

Author(s): August E. Evrard, PhD. 2010

License: Unless otherwise noted, this material is made available under the terms of the **Creative Commons Attribution-Non-commercial-Share Alike 3.0 License:**
<http://creativecommons.org/licenses/by-nc-sa/3.0/>

We have reviewed this material in accordance with U.S. Copyright Law **and have tried to maximize your ability to use, share, and adapt it.** The citation key on the following slide provides information about how you may share and adapt this material.

Copyright holders of content included in this material should contact open.michigan@umich.edu with any questions, corrections, or clarification regarding the use of content.

For more information about **how to cite** these materials visit <http://open.umich.edu/education/about/terms-of-use>.

Citation Key

for more information see: <http://open.umich.edu/wiki/CitationPolicy>

Use + Share + Adapt

{ Content the copyright holder, author, or law permits you to use, share and adapt. }


-  **Public Domain – Government:** Works that are produced by the U.S. Government. (17 USC § 105)
-  **Public Domain – Expired:** Works that are no longer protected due to an expired copyright term.
-  **Public Domain – Self Dedicated:** Works that a copyright holder has dedicated to the public domain.
-  **Creative Commons – Zero Waiver**
-  **Creative Commons – Attribution License**
-  **Creative Commons – Attribution Share Alike License**
-  **Creative Commons – Attribution Noncommercial License**
-  **Creative Commons – Attribution Noncommercial Share Alike License**
-  **GNU – Free Documentation License**

Make Your Own Assessment

{ Content Open.Michigan believes can be used, shared, and adapted because it is ineligible for copyright. }

-  **Public Domain – Ineligible:** Works that are ineligible for copyright protection in the U.S. (17 USC § 102(b)) *laws in your jurisdiction may differ

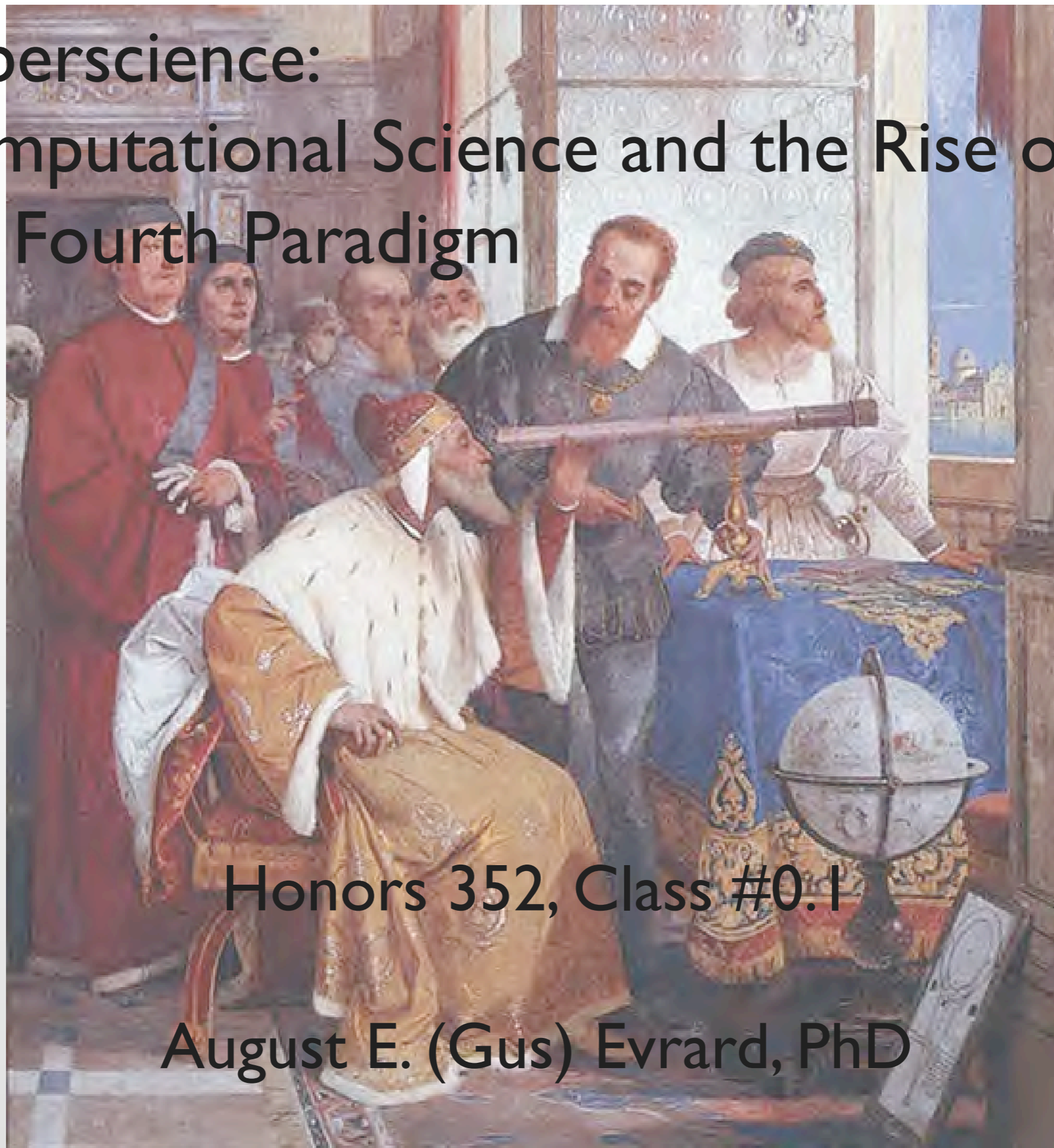
{ Content Open.Michigan has used under a Fair Use determination. }

