2. How to Think About Information

INFO 202 - 29 August 2007

Bob Glushko
Plan for INFO 202 Lecture #2

Defining Information

Models and Modeling

"Information Architecture"
What is Information?

Many different ways to define it; depends on your discipline or perspective or problem

A common sense is "something you don't know" or "news about something"

But news to one person might be "old news" to someone who already knows it

So if something happens that we expected, we get less information than if something unexpected happens
Svenonius

"Something retrieved or obtained through informing"

"The content of a message"
Information is "the pattern of organization of matter and energy"

So this exists independently of any observer or recipient, but each observer can construct, store, and act upon it in different subjective ways.

Environmental and evolutionary factors influence this so there are similarities in how information is experienced.

Anything human beings interact with or observe can be a source of information.

Living beings can assign meaning to information, but patterns of organization of matter and energy are not inherently meaningful.
Bates' Fundamental Forms of Information

- Natural
  - Represented
  - (no label)
- Embodied
  - Experienced
  - Enacted
  - Expressed
- Encoded
  - (no label)
  - Exosomatic
  - Embedded
  - Recorded
Taylor's Information Hierarchy

Taylor presents an information hierarchy of DATA, INFORMATION, KNOWLEDGE, and WISDOM as "different levels of comprehension of symbols"

- DATA
- INFORMATION
- KNOWLEDGE
- WISDOM
Information Hierarchy Process Model (Boisot & Canals, 2004)
Information Hierarchy Example

DATA - the value we collect from a measuring instrument is 25

INFORMATION - the instrument is a refractometer, which measures the sugar content of grapes on the Brix scale.

KNOWLEDGE - the sugar content of grapes predicts the potential alcohol content of wine fermented from them

WISDOM - If the summer weather is cooler than usual, grapes will contain more acid so growers usually pick later with a higher Brix to help balance the tart flavors from the acid

Is the Encyclopedia Britannica wisdom? Is Wikipedia? What about this encyclopedia?
Reddy -- The Conduit Metaphor [1]
The Conduit Metaphor [2]

Language functions like a conduit, transferring thoughts from one person to another.

In writing and speaking, people insert their thoughts or feelings in the words.

Words accomplish the transfer by containing the thoughts or feelings and conveying them to others.

In listening or reading, people extract the thoughts and feelings once again from the words.
Conduit Metaphor - Examples

It's hard to get that idea across to him.
I gave you that idea.
His words carry little meaning.
It's difficult to put my ideas into words.
That's not what I got out of what he said.
A definition of information that involves neither meaning nor people was proposed by Claude Shannon, considering information as the amount of bits that a communication channel can carry.

Shannon's Information Theory treats communication as producing the same pattern of bits at the destination as that sent from the source.

This makes communication a statistical problem, and communication failures result from "noise" on the channel that distorts the message.

The amount of information in a message is computed in terms of the probability that its (bit) parts occur.
Shannon - Over Space and Over Time
Bits vs. Atoms

Another way to think about information that is a descendant of Shannon is in the contrast between "bits" and "atoms" first expressed by Nicholas Negroponte of the MIT Media Lab.

Information encoded as bits can move several orders of magnitude faster than atoms can.

It can be in many places at once (broadcasting, networking) unlike atoms that have to be in one place.
Digital Information

Many distinctions between "types of information" disappear when we look past their tangible renditions or representations.

Rather than the appearance, the model that describes the structure and content of the information becomes paramount.

But another twist for multimedia information: most techniques for "interpreting" multimedia don't capture or process it at a semantic level.
Information Quality

Do you know the source of the information?

Is the source authentic?

Is the information authoritative or credible?
“I just feel fortunate to live in a world with so much disinformation at my fingertips.”
Instances and Classes / Categories

We can treat information as a description of, or as being attached to an INSTANCE of something, a particular realization or implementation or occurrence.

Or, we can treat information as a description of a CLASS or CATEGORY of things that we are treating as equivalent.

It is important to be precise about which descriptive perspective you're taking.
Information Content vs Container

We say "the document is about … the photograph is about… the movie is about"

We're expressing a distinction between information as conceptual or as content: and the physical container or medium, format, or technology in which the information is conveyed

But Svenonius defines "document" as "an information-bearing message in recorded form" -- which includes the container
Another Kind of "Message Container"
Abstracting Away the Container

It is very useful to think abstractly about "information content" or "information flows" without making any assumptions or statements about the "presentation" or "rendition" or "implementation".

