Ethics & Values of Data:

Coping with Complexity & Uncertainty

Joan E. Sieber

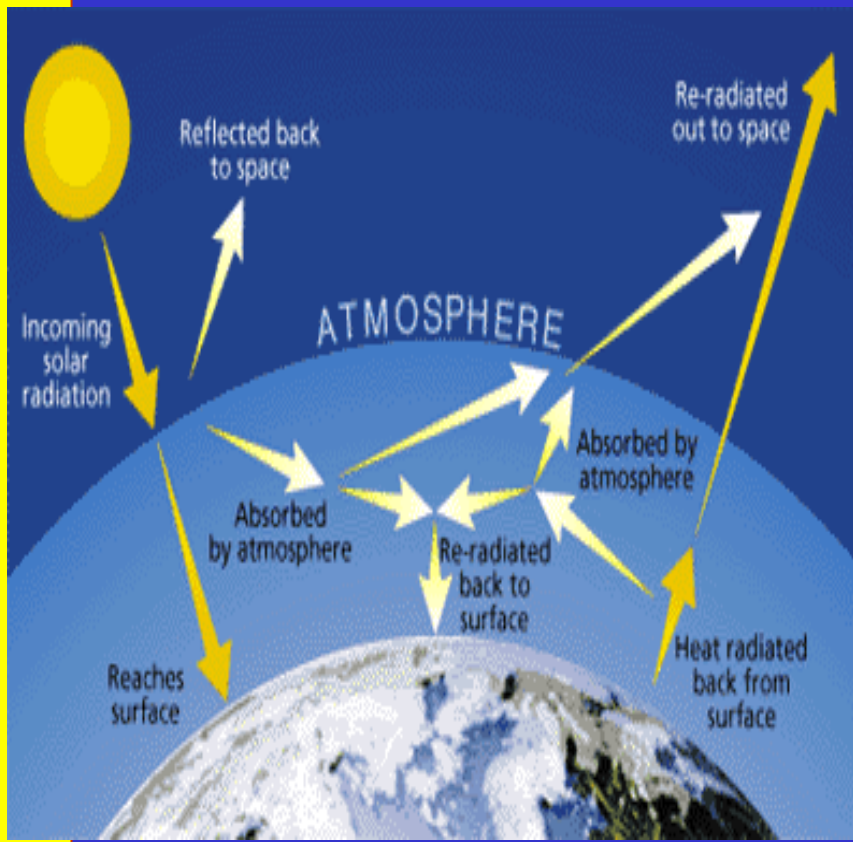
jsieber@bay.csuhayward.edu

To Do Practical Ethics:

- Thoughtfully select the values to be promoted.
- Minimize or balance conflicts among values.
- Consider how context can change priorities, nuances and values themselves.

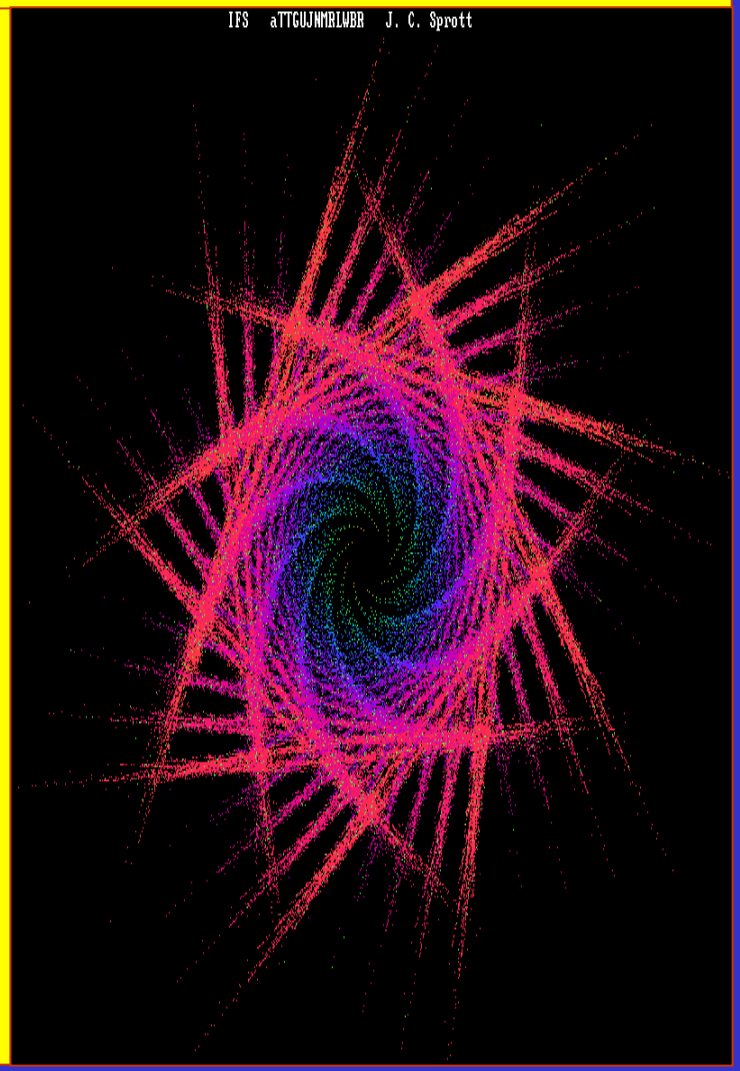
Why Ethics of Data is Complex:

