

**SOUTHAMPTON SOLENT UNIVERSITY**

**UNIT INFORMATION GUIDE**

**ADVANCED AND DISTRIBUTED DATABASES**

Unit Code: SWD301

School of Computing and Digital Communications

Faculty of Technology

2009 - 2010

## UNIT INFORMATION GUIDE

Contents	Page
1. Introduction by Unit Leader	2
2. Teaching Scheme	3
3. Unit Teaching Team	4
4. Software	4
5. Unit Descriptor	5
6. General and Administrative Unit Information	5
7. Assessment	5
8. Evaluating Technologies, Tools and Methods	5
Assessment Activities:	
Exploiting DBMS Data Models and Server Functionality	7
Accessing and Manipulating Data Programmatically	9
Improving Data Access by Data Distribution and Replication	11
Multi-Dimensional Modelling and Analysis for Decision Support	19
Mining Databases for Decision Support	21

Appendix - Assessment Specification (not page numbered)

### 1. Introduction by Unit Leader

Welcome to Advanced and Distributed Databases (A&DD). Databases are at the heart of most IT systems, for example the mobile networks and cash machines that we use everyday.

This unit takes a thematic approach in building on the database application development skills taught earlier on the courses to develop advanced database knowledge and development skills. The themes are:

- Exploiting DBMS Data Models and Server Functionality
- Accessing and Manipulating Data Programmatically
- Improving Data Access by Data Distribution and Replication
- Multi-Dimensional Modelling and Analysis for Decision Support
- Mining Databases for Decision Support

These themes are designed to provide a reasonably broad coverage of database systems including the design, access and management of both operational (including distributed) and decision support database applications. There is a brief narrative of each theme in myCourse.

Each theme includes both theory and practice. The practical activity involves the design and implementation of applications based mainly on the Oracle HROE (Human Resource - Order Entry) databases.

The teaching, learning and assessment approach, including the very important

need for you to do initial guided reading for each theme from the online learning resources in myCourse, is discussed in the "Teaching Scheme" section below. Key points from this guide are also presented in the "Studying A&DD" Powerpoint presentation.

Work hard and enjoy the unit!

## **2. Teaching Scheme**

The teaching scheme, briefly described below and fully detailed in myCourse, is a plan of the studies and work throughout the year.

The first week is essentially an introduction to the unit. This includes checking that you can access your account on the SSU Oracle database, and doing any necessary revision.

The unit is comprised primarily of five 4-week blocks of study, linked to the assessment, each based on one of the five key themes of the unit. Following each "Resources and Activities" link in myCourse will reveal the key practice and assessment activities of each block, and the underpinning narrative, topics, resources and initial guided reading. This flexible structure should help you to complete these activities (and achieve the intended learning outcomes) by the end of each block.

You should complete the initial guided reading for ALL the topics of a theme/block as soon as you possibly can, ideally before the beginning of each block of study. You should then complete the practice activities before tackling the assessment activities. The key focus during the block should then be on completing the key HROE-based activities of the theme, and uploading to myCourse the corresponding solutions/deliverables and evaluation (discussed in the assessment and later in section 8). You will need to draw upon the underpinning resources for each theme in order to help you to complete the activities.

The five blocks of study are followed by a research, application and evaluation of a chosen advanced topic (discussed in online forums in each of the five theme blocks) that extends from each of the themes.

The unit concludes with the compilation and completion of the evaluation report to be handed-in in week 25.

Normally, each timetabled class session throughout the year will start with a review of progress including clarifying any points arising from the guided reading and/or activities. (We may use the online question-based Powerpoint presentations to help clarify any points arising from the initial guided reading).

This review will be followed by working on the activities. This will be individually, in informal groups or the class as a whole, and guided where necessary by the tutor. It will be particularly helpful to discuss the uploaded myCourse solutions, and the content of Wikis for those activities that can be completed in groups.

The final classroom session of each theme will focus on an evaluation of the technologies, methods and tools of the theme. This session should help you to write this essential element to be included in each theme deliverable.

The blocks/themes summarised in chronological order (with details of the weeks on the main myCourse page) are:

1. Introduction
2. Exploiting DBMS Data Models and Server Functionality
3. Accessing and Manipulating Data Programmatically
4. Improving Data Access by Data Distribution and Replication
5. Multi-Dimensional Modelling and Analysis for Decision Support
6. Mining Databases for Decision Support
7. Research, Apply and Evaluate your Chosen Advanced Topic

You should use the scheme to organise your workload, in preparing for timetabled sessions, engaging in these sessions, completing any directed learning and reflecting on your learning through independent study. You are also expected to plan your studies in order to meet the assessment deadlines as specified on the teaching scheme.

A flexible mix of study modes including timetabled sessions, directed and independent learning is used. Therefore, week by week please:

- Check for notices in myCourse for any special arrangements.
- Familiarise yourself with the main topic(s), resources, initial guided reading and activities of each theme.
- Attend, engage in, and reflect upon all the timetabled sessions.
- Engage in the forums and other elearning activities.
- Complete any outstanding work associated with the timetabled sessions by the next week or theme.
- The amount of independent study time you need to put in each week to keep up with the requirements of the unit may vary considerably but as a bare minimum you should study for at least an equivalent amount of time outside of class as you have timetabled sessions.
- If you miss any session it is your responsibility to ensure you catch up with the work.

### 3. Unit Teaching Team

The unit teaching team are those individuals identified as having the main responsibility for the teaching and assessment of the unit.

Name	Role	Room	Telephone	eMail
Al Monger	Unit Leader	RM413	(023)80319385	al.monger@solent.ac.uk

### 4. Software

We will be using Oracle 10g (Release 2) throughout the year. The main development tools available will be SQL Developer and iSQLPlus. Up-to-date database connection and other information about this software is available in the first block in myCourse. We will also be using some other software which is indicated within the themes. (The first block in myCourse provides information on what to do if the Oracle database(s) or iSQLPlus clients appear to be down).

### 5. Unit Descriptor

The unit descriptor identifies the specific knowledge, skills and understanding a student should have acquired after studying and being assessed in this unit. (Please note that the case study is now named "HROE" rather than "mySports").

The current Unit Descriptor can be found in myCourse.

## **6. General and Administrative Unit Information**

Please refer to the unit in myCourse, and to the DATABASE portal, for up-to-date information such as reading lists, and all the essential learning resources.

Please also refer to the Student Course Handbook.

### **Unit Feedback**

The unit team will be interested in your feedback on the operation of the unit and ways that it might be improved. Feedback will be obtained via the use of questionnaires and/or informal small group discussions

### **Issues of Concern**

Any issues concerned with the academic content or delivery of this unit should firstly be raised with the lecturer concerned at the time they occur.

### **Student Attendance Reporting**

If your attendance at timetabled sessions falls below an acceptable level it will be reported through the faculty's attendance monitoring and reporting procedures.

## **7. Assessment**

Advanced and Distributed Databases has one in-course assessment. The assessment specification is in the appendix of this UIG. The assessment includes 5 assessment activities which are in the UIG and also in myCourse. You should also refer to block 9 of the main myCourse page for other important assessment information not contained in the Unit Information Guide.

It should be noted that the Faculty has a policy of **NOT HANDING BACK** student work submitted for assessment. However, you will always receive feedback on your work.

## **8. Evaluating Technologies, Tools and Methods**

You will be required in the assessment to:

1. Critically evaluate the technologies, tools and methods of all five themes against explicit criteria (eg functionality, ease-of-application, reliability, cost).
2. Identify and discuss any key issues that should be taken into account (eg skill shortages).
3. Make appropriate recommendations (eg on development methodologies).

In order to evaluate something, it is necessary first to establish criteria by which one might evaluate it. For example, one might evaluate a car by using criteria such as performance (acceleration, fuel consumption etc), reliability and cost

(purchase, servicing etc). For a database technology such as relational databases/DBMS, one might establish criteria such as functionality, reliability, usability etc.

Example evaluation criteria and related questions are indicated below.

#### *Functionality and Meeting Business Needs*

What does the technology actually do? Does it enable applications to be developed to meet business needs?

#### *Usability*

What is the learning curve for the technology? What is particularly challenging about the technology and why? Is the technology well supported by academic and industrial references/sources? Are the skills available in the market place?

#### *Reliability/Robustness*

Does the technology work? Is it reliable? Is it well grounded in database theory?

#### *Tools*

Are there tools available for developing applications using these technologies? Are the tools well supported by documentation? How do they contrast and compare to other tools?

