Digital Encoding in an Analog World: A Cautionary Tale

Prof. Jeff Sonstein  jeffs@it.rit.edu
Director, MS-IT Program
Department of Information Technology
Rochester Institute of Technology
Introduction

• **Basic Issues:** Some Definitions & Background

• **Questions:**
  
  Why is Digital Encoding Becoming the Norm?

  Can We Create New Meaning Through Digital Encoding?

• **Scenarios:** Blake & Tibetan Buddhist Sutras Use-Cases

• **Doing the Work:** the Model-View-Controller paradigm
Basic Issues

• Why is it that the digital encoding of data and information about analog realities has become the norm in academia and elsewhere?

• Are there dangers as well as advantages to this trend?
Basic Issues

1. Why use digital encoding?

“to facilitate the sharing of structured data across different information systems, particularly via the Internet. It is used both to encode documents and serialize data.”

Always remember that some analog reality is being encoded or modeled digitally.
2. What is analog reality?

“An analog or analogue signal is any time continuous signal where some time varying feature of the signal is a representation of some other time varying quantity. It differs from a digital signal in that small fluctuations in the signal are meaningful.”

[http://en.wikipedia.org/wiki/Analog_signal]

Remember, reality is continuous not discrete.
Basic Issues

Digital encodings are *discrete* measures, while analog reality is *continuous*.

Yes, this has implications for your work.
Basic Issues

3. What is data?

“data is a measurement that can be disorganized [...] when the data becomes organized it becomes information.”
4. What is information?

“Information is the result of processing, gathering, manipulating and organizing data in a way that adds to the knowledge of the receiver.”
Basic Issues

5. And what are texts?

What do you mean by “public texts”?
Will the following definition do:

“consisting solely of characters from a recognized character set”

No? Then what will?

[ Pause to argue here ]
Questions

Why is digital encoding becoming the norm?

- Nice Answer: Digital encoding schemes can facilitate:

  access to, analysis of, and the creation of commentaries about the analog realities within which human beings exist
Questions

Why is digital encoding becoming the norm?

- **Not-So-Nice Answer:** Digital encoding schemes can be made to produce **quantitative data** (for papers etc.), meaningful or not... but *do not forget* that this is a model or “map” of reality, this is *not* reality itself...

Do not confuse the map with the territory, the model with the thing being modeled.
Remember, this is not a pipe
Scenario I: Blake

- multimedia/multimodal text(s) and linked artifacts:
  Images, Characters & Meaning

  The William Blake Archive versus Blake at Project Gutenberg
Scenario I: Blake

The William Blake Archive
INTRODUCTION

Piping down the valleys wild,
Piping songs of pleasant glee,
On a cloud I saw a child,
And he laughing said to me:

‘Pipe a song about a Lamb!’
So I piped with merry cheer.
‘Piper, pipe that song again.’
So I piped: he wept to hear.

‘Drop thy pipe, thy happy pipe;
Sing thy songs of happy cheer!’
So I sung the same again,
While he wept with joy to hear.

‘Piper, sit thee down and write
In a book, that all may read.’
So he vanished from my sight;
And I plucked a hollow reed,

And I made a rural pen,
And I stained the water clear,
And I wrote my happy songs
Every child may joy to hear.
Scenario I: Blake

Introducing
Piping down the valleys wild,
Piping songs of pleasant glee,
On a cloud I saw a child,
And he laughing said to me:
Pipe a song about a Lamb;
So I piped with merry cheer,
Piper pipe that song again;
So I piped, he wept to hear.
Drop thy pipe, thy happy pipe,
Sing thy songs all happy cheer,
So I sung the same again,
While he wept with joy to hear.
Pipe sit thee down and write
In a book that all may read;
So he vanished from my sight,
And I plucked a hollow reed,
And I made a rural pen,
And I stained the water clear,
And I wrote my happy songs:
Every child may joy to hear.

Are These the Same Things?

[ Pause to argue here ]
Question: Can We Create New Meaning Through Digital Encoding?

1. yes, through developing structure & constraints
2. yes, through embedding data & commentary
3. we are condemned to create new meaning through recoding, whether we wish to or not
Creating New Meaning

1. through imposing structure & constraints
   • we impose a tree structure on the data (naming & positioning elements & attributes)... is everything in (textual) reality really a tree structure?
   • we specify content constraints for the data (delimiting what can go where)... do we always think of everything ahead of time?
Creating New Meaning

2. through embedding data & commentary
   • we embed attributes & stand-alone tags for comment and to add data (browsers can ignore them)
   • we embed hyperlinks & references for further information (scholars can follow them)
Creating New Meaning

3. we are condemned to create new meaning through the recoding process, whether we wish to or not...

choices made in encoding, structuring & constraining data inevitably will have effects upon data, perceptions & received meanings
Scenarios II: Sutras

• multimedia/multimodal texts and linked artifacts:

  commentaries, thoughts & root texts

  Tibetan Buddhist Sutras
Scenarios II: Sutras

[...] one of the sections of the “Medicine Buddha” meditation (which is used for healing) requires special hand movements (mudras), performed while imagining a dark blue Buddha figure as one chants the mantra of Medicine Buddha. [...] Instructions on how to perform the mudras and how to chant the mantra are not in the Tibetan text itself; the image of the blue Buddha is not in the text either. Traditionally, the mudras and chant melody must be learned from a Tibetan teacher, and the visualization image is usually memorized by focusing on a thangka painting in a shrine room. These additional dimensions are not simply contextual information necessary to understand the text being encoded: in an important sense, they are integral parts of the text and its meaning, without which it cannot be said to be truly preserved.

