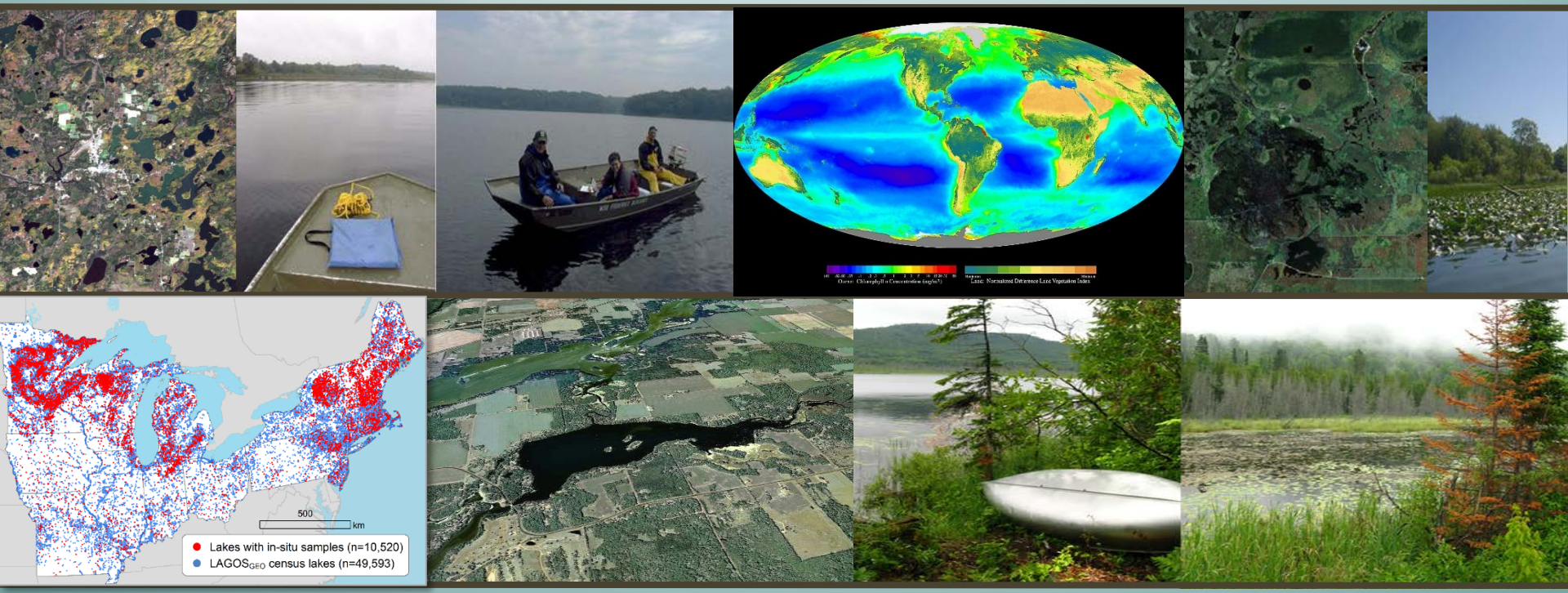
Ecology in the 21st century:
Team-based, data-intensive, and open



Patricia Soranno & Kendra Spence Cheruvelil

Dept. Fisheries & Wildlife, & Lyman Briggs College

Michigan State University

September 16, 2016

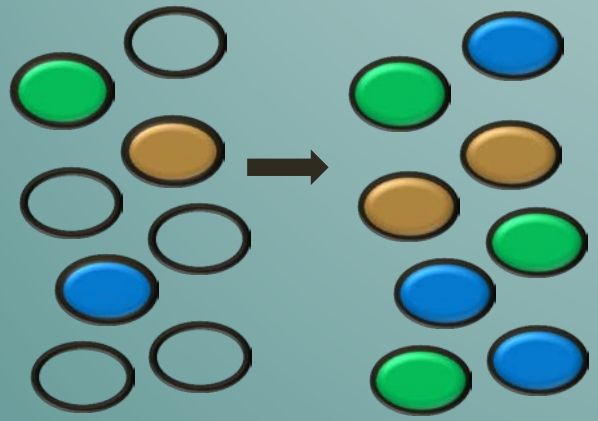
Some important challenges in ecology

(1) At broad scales -- ***Response of biota & ecosystem services to global change***

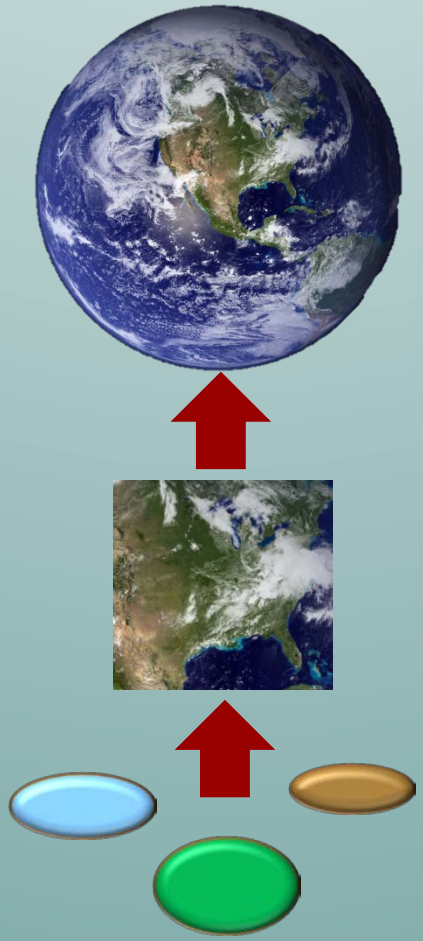
Some important challenges in ecology

*(2) At multiple scales: **Apply knowledge** from one scale, site or time period to another*

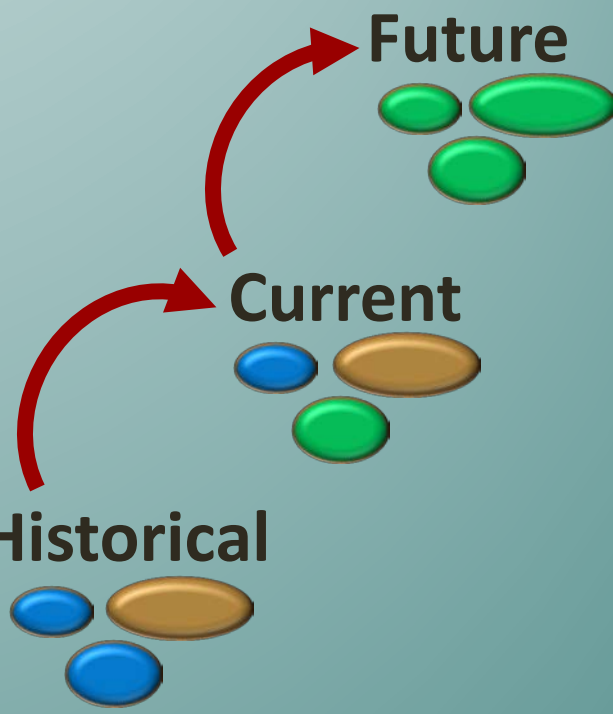
Extrapolation



Scaling up

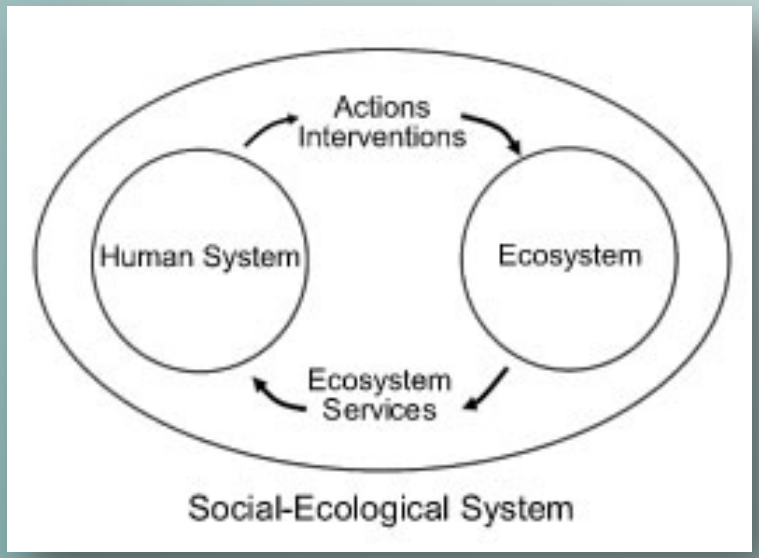


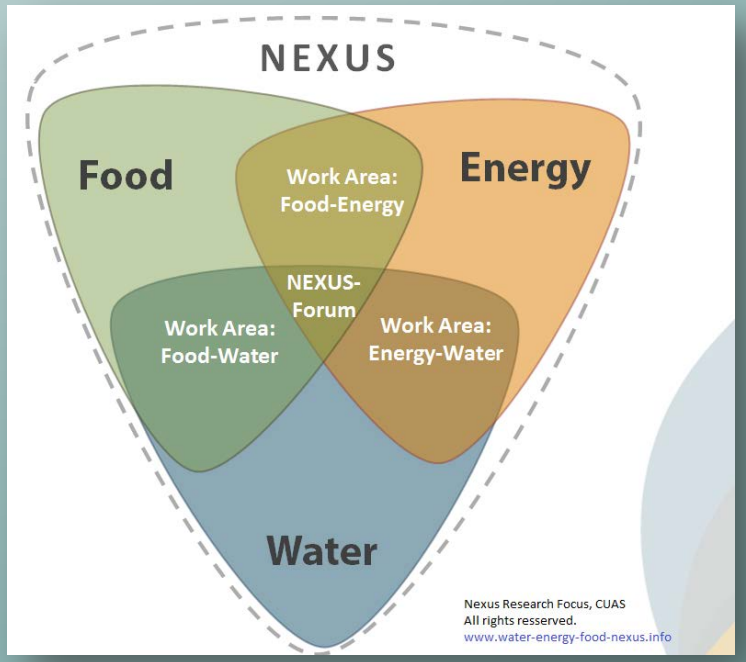
Forecasting



Some important challenges in ecology

(3) At more local scales – ***The role of ecological processes at interface of disciplines****: Socio-ecol. systems; FEW-Nexus*