#### *Methods*

Are there methods (and associated tools) available for developing applications using these technologies? How well established and universal are these methods?

#### *Cost*

Is the technology expensive to purchase and maintain?

The practical experience gained from having used some of these technologies, tools and methods in the practical activities should help inform this evaluation. You should therefore reference the design, code and other artefacts you have developed using these technologies and associated tools and methods.

The evaluation should be well-structured and concise. It should be targeted at a management audience who are reasonably IT-literate but perhaps somewhat out-of-date with respect to modern database technologies.

You must avoid just giving an opinion! You should base your evaluation on:

- a) Your experience of using the technology in developing the myCourse deliverables (and you should of course refer to these artefacts).
- b) Reference to academic and industrial/commercial references (using the Harvard referencing system).

# Exploiting DBMS Data Models and Server Functionality

## Scenario

There is a need to further develop product and price information, and its availability externally, for the OE database (data model below).

### 1. Development of Product and Price Information

Currently, there is no “price history” for a product. More specifically, it is important for the company to record that, for example, the price history of a product is:

1<sup>st</sup> December 2007 - £20.00 (product introduced)  
15<sup>th</sup> June 2008 - £25.00 (price increased to £25.00)  
5<sup>th</sup> January 2009 - £24.00 (price decreased to £24.00) etc

There is of course a need for full integrity in respect of the existing product and the new price information including, for example:

1. When a product is inserted, a row must be inserted in the price table  
(To test this requirement you should implement a temporary table (database or PL/SQL) to store any relevant data needed external to your script such as the updated price).
  2. Uniqueness (and not null) for the product name and as appropriate for other fields.
  3. When a product is deleted, all associated price records must also be deleted.
- The requirement is therefore to design and implement an improvement to the OE database in respect of product and price information.

### 2. External Availability of Product and Price Information

The company is increasingly getting requests from its customers (mainly other companies) to provide information about its products and previous prices in XML form.

The requirement is to compare and contrast solutions using the Oracle XML DB functionality. One solution should be based on using the existing “non-XML” tables, and the other solution should be based on implementing tables using the XML Type functionality. Examples of XML updates and retrievals should also be provided.

Please refer to “Access to the Case Study Databases” in myCourse for access to the HR and OE databases information.

## Required

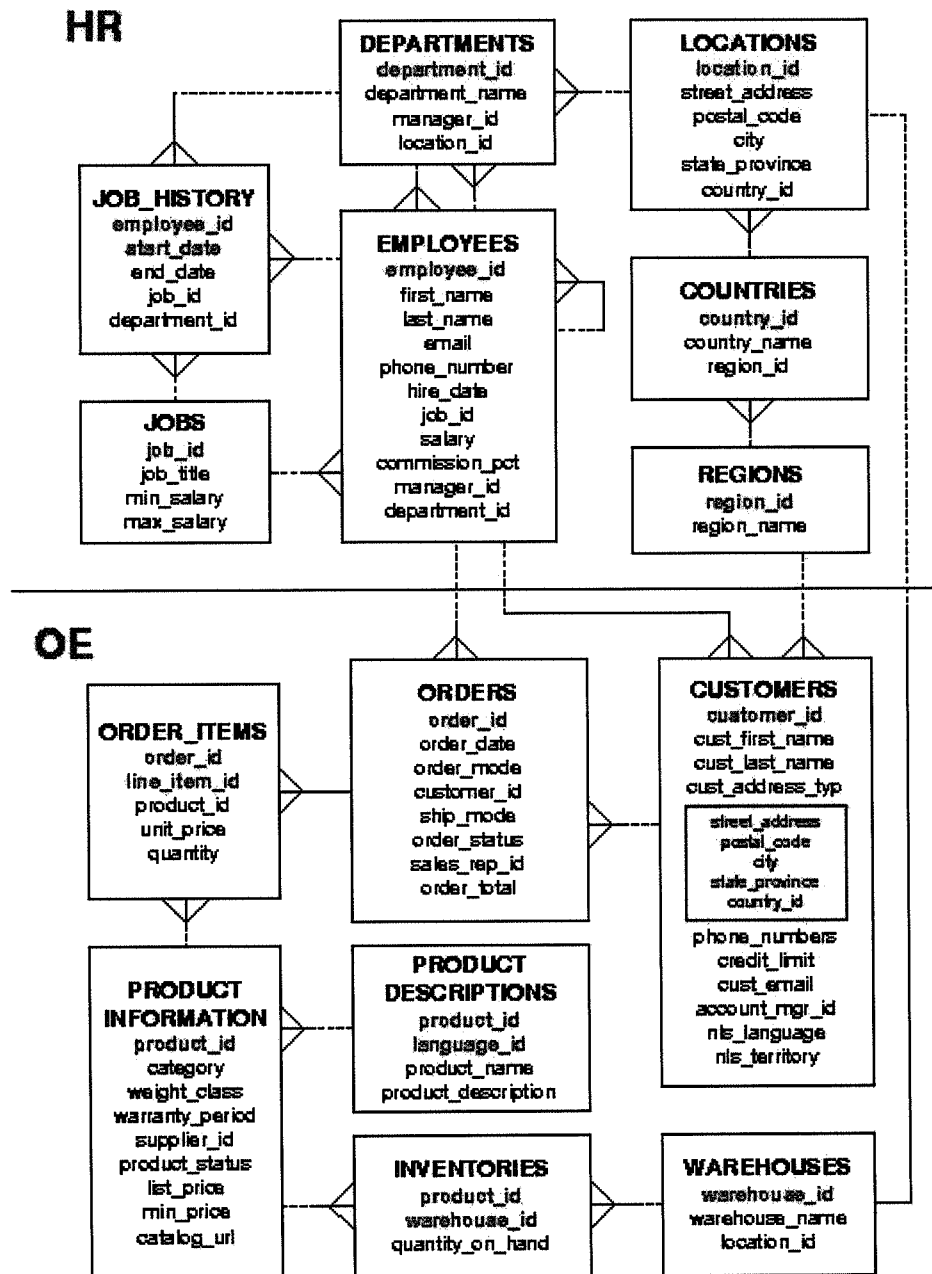
1. Implement and test (using iSQL\*Plus or SQL Developer) documented scripts which meet the above requirements. You should create the necessary tables (including a copy of the OE product information table) in your own Oracle account.

*Note - Evidence of the application of more complex Oracle XML XB functionality would be an important contribution to higher-level achievement.*

2. Critically evaluate the technologies, tools and methods relating to this theme as indicated in section 8.

You should upload the assessment activity (in the general content and format described in the assessment specification) by the end of week 5.

## HR and OE Data Model





# Accessing and Manipulating Data Programmatically

## Scenario

Some privileged customers have accounts. The ACCOUNTS table is modified according to instructions stored in the ACTION table. Each row in the ACTION table contains an account number, an action to be taken (I, U, D or X) for insert, update, delete or invalid operation), an amount by which to update the account, and a time tag used to sequence the transactions.

On an insert, if the account already exists, an update is done instead. On an update, if the account does not exist, it is created by an insert. On a delete, if the row does not exist, no action is taken.

*You will need to run the Accounts script (available in the theme block in myCourse) to create the additional tables required for this activity in your own Oracle account.*

## Input Tables

SQL> SELECT \* FROM accounts ORDER BY account\_id;

ACCOUNT_ID	BAL
1	1000
2	2000
3	1500
4	6500
5	500

SQL> SELECT \* FROM action ORDER BY time\_tag;

ACCOUNT_ID	O	NEW_VALUE	STATUS	TIME_TAG
3	u	599		07-DEC-06
6	i	20099		07-DEC-06
5	d			07-DEC-06
7	u	1599		07-DEC-06
1	i	399		07-DEC-06
9	d			07-DEC-06
10	x			07-DEC-06

## Output Tables

SQL> SELECT \* FROM accounts ORDER BY account\_id;

ACCOUNT_ID	BAL
1	399
2	2000
3	599
4	6500
6	20099
7	1599

SQL> SELECT \* FROM action ORDER BY time\_tag;

ACCOUNT_ID	O	NEW_VALUE	STATUS	TIME_TAG
3	u	599	Update: Success.	07-DEC-06
6	i	20099	Update: Success.	07-DEC-06
5	d		Delete: Success.	07-DEC-06
7	u	1599	Update: ID not found. Value inserted.	07-DEC-06
1	i	399	Insert: Acct exists. Updated instead.	07-DEC-06
9	d		Delete: ID not found.	07-DEC-06
10	x		Invalid operation. No action taken.	07-DEC-06

## Required

1. Design, implement and test a documented batch transaction processing program which will update the ACCOUNTS table according to the instructions in the ACTION table.

