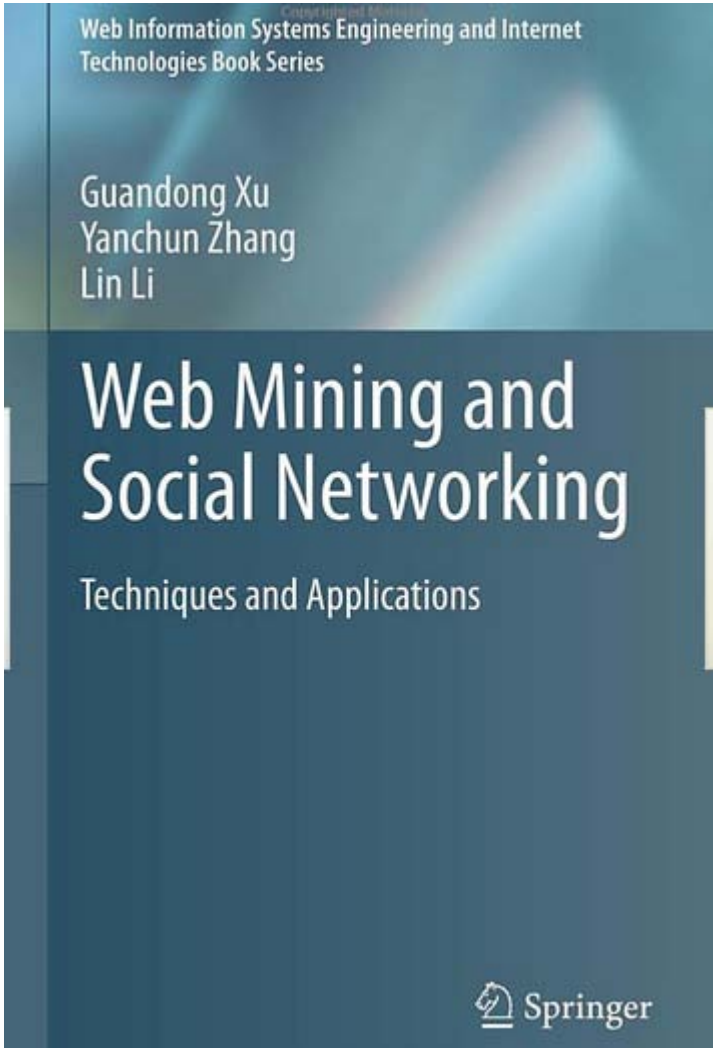
Source: http://www.fmsasg.com/SocialNetworkAnalysis/

Text Mining



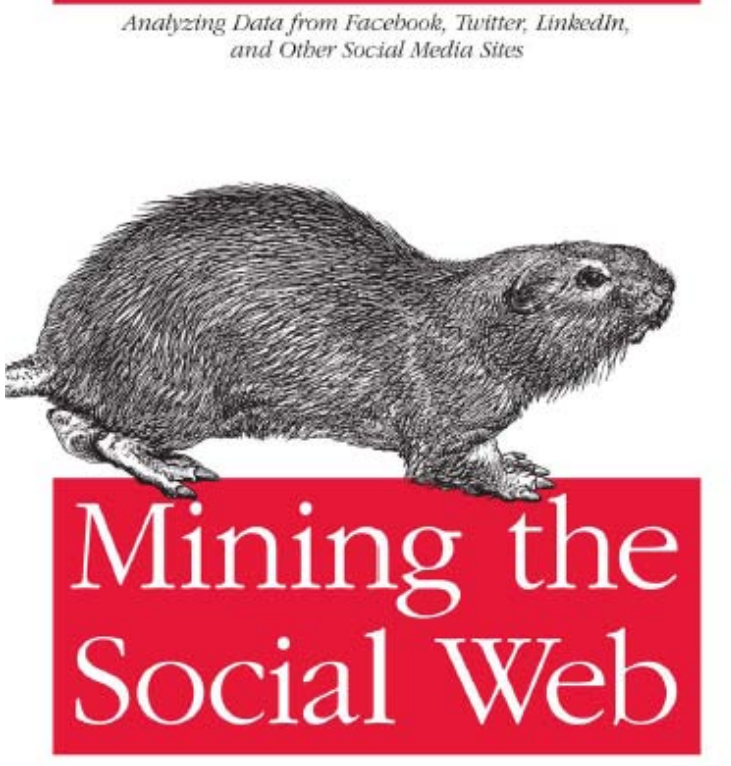
Source: http://www.amazon.com/Text-Mining-Applications-Michael-Berry/dp/0470749822/