Addressing these challenges will require research that is:

1. Conducted in **teams** of individuals with different expertise and sometimes disciplines

2. **Data-intensive** to help to address MANY of the challenges

3. Based on **sharing of data, models, and code**

BUT, the past norms and practices of ecology include…

1. The dominance of single-investigator science

2. Research efforts that are mostly data-limited *(i.e., good data, but not enough of it!)*

3. Data that are, for the most part, not shared outside of research teams

OVERVIEW: The new norms are…

1. Science is increasingly conducted by teams

2. Data-intensive research is on the rise

3. Data sharing is needed for emerging research areas

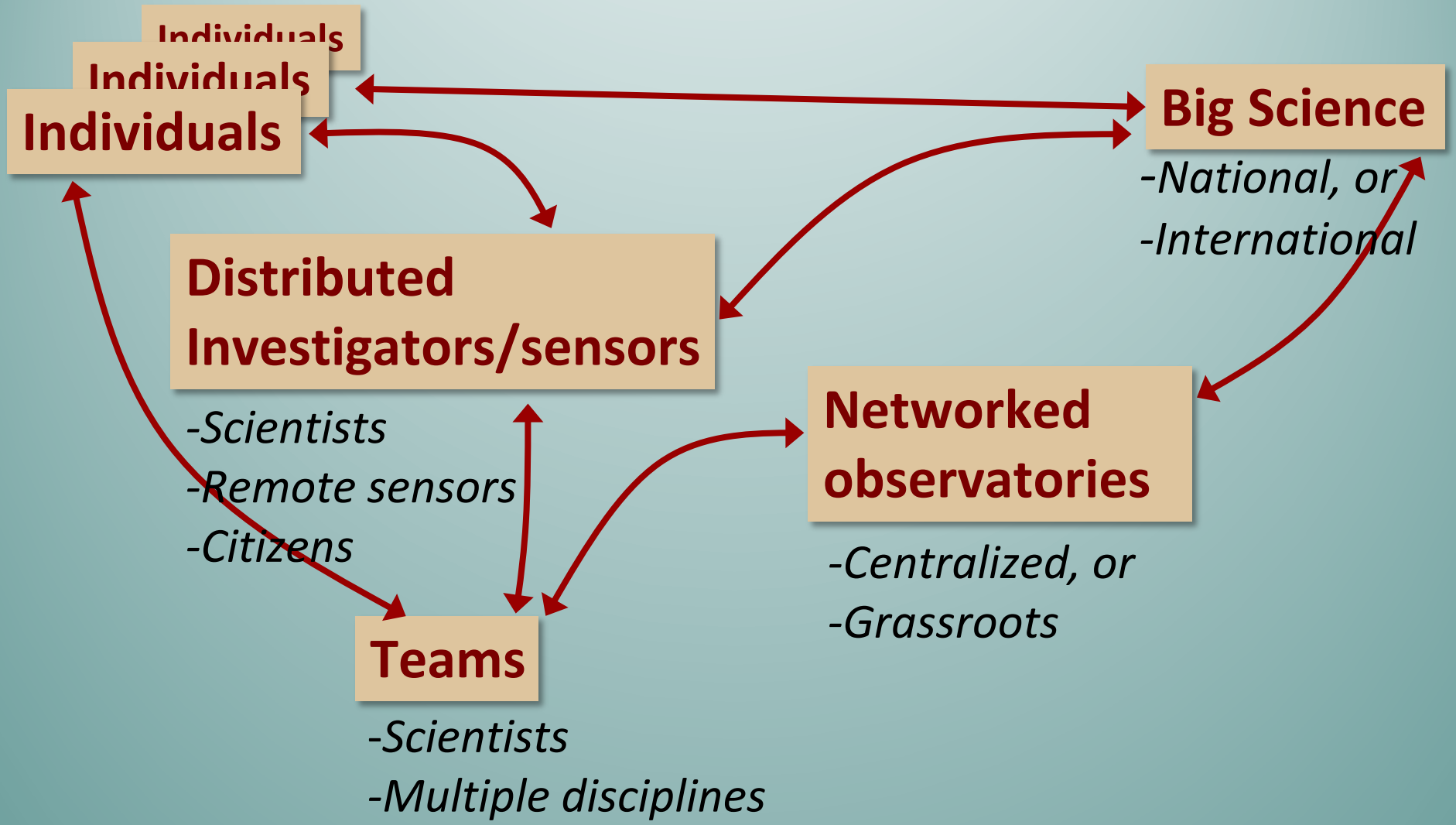
For each, lets consider the:

- *Context*
- *Resistance to adopting new practices related to the norms*
- *Recommendations for moving forward*

OVERVIEW: The new norms are…

1. **Science is increasingly conducted by teams**
2. Data-intensive research is on the rise
3. Data sharing is needed for emerging research areas

Science is predominantly conducted by teams



What makes an effective & productive team?

KS Cheruvelil

PA Soranno

KC Weathers

PC Hanson

SJ Goring

CT Filstrup

EK Read

Cheruvelil et al. 2014. Frontiers in Ecology and the Environment

What makes an effective & productive team?

What is related to group 'collective intelligence'?

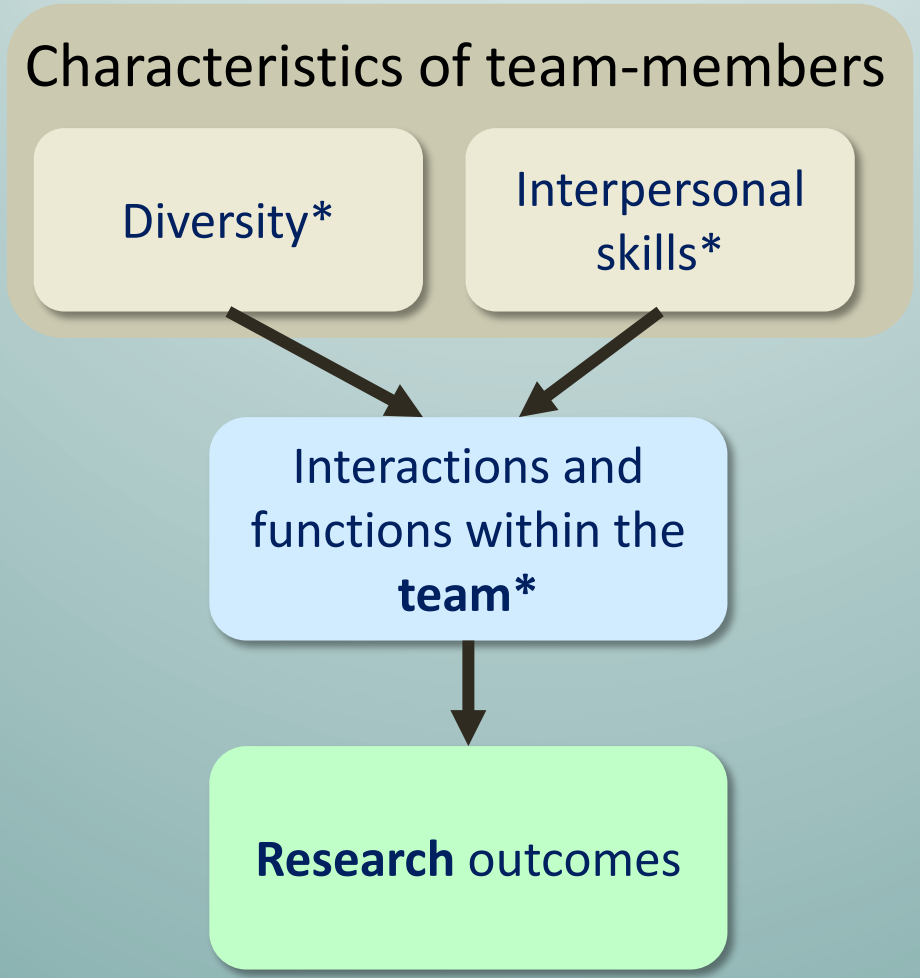
Woolley et al. 2010. *Science*

"**Collective intelligence**" of groups is NOT strongly correlated with the average or maximum individual intelligence of members.