*Note - It is important for a contribution to higher level achievement of 2.1 and above to take into account in your design the issue of potential migration to an alternative DBMS in the future.*

2. Critically evaluate the technologies, tools and methods relating to this theme as indicated in section 8.

## Notes

*You can choose from two distinct approaches to meeting this requirement. However, for both approaches you must develop and test the solution against the ACTION and ACCOUNTS tables created in your Oracle account.*

### 1. Using Java

*You may use any Java application development software (such as Eclipse or Netbeans) of your choice to develop a Java-based solution. Please also refer to **Some Guidance on a Java-based Solution to the Accounts Update Program** in myCourse. This approach should normally be attractive to Computing Programme students.*

### 2. Using PL/SQL

*You must develop an entirely Oracle-based solution using PL/SQL in iSQLPlus. This approach should normally be attractive to BIT (and some Computing Programme students). Please also refer to **Some Guidance on a PL/SQL-based Solution to the Accounts Update Program** in myCourse.*

**You should upload the assessment activity (in the general content and format described in the assessment specification) by the end of week 9.**

# Improving Data Access by Data Distribution and Replication

## Scenario

The company is rapidly expanding its international business. There are now many locations (ie sites) in four main regions of the world. However, its human resource (HR) and order entry/processing (OE) applications are still centralised at one site. Further information about these applications can be found below in the **HR and OE Data Model** and online at the Oracle® Database Sample Schemas.

In order to improve access to these applications the company has decided to distribute and/or replicate HR and OE data where appropriate across its international sites. There is, however, uncertainty about how this distributed database should be designed and implemented.

Each site operates in a largely autonomous manner in respect of its HR (employees, job histories etc) and OE (customers, orders etc) applications. It is particularly important, for example, that an order transaction is completed regardless of whether the site is connected to databases at other sites. Updates of orders (and, of course, corresponding items) to other sites does not need to be instantaneous, although consistency of product and price information is a high priority.

There will also be a requirement for fully transparent, distributed queries at HQ which will enable managers to easily retrieve information (such as a list of employees) for specific sites or for the entire company.

In order to better understand distributed database design and implementation issues, a pilot/test distributed database has been implemented for HR data at 3 sites, one each in the USA, UK and Germany. The USA site is also the company headquarters (HQ). Please refer to the following **HR Distributed Database Design** and the **HR Distributed Database Implementation** for details of the global database names, schema names, database links etc.

*Please refer to the “Access to the International Distributed Database” in the theme block in myCourse for access information.*

## Required

*Note - Question 1 is a practice activity and is not required in the myCourse upload.*

1. Using a user ID on STUDENT4 (USA) available from the tutor, test this distributed database as follows:

- 1.1 Check the availability of the public database links (in ALL\_DB\_LINKS) from the USA (STUDENT4) to Germany (STUDENT3) and to the UK (STUDENT5).
- 1.2 How many employees are there at each of the Germany, UK and USA sites?
- 1.3 Retrieve a list of ALL employees and their corresponding location and country.
- 1.4 Retrieve a list of the job history (including job title) of each employee at the USA and UK sites.
- 1.5 Retrieve a list of the departments including the city and country where they are located.

*Example syntax: SELECT last\_name FROM hrd.employees@STUDENT5;*

2. Create and test a documented SQL view which will HR staff (at HQ) can use to transparently retrieve a list of ALL the HR employees (including their corresponding locations and job histories). *Refer also to the **Distributed Query for Employees Illustration** in the resources in myCourse for syntax.*

3. Produce a model using an appropriate tool of your choice for a distributed database design solution to meet the needs of both the HR and OE applications. Provide an explanation (maximum 300 words) in support of the design decisions made. Innovative approaches are required for 2.1 and above.

*Note - This should be a straightforward activity to produce a design/architecture similar to the diagram style in 24.2 of Elmasri. However, I am not aware of any UML (or other notation) for representing a model such as this, so a well-researched, innovative and designed model (perhaps using stereotypes for horizontal fragments etc) would contribute to higher level achievement of 2.1 and above.*

3. Critically evaluate the technologies, tools and methods relating to this theme as indicated in section 8.

You should upload the assessment activity (in the general content and format described in the assessment specification) by the end of week 13.

#### Additional Optional Activities

It is recommended that you also do the following activities, although they are not required for the myCourse upload.

4. Implement code for the update of the replicated departments table. In PAIRS, proceed as follows.

1. *Refer to the following **Update of the Replicated Departments Table** which assumes that the master is held at STUDENT4 (HQ/USA Site) and the snapshot is at STUDENT5 (UK site).*

2. *Note that you also have an account on STUDENT5 of the name as your account on STUDENT4.*

3. *Logon to your account on Student4, and execute `CREATE TABLE departments AS SELECT * FROM hr.departments;` to create the (master) **Departments** table and then logoff. Repeat for your account on STUDENT5 using the same script to create the snapshot.*

4. *Create a snapshot log at the master site (STUDENT4) to store info. on updates.*

5. *Implement an update database trigger on the master table (which also updates the snapshot log).*

6. *Create a database link from STUDENT4 to STUDENT5 - refer to the following **Types of Users of the HR Distributed Database Links**.*

7. *Test by using the "fast" Refresh script (available in myCourse) which will update the snapshot site from the master site using this link.*

*Note - This refresh script is written in PL/SQL, but you could write your own Java/JDBC script!*

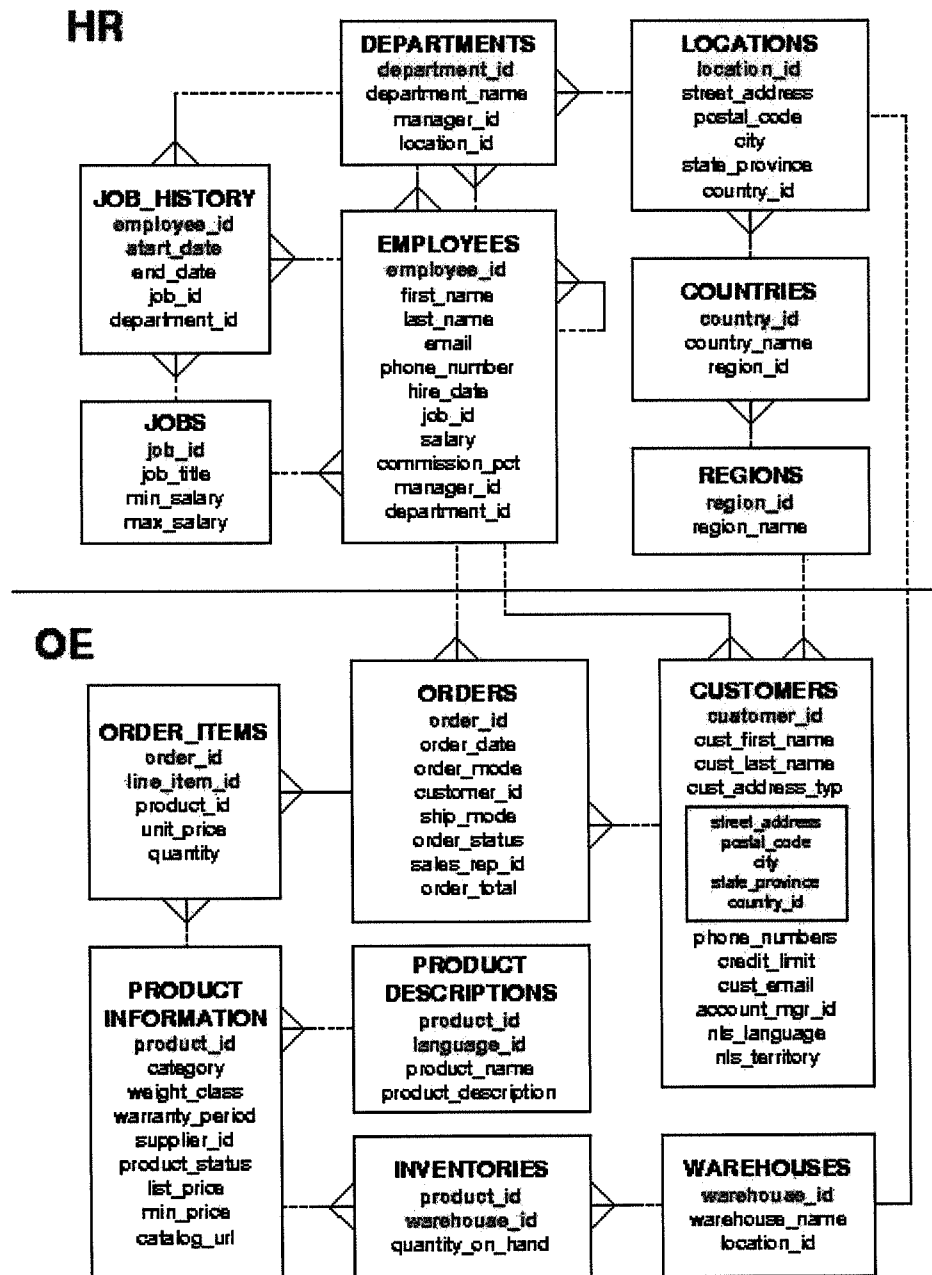
5. Discuss issues relating to distributed transactions.

