The Cornell Lab of Ornithology

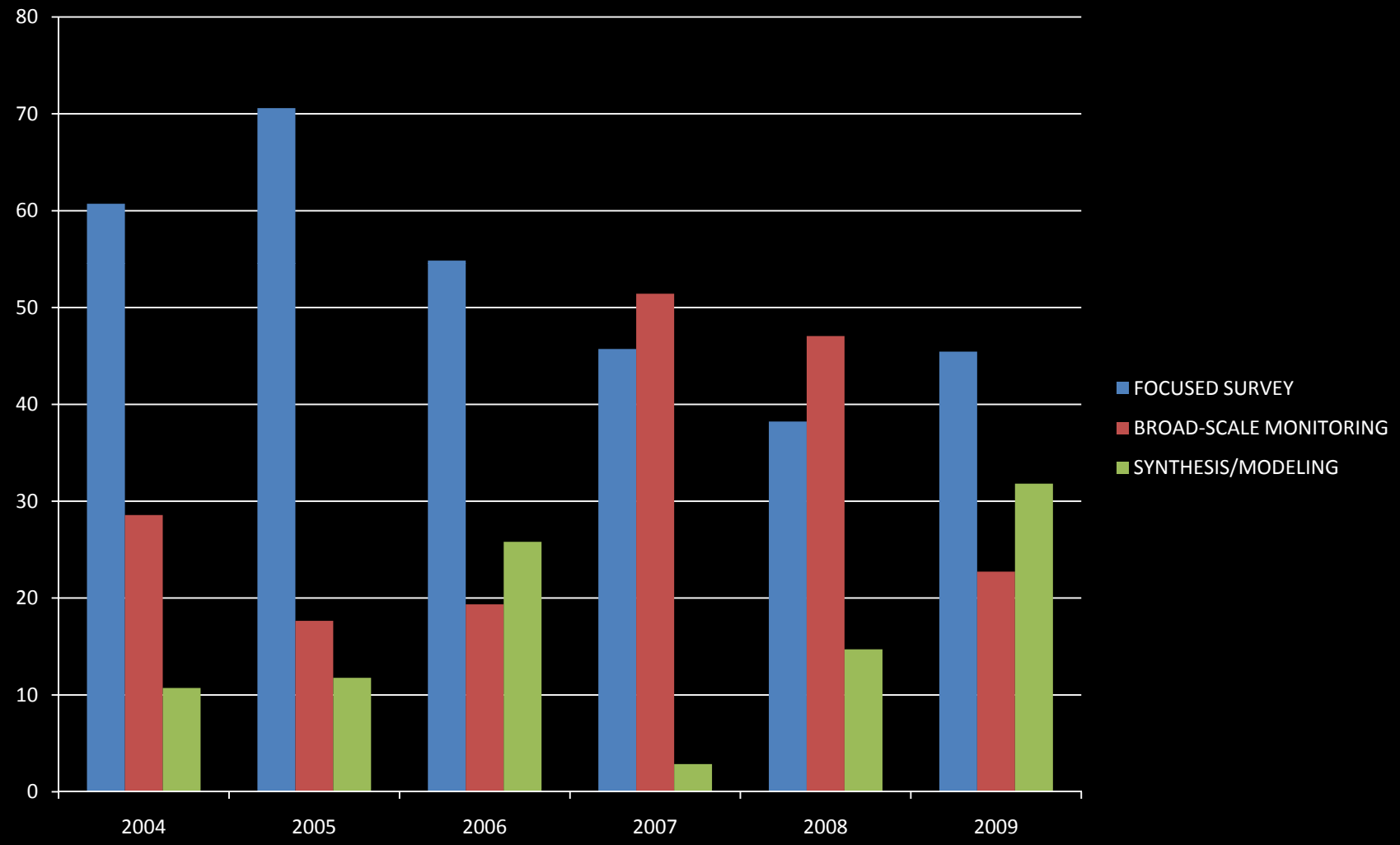
Biodiversity Research and Conservation in a Digital World

"Conservation Biology publishes groundbreaking papers and is instrumental in defining the key issues contributing to the study and preservation of species and habitats."

General Experimental Design

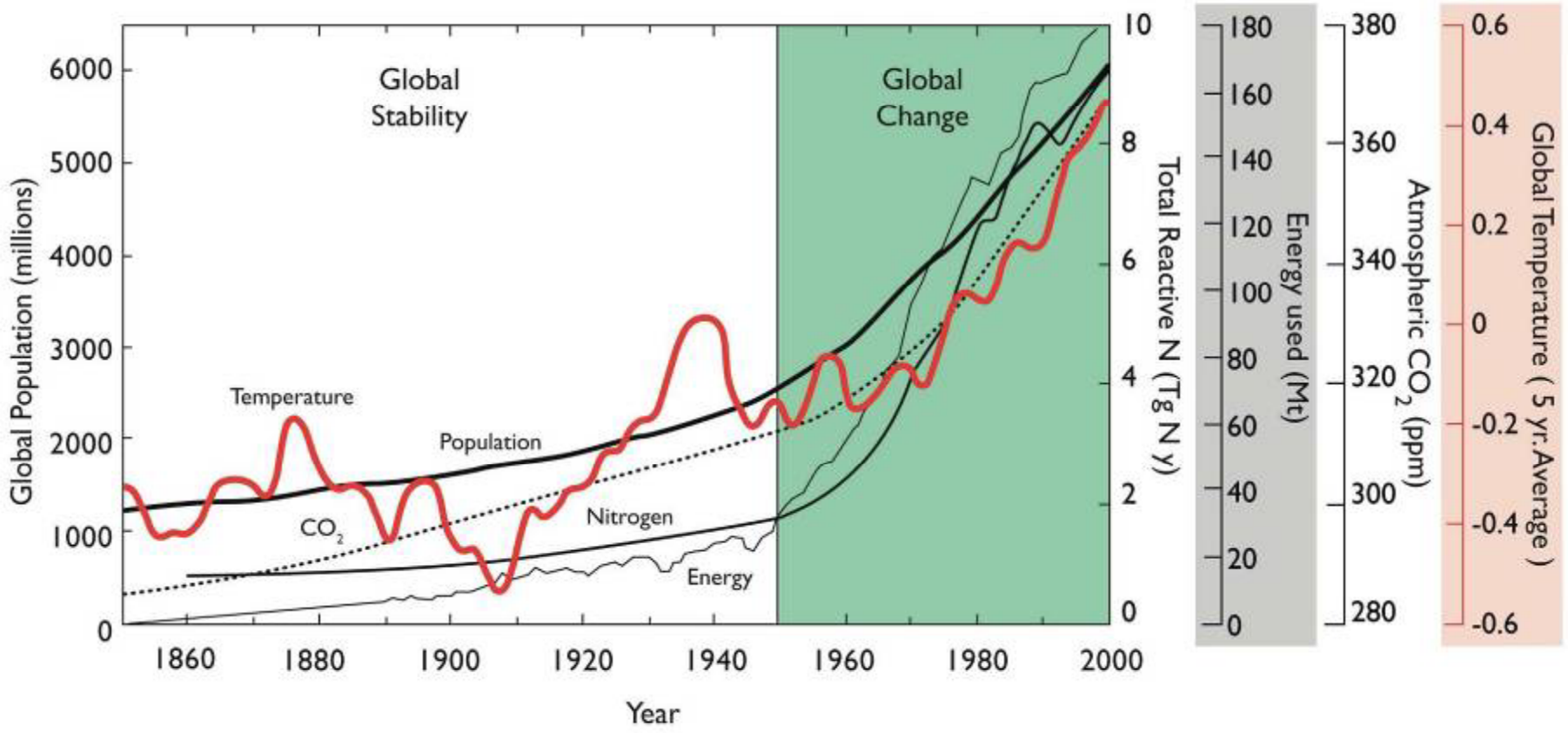
- Focused Surveys
- Broad-scaled Monitoring
- Synthesis and Modeling