It is correlated with the average 'social sensitivity' of group members – i.e., equality in distribution of speaking turn-taking

* AND, it works for face-to-face AND virtual teams!

What makes an effective & productive team?



** Each item supported by research on teams*

Redrawn from *Cheruvelil et al. 2014*

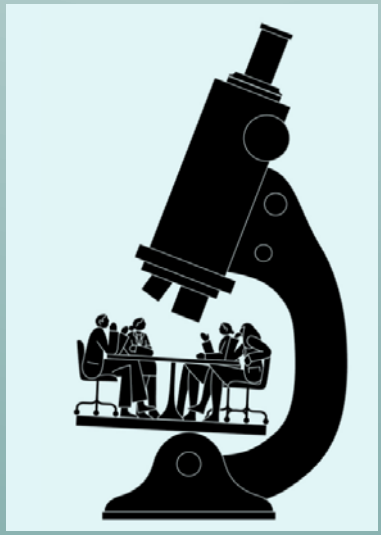
What makes an effective & productive team?

February 2016

QUARTZ

IMAGINE THAT

After years of intensive analysis, Google discovers the key to good teamwork is being nice



What makes an effective & productive team?

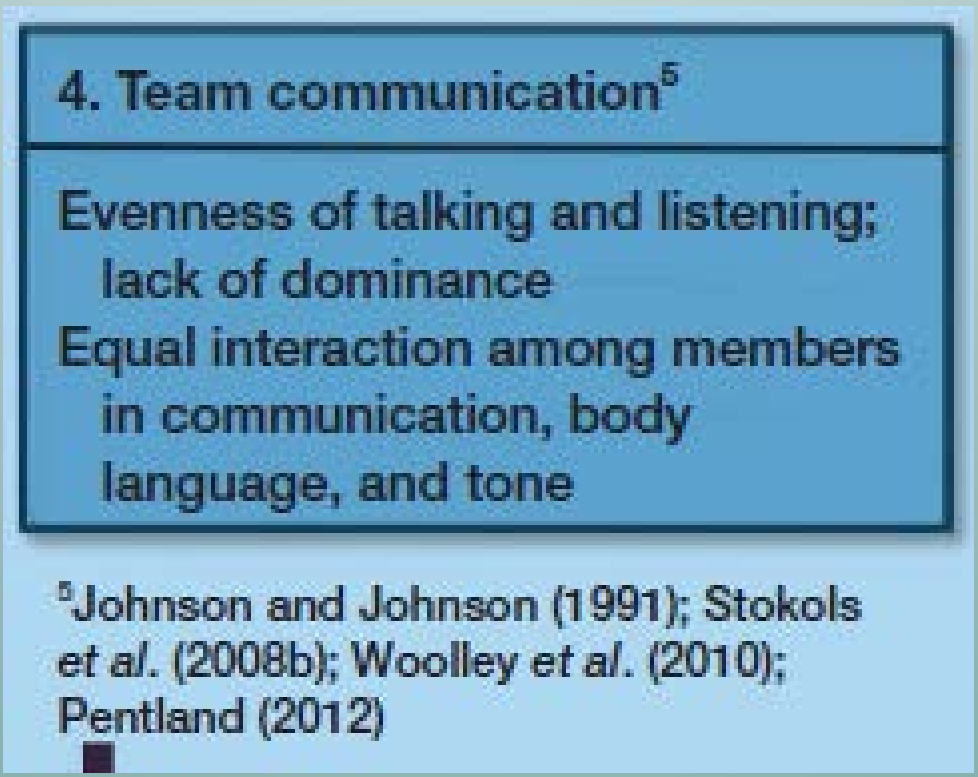
In total, there were five dynamics the researchers found to be significant for team effectiveness:

1. **Psychological safety**: Can we take risks on this team without feeling insecure or embarrassed?
2. **Dependability**: Can we count on each other to do high quality work on time?
3. **Structure & clarity**: Are goals, roles, and execution plans on our team clear?
4. **Meaning of work**: Are we working on something that is personally important for each of us?
5. **Impact of work**: Do we fundamentally believe that the work we're doing matters?

https://rework.withgoogle.com/blog/new-guide-and-tools-understand-team-effectiveness/

What makes an effective & productive team?

- Social sensitivity
- Emotional engagement
- Trust



Redrawn from *Cheruvelil et al. 2014*

How can you make and keep great teams?

How can you make and keep great teams?

<u>Exercises and discussions to:</u>

- *Appreciate & understand **team heterogeneity***
- *Build **social sensitivity***
- *Work on **team functioning***
- *Develop **written team policies***

Cheruvelil et al. 2014

I don’t do touchy-feely stuff like that.

Are you serious?

I am NOT doing that.

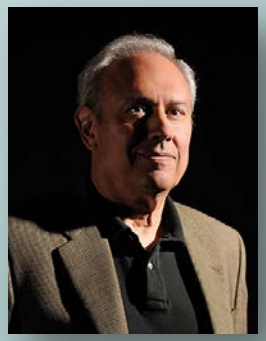
Better science outcomes

Scientists aren’t like other p
We are different.

Students are very comfortable talking to me; they can knock on my door any time.

“These days, scientists spend much of their time taking ‘professional selfies’—effectively spending more time announcing ideas than formulating them.” **August 2016**

Science in the age of selfies

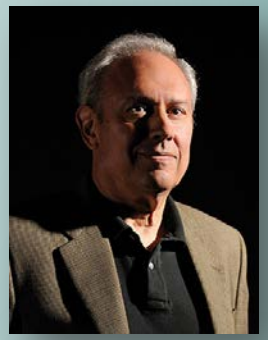


Professor of Applied Math, Johns Hopkins Univ.

James Manning Professor of Applied Math, Brown Univ.

Donald Geman, and Stuart Geman PNAS 2016;113:9384-9387

Arguments AGAINST new science norms



Donald Geman, and Stuart Geman PNAS 2016

In fact, maybe it has become too easy to collaborate.

Great ideas rarely come from teams.

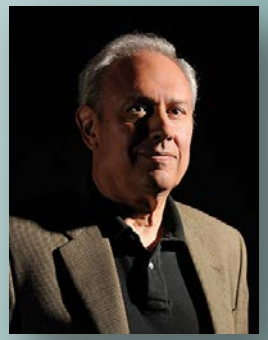
There is, of course, a role for "big science" (the Apollo program, the Human Genome Project, CERN's Large Hadron Collider), **but teamwork cannot supplant individual ideas**.

Science of the past 50 years seems to be **more defined by big projects than by big ideas.**

In a 2014 letter to *The Guardian* newspaper, 30 scientists, concerned about today's scientific culture, noted that **it was the work of mavericks like Feynman that defined 20th century science**.

* MAVERICK = An unorthodox or independent-minded person.

Arguments AGAINST new science norms



Donald Geman, and Stuart Geman PNAS 2016

SOLUTION:

Spend more time on each project, be less inclined to join large teams in small roles, and spend less time taking professional selfies. Perhaps we can then **return to a culture of great ideas and great discoveries.**

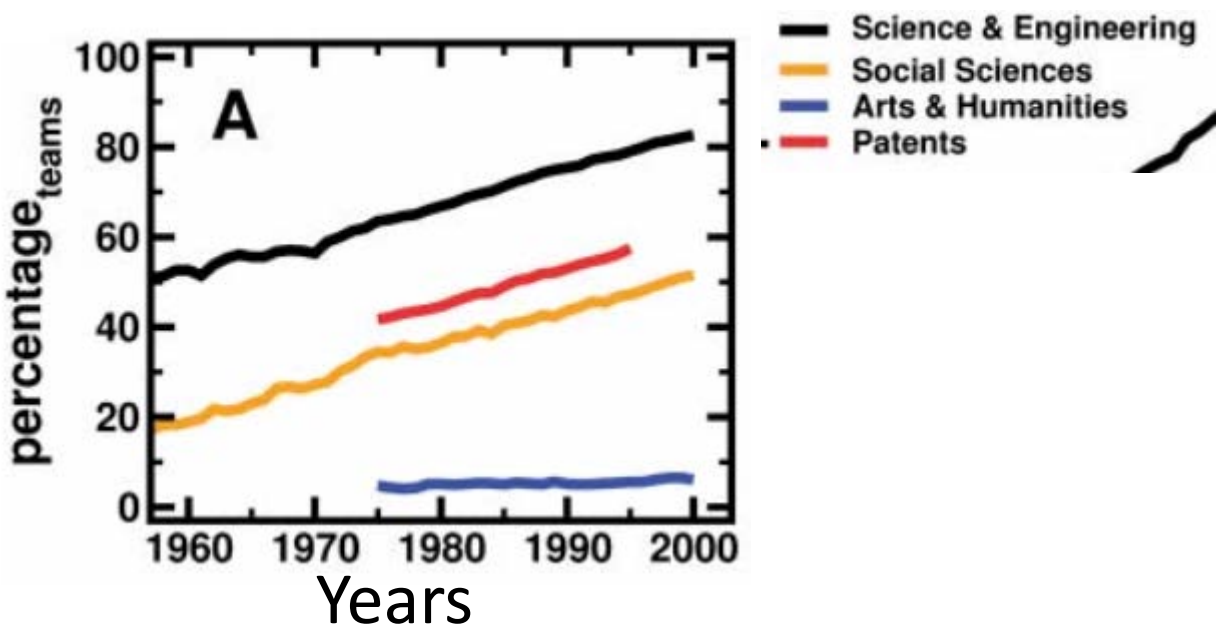
But, what do the numbers say???