-  **Fair Use:** Use of works that is determined to be Fair consistent with the U.S. Copyright Act. (17 USC § 107) *laws in your jurisdiction may differ

Our determination **DOES NOT** mean that all uses of this 3rd-party content are Fair Uses and we **DO NOT** guarantee that your use of the content is Fair.

To use this content you should **do your own independent analysis** to determine whether or not your use will be Fair.

Cyberscience: Computational Science and the Rise of the Fourth Paradigm



Honors 352, Class #0.1

August E. (Gus) Evrard, PhD

Fall 2010

WANT TO EXPLORE TEACHING?

The IDEA Institute has many
programs for you
including

**FUTURE
AND
SCIENCE SATURDAYS**

The **IDEA INSTITUTE** offers many different ways for you to explore science and math teaching! We have two different opportunities available this semester for dedicated science and math undergraduate students to collaborate with high school and middle school teachers.

If you enjoy learning science and mathematics and would love to share your knowledge with others, get involved in

FUTURE OR SCIENCE SATURDAYS

For more information and application forms, look online:

WWW.IDEAINSTITUTE.UMICH.EDU

Attend an **IDEA INSTITUTE**
informational meeting!

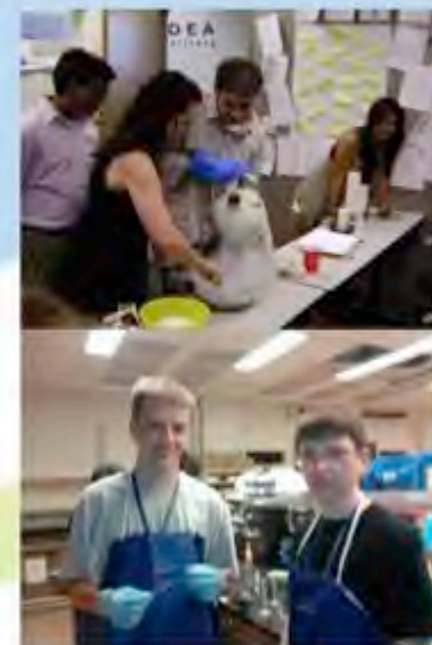
Monday, September 13, 4:00 - 5:30 p.m.

Tuesday, September 14, 4:00 - 5:30 p.m.