Conservation Biology



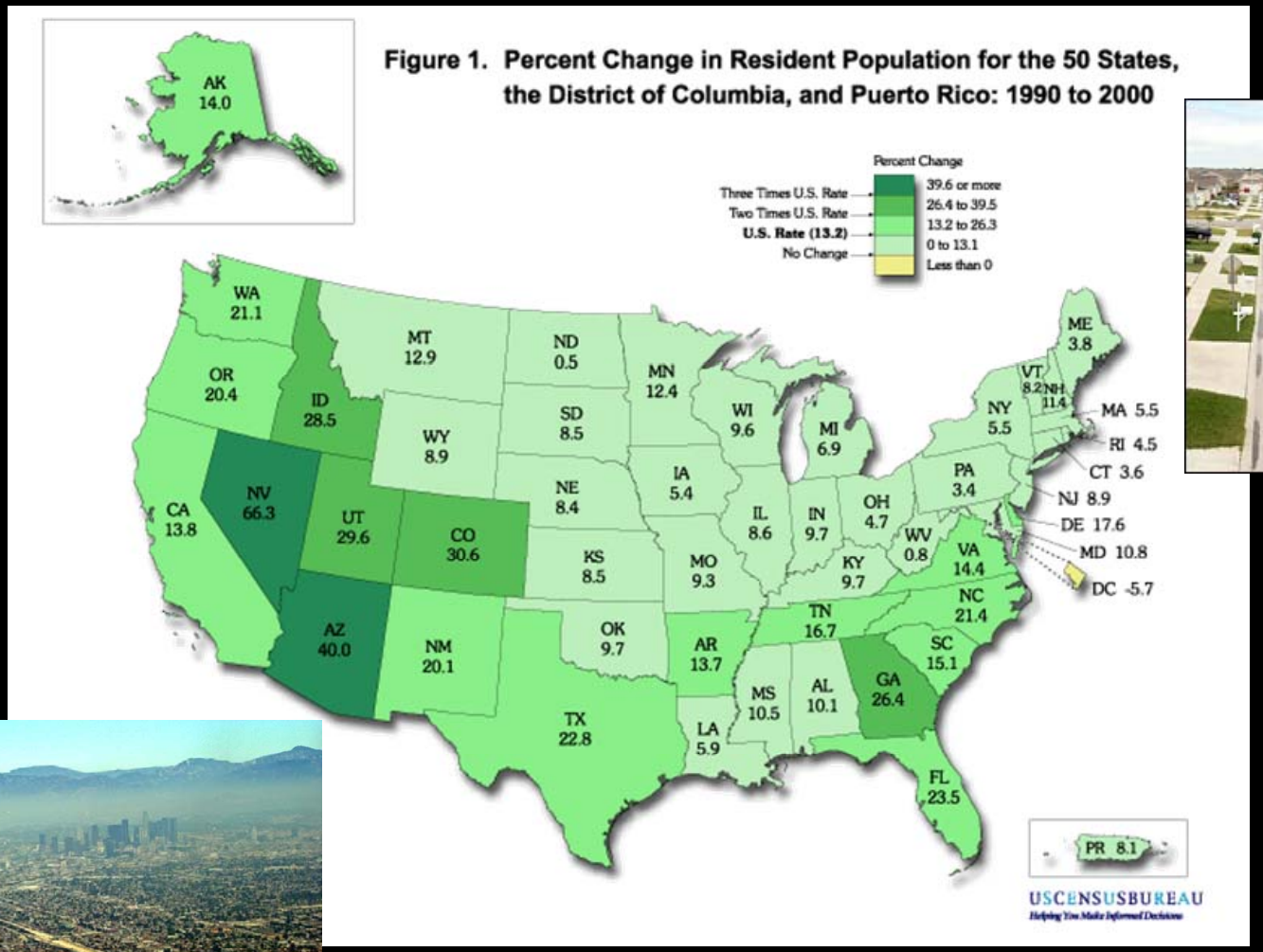

Living with Cancer
The changing science
Beyond Baghdad: Where The Enemy Has Its Own Surge
The Sopranos' Last Song: What Exit Will Tony Take?
TIME
SPECIAL DOUBLE ISSUE
The Global Warming Survival Guide
51 Things You Can Do to Make a Diffe
2ND-QTR SIZZLE PROFITS AT 900 COMPANIES (P. 74)
PAYING FOR COLLEGE BEWARE OF THOSE HIGH 529 FEES (P. 96)
TERRORISM WHAT COMPANIES STILL NEED TO DO (P. 26)
BusinessWeek
GLOBAL WARMING
Why Business Is Taking It So Seriously
BY JOHN CAREY (P. 60)
APRIL 3, 2006
SPECIAL REPORT GLOBAL WARMING
TIME
BE WORRIED. BE VERY WORRIED.
Climate change isn't some vague future problem—it's already damaging the planet at an alarming pace. Here's how it affects you, your kids and their kids as well
EARTH AT THE TIPPING POINT
HOW IT THREATENS YOUR HEALTH
INDIA CAN HELP WORLD—OR DESTROY IT
CRUSADERS
Adapted for A NEW GENERATION from the New York Times Bestseller
an inconvenient truth
the crisis of global warming
AL GORE
AUGUST 26, 2002
SPECIAL REPORT
TIME
HOW TO SAVE THE EARTH
The hot and wild weather is a sign of things to come. But fresh ideas and new technology can cool us down and make this a GREEN CENTURY
nature
Eocene global warming
Hydrothermal vents prompt methane release
TIME
007 at 25
Where's the Beach?
America's Vanishing Coastline
OCTOBER 19, 1987
$1.95
TIME
The Heat Is On
How the Earth's Climate Is Changing
Why the Ozone Hole Is Growing
SPECIAL REPORT Fighting for Global Markets
APRIL 9, 2001
GLOBAL WARMING
Climbing temperatures.
Melting glaciers. Rising seas.
All over the earth we're feeling the heat.
Why isn't Washington?
Dr. Bush's Rx for Health Care
TIME
VANISHING OZONE
THE DANGER MOVES CLOSER TO HOME
THE BIG DRY
W YORKER
LOBAL WARMING CARTOONS - 2007 - 210
CHINA
JAPAN
RUSSIA
PACIFIC OCEAN
TEXAS
UTAH
LAS VEGAS
ST LOUIS
CHICAGO
PENNSYLVANIA
ATLANTIC OCEAN
NOVEMBER 28, 2005
JOE KLEIN ON IRAQ ■ BROKEBACK MOUNTAIN: GIDDY-YEP, I'M GAY
TIME
New Orleans Blues
It's worse than you think. Three months after Katrina, the city still suffers
BY CATHY BOOTH THOMAS

Global change

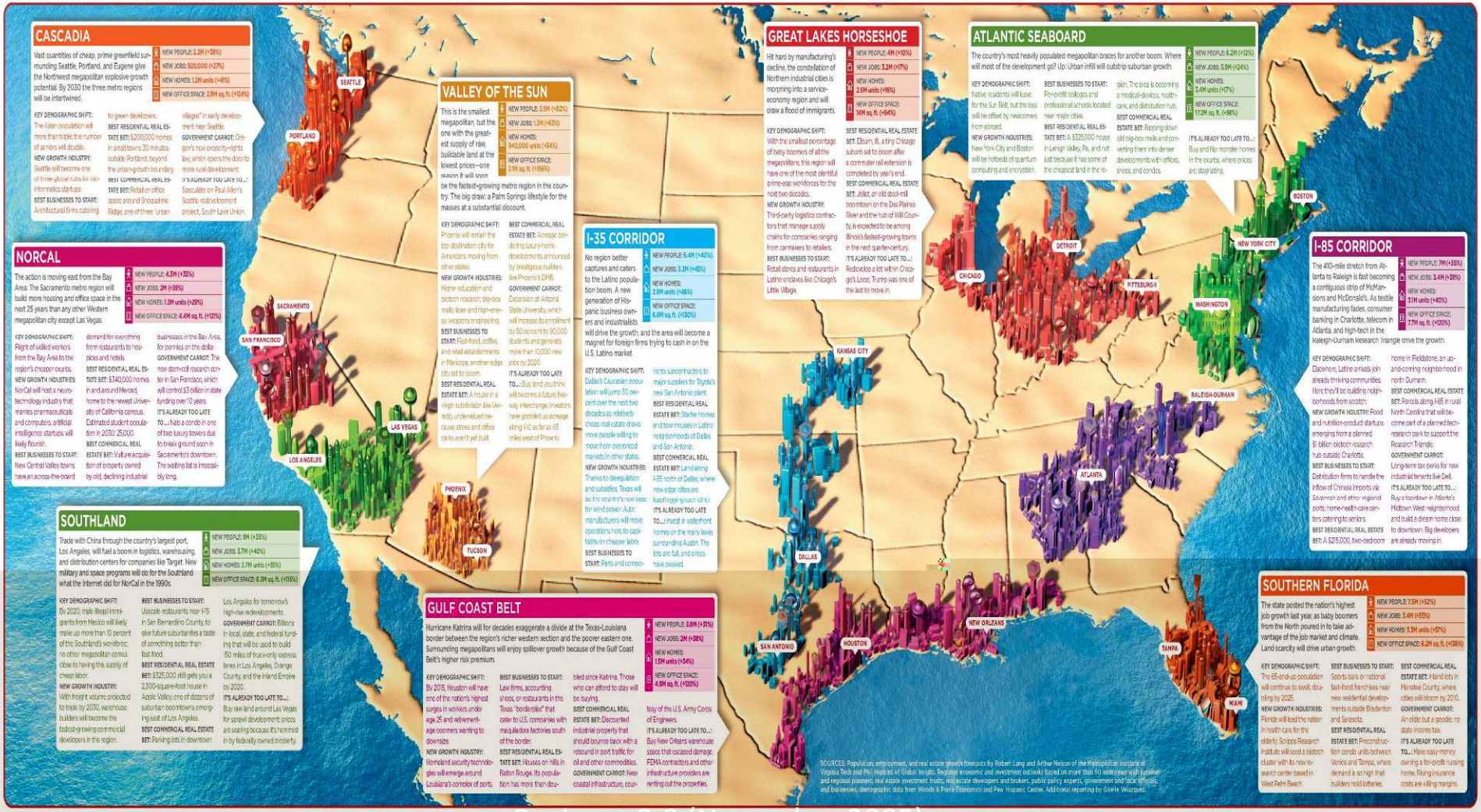


Smith, Knapp, Collins. In press.

Increasing Human Population



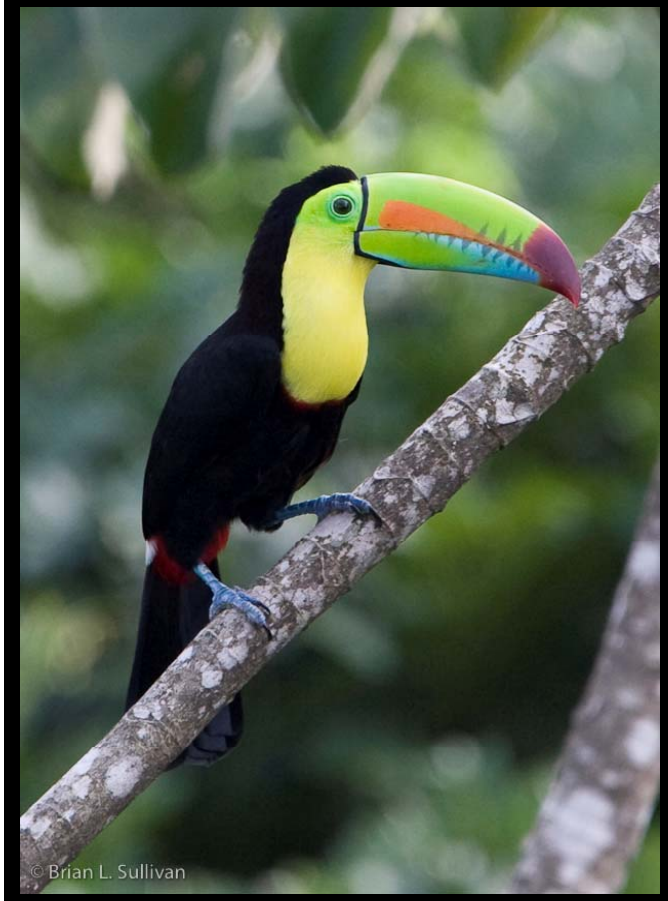
“Megapolitan”ization



Business 2.0 (November 2005)

Computation resources and a growing cyberinfrastructure is now an equal and indispensible partner for the advance of scientific knowledge.

Presentation Goals



The computational framework for biodiversity research.

The cyberinfrastructure for data curation and access.

Define environmental observational data networks.

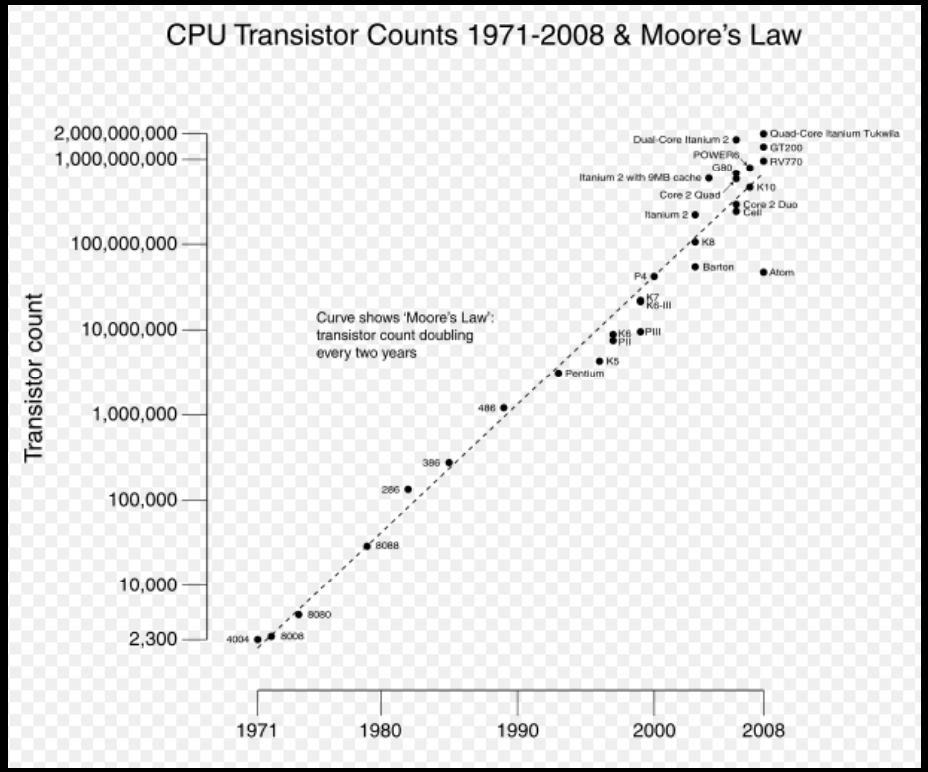
Describe the Data Intensive Science research paradigm.

Provide a domain example.

The computational framework for biodiversity research.

