

## **Comp 336/436 - Markup Languages**

---

Fall Semester 2018 - Week 1

Dr Nick Hayward

## Course details

---

### Lecturer

- Name: Dr Nick Hayward
- Office: Doyle 307 (LSC)
- Office hours
  - *Wednesday afternoon by appointment (WTC)*
- Faculty Page

## Course schedule

---

### Important dates for this semester

- Project outline and mockup
  - *presentation & demo: 26th September 2018*
- DEV week: 24th to 31st October 2018
  - *presentation & demo: 31st October 2018 @ 4.15pm*
- Thanksgiving break: 21st to 25th November 2018
  - ***n.b.*** no formal class: 21st November 2018
- Final class: 5th December 2018
  - *presentation & demo: 5th December 2018 @ 4.15pm*
- Exam week: 10th December to 15th December 2018
  - *Final assessment due on 12th December 2018 by 4.15pm*

## Initial course plan - part I

---

(up to ~ DEV Week)

- intro and consideration of **markup languages**
- their structure, usage, and implementation
- general manipulation...
- metadata
- semantic structure and usage
- ...

## Initial course plan - part 2

---

(up to the end of the semester)

- processing and manipulation of markup languages
- project usage and examples
- parsing and rendering of markup languages
  - standard parsing options
  - custom parsers - a consideration of language and parsers...
- stronger semantic structure and organisation
- analysis, visualisation &c. of markup languages
- ...

## **Assignments and coursework**

---

### **Course will include**

- weekly bibliography and reading (where applicable)
- weekly notes, examples, extras...

### **Coursework will include**

- exercises and discussions (Total = 20%)
  - *various individual or group exercises and discussions*
- project outline & mockup (Total = 15%)
  - *brief group presentation of initial concept and mockup*
- DEV week assessment (Total = 25%)
  - *DEV week: 24th to 31st October 2018*
  - *demo due 31st October @ 4.15pm*
- end of semester final assessment (Total = 40%)
  - *demo due 5th December 2018 @ 4.15pm*
  - *report due 12th December 2018 @ 4.15pm*

## Exercises & discussions

---

### Course total = 20%

- exercises

- *help develop course project*
- *test course knowledge at each stage*
- *get feedback on project work*

- discussions

- *sample projects and markup usage*
- *design topics and usage examples*

- extras

- *code and application reviews*
- *various other assessments*
- *peer review of demos*

## **Development and project assessment**

---

**Course total = 80% (Parts 1, 2 and 3 combined)**

### **Initial overview**

- combination project work
  - *part 1 = project outline & mockup (15%)*
  - *part 2 = DEV Week development & demo (25%)*
  - *part 3 = final demo and report (40%)*
- group project
  - *groups will be organised at the start of the semester*
- design and develop a markup language based app
  - *markup language/s has to form part of the core structure*
  - *clear demonstration of processing and usage of markup languages*
  - *combination of structural and presentational usage of markup languages*
  - *demonstration of semantic structure and usage*
  - *use and integration of metadata*

## Project outline & mockup assessment

---

Course total = 15%

- begin outline and design of a markup based project
  - *built from scratch*
  - *HTML5, XML...*
  - *builds upon examples, technology outlined during first part of semester*
  - *purpose, scope &c. is group's choice*
  - *presentation should include mockup designs and concepts*

## Project mockup demo

---

Assessment will include the following:

- brief presentation or demonstration of current project work
  - *~ 5 to 10 minutes per group*
  - *analysis of work conducted so far*
  - *presentation and demonstration*
  - *outline current state of project concept and design*
  - *show prototypes and designs, where applicable*

## **DEV Week assessment**

---

**Course total = 25%**

- project outline and introduction
- developed using a chosen markup language
- consider and apply metadata schemes and semantic organisation for chosen domain
- current working examples - what does and does not work...
- demo and project report
  - *due on Wednesday 31st October 2018 @ 4.15pm*
- anonymous peer review
  - *similar to user comments and feedback*
  - *chance to respond to feedback before final project*

## DEV Week Demo

---

DEV week assessment will include the following:

- brief presentation or demonstration of current project work
  - *~ 5 to 10 minutes per group*
  - *analysis of work conducted so far*
  - *e.g. during semester & DEV week*
  - *presentation and demonstration*
  - *outline current state of application/project*
  - *show prototypes, designs, outlines &c.*
  - *explain what works & does not work*
    - i.e. outline what has been completed to date...
  - ...

## Final assessment

---

Course total = 40%

- working final app
- modify and update DEV week project work
  - *include parsing and rendering of the data suitable for broader publication*
  - *develop and integrate stronger semantic organisation and manipulation*
  - *provide an opportunity for analysis, visualisation...*
- presentation and demo - live working app
  - *due on Wednesday 5th December 2018 @ 4.15pm*
  - *show and explain implemented differences from DEV week project*
  - *where and why did you update the app?*
  - *benefits of updates?*
- how did you respond to peer review?
- final report
  - *due on Wednesday 12th December 2018 @ 4.15pm*

## Goals of the course

---

A guide to developing and publishing interactive client-side web applications and publications.