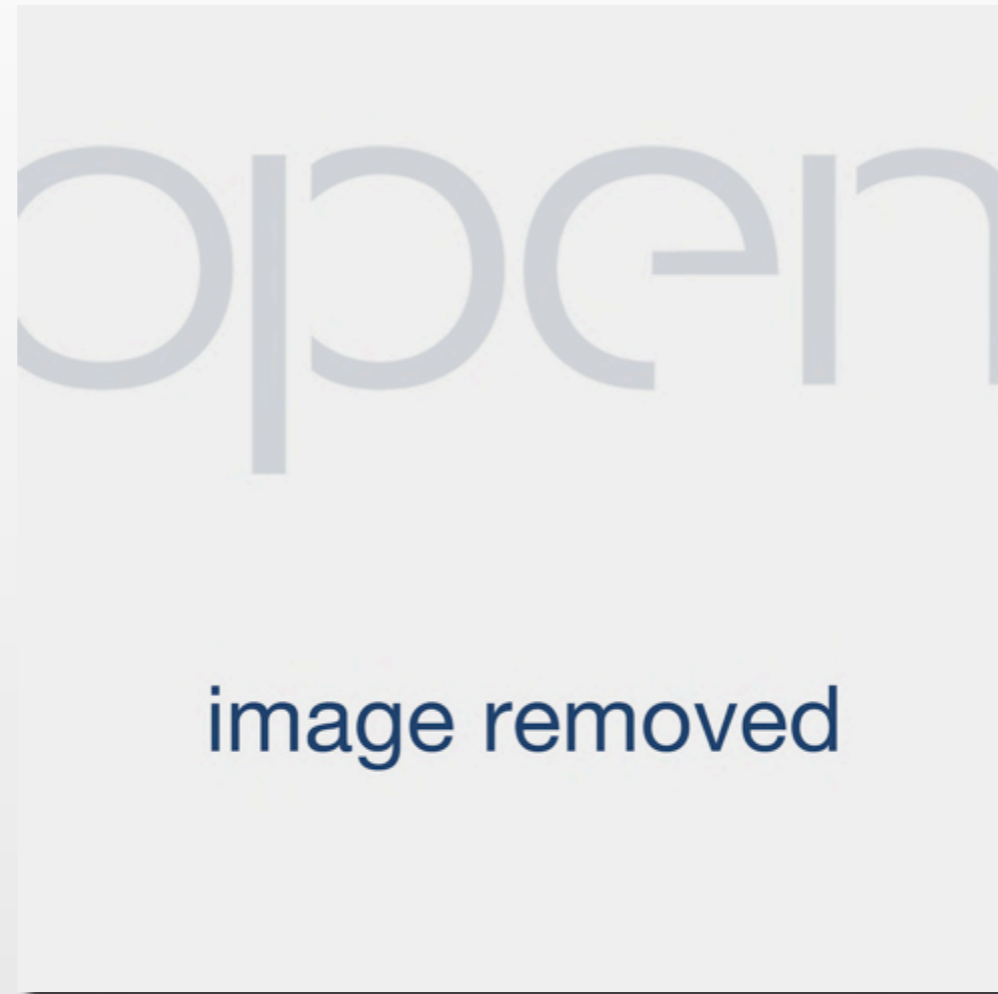
In the IDEA Institute:

3236 Undergraduate Science Building

Snacks provided



In today's news...



Please see original quote from New York Times regarding Google Instant at <http://bits.blogs.nytimes.com/2010/09/08/google-speeds-queries-with-instant-results/>.

Today

- **reading quiz**
- **short lecture (Prof. Gus) - Jim Gray's fourth paradigm = IBM's smarter planet?**
- **discussion: the participants and processes of scientific research**
- **blog / google site access**

Nth paradigm?

Thomas Samuel Kuhn (1922-1996) became one of the most influential philosophers of science of the twentieth century, perhaps the most influential—his *The Structure of Scientific Revolutions* is one of the most cited academic books of all time. His contribution to the philosophy of science marked not only a break with several key positivist doctrines but also inaugurated a new style of philosophy of science that brought it much closer to the history of science. His account of the development of science held that science enjoys periods of stable growth punctuated by revisionary revolutions, to which he added the controversial ‘incommensurability thesis’, that theories from differing periods suffer from certain deep kinds of failure of comparability.



Original biographical information of Thomas Samuel Kuhn at <http://plato.stanford.edu/entries/thomas-kuhn/>.

Nth paradigm?

The historian of science Thomas Kuhn gave paradigm its contemporary meaning when he adopted the word to refer to **the set of practices that define a scientific discipline at any particular period of time**. Kuhn himself came to prefer the terms **exemplar** and **normal science**, which have more precise philosophical meanings. However in his book *The Structure of Scientific Revolutions* Kuhn defines a **scientific paradigm** as:

- * what is to be observed and scrutinized
- * the kind of questions that are supposed to be asked and probed for answers in relation to this subject
- * how these questions are to be structured
- * how the results of scientific investigations should be interpreted

Alternatively, the Oxford English Dictionary defines paradigm as "a pattern or model, an exemplar." Thus an additional component of Kuhn's definition of paradigm is:

- * how is an experiment to be conducted, and what equipment is available to conduct the experiment.