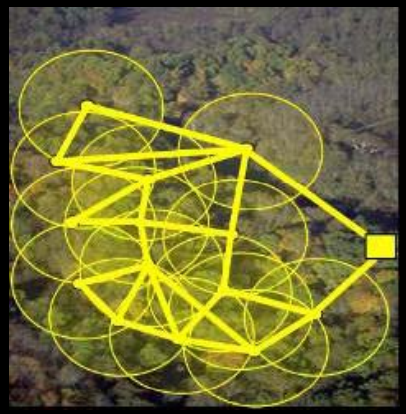
Moore’s Law

The number of transistors that can be placed inexpensively on an integrated circuit will increase exponentially, doubling approximately every two years.



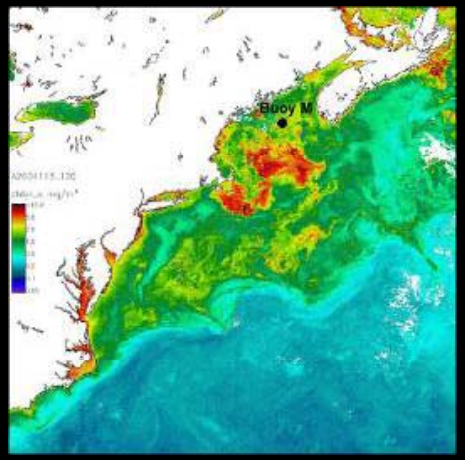
The computational framework for biodiversity research.

- Computational power



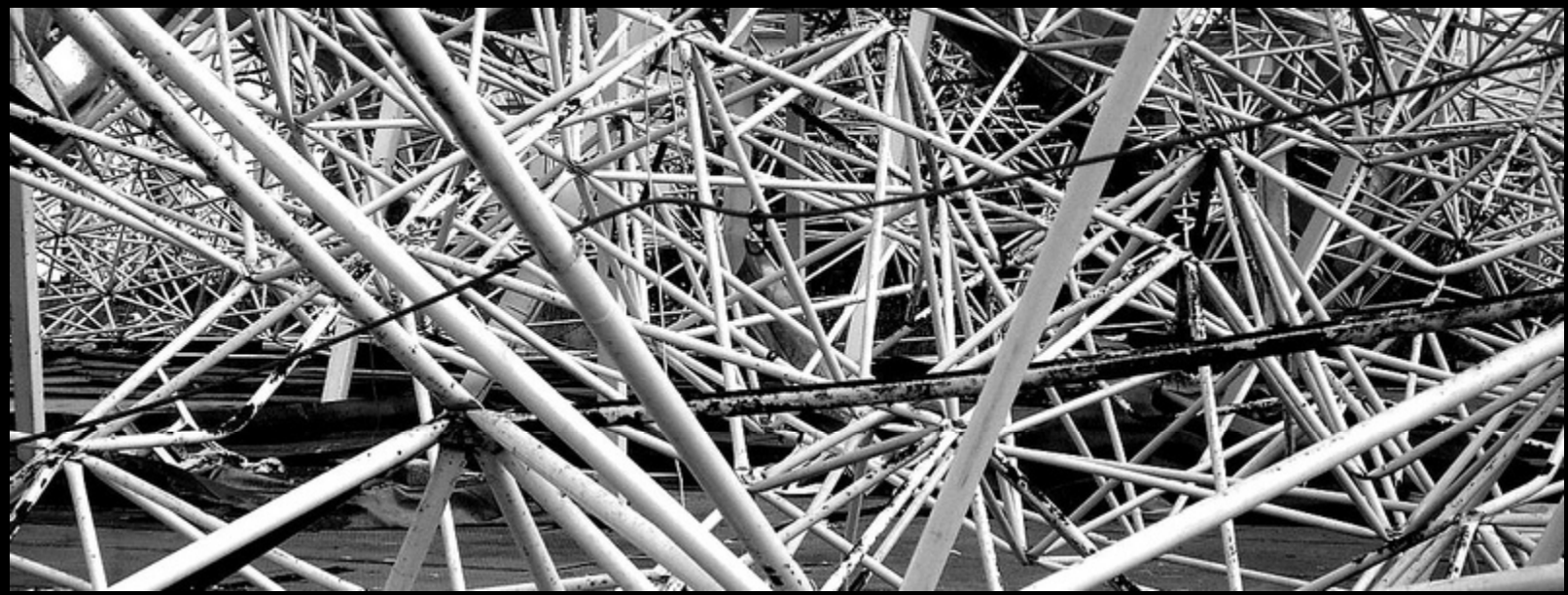
F16 $f(x)$ Σ =

	A	B	C	D	E	F
1						
2						
3	Date	Start time	End time	Pause	Sum	Comment
4	2007-05-07	9,25	10,25	0	1	Task 1
5	2007-05-07	10,75	12,50	0	1,75	Task 1
6	2007-05-07	18,00	19,00	0	1	Task 2
7	2007-05-08	9,25	10,25	0	1	Task 2
8	2007-05-08	14,50	15,50	0	1	Task 3
9	2007-05-08	8,75	9,25	0	0,5	Task 3
10	2007-05-14	21,75	22,25	0	0,5	Task 3
11	2007-05-14	22,50	23,00	0	0,5	Task 3
12	2007-05-15	11,75	12,75	0	1	Task 3



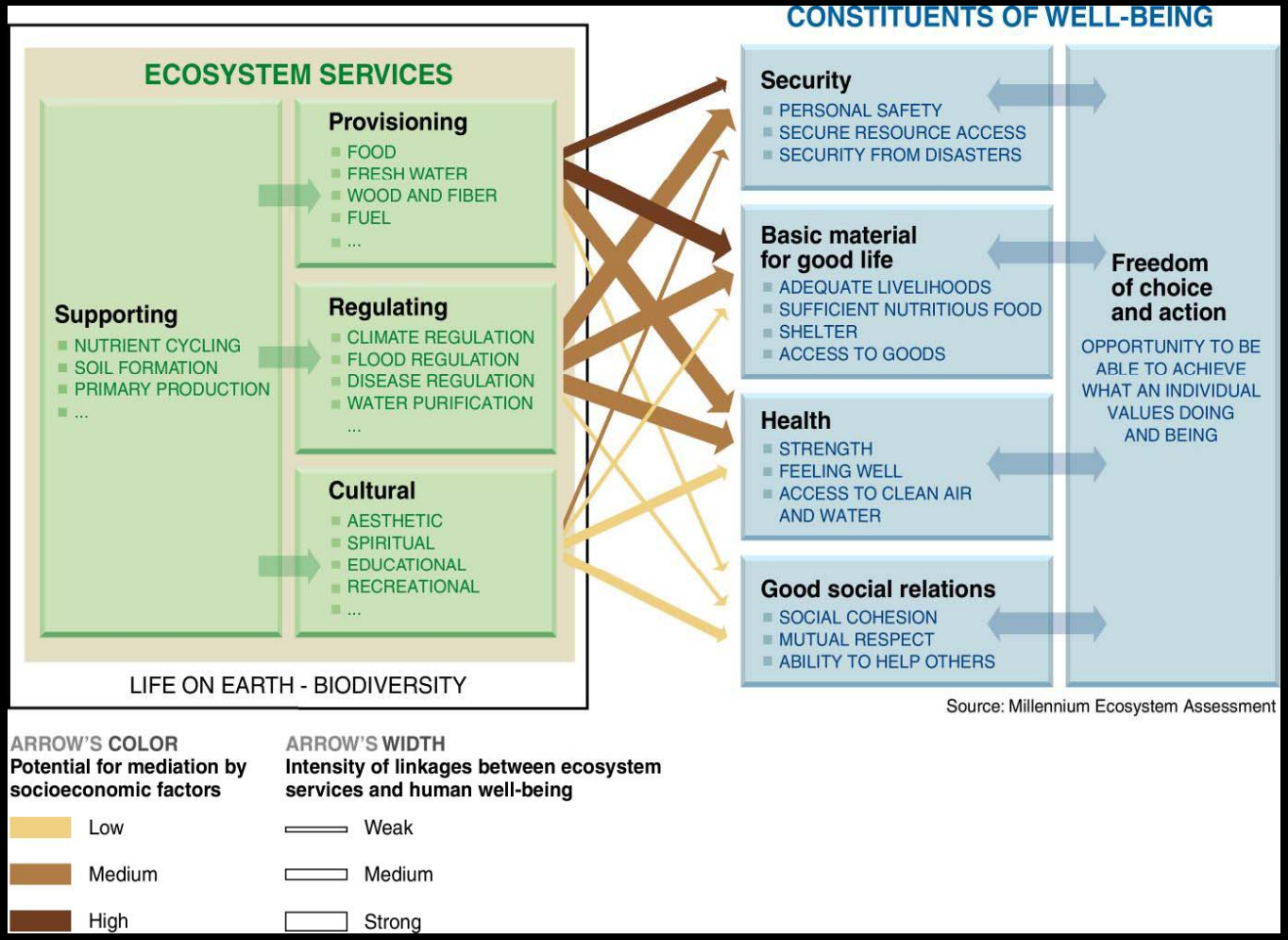
The computational framework for biodiversity research.

Multivariate Madness



The computational framework for biodiversity research.

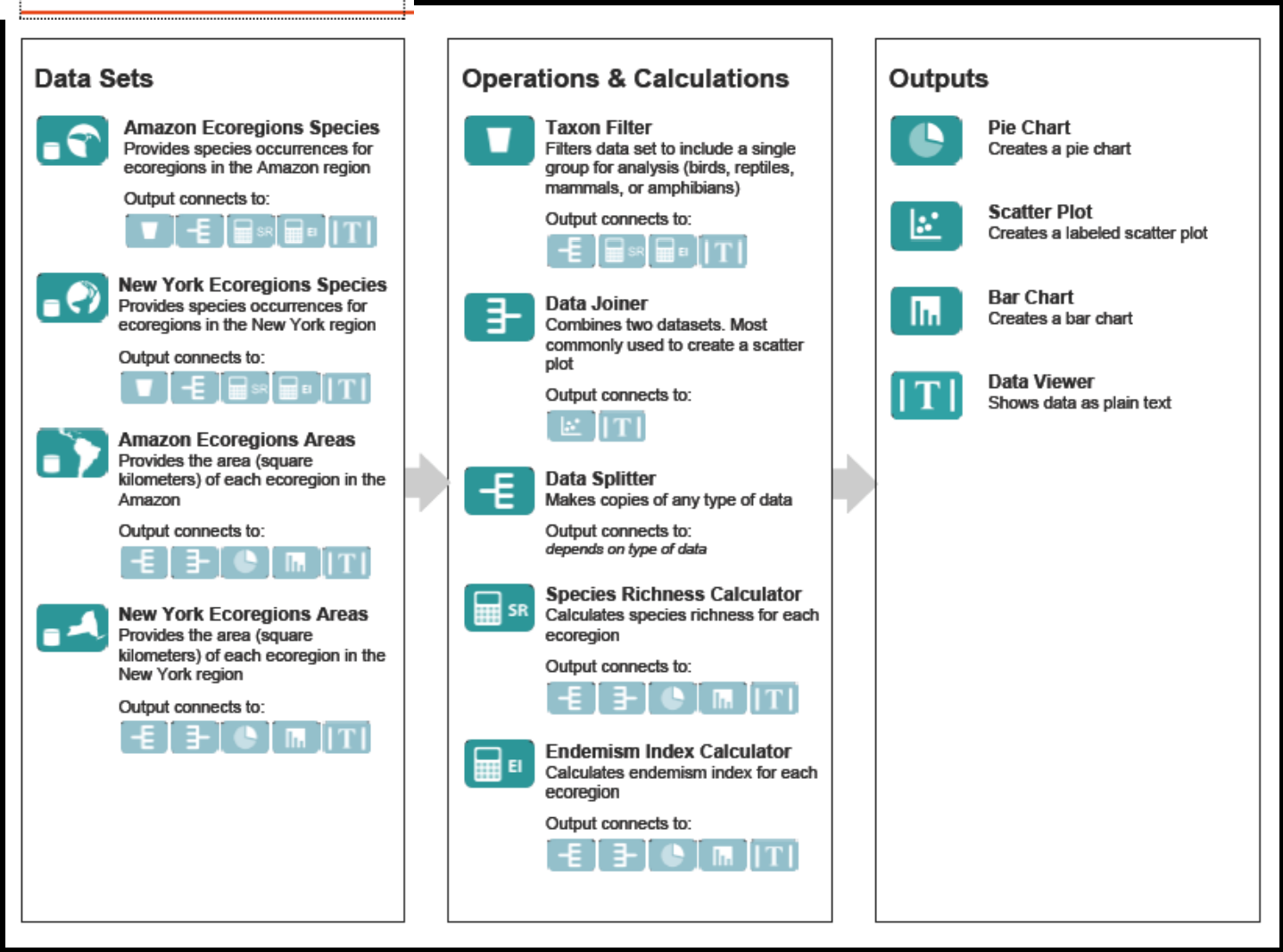
The coupling of human and natural systems.