- Data archives are for future use.
- Anticipate the future nature, problems and methods of science.
- Assemble data archives likely to be useful in the future.
- Anticipate possible ways of combining diverse kinds of data in informative ways.



The Course of Science is...

- Complex and unpredictable.
- Man-made, but its course is somewhat beyond our control.
- Possible to conceptualize and cope with via chaos theory.



Important Data Satisfy:

Micro-Ethics of Science:

- Build new knowledge
- Validity
- Transparency and appropriateness of methodology
- Adequate documentation
- Are shared with other scientists

Macro-Ethics

- Have broader social implications and uses.
- Foster important social values, policies.
- Address important current concerns: e.g., education, health, environment, building science infrastructure.

NSF's "Broader Impacts" Criteria:

The U.S. Congress and
the National Science Foundation
Require Funded Projects
to Seriously Address Macro-Ethical Issues

Find 5 pages of examples of "Broader Impacts" at:

Http://www.nsf.gov/pubs/2002/nsfo22/bicexamples.pdf

Normative Ethical Theory

- **DEONTOLOGY** - follow the most inviolable rule or value.
- **RULE UTILITARIANISM** - follow the rule most likely to lead to the most good for the most people.
- **ACT UTILITARIANISM** - do what seems, in the *particular* case, most likely to lead to the most good for the most people.

Values: Many, changing, complex, uncertain

Pre-September 11th - emphasis on openness

Post-September 11th - consideration of national security: e.g., should gene sequences of human pathogens be freely available?

Technology Licensing at Universities

20 years ago:

- Seen as stimulus to innovation.
- Cash cow.

Technology Licensing at Universities

Today:

- Inhibiting to productive young scientists.
- $ to Lawyers.

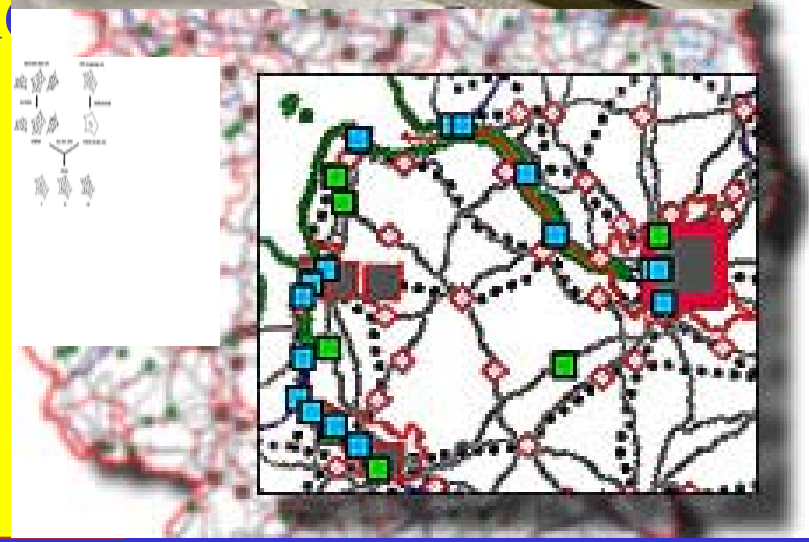
Office of
TECHNOLOGY
LICENS

Data Related Values...

- Evolve rapidly with new discoveries, priorities, technologies.
- Vary with context -- academe, industry, government laboratories.
- Change with changing reward structures in science.

Data … ?

- Raw, cleaned, digitized?
- Qualitative, descriptive?
- Cell lines?
- Samples of rock, sediment, i
  cores, DNA, bacteria?
- Fossils? Carbon dating?
- Financial records of the
  research administration?



Kinds of Archives

- Public data in public archives.
- Data of individual researchers shared informally via “invisible college.”
- Data of individual researchers shared via an organized archive.
- Privatized data:

  Produced by private industry.

  Publicly produced, value added.

Some Major Data-Related Values to Consider in Ethical Decision Making

Publicize. But How?

