

GIS-T Symposium 2003

Colorado Springs, CO

A thick, horizontal yellow brushstroke with a textured, painterly appearance, spanning across the width of the slide below the location text.

More Data Won't Make You Smarter

David Fletcher

Geographic Paradigm Computing, Inc.

Albuquerque, NM

What is this?



- ⌘ I have a box
- ⌘ The box is 3' wide, 3' deep and 6' high
- ⌘ The box is very heavy
- ⌘ The box has a door on the front of it
- ⌘ When you open the door, a light comes on
- ⌘ It is colder inside the box than outside
- ⌘ You usually find the box in the kitchen
- ⌘ There is a smaller box inside with ice in it
- ⌘ When I open the door, there is food inside

What is this?



- ⌘ Korpo restas senmova au en uniforma stato de movo krom se forto agas sur ĝin.
- ⌘ Forto egalas la maso obligita de la akcelo
- ⌘ Se du korpoj efikas per fortoj sur unun de la alia, tiuj ĉi fortoj havas la saman grandojn kaj mala direktojn.
- ⌘ Esperanto

Where's Waldo? ©



Copyright 2002 All Rights Reserved

David R. Fletcher

Where is the “best” data?

Entity Type	number of redundant databases
Person	53
Improvement Project	45
Contract/Grant	35
Location Reference Object	33
Vendor	33
Proposal/Bid	30
Work Item	30
Transportation System Component	27
Environmental Feature State	26
Employee	22
Material	21
Capital Equipment	15

Geospatial data overload??



- ⌘ Digital Earth - a virtual representation of our planet that enables a person to explore and interact with the vast amounts of natural and cultural information gathered about the Earth. [NASA]
- ⌘ The National Map - public domain core geographic data about the United States and its territories . . . [USGS]
- ⌘ The Geography Network - a global network of geographic information users and providers supporting the sharing of geographic information among data providers, service providers, and users around the world. [ESRI]

“Millions of pictures and no one to look at them.”
*Porter Goss, Chairman of the House Permanent Select
Committee on Intelligence*

Data glut



- ⌘ “There are very few political [or] social . . . problems that arise because of insufficient information . . . Yet [we] stand firm in believing that we need yet more information. Information has become a form of garbage.” *Neil Postman*
- ⌘ “Ironically, in this glut of information, we may know less, not more about the world around us. At the very least, surely we know more things with less certainty than we used to.” *Herald Journal*
- ⌘ “Information, traditionally viewed as an indispensable ally, has become an unremitting enemy, raising too many issues and posing too many choices.” *San Jose Mercury News*

What's going on?



- ⌘ Once you know something, more data doesn't add value.
- ⌘ If you don't know something, more data doesn't add value either.
- ⌘ Too much data hides information.
- ⌘ The more valuable the data, the more likely it will be redundant.
- ⌘ The **Info glut** tsunami is overtaking us.

The information paradox



⌘ more, cheaper, faster data leads to

☑ analysis paralysis

☑ scarcity of attention

☑ physical and mental stress

☑ more uncertainty

☑ poorer decisions

☑ wasted resources

Some observations

Most organizations have too much of, the wrong kind of, not very good and poorly understood geospatial data isolated in impregnable stovepipes just in case they might need it someday.

Better metadata won't fix this.

Better technology won't fix this.

Certificated GIS specialists won't fix this.

Insanity is doing the same thing over and over again and expecting different results. *Albert Einstein*



So how did we get here?



Some 'rational' data beliefs

- ⌘ Data is an objective view of the one, real world
- ⌘ Data captures inherent properties of objects
- ⌘ Data provides information creating knowledge
- ⌘ Since "being" is information, the data experience is indistinguishable from "being there."
- ⌘ More accuracy, more precision, more quantity, faster data is better



Aristotle

Some ideological beliefs



- ⌘ Data contains information
- ⌘ Information is meaning
- ⌘ Human beings are information processors
- ⌘ Technology can do information processing
- ⌘ Technological innovation is human progress

So, what is data?



- ⌘ Data is signal: Observed, quantified and quantifiable physical phenomena such as electrical, electromagnetic radiation or pressure generated by physical events.
- ⌘ Data is symbol: Something that represents something else by association, resemblance, or convention.
- ⌘ Data is information: A set of meaningful facts used in making decisions.