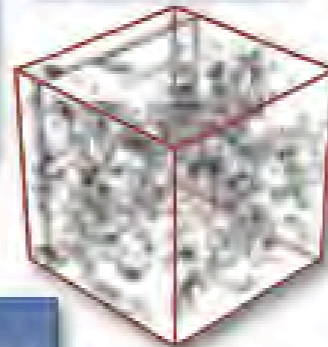
Jim Gray's four scientific paradigms / branches

Science Paradigms

- Thousand years ago:
science was empirical
describing natural phenomena
- Last few hundred years:
theoretical branch
using models, generalizations
- Last few decades:
a computational branch
simulating complex phenomena
- Today: **data exploration (eScience)**
unify theory, experiment, and simulation
 - Data captured by instruments or generated by simulator
 - Processed by software
 - Information/knowledge stored in computer
 - Scientist analyzes database/files using data management and statistics



$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G\rho}{3} - K\frac{c^2}{a^2}$$



1. empiricism

observe phenomenon and attempt to classify

Ptolemy's universe of concentric spheres

2. theory

describe above classifications with mathematical models

Newtonian/Einsteinian gravity

3. computation

build 'virtual' physical systems via solution of math models

Cosmic structure formation

4. data-driven synthesis (?)

unite empirical, theoretical and computational branches with data (X-info and Comp-X)

Matter/energy content of the universe

cosmic web of large-scale dark matter

image ~10 billion light-years wide derived from billion-particle N-body simulation

Evrard *et al.* (2002), *Astrophysical Journal*, vol. 573, 7 (231 citations)

data-rich research permeates all domains (X)

X-Info

- The evolution of X-Info and Comp-X for each discipline X
- How to codify and represent our knowledge



The Generic Problems

- Data ingest
- Managing a petabyte
- Common schema
- How to organize it
- How to reorganize it
- How to share it with others
- Query and Vis tools
- Building and executing models
- Integrating data and literature
- Documenting experiments
- Curation and long-term preservation

challenges

1. unifying the tiers

- data collections across the scales, from small labs to international consortia
- published literature with underlying data (*raw, derived*) and data processing algorithms/codes
- ironing the seams across disciplines
 - Disciplinary scientists and organizations (e.g., National Academy of Science, National Science Foundation)*

2. semantics

- describing objects, attributes, methods in a robust, scaleable manner
- curating and archiving collections
 - Disciplinary scientists, Librarians!*

3. funding

- recognize value of data-driven synthesis (DDS) infrastructure
- maintain `single investigator' support while growing new capabilities
 - Federal and state government agencies, scientific industry partners, universities!*

Building a smarter planet, for business (and science?)

The screenshot shows the IBM Smarter Planet website homepage. At the top, there is the IBM logo on the left, a search bar, and a navigation menu with links for Home, Solutions, Services, Products, Support & downloads, and My IBM. The main content area features a 'Welcome to a smarter planet' message with a globe icon and a 'View all topics' link. Below this is a 'This week's feature' section titled 'Do you understand how to build a smarter planet?' featuring three experts: Meg Selfe (Smarter Products), Jack Danahy (Security), and Paul Chang (Traceability). Each expert has a video thumbnail and a 'Watch the video' link. To the right, a 'Conversations' section lists these three experts with their respective video thumbnails and links. At the bottom, there are two main sections: 'Smarter healthcare' with a photo of medical files and the text 'On the mend' followed by a description of electronic medical records and a 'Learn more' link; and 'What is a smarter planet?' with a large '3 big ideas to build one smarter planet' headline and a list of three ideas: 1. Instrument the world's systems, 2. Interconnect them, and 3. Make them intelligent, followed by a 'Here's how we make it work' link and a globe icon.

© FAIR USE

Screen shot of IBM website from <http://www.ibm.com/smarterplanet/us/en>.

Discuss: Nature of Scientific Research

- who participates?
- what are the processes involved? (end-to-end view)
- how are the participants rewarded?

Lab next Tuesday

- three groups: PC / Mac / Linux
- bring cameras

Additional Source Information

for more information see: <http://open.umich.edu/wiki/CitationPolicy>

Slide 3: Giuseppe Bertini (1825–1898), "Galileo Galilei showing the Doge of Venice how to use the telescope"

Slide 4: Image of a flyer regarding an event that has already taken place. For more information, please go to www.ideainstitute.umich.edu.

Slide 5: Please see original quote from New York Times regarding Google Instant at <http://bits.blogs.nytimes.com/2010/09/08/google-speeds-queries-with-instant-results/>.

Slide 7, Image 1 (left): Please see original biographical information of Thomas Samuel Kuhn at <http://plato.stanford.edu/entries/thomas-kuhn/>.

Slide 9: Jim Gray's The Fourth Paradigm: Data-Intensive Scientific Discovery.

Slide 10: Evrard *et al.* (2002), *Astrophysical Journal*, 573(7).

Slide 11: Jim Gray's The Fourth Paradigm: Data-Intensive Scientific Discovery.

Slide 13: Please see original image of screen shot of IBM website at <http://www.ibm.com/smarterplanet/us/en>.