Web Mining and Social Networking



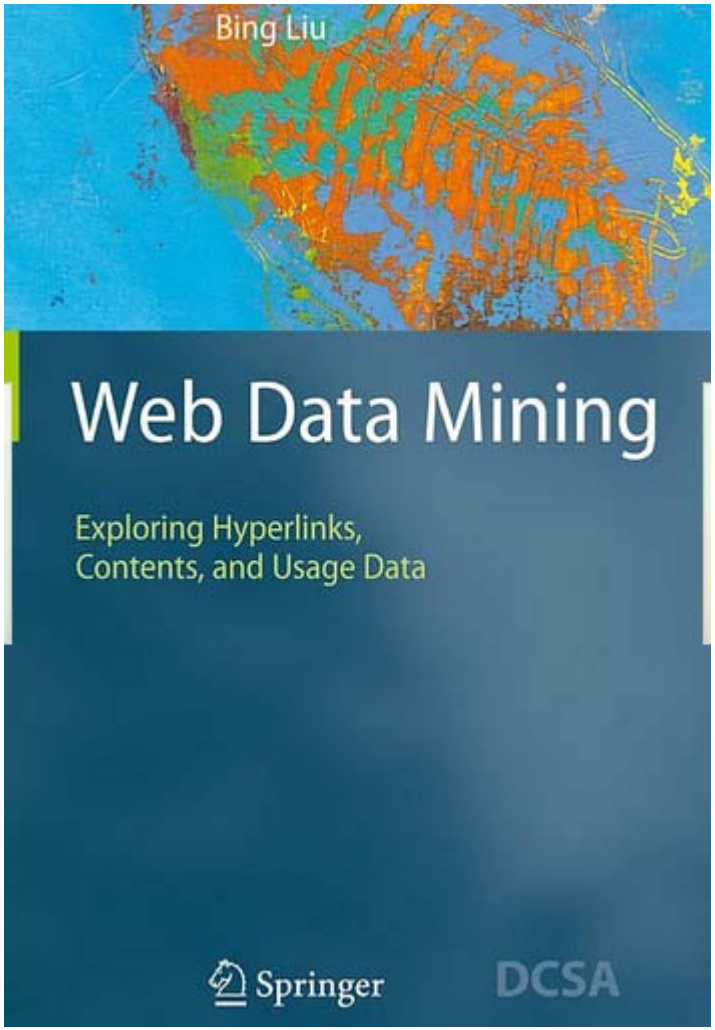
Source: http://www.amazon.com/Web-Mining-Social-Networking-Applications/dp/1441977341

Mining the Social Web: Analyzing Data from Facebook, Twitter, LinkedIn, and Other Social Media Sites