Illustrate your discussion by comparing execution strategies for a query executed at site 2 for a list of ALL HR employees and their corresponding department names. Assume employee and department records are 50 and 30 bytes long respectively. Express queries in the relational algebra. Refer to Elmasri 24.4.1 and 24.4.2 for a similar scenario.

#### *Notes*

- 1. Oracle system parameters have been set such that the name of the database link (or at least the first part of it) should match the global name of the database, ie a database link named **student5** (or **student5.solent.ac.uk** in full) would be OK for linking to the database **student5.solent.ac.uk**. If, however, you tried to name your link as (say) **fred** (ie **CREATE DATABASE LINK fred USING "student5.solent.ac.uk"**) then you will get an error when you try to use the link.*
- 2. You will be using both a fixed user, public database link from your given account, and a connected user database link.*
- 3. These are comparatively short activities. There is more emphasis in this theme on being able to identify and explain the key concepts of distributed (and replicated) databases rather than their practical application. Clearly, distributed databases are applied to large organisations only and are usually complex. It has only been possible to implement a small HR distributed/replicated database to provide you with at least some practical design/implementation activity.*

## HR and OE Data Model

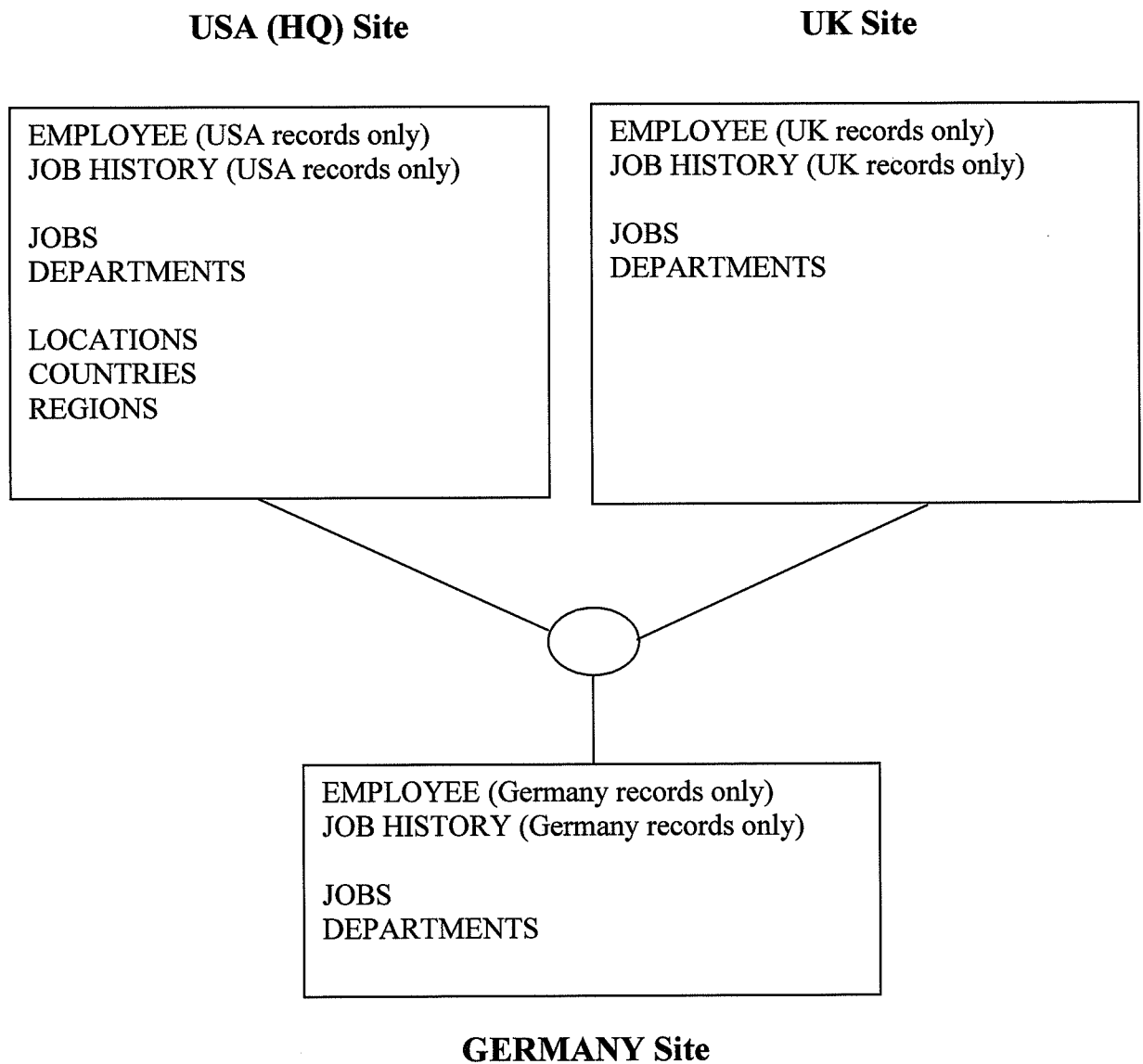


## HR Distributed Database Design

---

The HR database is distributed across three sites.

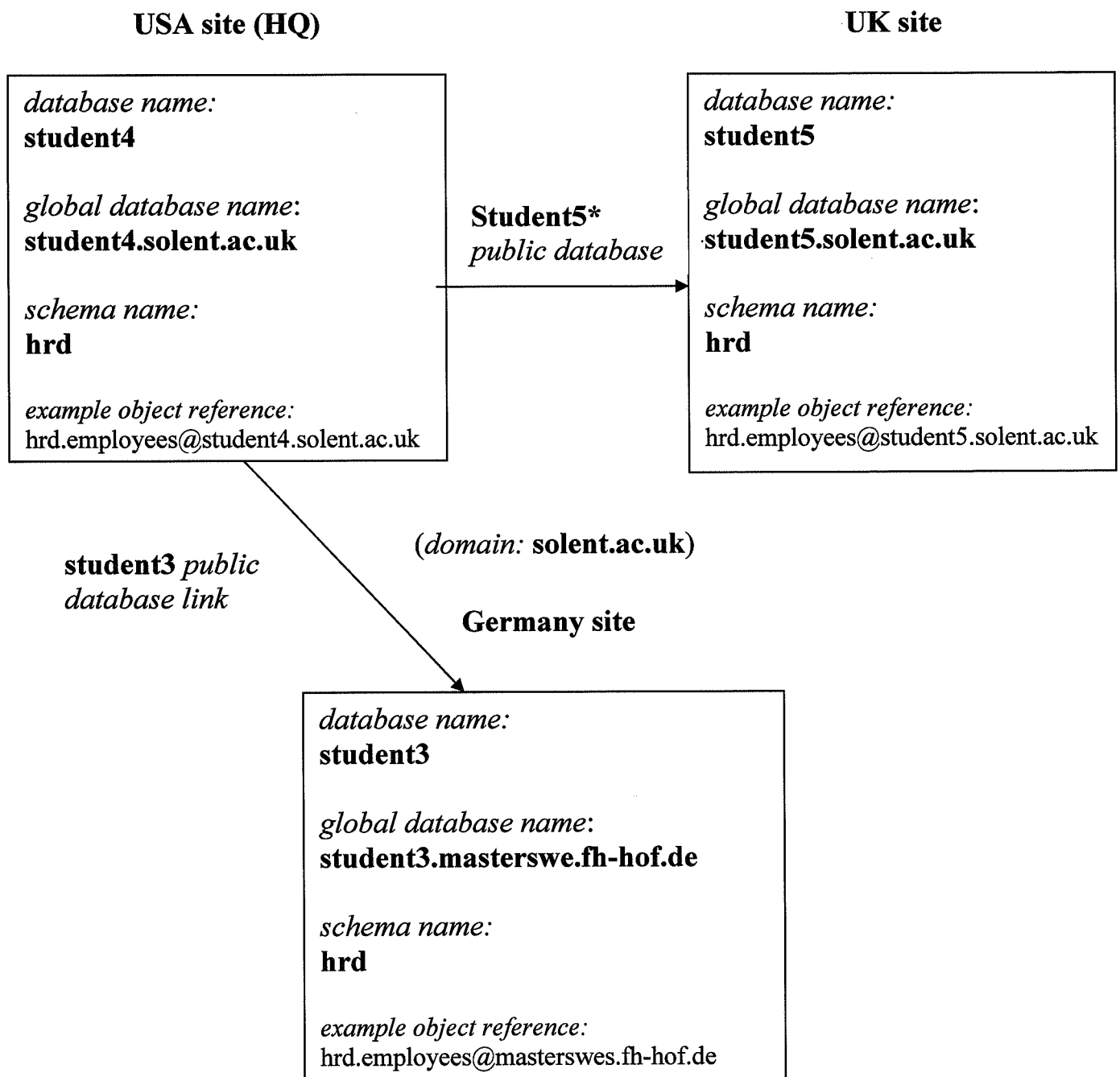
- EMPLOYEES and JOB HISTORIES are fully distributed with horizontal fragmentation
- JOBS and DEPARTMENTS are fully replicated
- LOCATIONS, COUNTRIES, REGIONS are partitioned at the USA/HQ site only