The computational framework for biodiversity research.

SCIENCE PIPES

http://sciencepipes.org

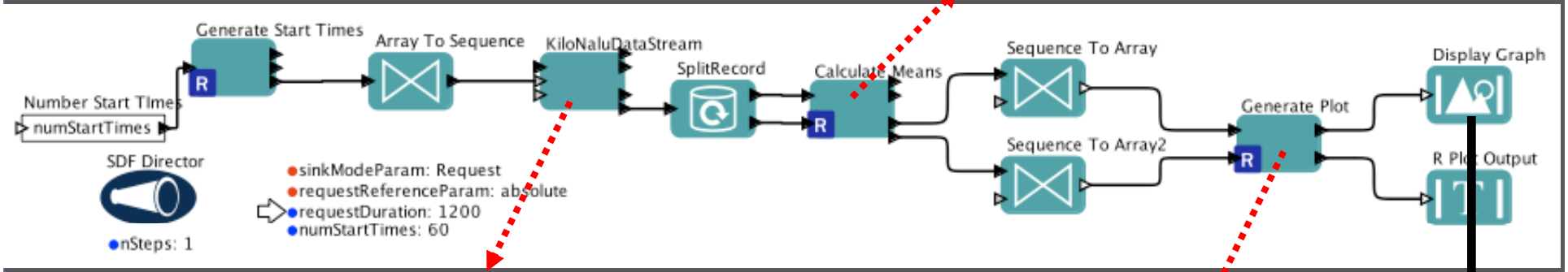


The computational framework for biodiversity research.

Kepler
Your Science. Enabled.

http://kepler-project.org

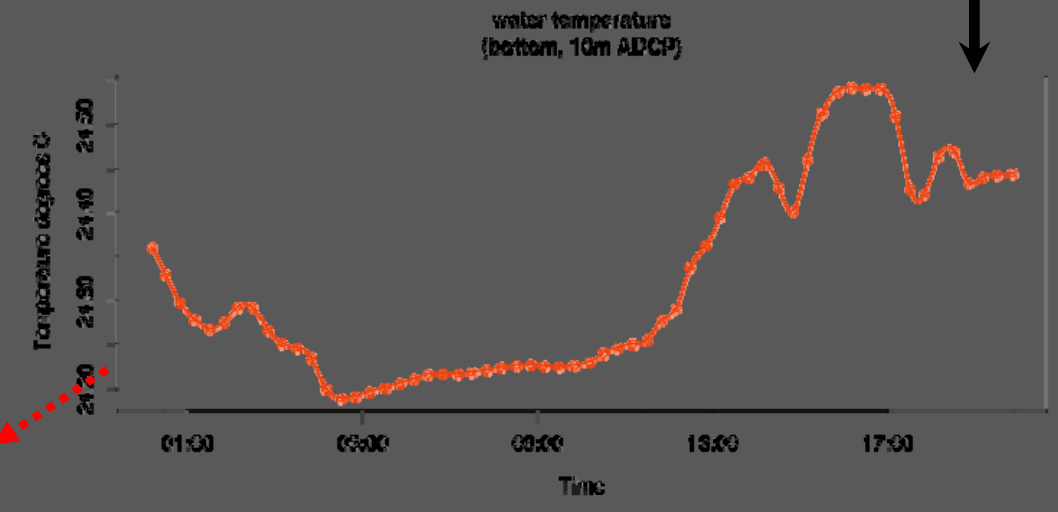
Support application scripts in R, Matlab, etc.



Streaming Data from observatory DataTurbine Server

Modular components, easily saved and shared

Publish to workflow repository with accession number
Documents the linkage between publication, analysis, and data



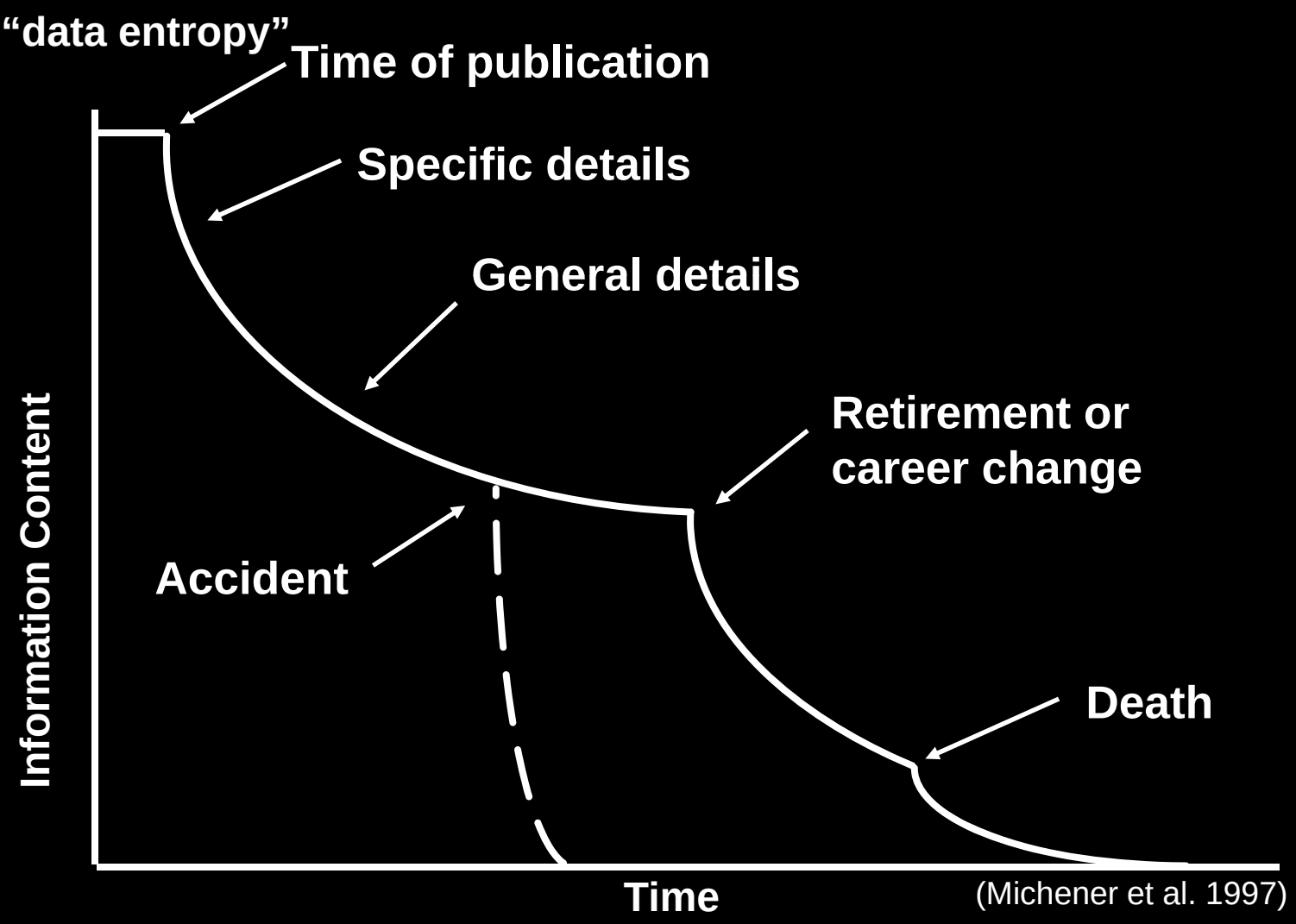
Graphs and derived data can be archived and displayed

The cyberinfrastructure for biodiversity research.

- Access
- Data organization
- Archive

The cyberinfrastructure for biodiversity research.

Poor data practice



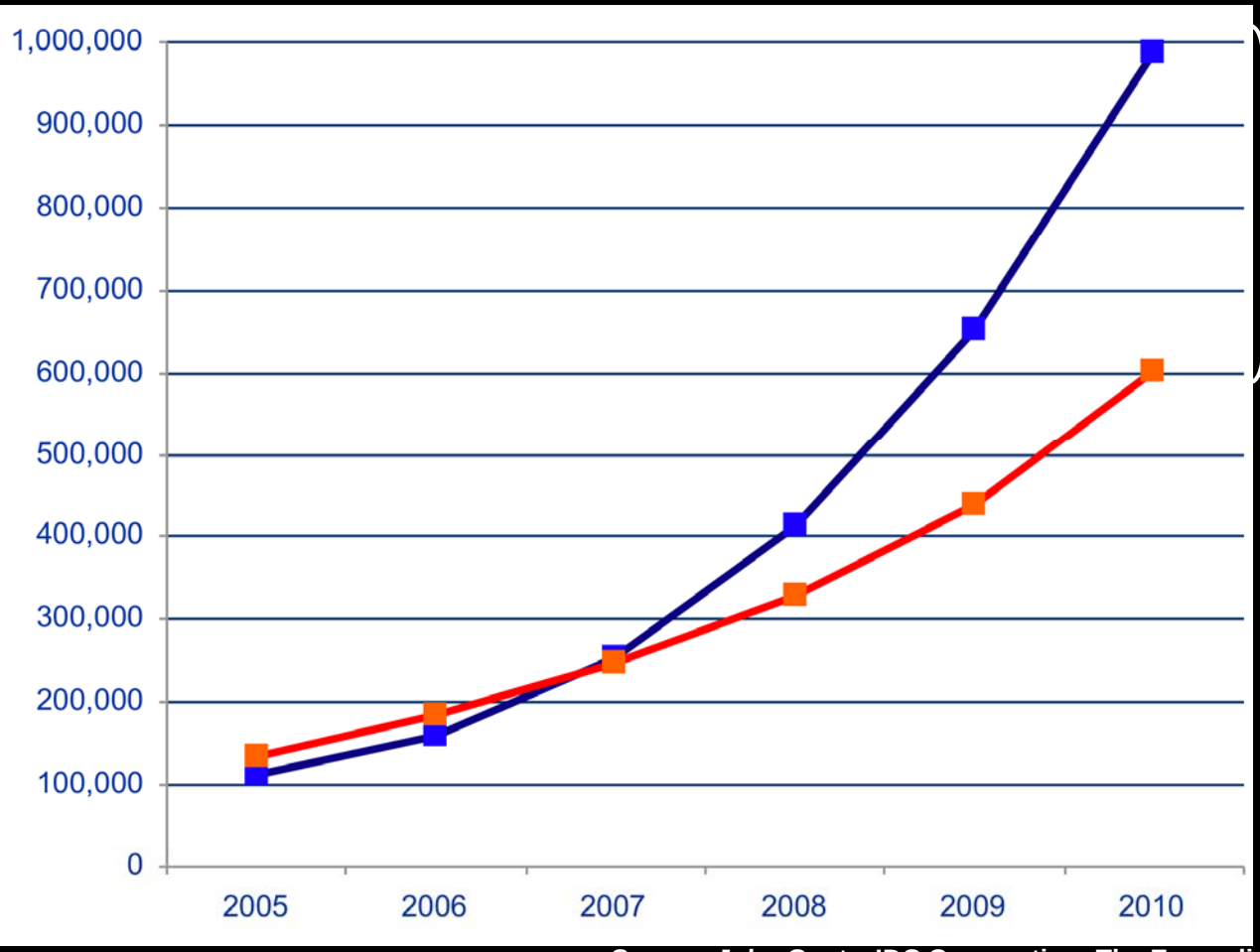
(Michener et al. 1997)

The cyberinfrastructure for biodiversity research.

