

Astroinformatics?

Paulo Penteado

pp.penteado@gmail.com

<http://www.ppenteado.net>

This presentation:

http://www.ppenteado.net/ast/pp_if_20121113.pdf



“We have created a new world, and it's amazing”

S. George Djorgovski

Evolving Science and Technology in Cyberspace

TEDx Calctech, January 14 2011

<http://www.youtube.com/watch?v=FB33pV2L0Vo>



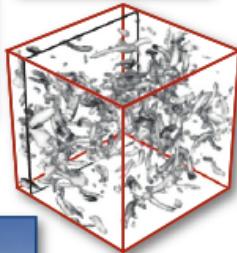
The fourth paradigm



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

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



The Fourth Paradigm

Data-Intensive Scientific Discovery

Edited by Tony Hey, Stewart Tansley, and Kristin Tolle

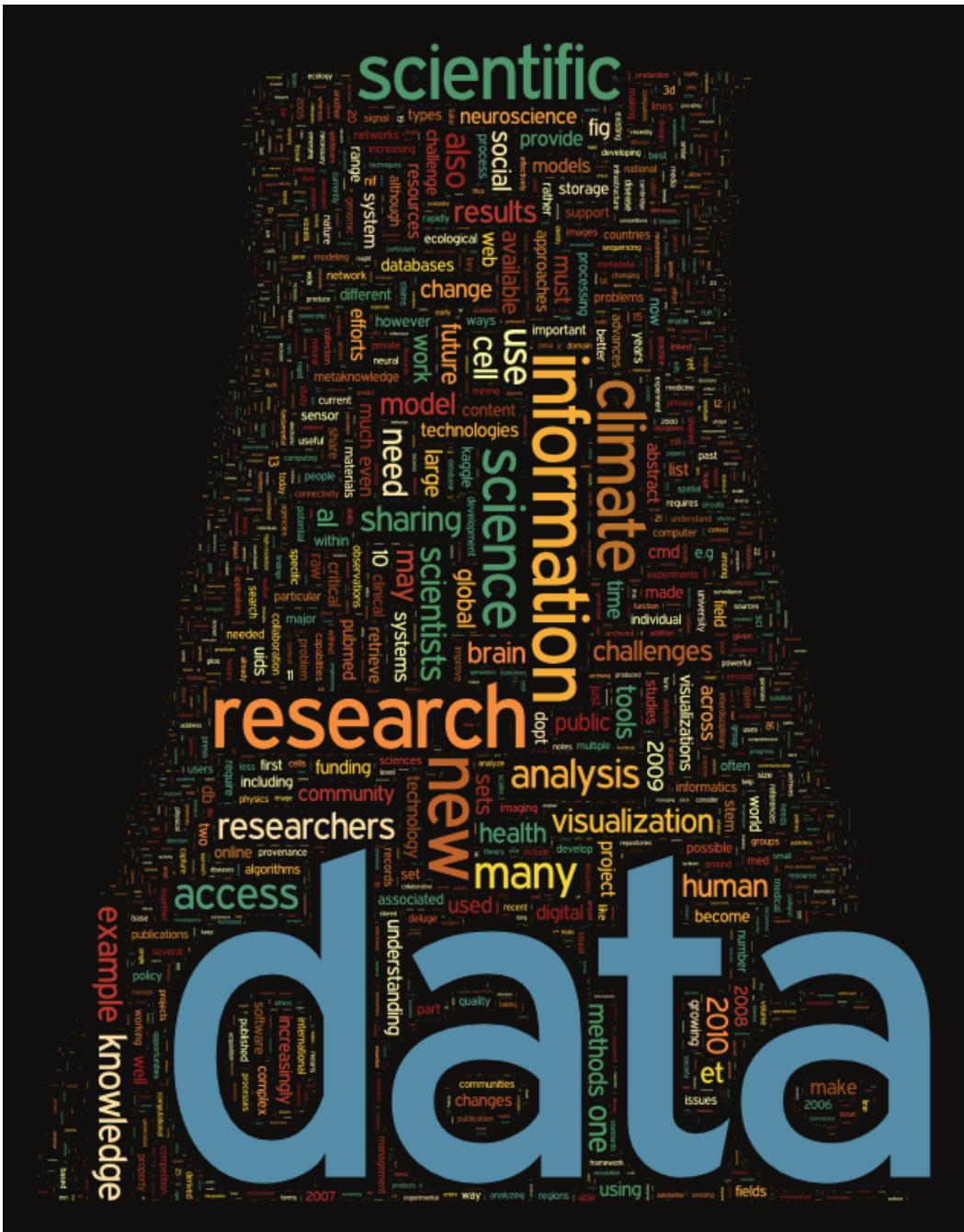
Tony Hey, Stewart Tansley, Kristin Tolle, (editors), 2009
Paperback: \$46.07 Kindle: \$0.99

