Digital Humanities

Digital Humanities for Medieval Philosophical Sources

0. Presentation

conf. dr. Mihai Maga

Babeş–Bolyai University, Cluj–Napoca
Master in Ancient and Medieval Philosophy

2nd semester, 2023–2024
HME2415

https://www.dhcluj.ro/dhm/

1. Course 1: Introduction to Digital Humanities
2. Course 2: Semantic encoding
3. Course 3: Critical editing
4. Course 4: Principles of TEI-XML
5. Course 5: Representation of manuscripts in TEI
6. Course 6: Representation of textual variation in TEI
7. Course 7: Representation of source apparatus in TEI
8. Course 8: Visualisation of digital editions
9. Course 9: Integration and digital processing
10. Course 10: Artificial intelligence

1. What does Digital Humanities mean?
   - a technique
   - a philosophy
2. History of Digital Humanities
3. Domains of Digital Humanities

1. Visual representation vs. semantic representation
2. The semantic paradigm in the digital world
3. Utility of semantic encoding: integration, interfaces, data processing

1. Methods of approach for critical editions
   - Best manuscript method
   - Eclectic method
   - Stemmatic (lachmannian) method
   - Unoriented (material) method
2. Types of critical editions
   - Facsimile type edition
   - Eclectic edition
   - Literary (critical) edition
   - Diplomatic edition
   - Material edition
3. Elements of a critical edition:
   Introduction, Text, Critical apparatus

1. The XML format
   - XML syntax: tag, attribute, text, declaration, comment
2. About TEI[1]
   - The TEI Consortium P5 Guidelines and the document structure
     - http://tei-c.org/

XML

```
<?xml version="1.0"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
    <text>
        <p xml:id="par01" lang="la">Hic est textus editionis.</p>
    </text>
</TEI>
```

[1] Text Encoding Initiative

1. Indication and description of manuscripts
   - manuscript identifier: place, libray, shelf number
   - description parts: physical, historical, contents, bibliography
2. Instruction set for describing manuscripts

<msDesc>	manuscript description block
<msIdentifier>	identifier: <settlement>, <repository>, <idno>...
<head>	heading
<msContents>	manuscript contents block: <msItem>, <locus>, <title>...
<physDesc>	physical description: <objectDesc>, <supportDesc>, <dimensions>, <layoutDesc>...
<history>	manuscript history: <origin>, <provenance>, <acquisition>...
<additional>	additional information
<msPart>	information about a section of the manuscript

1. Differences between manuscript copies
   - variants of redaction
   - copyist errors
2. Describing textual variations in TEI:

   <app> apparatus
   <lem> lemma
   <rdg> reading
3. Encoding methods:

   LRM Location-referenced Method
   DEPAM Double End-Point Attachment Method
   PSM Parallel Segmentation Method
4. Statistics and counters

1. Ontologies and trees
   - tree-structured data
   - objects, classes ontologies
2. Describing the elements of a textual reference
   - references to author, title, work, section...: <title>, <author>...
   - bibliographies: <bibl>...
   - other types of references
3. Attaching the references
   - inline
   - using pointers

1. Conversion to classic format
   - typesetting rules for editions
   - specialized DTP: LaTeX
2. Interactive interfaces
   - web interfaces: client-side, server-side
   - native applications
3. Inclusion of manuscript images
   - TEI instructions set: <facsimile>, <surface>, <zone>, <graphic>
   - image attachment: inline, pointer
4. Pitfalls of the visual

1. Indices and concordance tables
2. Query languages: `XQuery`, `XPath`
3. Search in text: simple, wildcard, stemmatized, lemmatized
4. Lemmatization, normalization, dictionaries
5. Digital corpora
6. Data-mining

1. What is artificial intelligence (AI) and machine learning (ML)?
2. Algorithms and models in ML
3. Training and prediction in ML
4. Classification of approach types in ML