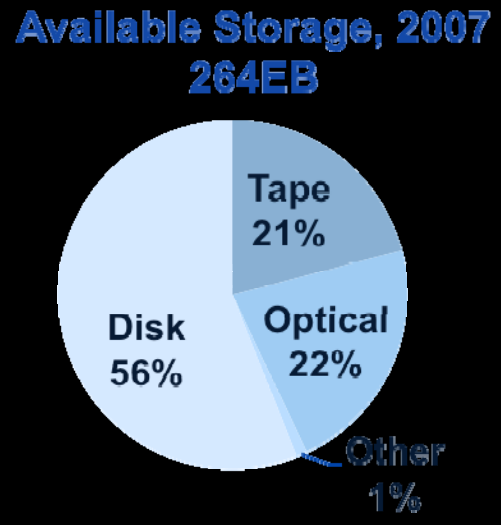
Data loss

- Natural disaster
- Facilities infrastructure failure
- Storage failure
- Server hardware/software failure
- Application software failure
- External dependencies (e.g. PKI failure)
- Format obsolescence
- Legal encumbrance
- Human error
- Malicious attack by human or automated agents
- Loss of staffing competencies
- Loss of institutional commitment
- Loss of financial stability
- Changes in user expectations and requirements

The cyberinfrastructure for biodiversity research.



Transient information or unfilled demand for storage



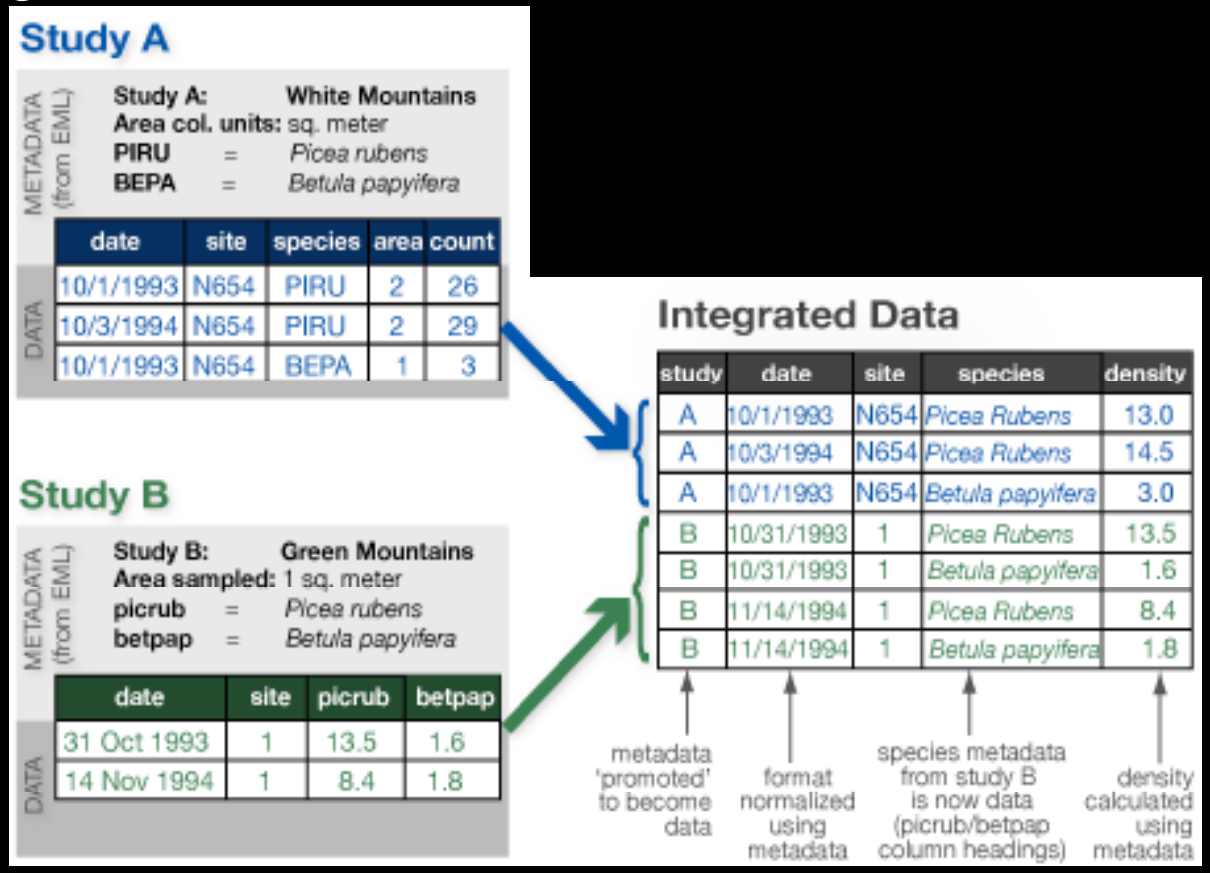
Source: John Gantz, IDC Corporation: The Expanding Digital Universe

The cyberinfrastructure for biodiversity research.

Data deluge

“the flood of increasingly heterogeneous data”

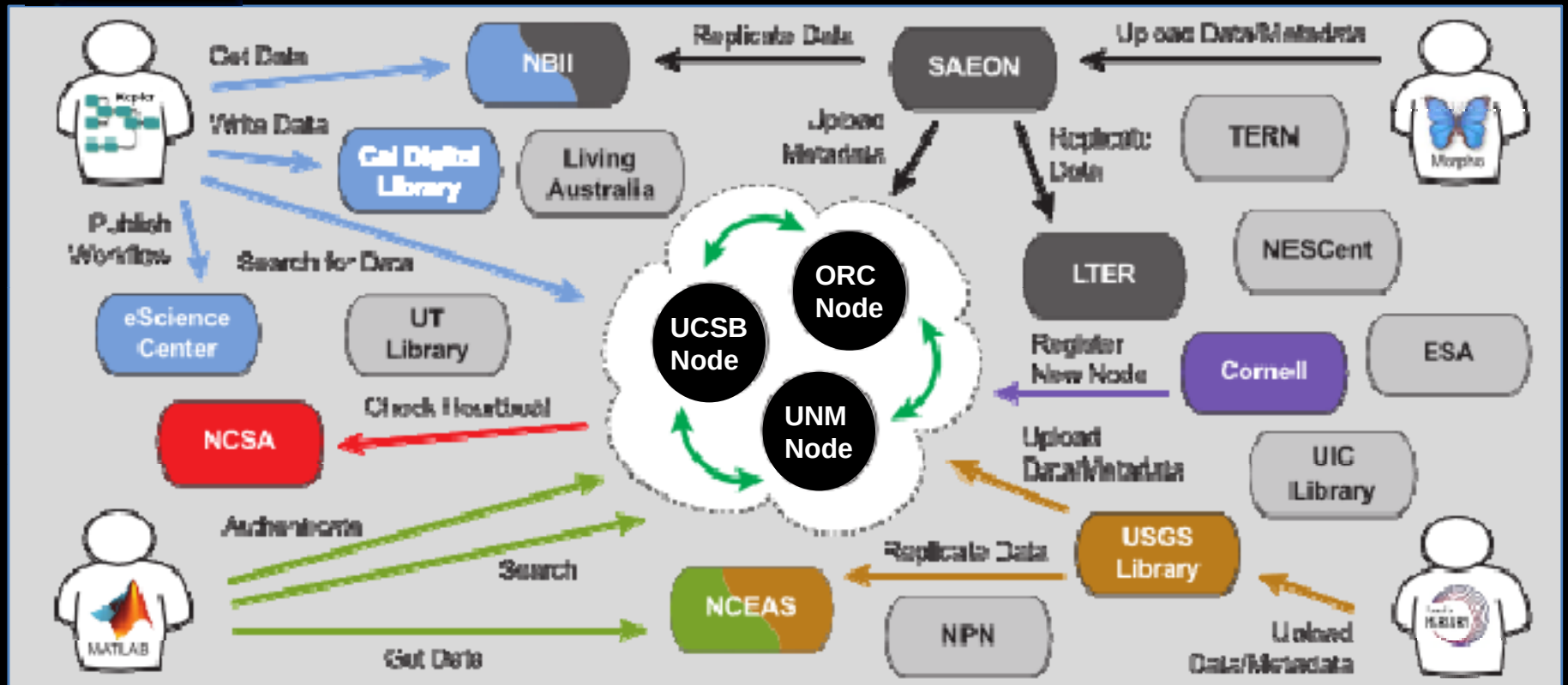
- Data are heterogeneous
  - Syntax
    - (format)
  - Schema
    - (model)
  - Semantics
    - (meaning)



Jones et al. 2007

The cyberinfrastructure for biodiversity research.