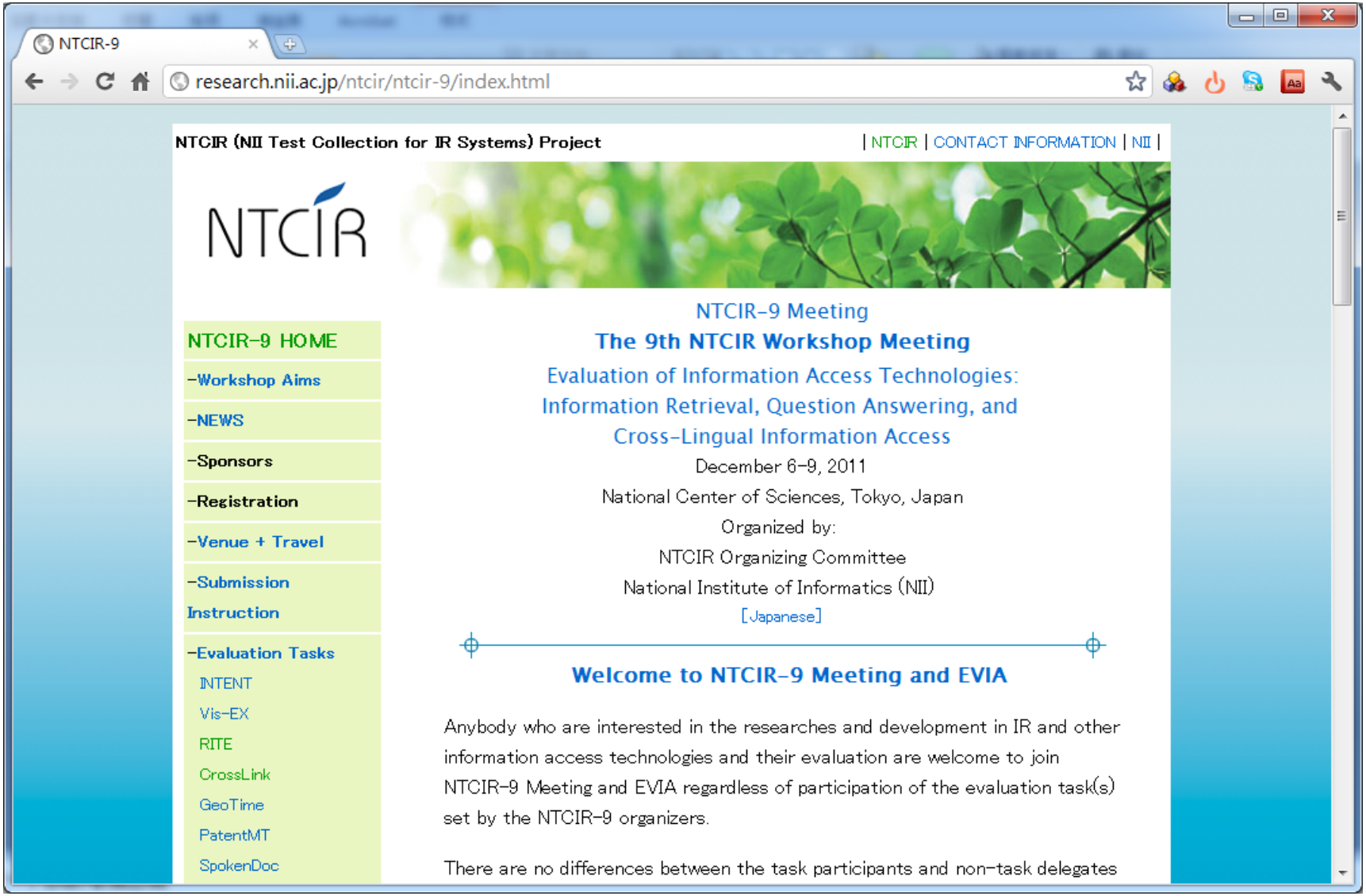
Source: http://www.amazon.com/Mining-Social-Web-Analyzing-Facebook/dp/1449388345

Web Data Mining:
Exploring Hyperlinks, Contents, and Usage Data



Source: http://www.amazon.com/Web-Data-Mining-Data-Centric-Applications/dp/3540378812

