What is Data?
MAT 259 Visualizing Information
(Working Notes – different from PPT – 1.25.2007)

Lecture 2: Systems of Classification

Presentation topics
. The collection of Data
. The organization of data
. The processing of data

Data
. Data is the product of research, collection, discovery, creation
. Data in itself is not useful
. Classification of data is closely related to classification of knowledge
. Its value is a result of how it is organized, transformed, and presented to give it meaning
. Context & relationships determines meaning for data

Some Datatypes (Spence)
. Numeric
. Ordinal: follow a particular order (days of the week)
. Categorical: Names of animals

Attribute Quality (Ware)
. Nominal (bus number, fruits into apples, etc.)
. Ordinal: Sequential
. Interval: Possible to measure drelative difference (train sched)
. Ratio: Object a is twice b

Organization Methods (Shedroff)
. Alphabet
. Spatial
. Time based, etc.

Attribute Dimensions
. Scalar: weight of a person
. Vector quality: Direction of a traveler
. Tensors: higher order: direction and force (stress)

Quantitative Data (with Attributes)
. Univariate data (cars, cost)
. Bivariate data
. Trivariate data
.

Metadata
. Data about data: Information that describes another set of data
. Examples: Library catalog card, address book, etc.
. Metadata is what allows the organization, storage, retrieval of data

Lo & Hi Metadata
. Example LO: Resolution, compression in a digital video file
. Example HI: Describes the structure of a media composition, ultimately its semantics
. Transition from metadata as tool to cultural form through semantic description

Systems of Classification
. We make sense of the world through organization
. We organize according to rules, systems (Linnaeus)
. but also according to experience (associative)

From Data to Information
. 1st step: explore its organization
. Organization affects interpretation and understanding
. Each organization of the same data set expresses different attributes and messages

Database/Data Structures
. Organized collection of data
. A collection of records stored in a systematic way
. Each record, a set of data elements, (basic unit of data such as name, street address, city, zip)
. Retrieval through any of the data elements
. Relational model: all data represented as mathematical relationships

Structures
. Mosaic display (eyes, hair color)
. Network Data (nodes)
. Hierarchical Tree Structure

Network Model
. Each record can have multiple parents and child records
. Organized in lattice structure consisting of links and nodes
. Lends easily to spatial visualization
. Example: Kohonen SOM map

Hierarchical Tree Structure Model
. Frequently hierarchical in structure, requiring parent/child relationship definitions
. Organization of computer hard drive
. Internet, WWW
. Dewey Decimal System
. Cladistics: evolutionary relationships (cladograms)

Standards
. **Library of Congress**: http://www.loc.gov/standards/
. **Dublin Core:** a metadata standard for describing digital objects (including webpages) to enhance visibility, accessibility and interoperability, often encoded in XML
. **Harmony Project**: research methods and models for

describing the variety of rich content
http://metadata.net/harmony/

Explorative & Innovation
. To see same data sets in different organizations reveals unexpected patterns in the relationship of things
. To invent new forms of organization based on personal, idiosynchratic rules enhances novelty of experience
. Nonetheless there needs to be some cultural common ground

Arts Examples
. **Associative**: Lisa Jevbratt's 1:1
. **Biographical**: Daniel Spoerri's Anecdoted Topography of Chance
. **Affect**: Melanie Wein's http://www.the-fleetingness-of-bits.de/

Media Based Exploration Areas (Manovich)
. Large **Scale** data sets
. New **Structure**: data with high-metadata content
. New **Interface**: Navigation & efficient access
. New **Image**: New forms of visualization
. Experimentation: Trial & error process

References
. Savage Mind (Science of the Concrete), Levi-Strauss
. http://time.arts.ucla.edu/AI_Society/manovich.html
. http://time.arts.ucla.edu/AI_Society/daniel.html
. "Metadating the Image" Manovich
. Library of Congress, Marc System: http://www.loc.gov/marc/