## HR Distributed Database Implementation

This shows physical implementation details of the HR distributed database including database and schema names, and public database links. The Oracle server is running at each of the three sites. The implementation can be tested with queries (run at HQ) such as **select \* from hrd.employees@student5;**

Physically, the STUDENT4 and STUDENT5 servers are located in SSU, and STUDENT3 at the fh-hof in Germany.

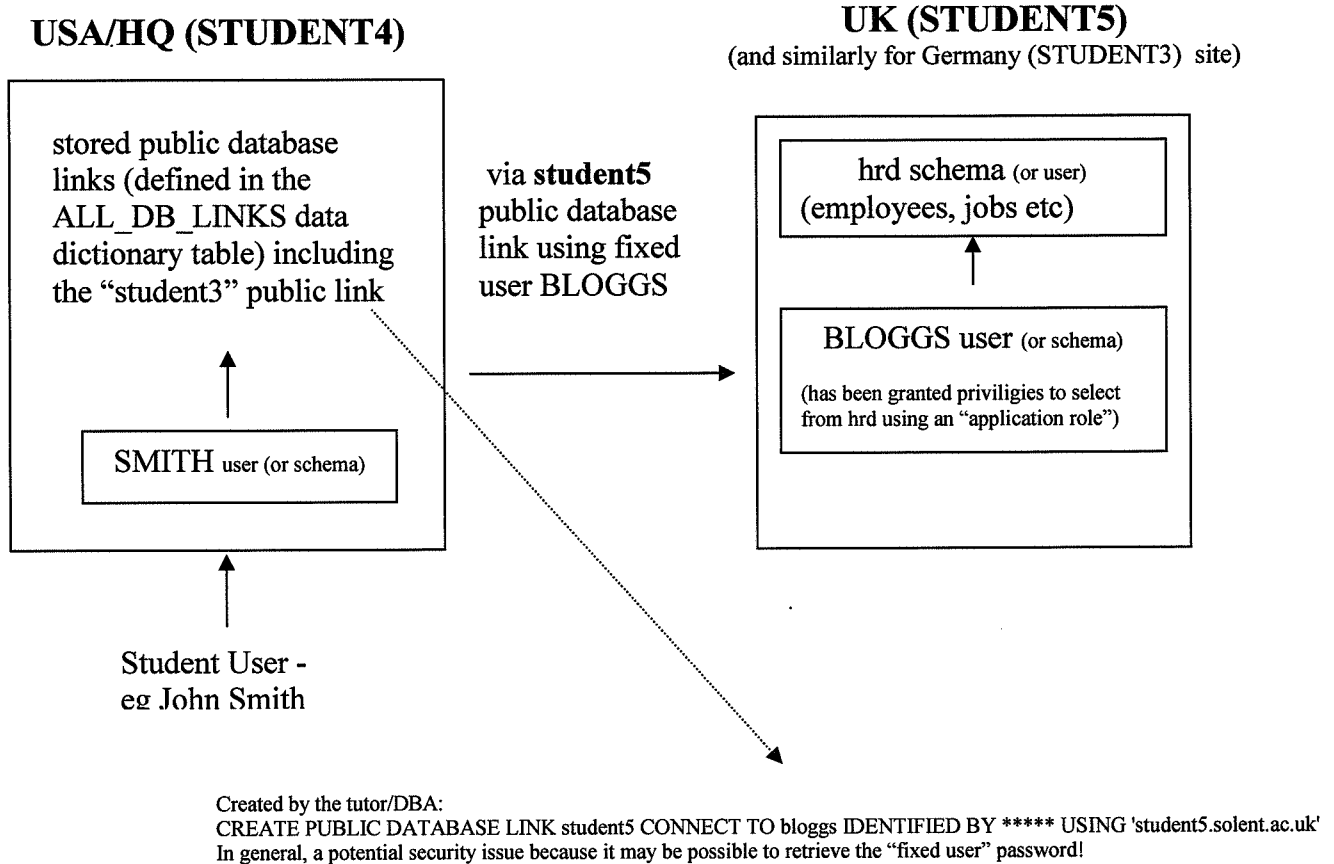


\* Full link name is **student5.solent.ac.uk** - but can be abbreviated to **student5** in this context