Wuchty et al. 2007. Science

Tested the 'lone genius (maverick??)' hypothesis

- 19 million articles, 2 million patents, 50 yrs

% of papers written by teams is increasing

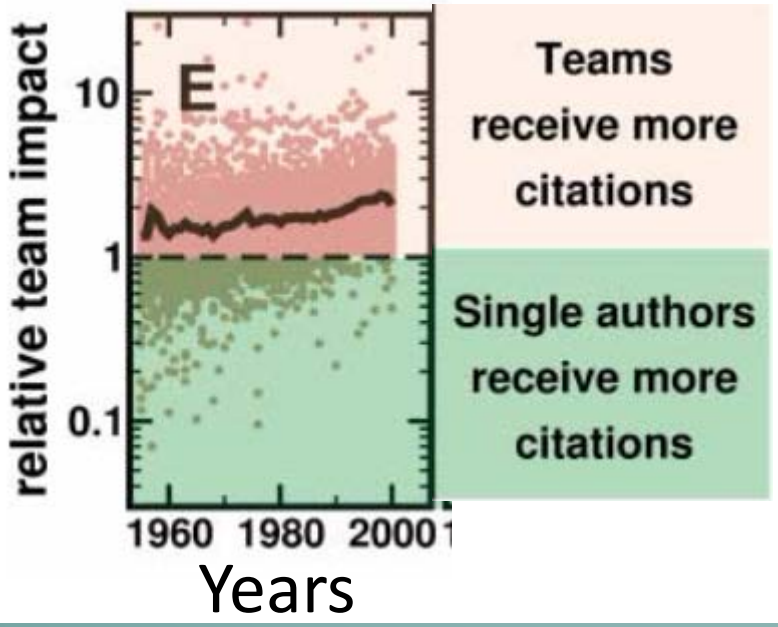


But, what do the numbers say???

Wuchty et al. 2007. Science

Tested the 'lone genius (maverick??)' hypothesis

Team papers are/have been cited at a higher rate compared to single-author papers Science & Engin.

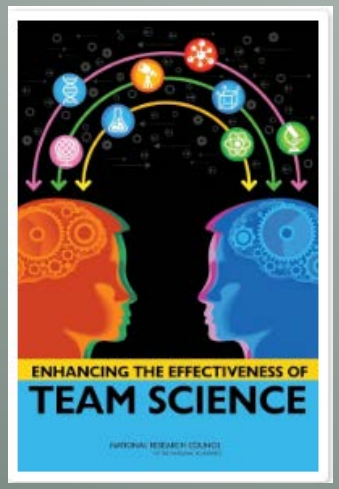


- Also happens in math, economics, sociology

Recommendations: Team Science

- **Training in collaboration**
  - Practices for true inclusion
  - Constructive & effective conflict resolution
  - Communication skills
  - Practices for maximizing creativity

- **Develop best-practices for your teams**
  - Copy what works, and share



SciTS | Building the knowledge base for effective team science

http://www.scienceofteamscience.org/

The new norms are…

1. Science is increasingly conducted by teams
2. **Data-intensive research is on the rise**
3. Data sharing is needed for emerging research areas

Data-intensive science - *Defined*

Research in which the capture, curation, and analysis of large volumes of data are central to the scientific question.

OR…research that uses datasets so large or complex they are hard to analyze using traditional methods

The Opinion Pages | OP-ED CONTRIBUTORS

Eight (No, Nine!) Problems With Big Data

By GARY MARCUS and ERNEST DAVIS APRIL 6, 2014

Professor of Psychology, NYU

Professor of Computer Science, NYU

Data-intensive science - *PROBLEMS*

- **Atheoretical**, only data-driven
- **Biased**
- **Multiple comparisons** problem
- **Expertise** missing
- Issues with using **other people's data**

Historical perspective – *Debate over hypo-driven rsch*

1600's

Hypotheses **Inductive reasoning**

"Investigators could be easily led astray if they **proposed bold conjectures** rather than working inductively from the available evidence."

