UFCEKG-20-2
Data, Schemas & Applications

Lecture 1
Module Introduction & Outline
N. H. N. D. de Silva
(Slides adapted from Prakash Chatterjee, UWE)
Teaching and learning:

• 2 hour lecture
• 1 hour workshop sessions
• emphasis on both individual & working with others in workshop sessions
• Preparation will be required – readings and other material listed for each week (in the moodle page) should be done at home in preparation for the following week (3-4 hours home based study each week)

Assessment:

• Coursework - 1 piece – weighting 50% - will be an group & individual design and programming task. After hand-in, you will also be required to do a 15-20 minute presentation to your tutor illustrating and explaining your application.

• Examination – weighting 50% - 2 hours - compulsory part 40% - selective part 60%.
Basic Knowledge (prerequisites):

This module makes some assumptions about the student's understanding of computing, programming and web technology.

- **Basic programming control structures**: assignment statement; conditional expression (if then); loops; functions
- **Basic programming data structures**: integers; real numbers; strings; arrays and hashes (dictionaries); objects
- **Basic HTML mark-up**
- **Basic Browser usage**: URLs and their structure; HTTP
- **Basic Operating system terminology**: files; directories; paths; permissions
- **Basic Systems concepts**: data flow diagrams, entity-relationship diagrams

We aim to provide at least links to material in these areas so that you can brush up your understanding, but you will have difficulties on the module if you do not have a basic knowledge of these areas.
Course materials:

No specific set textbook is required – all course materials, lecture notes, workshop notes, readings and other resources are pointed to and made available via the moodle page.

However it is strongly recommend that you acquire one of the following texts for help with the assignment:


Jan-2013

N. H. N. D. de Silva
Techniques for knowledge acquisition:

• Develop a glossary of terms and their meanings, with examples of use
• Relate terms with a mind-map or data model
• Find illustrative examples
• Find good references and tutorials
• What else?

The same techniques apply to learning technology areas.
Data, Schemas and Applications:

- **Data** - raw data available in local databases and the global web

- **Schema** - the way data is organised and accessed - aka pattern; coding; format

- **Application** - the use that is made of that data and mechanisms for creating and maintaining data
Example:

Illustrating

Data on the web and how a standard schema allows data sharing and code reuse

Separating data from use allows multiple presentations of the same basic data

Different representations support different usage and the need to be able to transform from one to another
Example : Weather Data

Weather Data

**Data** : Data collected by weather stations is stored as a single record e.g. [Frenchay weather](#).

**Schema** : To use this data, you need to know how the data is structured (a single record with space separated values) and what each value represents. [Decoding guide](#)

**Applications** : This standard format is used all over the world by amateur weather observers so that standard software can provide a display. More significantly, data for around the world can be brought together into a single application: [Weather Underground](#)
Representations:

**Music**
We see here multiple representations of the 'same' thing which differ in what you can do with the representation. The name of the piece

**MP3** each item is a sampled sound amplitude, binary representation
Representations:

**Music**
We see here multiple representations of the 'same' thing which differ in what you can do with the representation. The name of the piece

**MIDI** tracks and notes, binary
Representations :

**Music**
We see here multiple representations of the 'same' thing which differ in what you can do with the representation. The name of the piece

*XML markup language* for a Score – readable
Representations:

**Music**
We see here multiple representations of the 'same' thing which differ in what you can do with the representation.
The name of the piece

**Sheet Music** textual/graphical music score
Text for track name, date, line-up, descriptions, reviews

`Für Elise in 6/8`

Jan-2013
N. H. N. D. de Silva
Transformations:

Some *transformations* between *representations* are harder than others. How would you rate the difficulty of these transformations?

- From a video to an MP3
- From a performance to an MP3
- From MIDI to MP3
- From MP3 to Score
- From a Score to a MIDI file
- From an MP3 to the name
The Web of Data:

Data is created on and off the web at an increasingly fast pace. It is estimated that as of March 2009 there were 25.21 billion pages indexed by Google. In July 25, 2008, Google software engineers Jesse Alpert and Nissan Hajaj announced that Google Search had discovered one trillion unique URLs.

• What data are out there?
• How can I understand the data?
• How can I use this data for new purposes?
• How can I create data which is usable on the web?
• How do I know which data to trust?
• What rights do I have to use the data?
Data everywhere:

- Beaches around the UK
- Ship movements in the English Channel
- Weather buoy network
- BBC News
- Twitter stream .. as RSS
Visualization as a transformation:

Transforming numbers into a visual representation helps viewers to see the deeper patterns and understand the meaning of the data.

- Search trends
- Olympic Medals
- Baby Names
- Google Fight
- FTSE 100

What data and technologies were used in these visualisations?
Yahoo Pipes:

Yahoo Pipes provides a visual programming environment to create applications which transform data into new representations.

Workshop 2 introduces this environment as a way to combine and search news feeds based on the feed format RSS, and XML standard.

In Pipes, simple transformations (like constructing a new value from existing data) are combined in a sequence of operations to create a new custom transformation (which can in turn be used in other pipes)
Questions? & Answers!