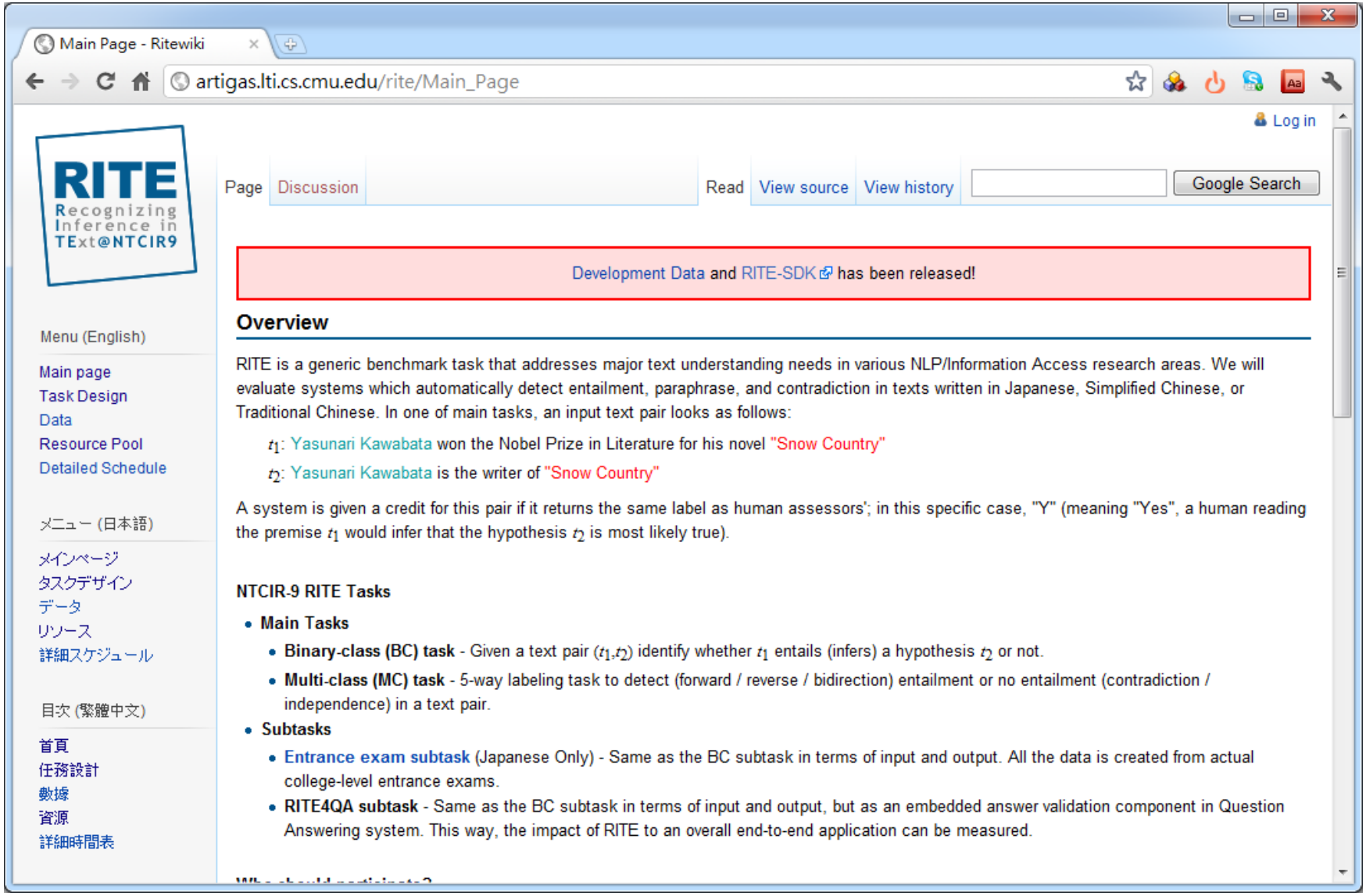
NTCIR Project
(NII Test Collection for IR Systems)



Source: http://research.nii.ac.jp/ntcir/ntcir-9/index.html

NTCIR-9 RITE

Recognizing Inference in TExt @NTCIR9



Source: http://artigas.lti.cs.cmu.edu/rite/Main_Page)

NTCIR-9 RITE

Recognizing Inference in TExt @NTCIR9

Source: http://artigas.lti.cs.cmu.edu/rite/Main_Page_(TC)

Contact Information

戴敏育 博士 (Min-Yuh Day, Ph.D.)

專任助理教授
淡江大學 資訊管理學系

電話：02-26215656 #2347
傳真：02-26209737
研究室：I716 (覺生綜合大樓)
地址： 25137 新北市淡水區英專路151號
Email： myday@mail.tku.edu.tw
網址：http://mail.im.tku.edu.tw/~myday/