- Newton

Conceptions of good science in our data-rich world

Kevin Elliott, Philosopher

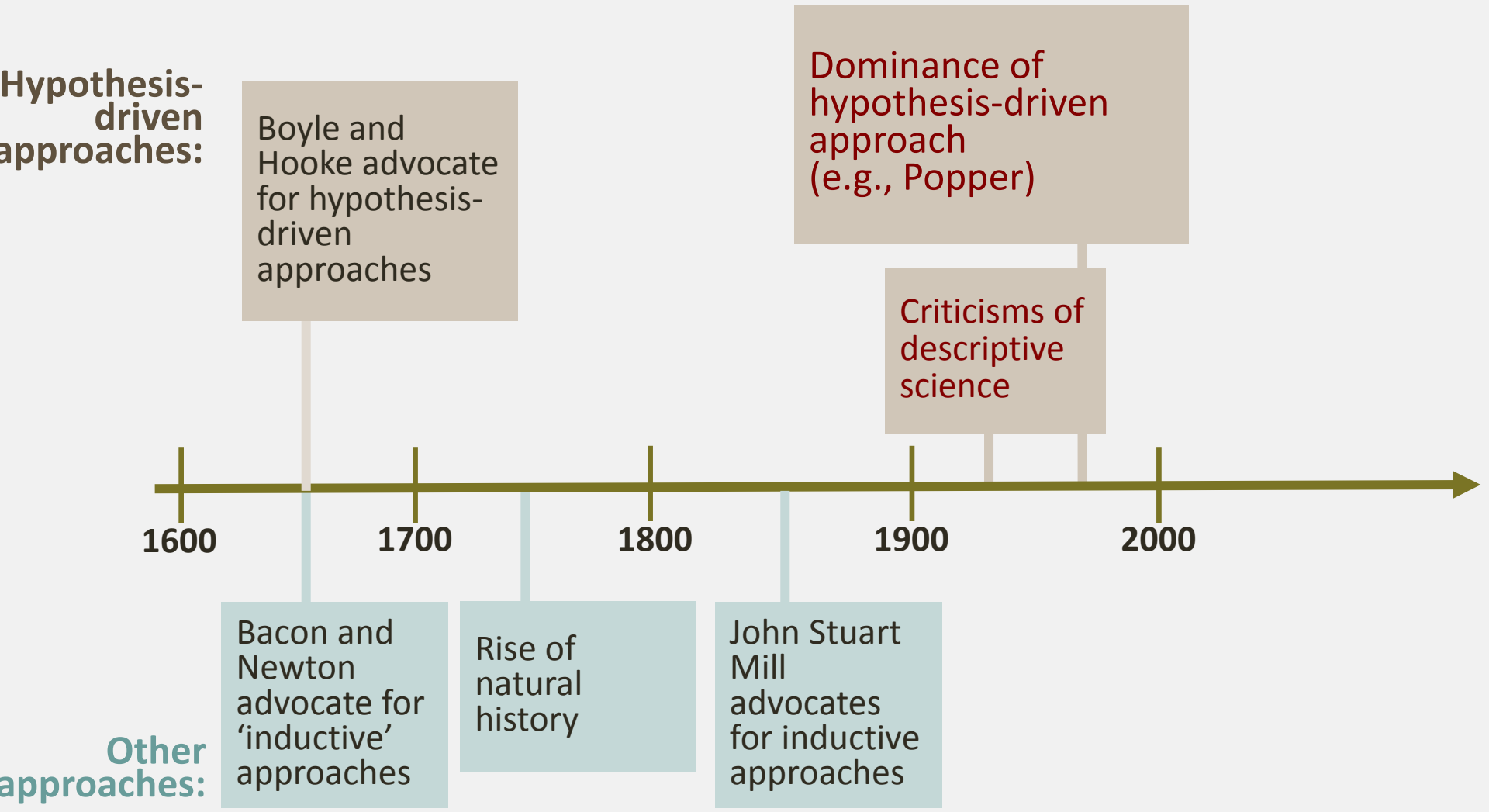
Kendra Cheruvelil, Ecologist

Georgina Montgomery, Historian

Patricia Soranno, Ecologist

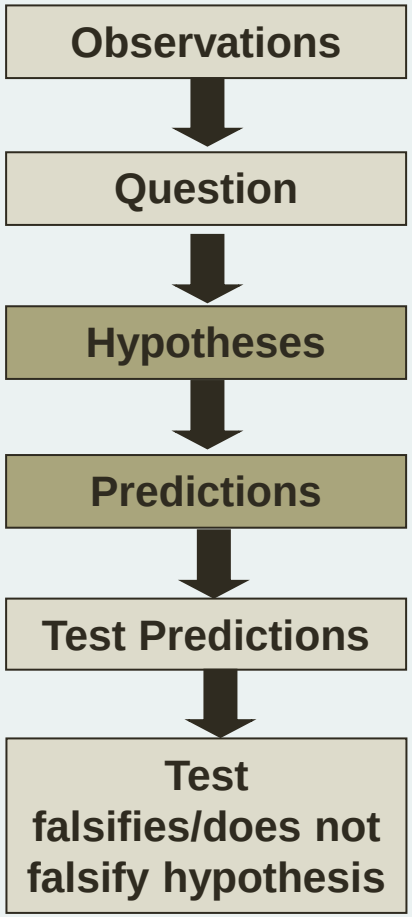
Elliott et al. in press. BioScience

Historical context of hypothesis-driven science

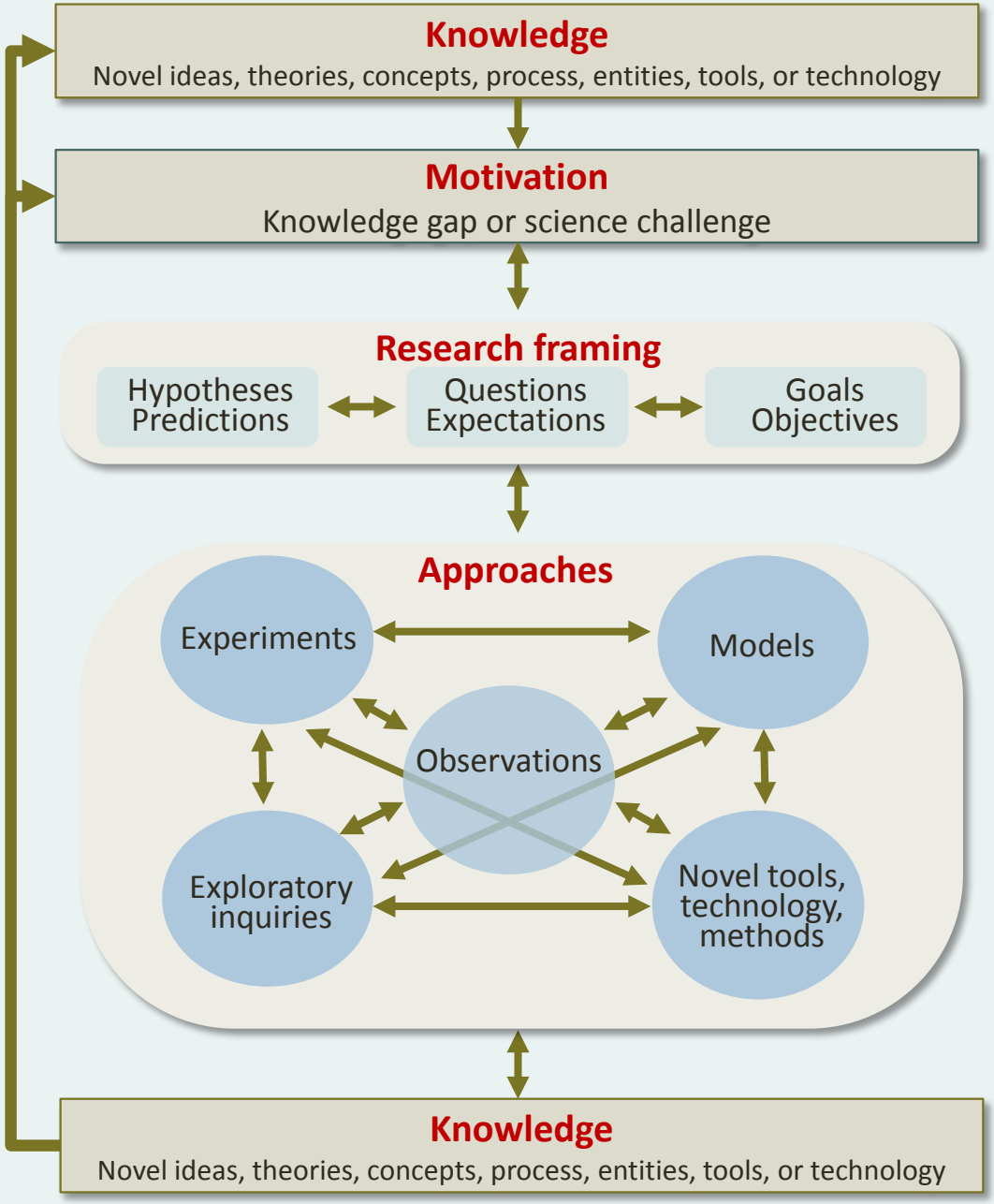


Elliott et al. in press. BioScience

Hypothesis-driven Research

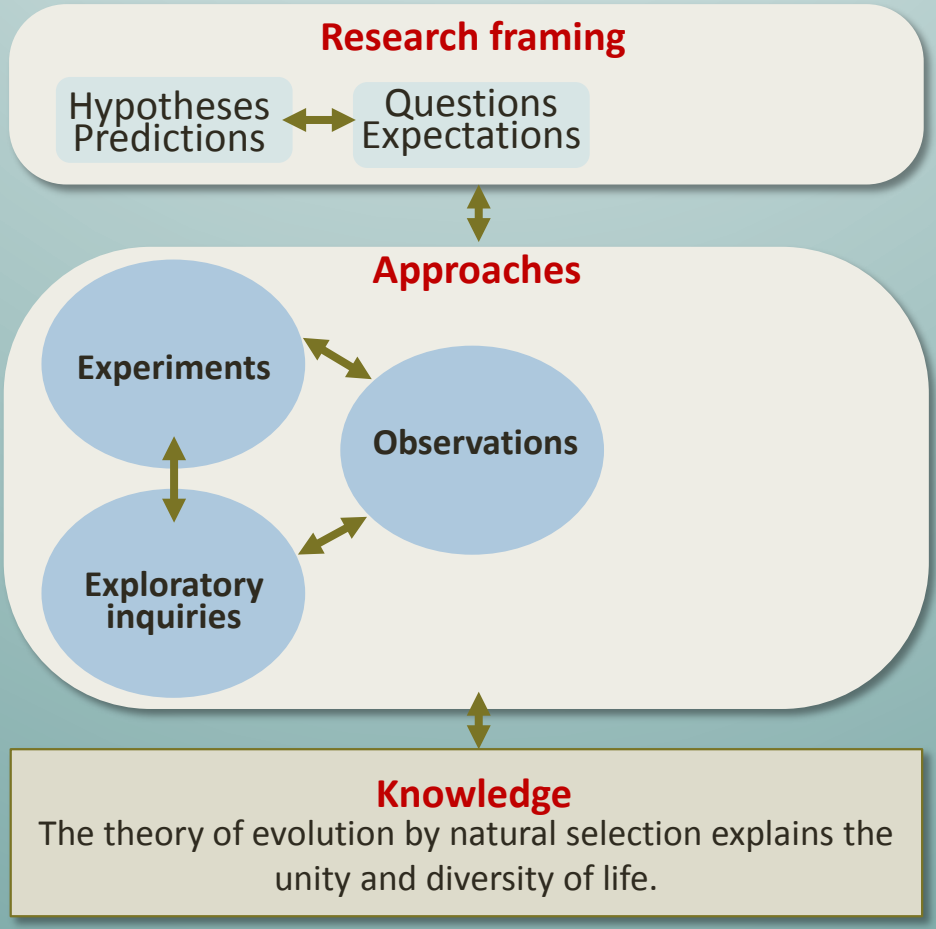


Iterative Research

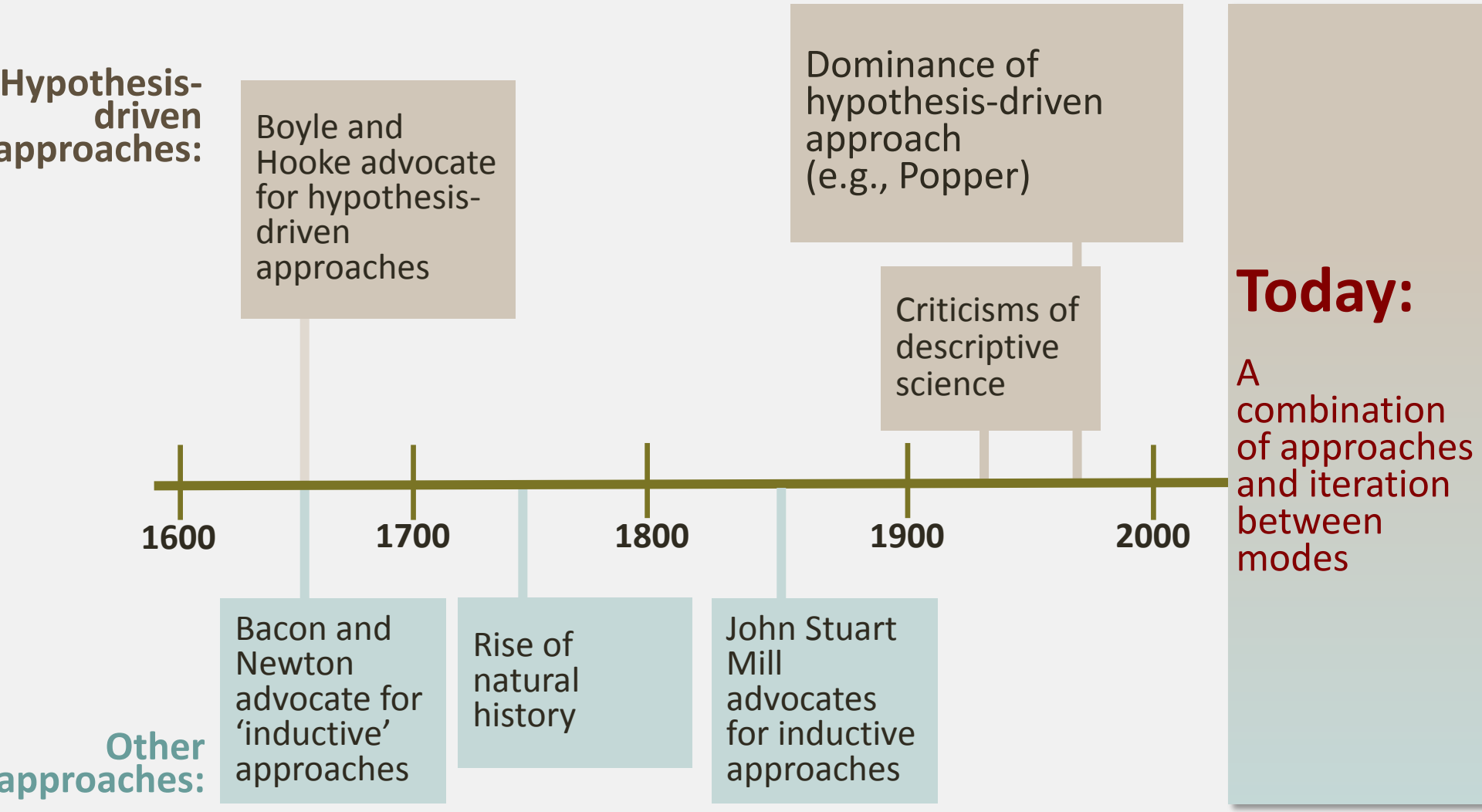


EXAMPLE: The study of evolution by natural selection

* Darwin was ‘forced’ to conform to the scientific norms of the time, with strongly inductive arguments, but he did eventually arrive at successively more general causal laws.



Historical context of hypothesis-driven science



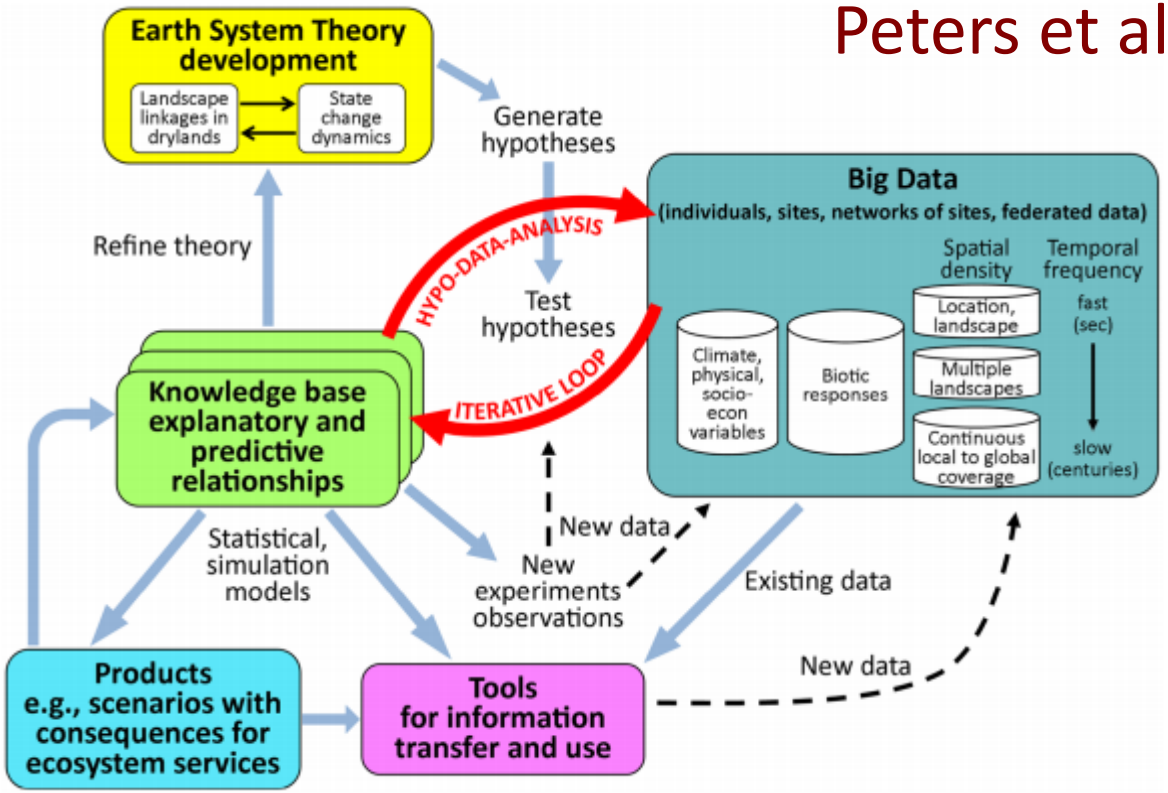