- Hard copy or electronic?
- Early and incomplete, or
  Later after elaboration?

The file drawer problem ----->
What about null results?

The role of peer review,
especially with null results
and electronic publication.

Sharing Useful Data

Avoiding data graveyard

- **Serving methods of data integration:**

-- **Meta-analysis**

-- **EITM**

- **Assembling panel and longitudinal data in useful formats.**
- **Otherwise deciding what's useful.**

Responsibilities for Sharing

- Who releases data for sharing? How soon after publication?
- Who operates archive, answers users' questions, updates the archive?
- How much sharing-related service is expected from the individual researcher? The public archive? The private archive?

Costs of Sharing

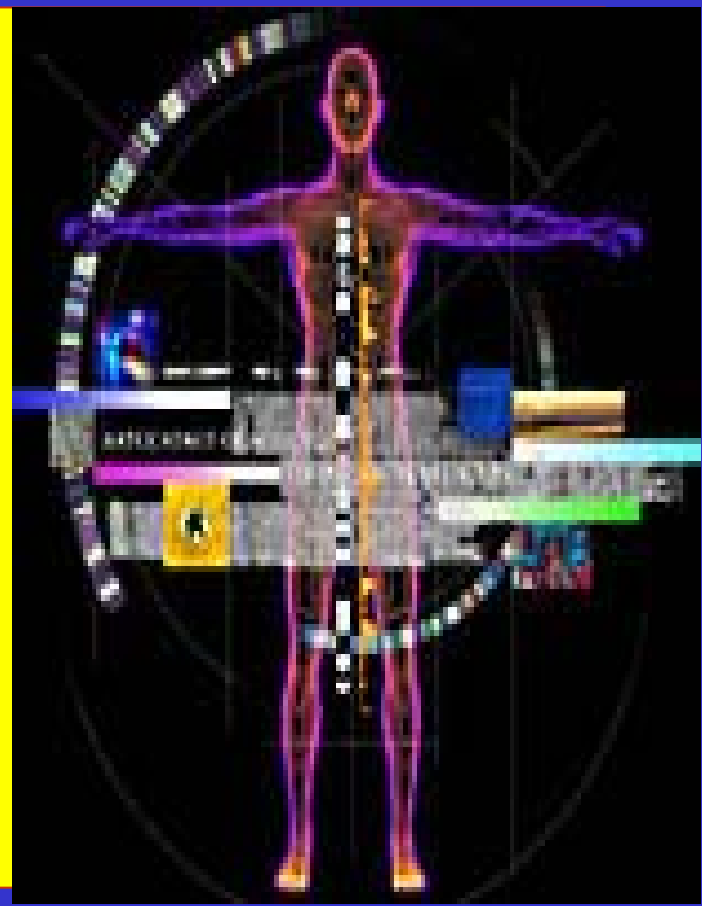
- Who pays for sharing services?
- What pricing formulas are used?
- Trade-offs between funding
  - -- new research
  - -- quality documentation
  - -- quality services
  - -- quality state-of-art technology

Protections Proportional to Main Goal

- **Academe** - Main Goal: Education - Mostly Open
- **Government** - Main Goal: Public Service - Mostly Open
- **Business** - Main Goal: Production & Profit - Mostly Protected Intellectual Property

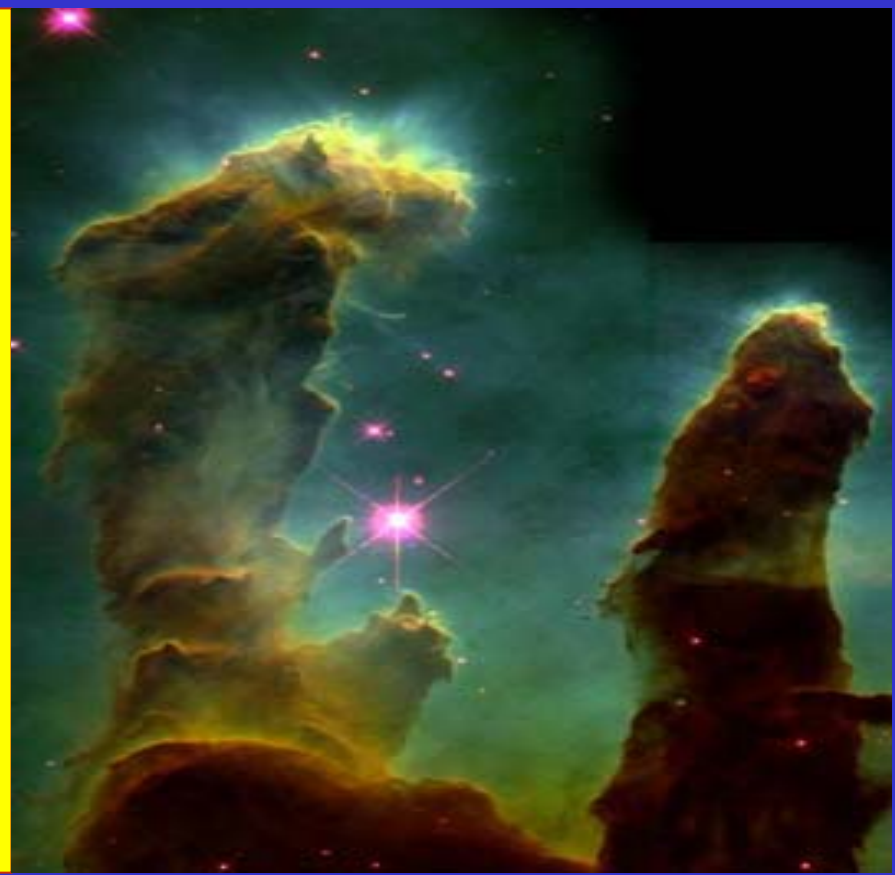
Protections Proportional to Immediacy of Application