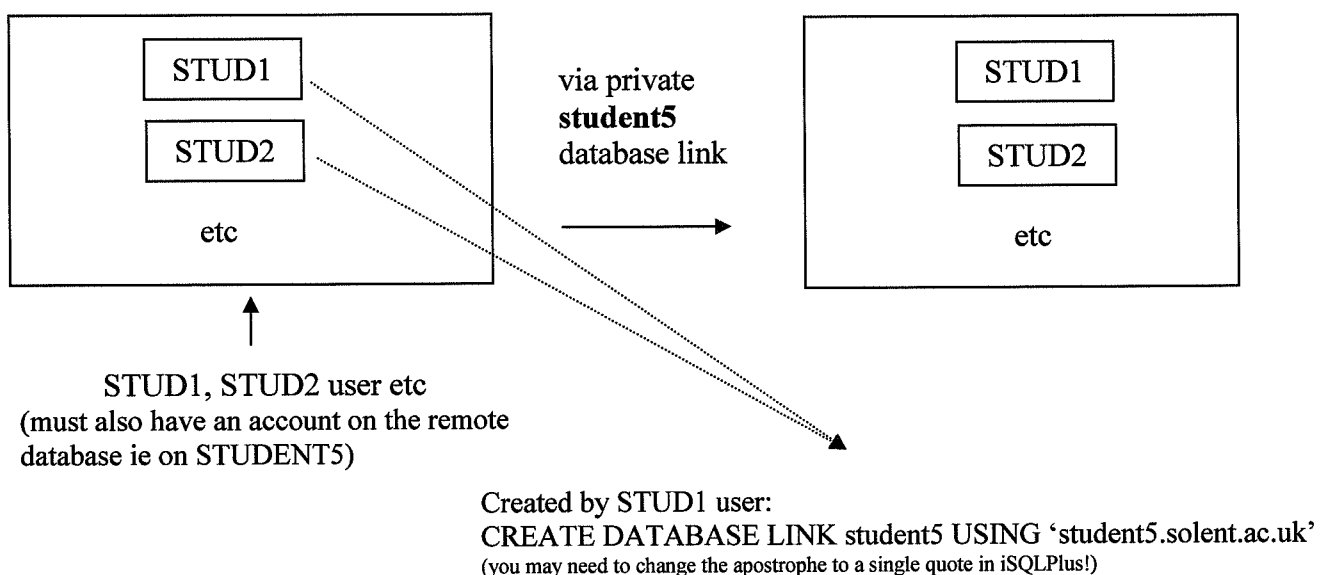


# Types of Users of the HR Distributed Database Links

## “Fixed User” Link (use for scenario requirements 1 and 2)

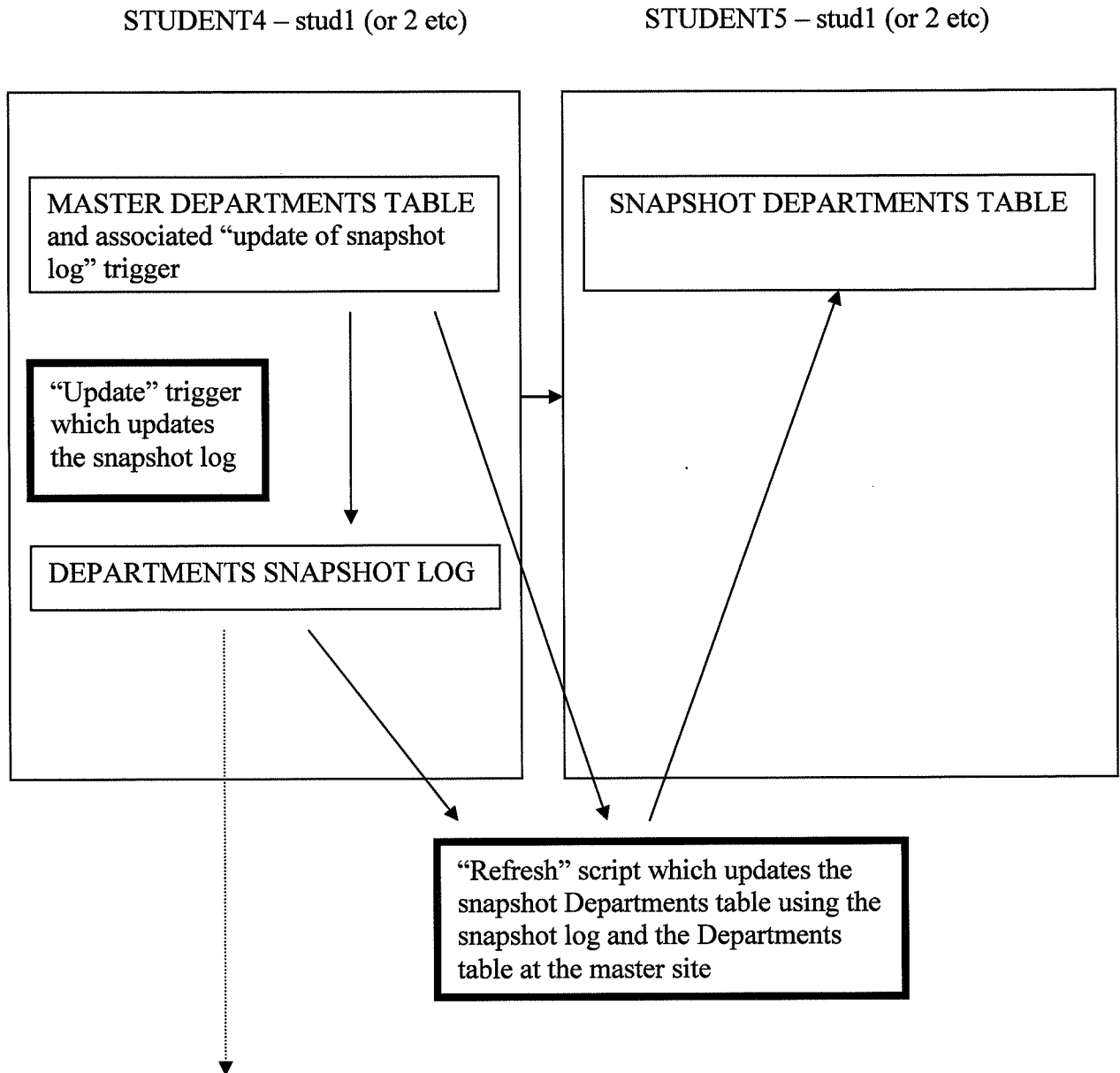


## “Connected User” Link (use for scenario requirement 4 and to UK site only)



## Update of the Replicated Departments Table

---



*Structure of the DEPARTMENTS SNAPSHOT LOG table:*

log\_id – just a sequence primary key for this table

updated\_master\_primary\_key – the departmentid of the updated master department record

updated\_master\_timestamp – the date when the master department record was updated

refresh\_snapshot\_timestamp – the date when the snapshot was updated

# Multi-Dimensional Modelling and Analysis for Decision Support

## Scenario

As the business and operational database grows, management are starting to realise that it represents an increasingly valuable information resource for the business. Identifying changing patterns of customer demand for the products, for example, would be valuable in informing decisions of which products and customers to focus on. Some information is currently being queried from the OE operational database, but this is giving rise to increasing performance and other issues. An OLAP database SH (data model below) has therefore been designed and implemented to meet this need.

Please refer to “Access to the Case Study Databases” in myCourse for access to the SH database information.

This OLAP cube is of significant size - one table, for example, has over a million records - so be careful how you construct your queries!

## Required

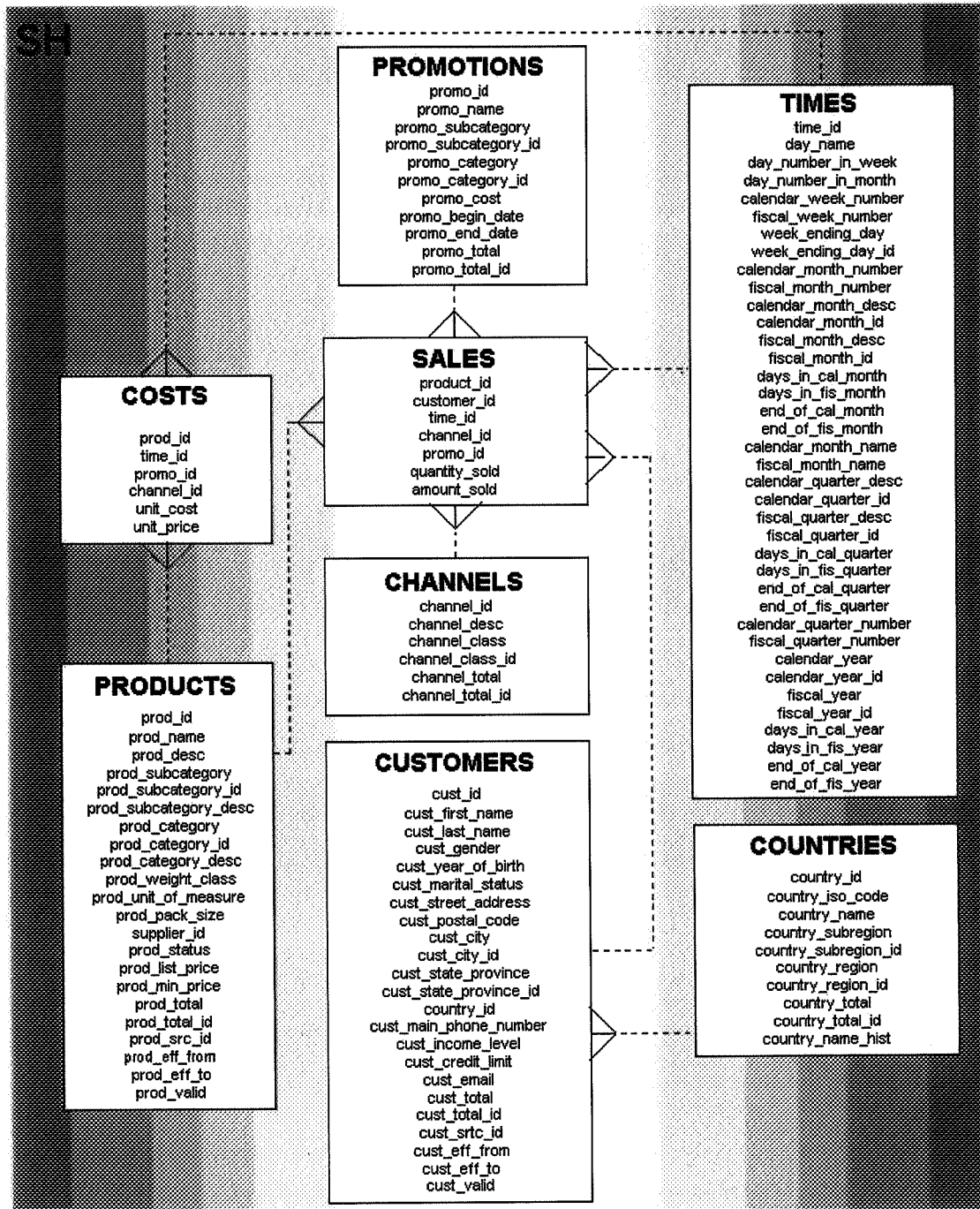
1. Implement an example of analysing data for decision support (refer to the examples in the “Analyse Multi-Dimensional Data Practical” practice activity) that could be used in a presentation to management. The example should be adequately described in respect of its analysis (ie what decision support it is providing), design and implementation. Screen shots and documented code should be provided. Example recommendations to management should be made.
2. Develop a dimension model for the HR database based on any reasonable HR decision support requirements. For example, it would be useful to know the total number of employees by country and job title. Use any appropriate diagrammatic modelling notation including UML or "crow's foot". Provide an explanation (maximum 300 words) in support of the design decisions made.