Elliott et al. in press. BioScience

Other examples recognizing the rise of data-intensive and iterative research:

Big

Knowledge Learning and Analysis system (KLAS)

Peters et al. 2014



-- Focus on **iteration between methods/tools, use of large datasets, hypotheses, theories**

RECOMMENDATIONS: Data-intensive research

- **Replace the requirement of hypothesis-testing** (in articles & proposals) with evaluation of the **alignment between knowledge gap and approach & conclusions**

- **Publish (and value) the different components of the research cycle**

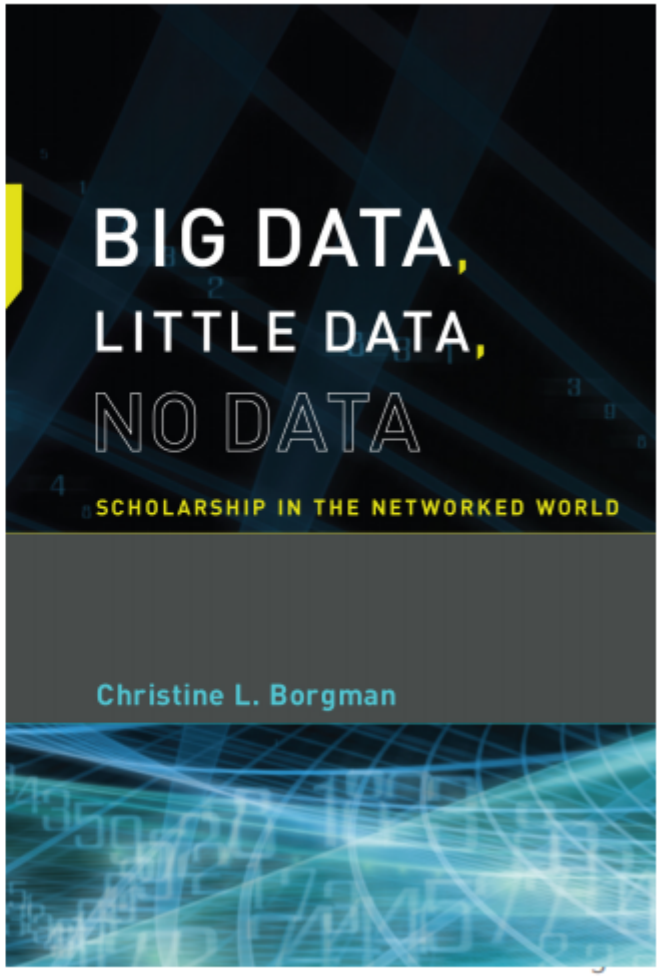
- **Teach the iterative mode of scientific practice at all levels of schooling**

The new norms are…

1. Science is increasingly conducted by teams
2. Data-intensive research is on the rise
3. **Data sharing is needed for emerging research areas**

Why share research data?