And we can apply abstract patterns of information exchange to new domains or technologies.

Abstract patterns of information content or structure are sometimes called architectures.
Not All Containers Are Equally Separable, Valuable or Usable

Not all information content can be separated from its container (sometimes the medium is the message)

But it is important to think of the information content abstractly if you can because that's the key to representing the same information in multiple formats, media, or technologies

Some information formats or representations are inherently more reusable or adaptable than others
"Information IQ"

- Printed text
- Unstructured electronic text: ASCII
- Formatted electronic text: HTML, EDI, word processing files
- Content/structure-based text objects: XML, SGML, databases

More "processability" / reusability

Easier to translate to
Models and Modeling [1]

Separating information content from the structure and presentation in which it occurs in "artifacts" is a core skill in all the overlapping disciplines in IO.

This is a key activity in what is called {data, document, information, conceptual} MODELING.

Models are simplified descriptions of a subject that abstract from its complexity to emphasize some features or characteristics while intentionally de-emphasizing others.
A model can represent a human activity, a natural system, or a designed system.

We can model structures – objects, their characteristics, their static relationships with each other like hierarchy, and reference.

We can model functions, processes, behaviors – dynamic activities that create and affect structures.
An Everyday Model: The Recipe

A recipe describes both objects and structures (ingredients) and the processes (instructions) for creating a food dish.

You can communicate a recipe to someone else who can then create the same dish.

You can use the recipe as a guide for experimentation with the objects or the processes in the recipe to create alternative dishes.

A set of recipes may contain common or re-usable objects (e.g. flaky pastry) or processes (stir-fry).
Not A Model of Wonton Soup
A Model of Wonton Soup

**Wontons**
- ½ lb. ground chicken, veal or lean pork
- 2 green onions, finely chopped
- 2 tsp. grated fresh ginger root
- 1 tsp. cornstarch (corn flour)
- 1 tsp. lite soy sauce
- 6 drops hot pepper sauce
- 18 wonton wrappers

**Soup**
- 4 cups chicken broth
- 2 tsp. lite soy sauce
- 1 tsp. sesame oil
- ¼ tsp. hot chili paste
- Salt and freshly ground black pepper
- 1 cup shredded fresh spinach or watercress
- 1 green onion, thinly sliced

In bowl, combine ground chicken, green onions, ginger root, cornstarch, soy sauce and hot pepper sauce; mix until well blended. Lay wonton wrappers on work surface. Place spoonful of ground chicken mixture in center of each one.

Moisten edges with water. Fold each wrapper in half to form triangle; press edges to seal. Arrange in steamer basket. Cover and steam over boiling water for 3 minutes.

Meanwhile, in large saucepan, combine broth, soy sauce, sesame oil, and hot chili paste. Bring to boil. Season to taste with salt and pepper.

Just before serving, stir in spinach until wilted. Place 3 wontons in each bowl. Ladle soup over top. Garnish with green onion rings.
The Classical Modeling Approach [1]
The Classical Modeling Approach [2]

We start with the EXTERNAL view of actual instances, observations, or artifacts.

We then develop a PHYSICAL model that describes the structures and characteristics of the artifacts as they exist.

From this PHYSICAL model we create a CONCEPTUAL model that abstracts an implementation-independent view to capture the more fundamental structures, characteristics, or relationships on which the artifacts are based.

From this CONCEPTUAL model we can create PHYSICAL models that re-implement the (presumably more rational, robust) model.
There is an essential difference or gap between the domain being modeled and the model.

So much of the skill of modeling is knowing what to ignore or throw away when you study something.

So whenever we try to describe anything, whether it is an instance or a class of instances that we are treating as equivalent, we are making choices about which "information" to represent.

Once we make these choices they shape or constrain how we think about what we've described.
"Information Architecture" [1]

An architecture describes a system's components (or "building blocks") and their relationships with each other, using hierarchical and compositional structure to define the component boundaries.

This gives an "architected" system an explicit model, in contrast with systems that are implemented incrementally without a master plan or without the effort to create reusable abstractions and components.

IA = (((content + information structure) + navigation structure) + presentation structure) + presentation design
Information Architecture [2]
Readings for INFO 202 Lecture #3

Weinberger, Chapters 7, 8, & 9

Svenonius, Chapter 2