Scholarly reasoning and writing in a digital library

Knuth: "literate programming"

Programmed literature?

You can't make this stuff up

AHA recommends 6-year embargo on digital dissertations

Typifies instrumentalist thinking

Digital humanists also susceptible?

Abstract model or specification

A broad example: Particularity of humanistic scholarship implies primacy of citation

A machine-readable notation: the CTS URN
discrete objects: the CITE Object URN
extended citation: image extension to CITE URN

Examples

A CITE URN:

urn:cite:hmt:vaimg.VA012RN-0013

A CTS URN:

urn:cts:greekLit:tlg0012.tlg001.msA:1.1

A more detailed model

OHCO2 model of citable text as a set of citable nodes that...

have an order
belong to a passage hierarchy
belong to an ontological work hierarchy
may include rich content

OHCO2 implementations

tree (XML document)
tabular structure
graph (RDF)

Choose best implementation for task

Applications in HMT project

Texts edited in XML (Other featured recorded in CITE Collections)
Tokenized analysis from tabular form of text
Integration with non-textual material in RDF graph

Single-task builds

http://www.homermultitext.org/nightlies/:
zip archives of all tabular and XML data
graph of ~800K RDF triples

So how should we write?

Writing is not editing
Minimal structure allows light-weight solutions

Markdown

Good solution apart from absence of citation
So extend to create citedown

Syntax: replace URL citation in square brackets with URN citation in curly brackets

Some markdown tools

- HTML, pdf generators
- e-book generators (lean pub)
- Slide-show generators

As an example: write a pure citedown document

A second extension to markdown?

In addition to URL vs. URN references, citation vs. quotation already present in image quotation (embedding)
extend to texts and data objects?

New standards for writing

Draft standards and research resources on material should observe unifying standards

A distinctly digital wikicite?

Could we exploit digital technologies to

reverse Wikipedia editorial principle: no secondary sources
open access to anyone who can directly cite appropriate evidence

Thank you

Neel Smith
nsmith@holycross.edu
nsmith@holycross.edu