Some genome sequencing work involves large investment and has immediate pharmaceutical applicability.



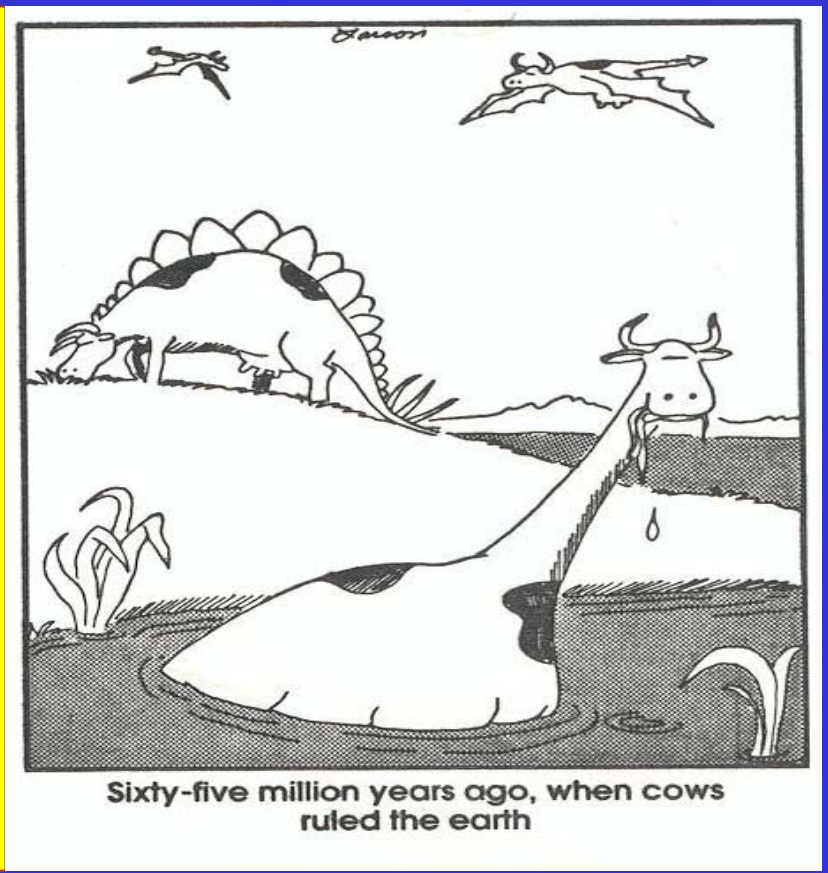
Astronomy

- No immediate applicability.
- Total openness of technology and data seems appropriate.



Natural Sciences

- Less business applicability.
- Educational value.
- Understanding of environment



NSF's Broader Impact Criteria

- Advance discovery while promoting teaching / learning at all levels (K-post doc).
- Broaden participation in science.
- Enhance scientific infrastructure
- Disseminate broadly via many media.
- Benefit society; educate non-scientists, partner with all kinds of institutions.

How to "do ethics" with this kaleidoscopic array of values and contexts?

Use of Concepts from Chaos Theory

Use complexity, change and uncertainty as a context for creative problem solving.

Experiment with small local changes, which can have big impacts. Avoid sweeping changes.

Chaotic structures self organize. See what scientists want in order to be productive.

Complex structures can have rich simple subtleties. Beware of stereotyped ideas.

Show-case and enjoy
the elegant
self-organization
you find in
scientific institutions.

Time may be regarded as a process. Use it that way to think about new data practices.

Reductionistic notions of dissection and control are only one approach. Accept chaos; be creative.