- To reproduce research
- To make public assets available to the public
- To leverage investments in research
- To advance research and innovation



MIT Press, 2015

Christine Borgman
Distinguished Professor & Presidential Chair in Information Science

https://works.bepress.com/borgman/390/

Arguments FOR data sharing:

"The Science *family of journals is committed to the sharing of data relevant to* ***public health emergencies****, and therefore… we endorse the statement below".* *- McNutt 2016*

- **Journal content on the topics are open-access**
- **Require data sharing as rapidly and widely as possible**

Science, Editor-in-chief

Arguments FOR data sharing:

"Sharing data and crediting sources are among the most basic of scientific ethical principles."

- Duke and Porter 2013

Director of Science Programs, ESA

Rsch Assoc. Prof., Univ. Virginia

Arguments FOR data sharing: *Ethical considerations*

Patricia Soranno, Ecologist

Kendra Cheruvelil, Ecologist

Kevin Elliott, Philosopher

Georgina Montgomery, Historian

Soranno et al. 2015. BioScience

Arguments FOR data sharing: *Ethical considerations*

The norms about data must better align with sciences' growing emphasis on **inclusion**

Science AAAS

Diversity: Promoting New Perspectives

By Chris Tachibana | Jul. 20, 2012 , 4:00 AM

npr

Top Medical Journals Give Women Researchers Short Shrift

Science AAAS

EDITORIAL

Implicit bias

Marcia McNutt

+ Author Affiliations

"...successful organizations and people welcome those who do not...act like they do."

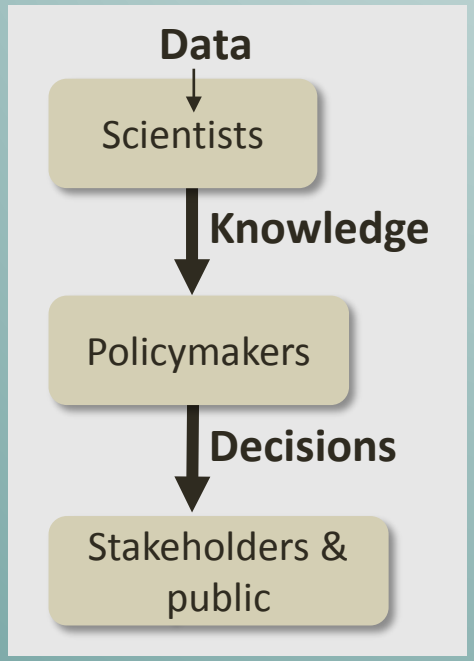
Science 27 May 2016:

Soranno et al. 2015. BioScience

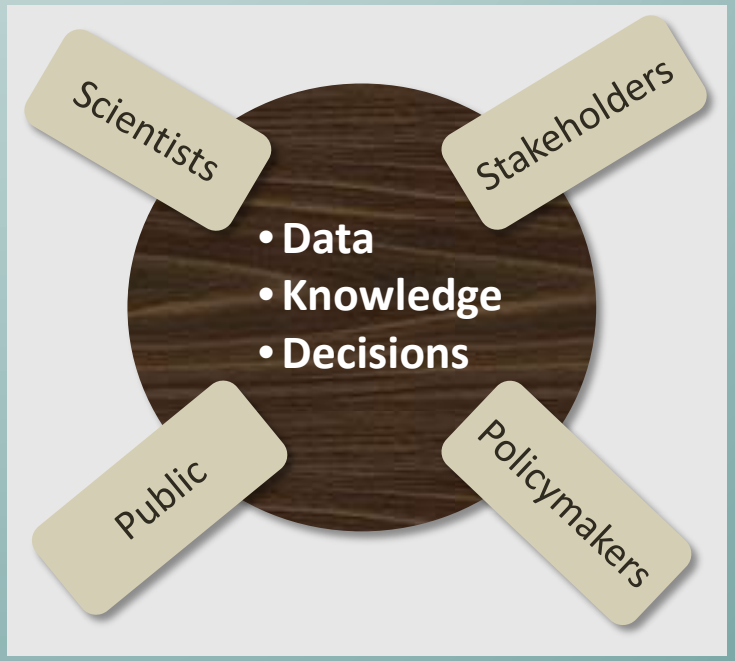
Arguments FOR data sharing: *Ethical considerations*

Data sharing is essential for ensuring successful efforts at the science-policy interface

Deficit-linear model



Round-table model



Soranno et al. 2015. BioScience

Arguments FOR data sharing: *Ethical considerations*

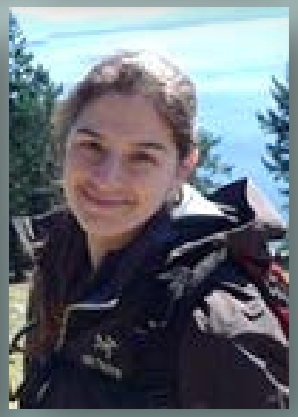
The combination of **public sponsorship of research**, '**inclusive**' **participation in research**, and **connection with policy** creates a set of circumstances that push environmental scientists—and particularly those who seek to broaden participation in science—toward data sharing as an ethical obligation.

- Soranno et al. 2015

Arguments FOR data (**and other**) sharing in **ECOLOGY**:

2005 - 2015

- Parr and Cummings 2005
- Zimmerman 2008
- Reichman, OJ et al. 2011
- Wolkovich et al. 2012
- Hampton et al. 2013
- Duke and Porter 2013
- Hampton et al. 2015
- Michener 2015
- Gries et al 2016



Open Science

Defined as science that has:

1. *Fully accessible publications*
2. *Fully accessible data*
3. *Transparent and reproducible methods*

Good blog by Michael Nielsen about open science:
http://michaelnielsen.org/blog/the-future-of-science-2/

Open-science practices

(1) Publishing: **Open Access** *journals & publishers*

Open-science practices

(2) Data: **Data repositories**

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

LTER Network Data Portal

DRYAD

Open-science practices

(3) Data, continued: **Journal policies requiring public access to data**

Announcement: Where are the data?

07 September 2016

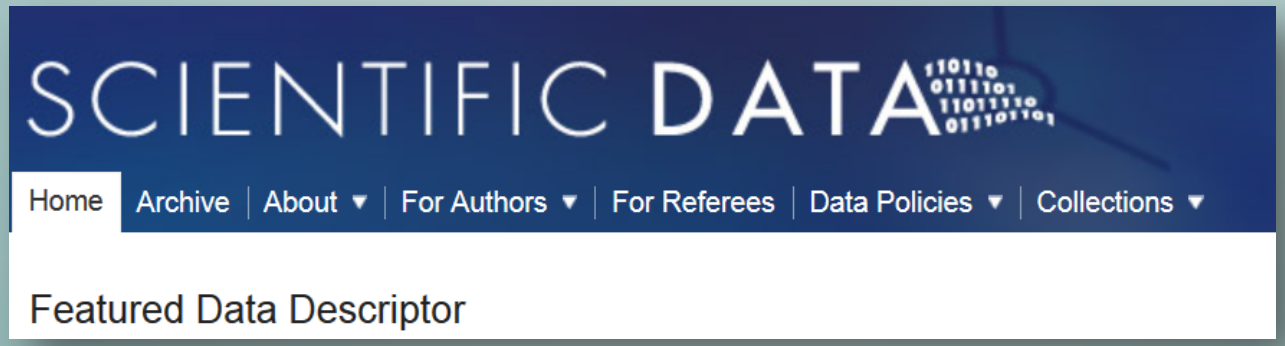
PDF | Rights & Permissions

nature International weekly journal of science

As the research community embraces data sharing, academic journals can do their bit to help. Starting this month, all research papers accepted for publication in *Nature* and an initial 12 other Nature titles will be required to include information on whether and how others can access the underlying data.

Open-science practices

(4) Transparency: **Data papers, methods papers, metadata, etc.**



Arguments AGAINST data sharing:

Some barriers that have been highlighted (Hampton et al. 2015)

- Time to adopt new practices, learn new tools
- High-stake issues for scientists -- *Precedence, attribution, investment, and payoff*
- Relinquishing control
- Mindset of 'data ownership'

Arguments AGAINST data sharing:

- Should **not have a 'blanket' policy** because scientists should be using 'proprietary' data of the 'data economy':

  *"Requiring data to be…open access **may feel right** but could have **perverse consequences** for the future of science."*

  - Fenichel and Skelly 2015

Yale Univ.

F.R. Oastler Professor of Ecology, Yale Univ.

Arguments AGAINST data sharing:

- A global survey of 73 researchers engaged in long-term ecological and evolutionary studies.

  "Positive attitudes towards sharing data with agreement or involvement of the PI… only 8% were in favor of uncontrolled, open access to primary data. AND,

  A more balanced viewpoint is necessary to allow a discussion to emerge on a code of ethics and ways to preserve and protect the data.

  - Mills et al. 2015. Trends in Ecology and Evolution

Arguments AGAINST data sharing:

- Scientists should not be expected to share because they have **intimate knowledge of their systems**, and:

  *"There is also the emerging issue of a generation of what we term as '**parasitic scientists**' who will never be motivated to go and gather data because it takes real effort and time and it is simply easier to use data gathered by others."*

  - Lindenmayer and Likens 2013

Professor, Austr. Nat'l. Univ.