*Note - A good implementation of the design would contribute to higher level of achievement of 2.1. If you do this, then include details and evidence of the implementation with the uploaded dimension model.*

3. Critically evaluate the technologies, tools and methods relating to this theme as indicated in section 8.

You should upload the assessment activity (in the general content and format described in the assessment specification) by the end of week 17.

# SH Data Model



# Mining Databases for Decision Support

## Scenario

Management have already recognised the need for an OLAP database to better inform decision making. However, management are now beginning to realise that there may be "undiscovered" knowledge in the OE operational databases (data model below) that has the potential to benefit the business. The main requirement at this stage is to understand what is involved in a data mining project and some of the key issues. In order to achieve this, it is suggested that an example "market basket analysis" data mining model is developed. This is clearly relevant to management, and should be sufficient to demonstrate the data mining life-cycle, potential benefits and some key issues.

You have read access to the OE database (schema below) which is of sufficient size to enable development of a realistic model, and to expose some of the key issues.

Please refer to "Access to the Case Study Databases" in myCourse for access to the OE database information.

## Required

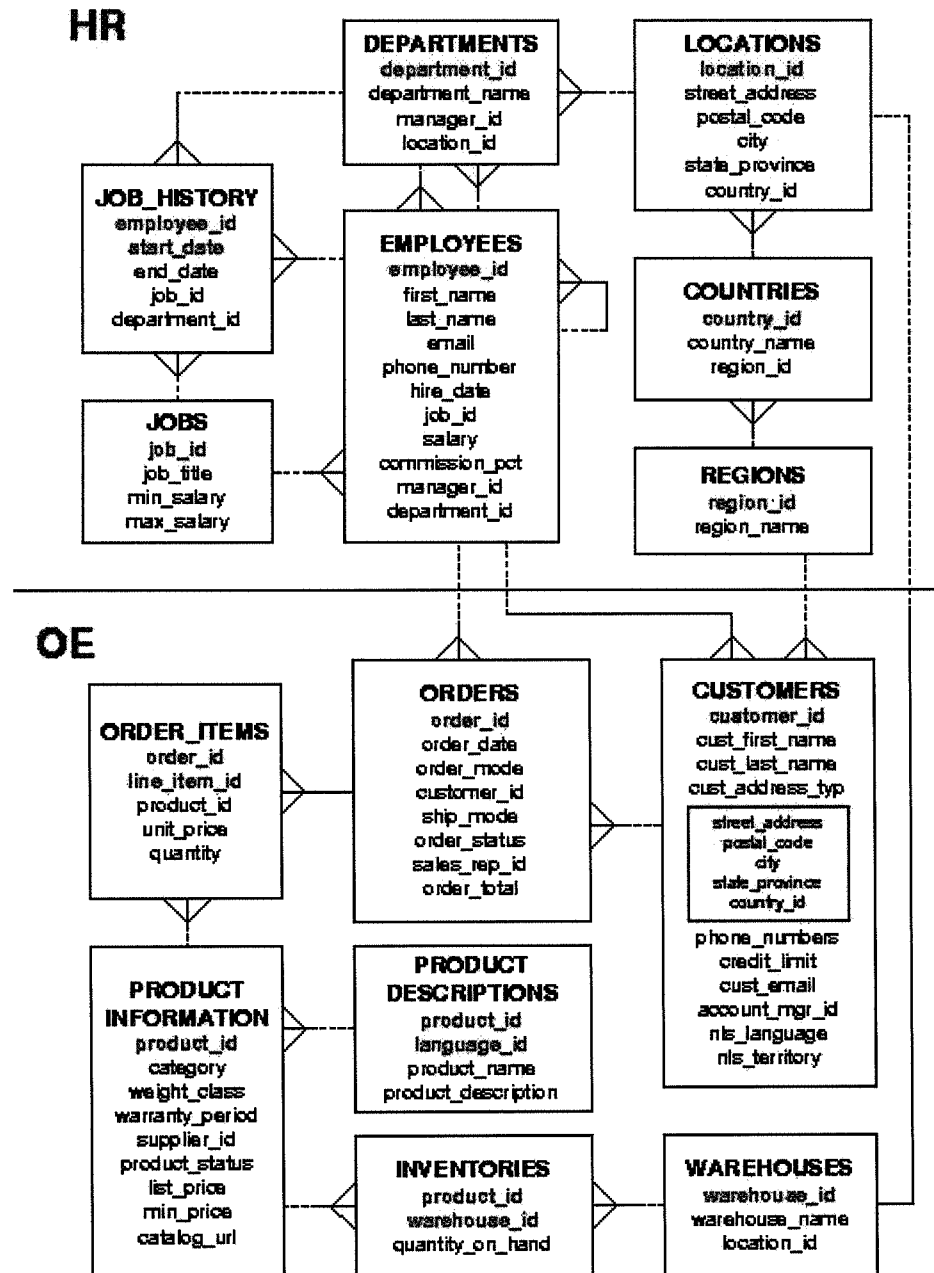
1. Design and implement an example market basket analysis model based on the OE database. You can follow closely the example provided in "A Practical Introduction to Data Mining". Your example should clearly identify the stages you have gone through. You should illustrate each stage with appropriate screen shots, and a corresponding brief description including issues you have encountered (for example - large numbers of products). Actually discovering knowledge in the OE database would be great, but is not really expected!

*Note - More comprehensive application of techniques as described in Berry, for example, would contribute to higher-level achievement of 2.1 and above.*

2. Critically evaluate the technologies, tools and methods relating to this theme as indicated in section 8.

You should upload the assessment activity (in the general content and format described in the assessment specification) by the end of week 21.

## HR and OE Data Model



**SOUTHAMPTON SOLENT UNIVERSITY**

**FACULTY OF TECHNOLOGY**

**Course Code: BSc Courses on the Computing and BIT Programmes**

**Unit Code: SWD301**

**UNIT: ADVANCED AND DISTRIBUTED DATABASES**

**Assessment No.: 1 (of 1)**

**Assessment Type: Project Report**

**Weighting %: 100%**

The current version of "The Students Guide to Assessment" should be consulted for the regulations that apply to all assessments. This guide was handed out at enrolment and is also available from the Faculty Office RM304.

**Hand-out Date to students: w/c 29<sup>th</sup> September 2009**

**Latest Hand-in Date (by student):**

Evaluation Report: 30<sup>th</sup> April 2010 (Week 25)

*Very Important Note - It is strongly recommended that you have uploaded each theme deliverable by the week indicated in the assessment activity and in myCourse*

[Note: Extensions to the published hand-in date will not be given, but refer to the current guidelines and form for extenuating circumstances, available from the Faculty Office RM304.]

**Planned Feedback Date (by Tutor and in myCourse):**

Evaluation Report w/c 31<sup>st</sup> May 2010

[Note: Normally, you will not receive your original work back after marking, therefore, always keep a copy of what you hand in. However, you should always receive feedback on your performance. The work will not be marked anonymously.]

**Learning Outcomes Assessed:**

[Note: Cross referenced with the Unit syllabus.]

*Knowledge and Understanding*

1. Discuss the tools, technologies, methods and techniques associated with business database application development.

*Cognitive Skills*

2. Analyse, design and evaluate elements of centralised, distributed and decision support database applications.

*Practical and Professional Skills*

3. Undertake research of advanced database technologies.

4. Apply tools for the development of elements of centralised, distributed and decision support database applications.

**Assessment Criteria:**

Refer to the separate assessment criteria descriptor sheet.