Course will provide

- examples of processing and manipulation of markup languages
- examples of structure and rendering of markup languages
- clear understanding of underlying technologies used to
  - *develop and publish markup languages based projects*
- provide a deeper understanding of example markup languages
- relation of markup languages to metadata and semantic organisation and usage
- opportunity to consider and develop project-based applications
- ...

## **Course resources - part I**

---

### **Website**

Course website is available at <https://csteach436.github.io>

- timetable
- course overview
- course blog
- weekly assignments & coursework
- bibliography
- links & resources
- notes & material

### **NO Sakai**

## **Course resources - part 2**

---

### **GitHub**

Course repositories available at <https://github.com/csteach436>

- weekly notes
- examples
- source code (where applicable)

### **Trello group**

Group for weekly assignments, DEV week posts, &c.

- Trello group - COMP 436
  - <https://trello.com/csteach436>

### **Slack group**

Group for class communication, weekly discussions, questions, &c.

- Slack group - COMP 436
  - <https://csteach436-2018.slack.com/>

## Group projects

---

- add project details to course's Trello group, *COMP 436 - Fall 2018 @ LUC*
  - Week 1 - Project Details
  - <https://trello.com/b/YOu07xPZ>
- create channels on Slack for group communication
- start working on an idea for your project
- start planning weekly development up to *Project Outline & Mockup*
  - demo on Wednesday 26th September 2018 @ 4.15pm

## Intro to markup languages - part I

---

- a simple consideration of a markup language
  - *highlight, draw attention to, content & parts of a given document*
- markup might provide
  - *instructions, comments, notes*
  - *or simply background information for content or data*
- encounter markup languages & rendered output on a daily basis
  - *access a web page, news feed, or other online source...*
- use of HTML - HyperText Markup Language
  - *simplicity, ease of development & usage*
  - *helped to drive ongoing development & popularity of web*
- HTML has faced shortcomings and limitations

## Intro to markup languages - part 2

---

- markup may also be construed in slightly different terms
  - *as we consider a given target audience*
- e.g. markup of a HTML document for a web browser
- XML, SVG, and many other examples
  - *may have different intended recipients and interpretations*
- HTML5 markup needs to be syntactically precise intended web browser
- XML - eXtensible Markup Language
  - *popular option for storing data*
  - *long time option for the transmission and sharing of data*
  - *now challenged by other options such as JSON*
- HTML was designed to display data - ie: how it looks
- XML was designed to carry data - ie: how it is stored
- presentational vs structural

## Markup languages - literary studies

---

- many uses of markup languages include structured data and metadata
  - *digitised artifacts and objects*
- consider *literary studies*
- assumption of automatic means for analysis and testing
- application of strict, detailed textual preparation if no automated process
- markup as an external intervention
- markup allows a critically interpretative process
- open criteria for encoding textual material

## Markup languages - literary studies

---

- Electronic Scholarly Editions
- TuStep
  - *example*
  - *project homepage*
- personalised markup
- modification of encodings
  - *to match software analysis requirements*

## Markup languages - TEI

---

- The Wonderful World of TEI...
  - read a potted introduction and history at the TEI website
    - <http://www.tei-c.org/About/history.xml>
- Malory Project - TEI XML based project
  - Caxton example - <http://www.maloryproject.com/xml/caxton/Caxton.xml>

## Digitisation - intro

---

- crossroads of research and development
- consider options for digitisation
  - e.g. *current analogue record*
  - *historical, literary*
  - *structural*
  - ...
- inherent benefits of digitisation
  - *accessibility*
  - *flexibility*
  - *ease of manipulation*
  - *aggregation*
  - *storage....*

## **Digitisation - examples (good and bad)**

---

- many fine examples now exist of digital archives, projects, editions...
  - *Papers Past*
  - <https://paperspast.natlib.govt.nz/>
  - *Library of Congress*
  - <https://www.loc.gov/>
  - *British Library*
  - <https://www.bl.uk/>
- Google Books project
  - *interesting initial 'History' of the project*
- EEBO, ECCO, NCCO...
- others such as Internet Archive
  - <https://archive.org/>
- EEBO-TCP
  - *intro*
  - <http://www.textcreationpartnership.org/tcp-ebo/>

## Digitisation - why do we bother?

---

### benefits and costs

- loss of analogue information to precise digitisation

*"We should be cautious about letting the radiance of the bright future blind us to the limitations of this new technology" (Smith, Abby. 1999. "Why Digitize?")*

## **Digitisation - why do we bother?**

---

*"analogue information can range from the subtle tone and gradations of the chiaroscuro in a Berenice (bernice) Abbott photograph of Manhattan in the early morning light, to the changes in the volume, tone, and pitch recorded on a tape that might, when played back on equipment, turn out to be the basement tapes of Bob Dylan."* (Smith, Abby. 1999. "Why Digitize?")