Supporting the data lifecycle



The data lifecycle

1. Deposition/acquisition/ingest
2. Curation and metadata management
3. Protection, including privacy
4. Discovery, access, use, and dissemination
5. Interoperability, standards, and integration
6. Evaluation, analysis, and visualization

Building global communities of practice: … creating long-lived CI enterprises,

- Broad, active community engagement
  - Involvement of library and science educators engaging new generations of students in best practices
  - Existing outreach and education programs
- Transparent, participatory governance
- Adoption/creation of innovative and sustainable business and organizational models

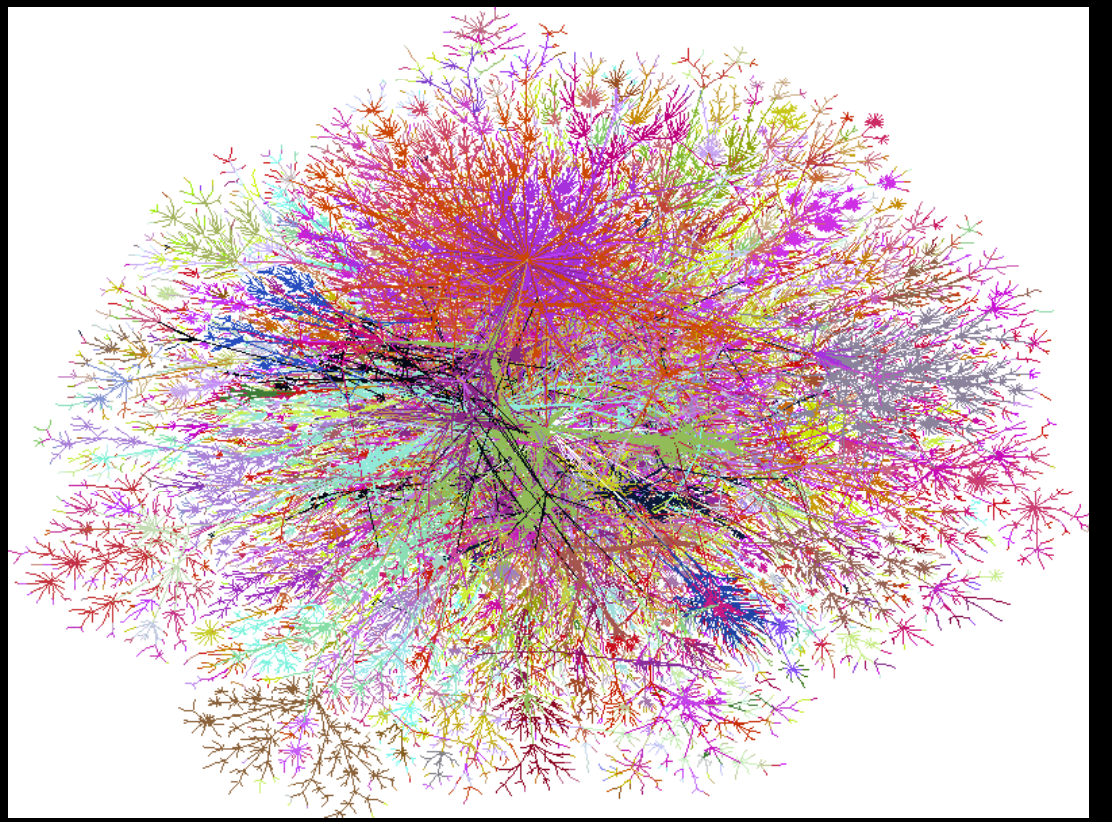


The Earth Observation Network

Metcalf's Law

The value of a network grows by the square of the size of the network.

- Sensors
- Sensor Networks
- Observational Data



Global Internet Network Image from the Lumeta Internet Mapping Project

The Earth Observation Network
Sensors, sensor networks, and remote sensing gather observations.

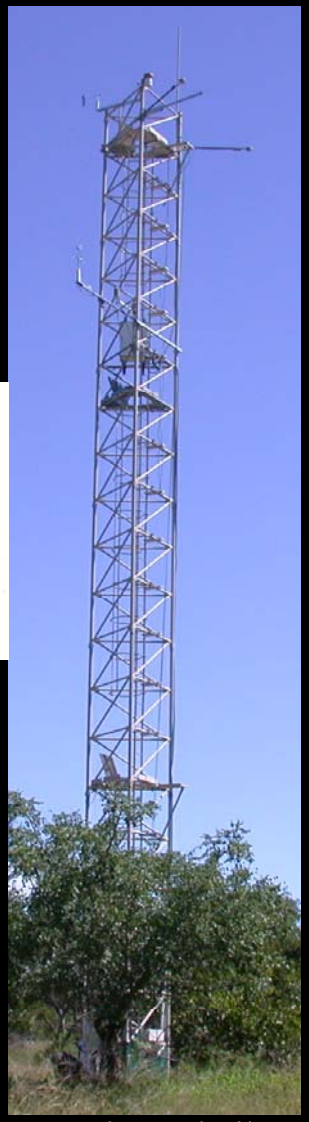
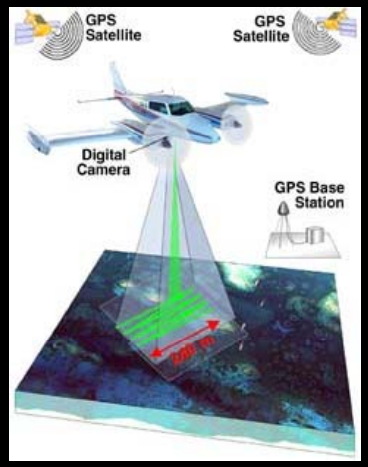
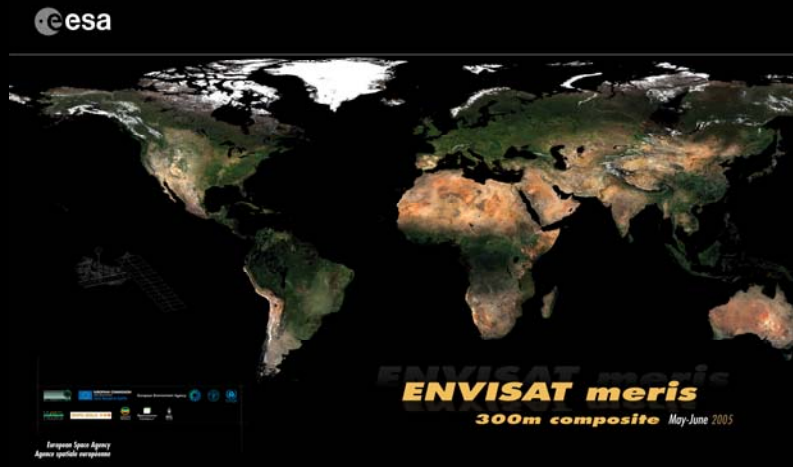
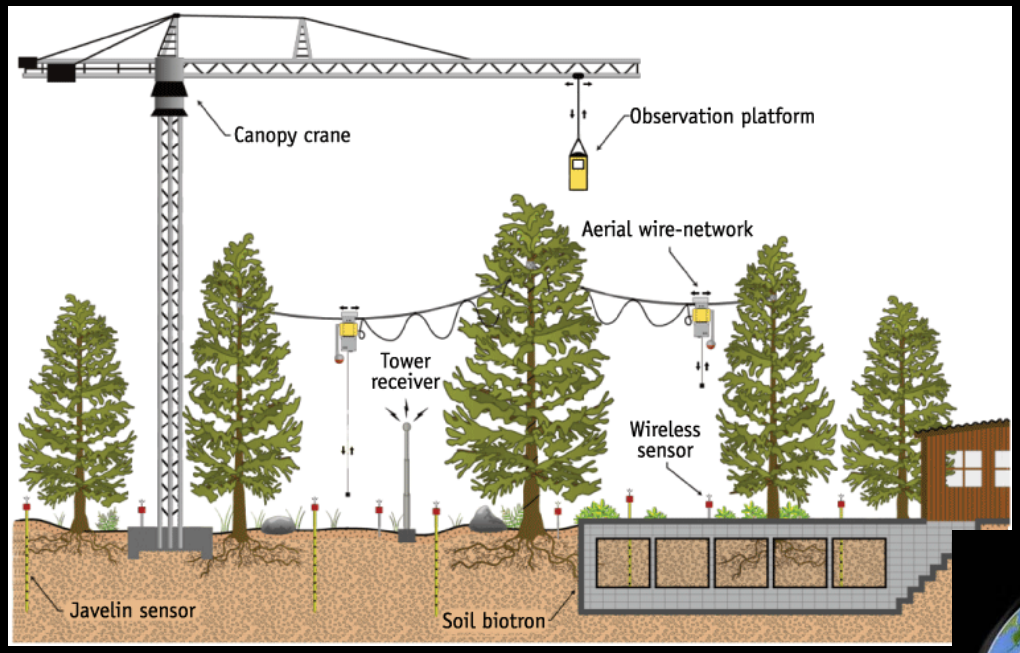


Photo courtesy of www.carboafrica.net

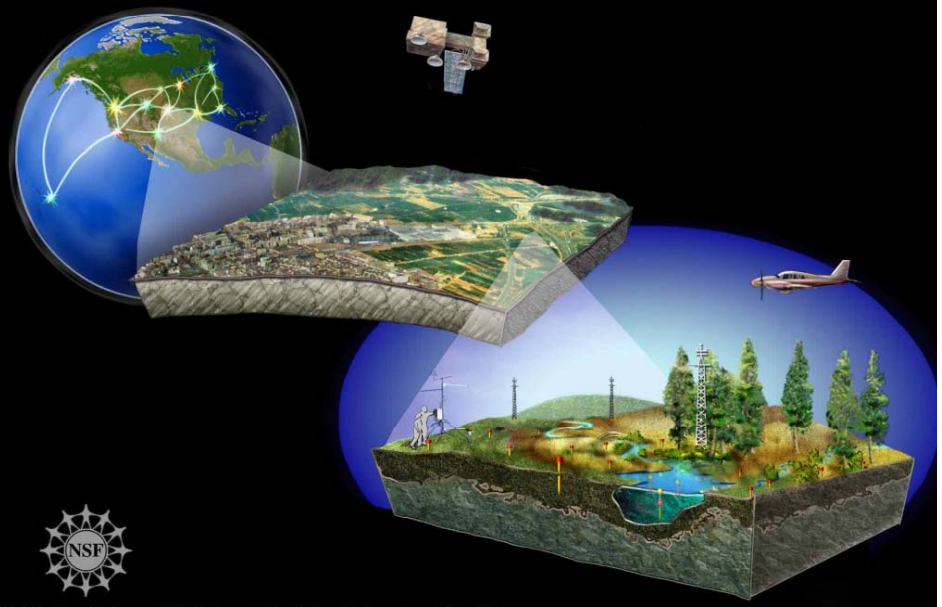




The Earth Observation Network
Sensors, remote sensing, sensor networks, and observational data



neon

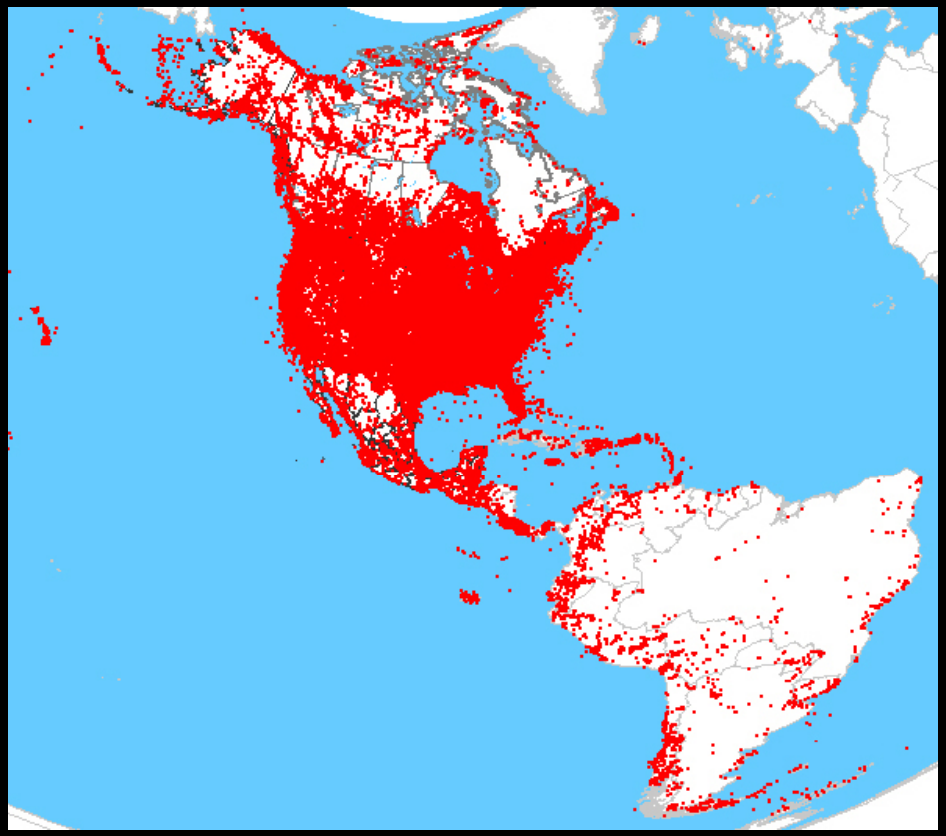


Modified from the image of Nicolle Rager Fuller, National Science Foundation, 2007