Assessment Prepared by: A.Monger	Signature:	Date: 11-Sep-09
Peer Partner: S.Baron	Signature:	Date: 11-Sep-09

## Introduction

You have been asked as the senior IT applications manager for the company to undertake a project to evaluate advanced database technologies, methods and tools that may benefit the company. The primary themes of interest to the company are:

- Exploiting DBMS Data Models and Server Functionality
- Accessing and Manipulating Data Programmatically
- Improving Data Access by Data Distribution and Replication
- Multi-Dimensional Modelling and Analysis for Decision Support
- Mining Databases for Decision Support

Fundamental to informing the resulting evaluation report is that you apply these technologies, methods and tools in the context of the scenario for each of these themes. You are also required to discuss, choose, research, apply and evaluate one of the advanced topics that extends from each theme.

This evaluation project comprises activities and corresponding deliverables/artefacts scheduled throughout the year as detailed below.

## Activities and Deliverables

### **1. Analyse, design, implement, test and evaluate as required for each theme (Weeks 2-21)**

There is an activity for each theme which comprises a scenario with corresponding requirements and deliverables. These activities are in the UIG and in myCourse.

You should upload to myCourse the deliverable(s) at the end of the final week of each theme (as scheduled in myCourse). This helps to assure the tutor that it is your own work and that you are completing the work on a regular basis. It also makes it easier for the tutor to provide formative feedback in the classroom on both an individual and group basis. The general content and format of the upload is detailed in the next section.

The version of any particular required deliverable (eg a design) that will be marked is the one that is handed-in, and so you can continue to improve these deliverables until then. Please note that myCourse will allow you to submit a later version after having previously uploaded an earlier version.

Although this is primarily an individual activity, there are some indicated activities and deliverables that can be completed in groups using a wiki. Please note the "Identification of Group Work" section below.

**It is strongly recommended that you have uploaded your final versions to myCourse by the end of each of theme**

Each theme includes an emerging or advanced topic that is not covered explicitly in the learning resources and core activities. This year, these topics are respectively for each theme:



- Using XQuery in Oracle
- Using Oracle APEX for Advanced Application Development
- Advanced Oracle Replication
- Using the Mondrian OLAP Tool
- Using the Weka Data Mining Tool

You will need to choose one of these topics for further work as described in 2. below. However, in order to inform this choice, and to provide a shareable base of knowledge and research, you should contribute to the forum set up in myCourse for each advanced topic. This includes contributing relevant information, initiating and contributing to discussion threads and summarising forum content.

Although it is preferable that the forum concludes by the end of the final week of each theme, the forums will remain open throughout the year. It is particularly important that initial practical issues (eg access to software tools) have been resolved before 2. below starts (ie before week 22).

Although contributions to the forums are not explicitly included in the assessment criteria, your contributions may be scrutinised in a borderline case for further evidence of having achieved the learning outcomes (particularly P&PS3), and of higher level achievement in P&PS3 (particularly 2.1 and above).

## **2. Research, apply and evaluate your chosen advanced topic (Weeks 22-24)**

1. Write a detailed and referenced explanation of the key elements (eg functionality of the technology), and discussion of key issues (eg usability), relating to your chosen advanced topic from 1.
2. Devise a scenario, linked to this research, that enables you to apply your chosen topic. This scenario should be written in a manner similar to any of the five theme scenarios, including requirements. This scenario must enable you to produce analysis, design, implementation and test artefacts, again in a manner similar to 1.
3. Analyse, design, implement and test as appropriate to your requirements, and upload these artefacts to myCourse.
4. Evaluate your chosen topic.

Provided all members have contributed meaningfully to the forum, you may work in groups using a wiki to devise a shared scenario, and corresponding analysis and design artefacts. However, the implementation and testing must be done individually. Please note the “Identification of Group Work” section below. Also, there is no myCourse upload for this deliverable.

Management is considering extending its use of Oracle as its main corporate database environment, and so this research and application should be undertaken in the context of Oracle wherever appropriate.

***Note - This activity is important as an indicator of higher-level achievement.***

## **3. Compile the evaluation report and hand-in (Week 25)**

The report (of maximum 2500 words) must be page numbered and include a contents page with the following sections:

1. Exploiting DBMS Data Models and Server Functionality

Page No.

\*

2. Accessing and Manipulating Data Programmatically \*
3. Improving Data Access by Data Distribution and Replication \*
4. Multi-Dimensional Modelling and Analysis for Decision Support \*
5. Mining Databases for Decision Support \*
6. Research, Application and Evaluation of << insert topic of chosen advanced topic here >> \*

There is no need for an introduction, conclusion or other sections. It should therefore take no more than an hour to compile the report from the final versions of the 5 myCourse uploads (assuming you have kept to the schedule!).

### General Content and Format of the 5 myCourse Uploads

Each of the 5 theme uploads must be in a .pdf or .doc file and should include:

1. Documented and described analysis, design and implementation artefacts that meet the scenario requirement.  
(Note - Do not submit zipped or other files for more comprehensive solutions - you can refer to these and retained as possible evidence of higher-level achievement).
2. Evidence of testing against the University's databases (including screen shots etc).
3. Accompanying explanation and commentary about the development as appropriate.  
(Note - Design models and decisions should always have an accompanying explanation, and code should always be documented).
4. The Evaluation of the Technologies, Tools and Methods for that theme (including any key issues and recommendations).
5. Academic and industrial references relating to the theme.

### Identification of Group Work

Where it is permissible to work in groups on an activity and deliverable, you must use a wiki and indicate clearly in your report and the wiki the other members of the group.

### Demonstration of Practical Work

You must be able to assure the tutors that the programs, forms and other application components are your own work. You may be asked to demonstrate them and you may be asked straightforward questions (as part of a viva/interview) about their development and implementation. You must be able to demonstrate them against the University's Oracle databases.

A link to this page will be published on the scheme of work if any demonstrations are required. The page will provide details about who, when and where in respect of demonstrations.

### Extenuating Circumstances

The University's Extenuating Circumstances procedures are in place if there are genuine circumstances that may have affected your academic performance. Remember however you need to be 'fit to study', this means that you can either submit your assessed work or declare extenuating circumstances, but you cannot do both.

A summary of guidance notes for students is given below:

<http://blade2-5.solent.ac.uk/DocMan8/rns?RNSEexact=ASQS/PPG/1234570925>

### Academic Misconduct

Any submissions must be your own work and, where facts or ideas have been used from other sources, these sources must be appropriately referenced. The University's Academic Handbook, includes the definitions of all practices that will be deemed to constitute academic misconduct. You should check this link before submitting your work.

Procedures relating to student academic misconduct are given below:

<http://blade2-5.solent.ac.uk/DocMan8/rns?RNSEexact=ASQS/PPG/1234570157>

### Ethics Policy

The work being carried out by the student must be in compliance with the Ethics Policy. Where there is an ethical issue, as specified within the Ethics Policy, then the student will need an ethics release or an ethical approval prior to the start of the project.

## Assessment Criteria and Feedback Sheet

Higher levels of achievement are described on the RHS of the grid. Each level subsumes the previous level. Each of the three criteria below contributes a third to the overall mark.

Primary, Analysis and Design, Learning Outcome CS1				
Does not reach required threshold.	Analysis and design artefacts meet most scenario requirements, but with some deficiencies.	Described analysis and design artefacts meet scenario requirements and using appropriate tools and methods.	Rationale for design decisions explained (inc. most higher-level requirements).	Comprehensive, fully justified and documented analysis and design artefacts meet all requirements in full.
Primary, Implementation and Learning Outcome CS2				
Does not reach required threshold.	Implementation artefacts meet most scenario requirements, but with some deficiencies.	Documented implementation and test artefacts meet scenario requirements.	Discussion of encountered implementation issues and solutions (inc. most higher-level requirements).	Comprehensive and fully documented implementation and test artefacts meet all requirements in full.
Primary, Evaluation Report and Learning Outcomes CS3, CS4 and CS5				
Does not reach required threshold.	Evaluation report meets most requirements but with some deficiencies. Evidence that relevant background reading has been undertaken.	Richer and referenced evaluation against some appropriate criteria. Some logical recommendations developed.	Well written evaluation report against full set of criteria and full set of recommendations developed. Most higher-level requirements addressed. References clearly underpin the evaluation throughout.	Comprehensive and critical evaluation report addressing all requirements, and based on extensive research.

**Notes :**

1. Normally, artefacts for all of the 5 themes must be submitted and reach the required threshold.
2. Each of the activities indicates work which could contribute to higher-level achievement.