## Image - Berenice Abbott, New York

---



Berenice Abbott, Pike and Henry Street, New York, 1936

- Source - Lumiere
- <http://lumieregallery.net/wp/167/berenice-abbott/>

## **Digitisation - potential for loss**

---

- dependent upon various factors
  - e.g. *how much information you actually gather*
- whilst digitising we may consider
  - *density of data*
  - *frequency of sampling*
- breadth or depth of information gathered is also a consideration

## Digitisation - analogue vs digital

---

- ongoing saga of 'Analogue vs Digital'
- consideration of cost, time, transmission and storage
- loss of information from analogue to digital
- benefits of digital
- perceived benefits of analogue

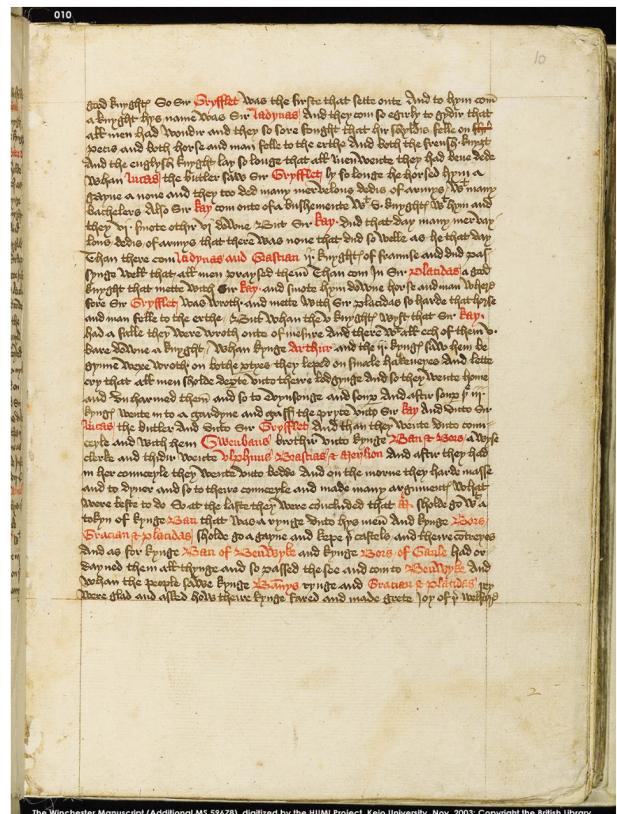
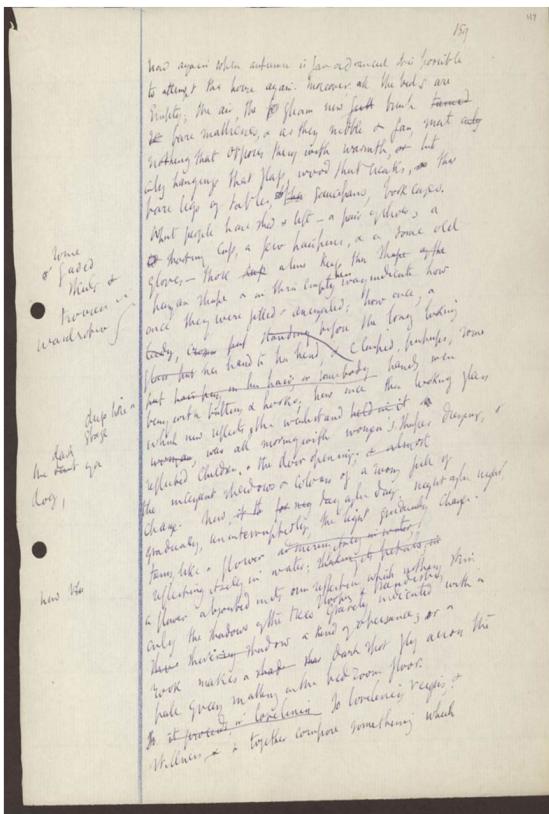
*"...reducing it to 0s and 1s is not unlike a root canal: by extracting the nerves, the tooth is killed in order to save it." (Unknown. 2004. "The Great Analog Versus Digital Debate". VoicePrint Online)*

## Digitisation - textual considerations

---

- comparable concerns with music in textual digitisation
- *density of data* is still a concern with text
- should we encode texts with style, punctuation, notes...
- how much do we encode, in particular with manuscripts
- considerations of discipline, project, and personal preferences
- variation in requirements dependent upon manuscript
  - e.g. *modern vs manuscript*

# Image - Modern vs Medieval



## Modern and Medieval texts

### Source

- Left = Page 159 of *To the Lighthouse* by Virginia Woolf
- Woolf Online
- Right = Page 10 of the Winchester manuscript of Thomas Malory's 'Morte Darthur'
- Malory Project

## References

---

- British Library
- Library of Congress
- Lumiere - Berenice Abbot
- Malory Project
- Papers Past
- TEI - History
- Woolf Online