BIRKBECK COLLEGE
University of London

a. Module Specification

Module number
/ /
(For Registry Use)

1. Faculty	2. School	3. Degree(s) of which the module forms part
Social Sciences	Computer Science and Information Systems	MSc in Advanced Information Systems Masters by Research in Computer Science

4. Module title
Data Warehousing and Data Mining

5. Module value	6. Date from which the module will operate	7. Number of students per intake
Half-unit	2002/2003 academic year	About 20 FT and 20 PT

8. Pre-requisites and co-requisites to the module

Prerequisites: A first course in Database Systems (e.g. as taught in a typical U.K. undergraduate degree in computer science, or on a conversion MSc, or equivalent)

9. Does the module supersede an existing module?

No

10. Main aims of the module

To study advanced aspects of data warehousing and data mining, encompassing the principles, research results and commercial application of the technologies.

11. Learning outcomes

Knowledge and understanding in the context of the subject

To understand the technology underlying data warehousing and data mining.

Cognitive skills

Ability to analyse problems amenable to data warehousing and data mining approaches and identify appropriate technical solutions.

Subject-specific practical/professional skills

Familiarity with using the data warehousing and data mining functionality of the SQL*Server 2000 and Oracle 9i database management products.

General/transferable skills (including key skills)

To understand how data warehouses and data mining technologies are deployed within modern-day information systems.

12. Module structure, syllabus and assessment method

One-term course, 11 weeks, 3 hours per week

Assessment: by written examination and practical coursework. The final course mark will be the exam mark attained. Passing the practical coursework component will be compulsory in order to pass the course overall.

Syllabus:

- Review of basic concepts of data warehousing and data mining, reasons for their use and benefits and problems arising.
- Data warehouse logical design: star schemas, fact tables, dimensions, other schemas, materialized views.
- Data warehouse physical design: hardware and i/o considerations, parallelism, indexes.
- Data warehousing technologies and implementations: data extraction, transportation, transformation, loading and refreshing. Data warehouse support in SQL Server 2000 and Oracle 9i.
- From data warehousing to data mining: OLAP architectures, design and query processing. SQL extensions for OLAP.
- Data mining approaches and methods: concept description, classification, association rules, clustering.
- Mining complex types of data.
- Research trends in data warehousing and data mining.

Indicative Reading:

J Han, M Kamber,
Data Mining Concepts and Techniques,
Morgan Kaufmann, 2001, ISBN 1-55860-489-8

C Seidman,
Data Mining with Microsoft SQL Server 2000 Technical Reference
Microsoft Press, ISBN 0-7356-1271-4

Reading will be supplemented by papers from the research literature.

13. Workload

Indicate the number of hours the student will spend in:

Lectures 23

Seminars:

Tutorials:

Field Work:

Project Work:20 hours (2 hours per week coursework)

Laboratories: 10

b. Resources Specification

1. Teaching staff required

Name	Department	% of total teaching
N J Martin	School of Computer Science and Information Systems	100%

2. Additional resources required

Accommodation

See course proposal forms

Library

Have you discussed library provision for the course with your subject librarian?

Yes

Other Library Resources required e.g. computing, a-v equipment.

None

Computing

Have you discussed any requirements for the use of specific software packages with CCS technical support staff? **Yes** / **No**

No CCS support required.

CCS No implications

Department No outstanding problems

Part-time teaching

Nil

Other

Nil

3. Reading list

a) Books of which students are expected to own copies

None

b) Books for which a high level of duplication within the library will be needed

None

c) Other required reading

None – several background reading texts will be suggested (see Syllabus above), but the course material – both textual and on-line resources – will be self-contained.

4. Recommendations

AGREEMENT

Dean of Faculty _____ **Date**_____

Head of School _____ **Date**_____

Librarian _____ **Date**_____

Comments _____

CCS Manager _____ **Date**_____

Comments _____

CLOSING DATE: **1 FEBRUARY** PRECEDING THE SESSION IN WHICH TEACHING WOULD BEGIN. YOU MUST ALSO OBTAIN **ALL** OF THE ABOVE SIGNATURES BEFORE YOU SUBMIT THE FORM.

On completion please return to the Dr Brian Harwood, Registrar.