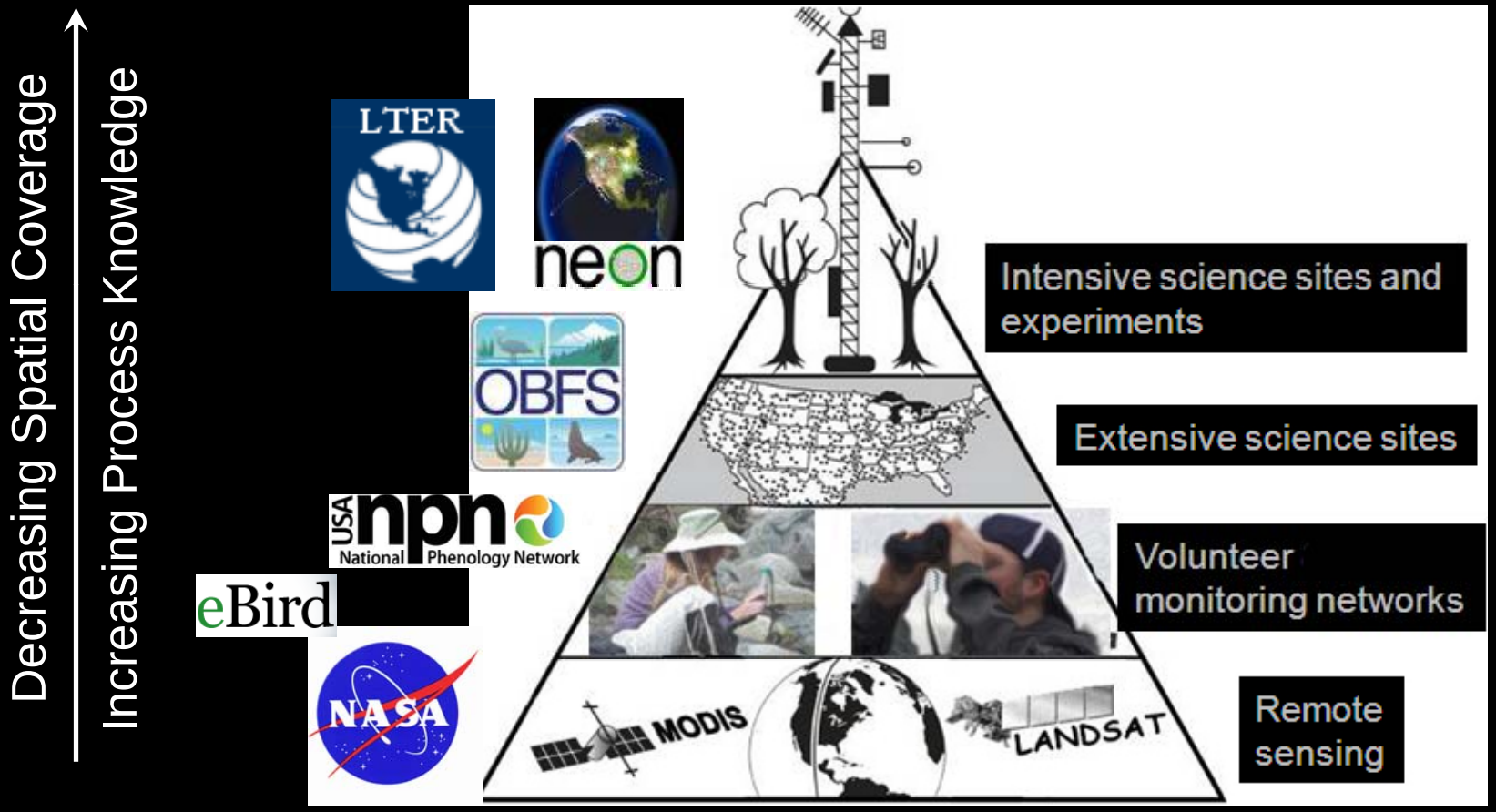
The Earth Observation Network

OBFS Organization of Biological Field Stations

The Earth Observation Network

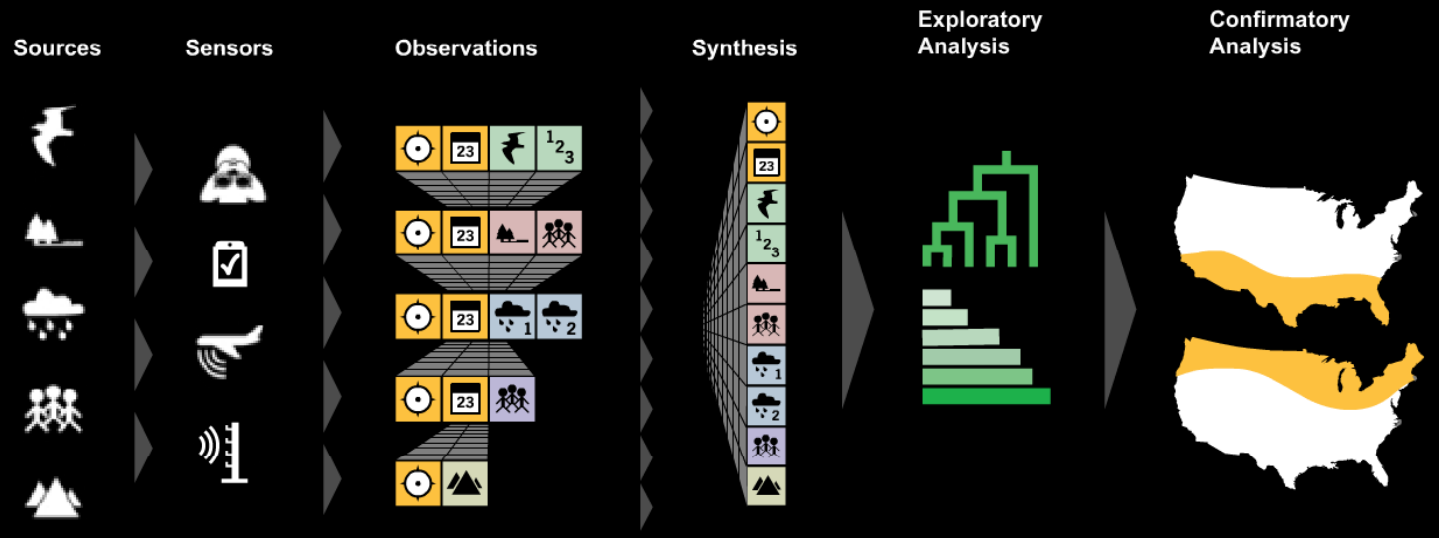


The Earth Observation Network



Adapted from CENR-OSTP

Data Intensive Science

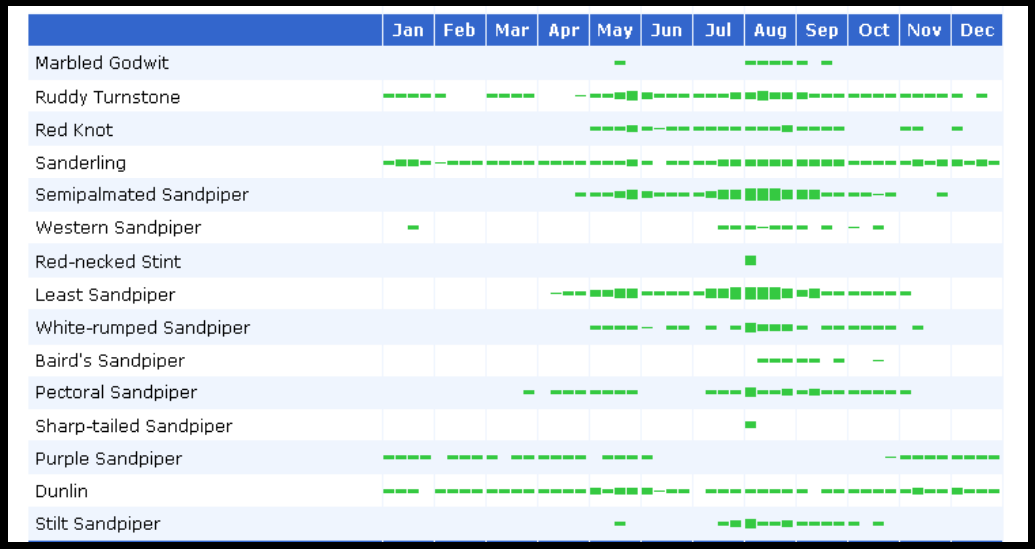


Avian Knowledge Network http://avianknowledge.net

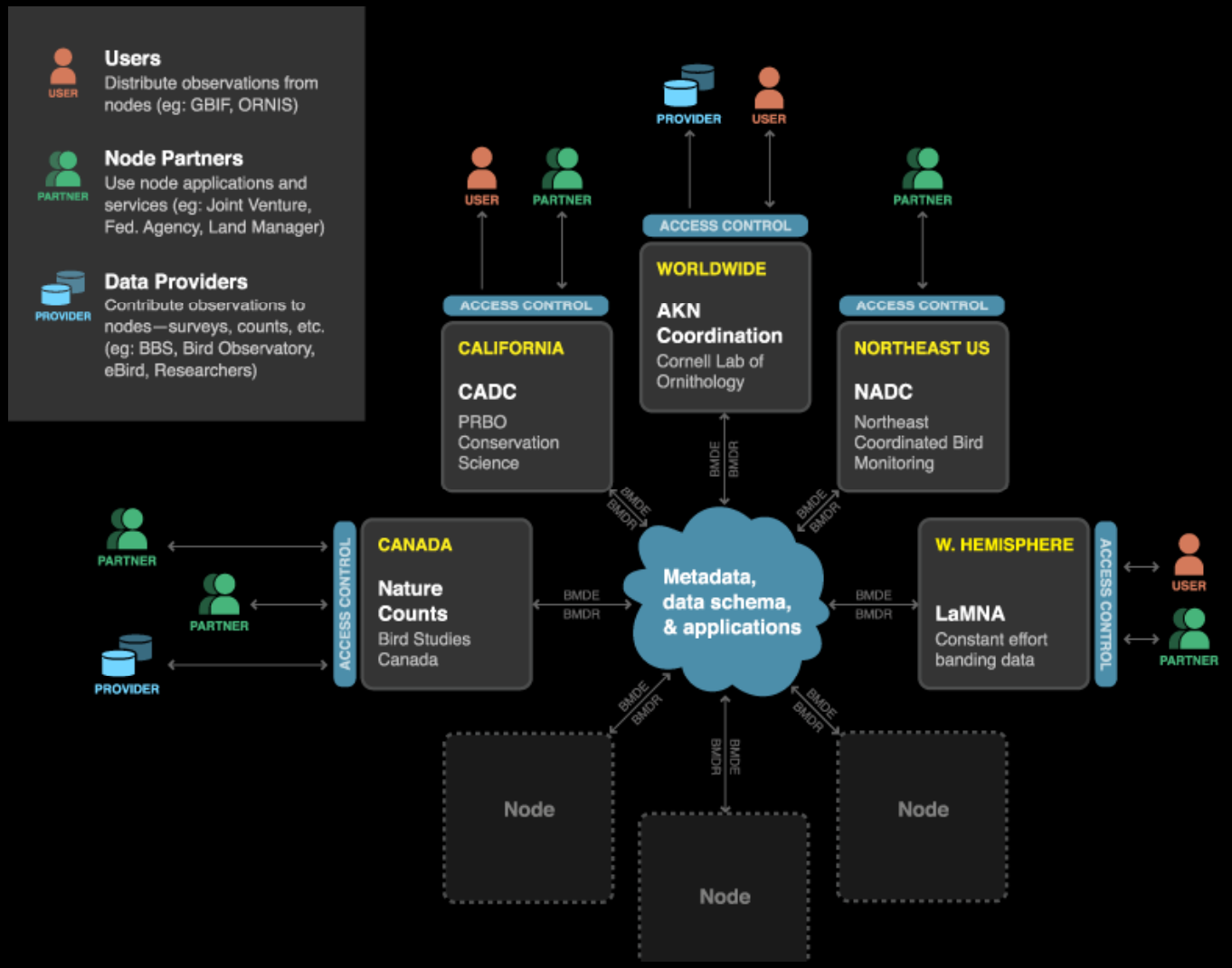
Access to data in a standardized format

Tools to explore and visualize data

New analysis techniques to discover patterns of species occurrence



Avian Knowledge Network



Avian Knowledge Network

55 Projects

have contributed

~ 50 million Observations

and

~ 10 million Banding Records

from

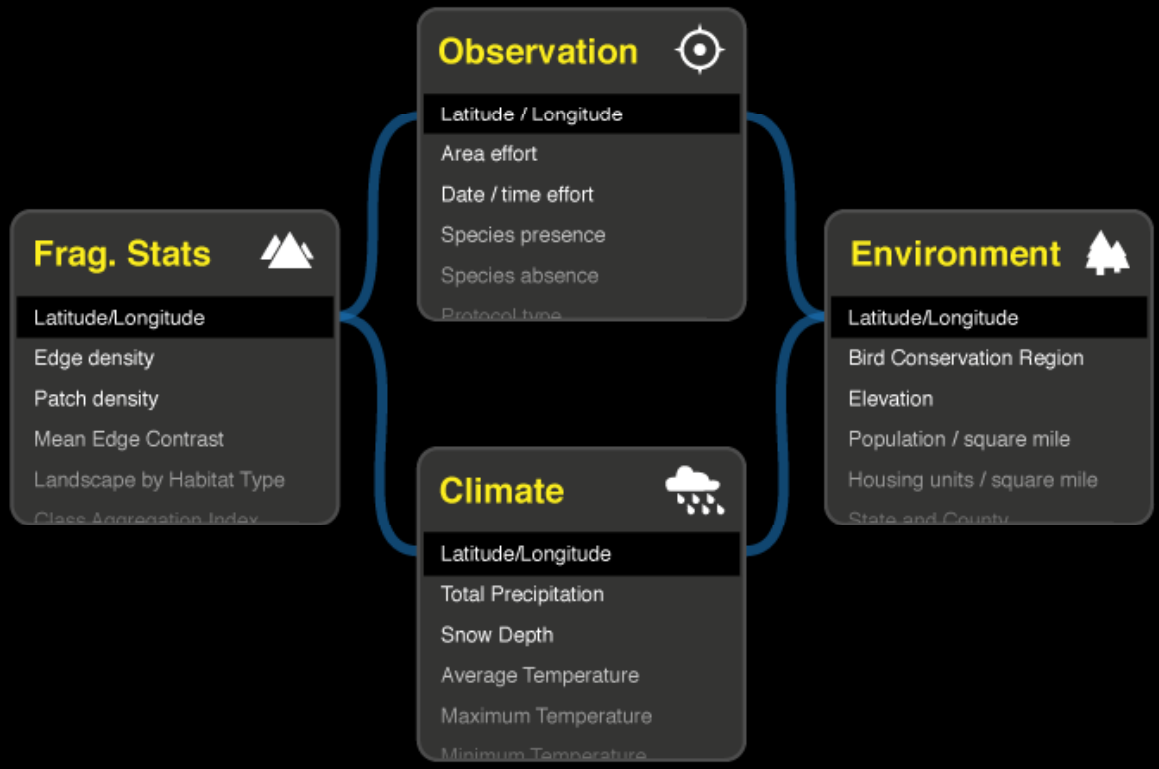
750,000+ Locations across North America

with each location linked to

1300 Climate, Land Cover, Anthropogenic, and Geographic Features

Avian Knowledge Network

Data Synthesis and Access



The eBird Reference Dataset

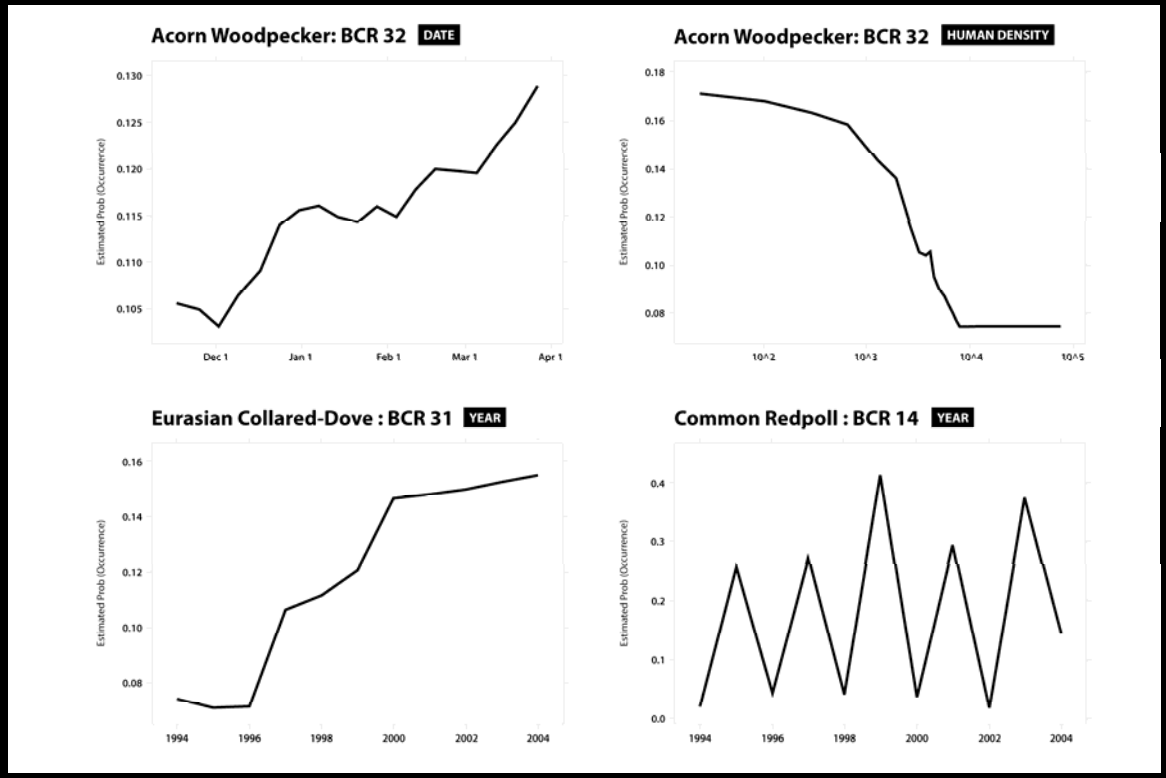
M. Arthur Munson Dan Sheldon Kevin Webb

May 20, 2009

http://www.avianknowledge.net

Avian Knowledge Network

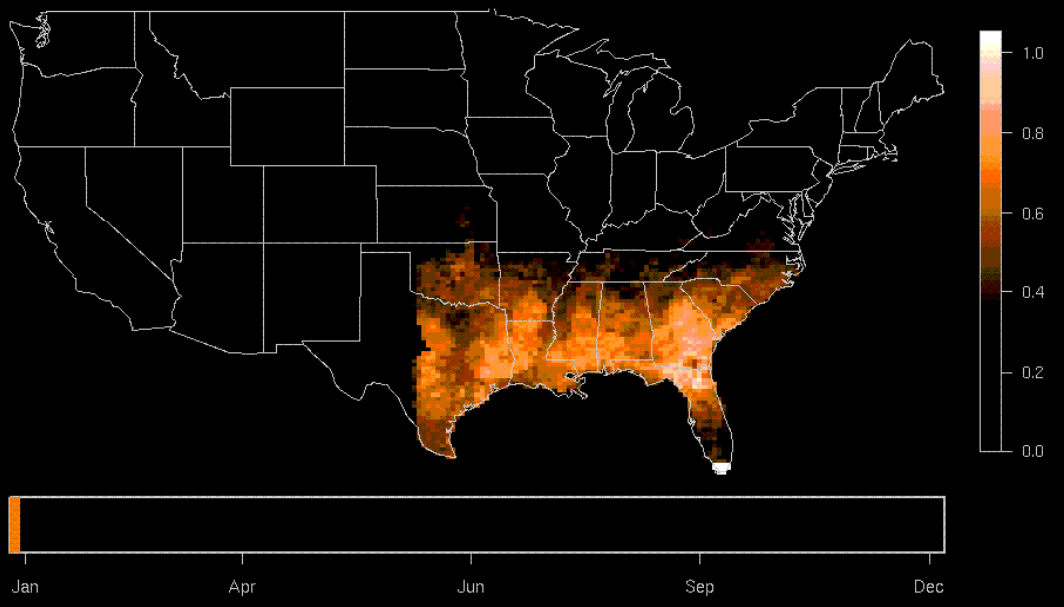
Exploratory Analysis: Partial Dependency Plots using Bagged Decision Trees



Avian Knowledge Network

Exploratory Analysis: Modeling Dynamic Patterns of Species Occurrence