In all cases, data is NOT a property of the world,
but of our interacting with something in the world!

Maps, data, and the world



- ⌘ The map is NOT the territory
- ⌘ The map does not show all of the territory
- ⌘ The map reflects the map maker's POV
- ⌘ The data is NOT the object
- ⌘ Data describe processes, NOT objects
- ⌘ Data reflects the designer's POV, not the user's
- ⌘ Measures of the earth cannot be equated to knowledge of the world.
- ⌘ The earth supports many "worlds"
- ⌘ These worlds reflect the measurer's POV

What is knowledge?



- ⌘ Knowledge is an encapsulated, uniquely human (personal and social) not technological experience
- ⌘ Knowledge is constructed out of use, not symbols (e.g., data).
- ⌘ Knowledge emerges from human interactions with the world
- ⌘ Knowledge is *person*, not object or data centered.
- ⌘ Knowledge leads to data.
- ⌘ The earth is *not* the world.

Genuine, practical, knowledge is forged by involvement in the world.

Involvement in the world



- ⌘ Direct experience of objects or events
 - ☒ direct sensation (e.g., seeing, hearing, touching, manipulating)
 - ☒ counting (i.e., naming or listing a group of objects)
 - ☒ measuring (i.e., a quantitative comparison of a thing to a standard)
- ⌘ Reasoning: Making conclusions, inferences and judgements.
- ⌘ Communicating: Reporting or receiving verifiable notations or descriptions of objects or events that we have seen, heard, felt, counted, measured, witnessed or experienced with another person.

Issues with data sharing



⌘ Access & Awareness

- ⊞ technology barriers
- ⊞ cultural, institutional, social spaces

⌘ Understanding

- ⊞ physical environment (data environment)
- ⊞ structure (data organization)
- ⊞ format (data type)
- ⊞ syntax (data coding)
- ⊞ semantics (meaning)

Technical metadata

Lexical metadata

⌘ Trust

- ⊞ social capital

⌘ Desire

Issues with metadata



- ⌘ Definitions are about things (i.e., truth-in-labeling) and not about use
- ⌘ Definitions created by data developers may not be understood by downstream consumers
- ⌘ Standard definitions are not the same as shared meaning
 - ⊞ Definitions of terms use (undefined) terms
 - ⊞ Meaning is established by knowledge, context and use.
- ⌘ Data never has exactly the same meaning twice.

So what can I do differently?



Copyright 2002 All Rights Reserved

David R. Fletcher

Adopt “irrational” data beliefs

- ⌘ Data is not the object
 - ☒ Whatever the data says, the object isn't
 - ☒ Data is less important than the object it stands for
- ⌘ Data has no intrinsic value to anybody
- ⌘ Data is the raw input to or the lifeless consequence of human effort. (process ore-to-spoor)
- ⌘ Data is a consequence of knowledge
- ⌘ Information results from knowledgeable persons interacting with data
- ⌘ Just-enough, just-good-enough, just-in-time is better than heaps of just-in-case data



Fletcher

Change your behavior



⌘ Be skeptical!

☑ "Why am I collecting this data?"

☑ "What decisions will I make differently if I had {faster, better, more} data?"

☑ "Is this technology a solution or a distraction?"

⌘ Remember that most of the world is low-tech and high-touch. Act accordingly and involve them.

⌘ Know what you don't need to know!

Final thoughts



- ⌘ Human concerns involving trust, values, knowledge, meaning, uncertainty, judgement, perspective or language cannot be resolved with more data or more technology.
- ⌘ My priorities:
 - ☑ People come first. Social capital is far more valuable than metadata.
 - ☑ Processes bring knowledgeable people together. Figure out what that society is and you'll find a world.
 - ☑ Get just enough data and no more.
 - ☑ Worry about technology last; don't be distracted.

Some readings



- ⌘ Coyne, R., *Designing Information Technology in the Postmodern Age*
- ⌘ Hayakawa, S. I., *Language in Thought and Action*
- ⌘ Heidegger, M., *Being and Time*
- ⌘ Korzybski, A., *Science and Sanity*
- ⌘ Postman, N., *Technopoly*
- ⌘ Setzer, V., "Data, Information, Knowledge and Competency"
- ⌘ Shenk, D., *Data Smog*