President of Cary Institute-Emeritus

Arguments AGAINST data sharing:

Scientists NOT involved in the study **cannot understand/know the choices the scientists made**; and:

> *"There is concern among **some front-line researchers** that the system will be taken over by what some researchers have characterized as "**research parasites**".*
>
> *- Longo and Drazen 2016*

NEJM & Harvard Med. School

NEJM, Editor-in-chief

Twitter backlash to Longo and Drazen .... *Jan 21, 2016 – Mar 29, 2016*

#IamAResearchParasite

… and some repenting by one of the authors:
http://www.nejm.org/doi/full/10.1056/NEJMe1601087

Twitter backlash to Longo and Drazen

Jan 21– Mar 29, 2016

DJ Patil,
U.S. Chief Data Scientist

DJ Patil
@DJ44

Follow

#IAmAResearchParasite. The best science is done as in collaboration not in silos.
Data is a team sport. twitter.com/atulbutte/stat…

7:37 AM - 5 Mar 2016

63 90

DJ Patil @DJ44 · Mar 29

The open data culture shift is real and it will increase our speed of innovation as a country

> **Taha Kass-Hout** @DrTaha_FDA
> It's fascinating to see the growing number of peer-reviewed publications using data from @openFDA #Reproducibility #OpenScience cc:@DJ44

24 35 •••

Moving forward? **The discussion is NOT over, so lets keep working**

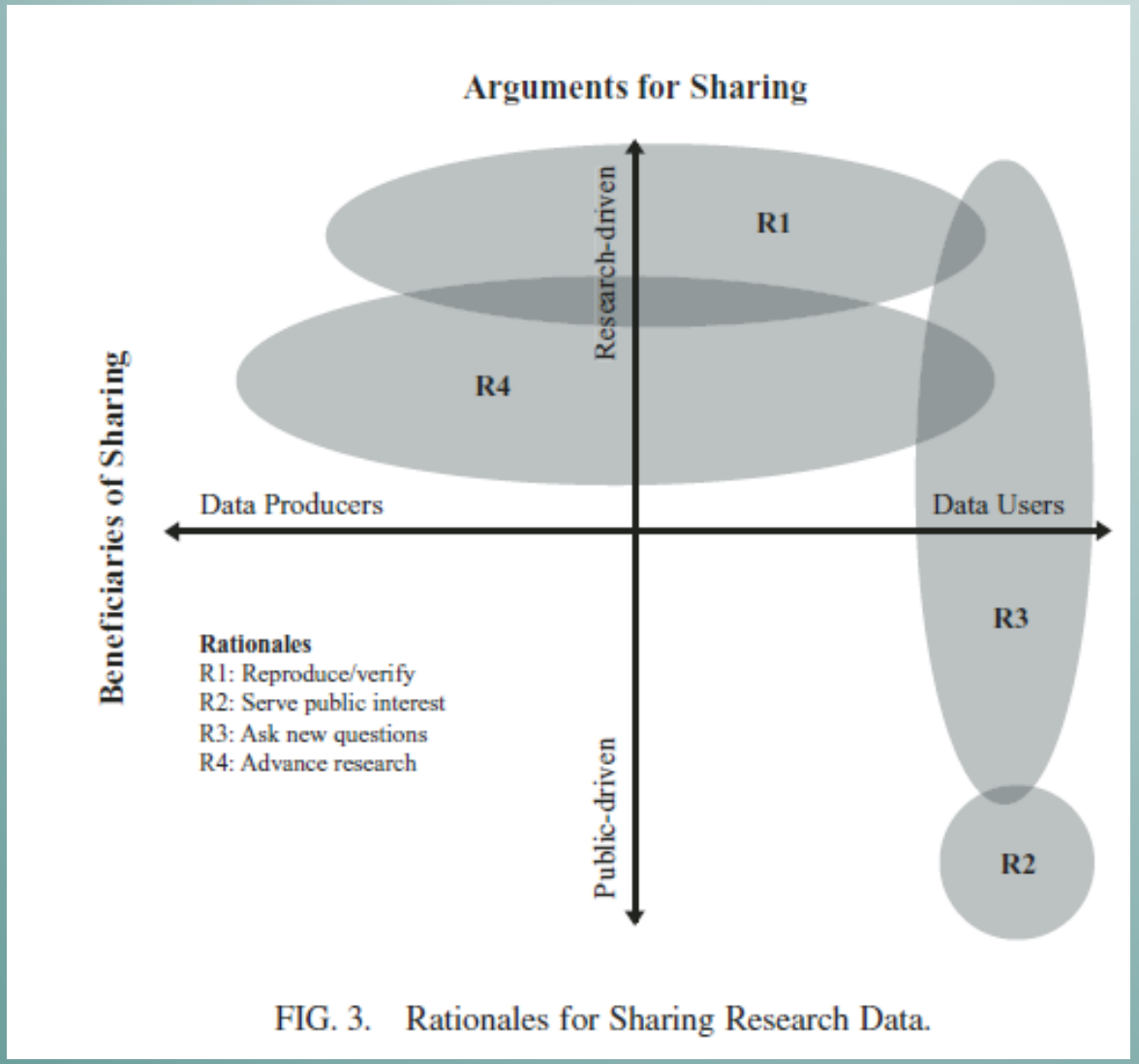


FIG. 3. Rationales for Sharing Research Data.

Christine Borgman
Distinguished Professor & Presidential Chair in Information Science

From: Borgman 2011. The conundrum of sharing research data

RECOMMENDATIONS: Data-sharing and open science

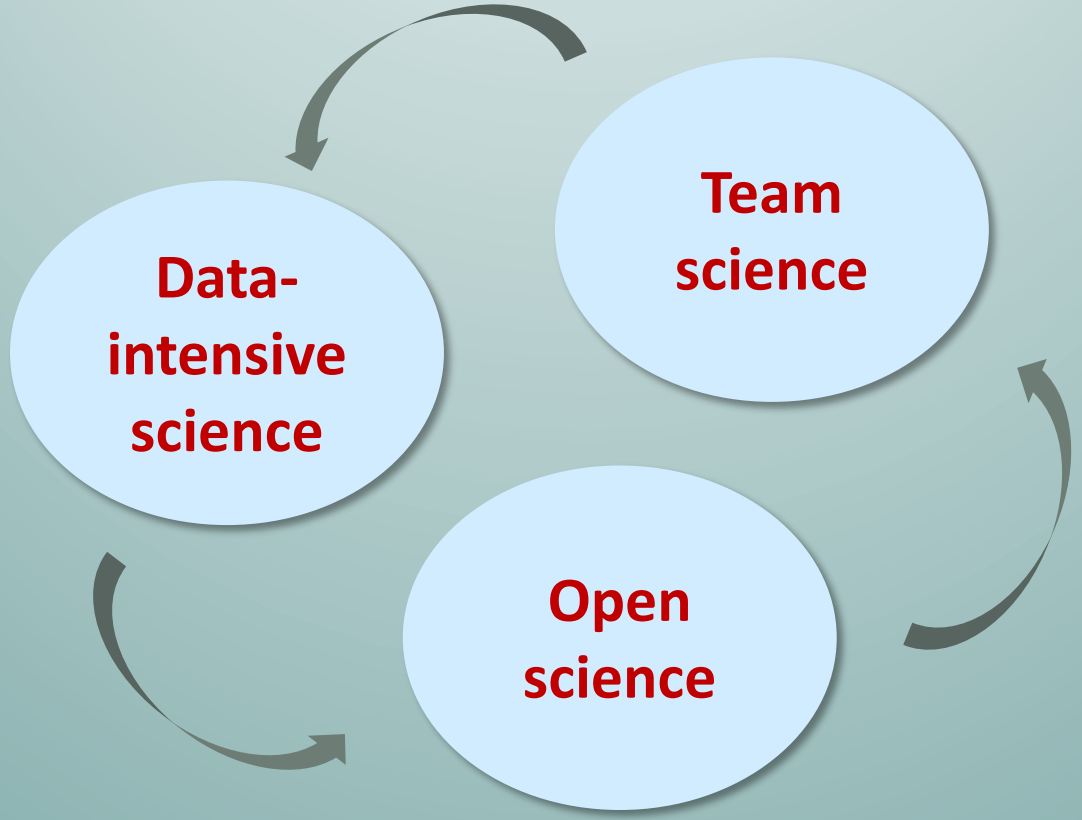
1) Provide **training and tools** for ALL scientists

2) Develop **incentives and rewards** for sharing

3) Change the **culture** of science from closed to open

4) Have more conversations about **what scientists' ethical obligations and roles in society** really are

Addressing many important challenges in ecology



But to incorporate these modes, will the culture get in the way?

Question for you

- Are we on the cusp of making data-inspired discoveries to answer societally-relevant and scientifically-interesting questions, in an inclusive manner?

- If so, are the efforts currently hampered by our the norms and culture in ecology?

vs