Eastern Phoebe



Sullivan et al *Biological Conservation* 2009

Biodiversity Research and Conservation in a Digital World

Gaining insight into the complexities and processes of natural systems is no longer an exclusive realm of theory and experiment; computation is now an equal and indispensible partner for advances in scientific knowledge, land management, and informed decision making.

Biodiversity Research and Conservation in a Digital World
Acknowledgements:

AKN

Art Munson - *CU*

Daniel Fink - *CU*

Wesley Hochachka - *CU*

Grant Ballard - *PRBO*

Denis Lepage - *BSC*

Rich Caruana - *MS*

Mirek Riedewald - *NEU*

Daria Sorokina - *CMU*

Kevin Webb - *CU*

Giles Hooker – *CU*

CJ Ralph – *USFS*

Brian Sullivan – *CU*

Will Morris - *CU*

Computational Sustainability

Carla Gomes - *CU*

Tom Dietterich - *OSU*

Daniel Sheldon - *CU*

Ken Rosenberg – *CU*

Rebecca Hutchinson – *OSU*

Weng-Keen Wong – *OSU*

Megan MacDonald – *CU*

Stefan Hames - *CU*

DataONE

Bill Michener - *UNM*

Suzie Allard – *UT*

John Cobb – *ORNL*

Bob Cook – *ORNL*

Patricia Cruse – *CDL*

Mike Frame – *USGS*

Stephanie Hampton – *UCSB*

Viv Hutchison – *USGS*

Matt Jones – *UCSB*

Kathleen Smith - *Duke*

Carol Tenopir – *UT*

Bruce Wilson – *Joint ORNL – UT*

NSF nbii LEON LEVY FOUNDATION Wolf Creek Charitable Foundation