[from:“Encoding for Endangered Tibetan Texts”, Linda E. Patrik <patrikl@union.edu>, Department of Philosophy, Union College]
Scenarios II: Sutras

How do we adequately represent in digital form the non-character dimensions of analog realities which interpenetrate texts?

[ Pause to argue here ]
Doing the Work

• the Model-View-Controller paradigm

“Model-view-controller (MVC) is an architectural pattern used in software engineering. In complex computer applications that present a large amount of data to the user, a developer often wishes to separate data (model) and user interface (view) concerns, so that changes to the user interface will not affect data handling, and that the data can be reorganized without changing the user interface. The model-view-controller solves this problem by decoupling data access and business logic from data presentation and user interaction, by introducing an intermediate component: the controller.”

[from: http://en.wikipedia.org/wiki/Model-view-controller ]
Doing the Work

- using the M-V-C paradigm

  model of the data

  “The domain-specific representation of the information[...] Domain logic adds meaning to raw data”

  [from: http://en.wikipedia.org/wiki/Model-view-controller ]

XML uses a tree-structure... is this always appropriate to all data?
Doing the Work

Encoding a CD Catalog

```xml
<?xml version="1.0" encoding="ISO-8859-1"?>
<?xml-stylesheet type="text/css" href="cd_catalog.css"?>
<catalog>
  <cd>
    <title>Empire Burlesque</title>
    <artist>Bob Dylan</artist>
    <country>USA</country>
    <company>Columbia</company>
    <price>10.90</price>
    <year>1985</year>
  </cd>
  <cd>
    <title>Hide your heart</title>
    <artist>Bonnie Tyler</artist>
    <country>UK</country>
    <company>CBS Records</company>
    <price>9.90</price>
    <year>1988</year>
  </cd>
</catalog>
```
Doing the Work

• using the M-V-C paradigm

view(s) of the data

“Renders the model into a form suitable for [use...] Multiple views can exist for a single model for different purposes.”

[from: http://en.wikipedia.org/wiki/Model-view-controller ]

Many browsers now have native support for XML technologies, and cross-platform Javascript implementations of supporting technologies are freely available
Doing the Work: XML

Styling a CD Catalog

```xml
<?xml version="1.0" encoding="ISO-8859-1"?>
<?xml-stylesheet type="text/css" href="cd_catalog.css"?>
<catalog>
  <cd>
    <title>Empire Burlesque</title>
    <artist>Bob Dylan</artist>
    <country>USA</country>
    <company>Columbia</company>
    <price>10.90</price>
    <year>1985</year>
  </cd>
  <cd>
    <title>Hide your heart</title>
    <artist>Bonnie Tyler</artist>
    <country>UK</country>
    <company>CBS Records</company>
    <price>9.90</price>
    <year>1988</year>
  </cd>
</catalog>
```
Doing the Work: CSS

catalog {
    background-color: #ffffff;
    width: 100%;
}

cd {
    display: block;
    margin-bottom: 30pt;
    margin-left: 0;
}

title {
    color: #FF0000;
    font-size: 20pt;
}

artist {
    color: #0000FF;
    font-size: 20pt;
}

country, price, year, company {
    display: block;
    color: #000000;
    margin-left: 20pt;
}
Doing the Work

• using the M-V-C paradigm

controller (code)

“Processes and responds to events, [...] and may invoke changes on the model.”

[from: http://en.wikipedia.org/wiki/Model-view-controller ]
Doing the Work

• the M-V-C paradigm allows us to break up complexity into more easily managed “chunks”, separating the domain logic (model) from the presentation (view), and both from the code mediating between them (controller).
Doing the Work

• a familiar example of the M-V-C paradigm:
  ‣ (x)HTML for a tree-structured **model** of the data
  ‣ CSS & XSL to construct **view(s)** of the data
  ‣ Transformations, XPointer & XLink, and programming languages to provide **controller** code mediating between model(s) & view(s) [note: Google ajaxslt project]
Summary & Conclusions

1. Digital encoding schemes can facilitate access to, analysis of, and the creation of commentaries about the analog realities within which human beings exist, but they also make it easy to confuse the map with the territory.
2. Choices made while implementing digital encoding schemes will *inevitably* result in new (and possibly unintended) meanings being attached to the original data.
Summary & Conclusions

3. The Model-View-Controller paradigm can be a useful way to think about digital encoding processes.
Q & A Time

[ Pursue random associations here ]