<http://www.amazon.com/The-Fourth-Paradigm-Data-Intensive-Scientific/dp/0982544200/>

PDF: free

<http://research.microsoft.com/en-us/collaboration/fourthparadigm/>

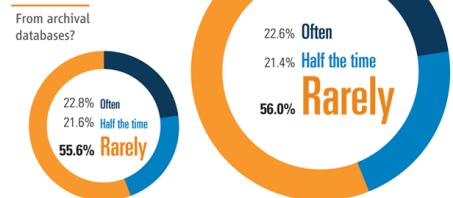
Science special edition (11 February 2011)



<http://www.sciencemag.org/site/special/data/>

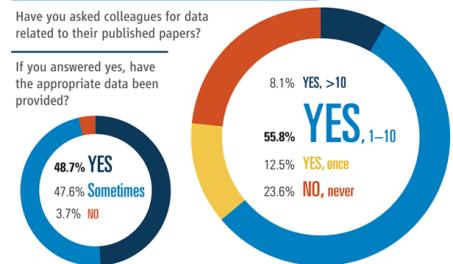


How often do you access or use data sets from the published literature for your original research papers?

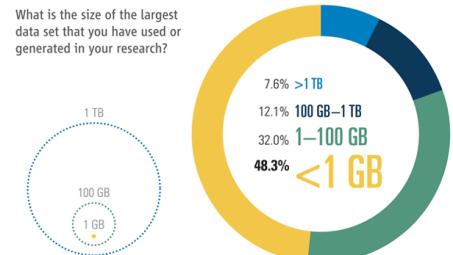


Have you asked colleagues for data related to their published papers?

If you answered yes, have the appropriate data been provided?

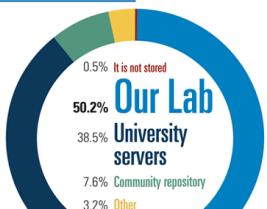


What is the size of the largest data set that you have used or generated in your research?



Where do you archive most of the data generated in your lab or for your research?

“ Even within a single institution there are no standards for storing data, so each lab, or often each fellow, uses ad hoc approaches



Do you have the necessary expertise in your lab or group to analyze your data in the way you want?

“ The next few years [particularly in medicine] the volume of data we

"Most universities still operate on a medieval model of how we teach"



>

>

...SCIENTISTS AND THEIR SOFTWARE

A survey of nearly 2,000 researchers showed how coding has become an important part of the research toolkit, but it also revealed some potential problems.

> **45%** said scientists spend more time today developing software than five years ago."

> **38%** of scientists spend at least one fifth of their time developing software.

> Only **47%** of scientists have a good understanding of software testing.

> Only **34%** of scientists think that formal training in developing software is important.

"We invent technology but technology changes us"

C:\lab>

f77 -o

data.exe

>

>

...ERROR

....why scientific programming does not compute

>

BY ZEEYA MERALI

The new technology enables new science

"It has empowered anybody, anywhere with an internet connection to have access to first-rate data, tools, literature, and, if they have the expertise and the talent, to do first-grade science..."

"They do not have to be members of some prestigious institution, they do not have to have access to expensive facilities."

<http://dx.doi.org/10.1038/463416a>



NEWS FEATURE

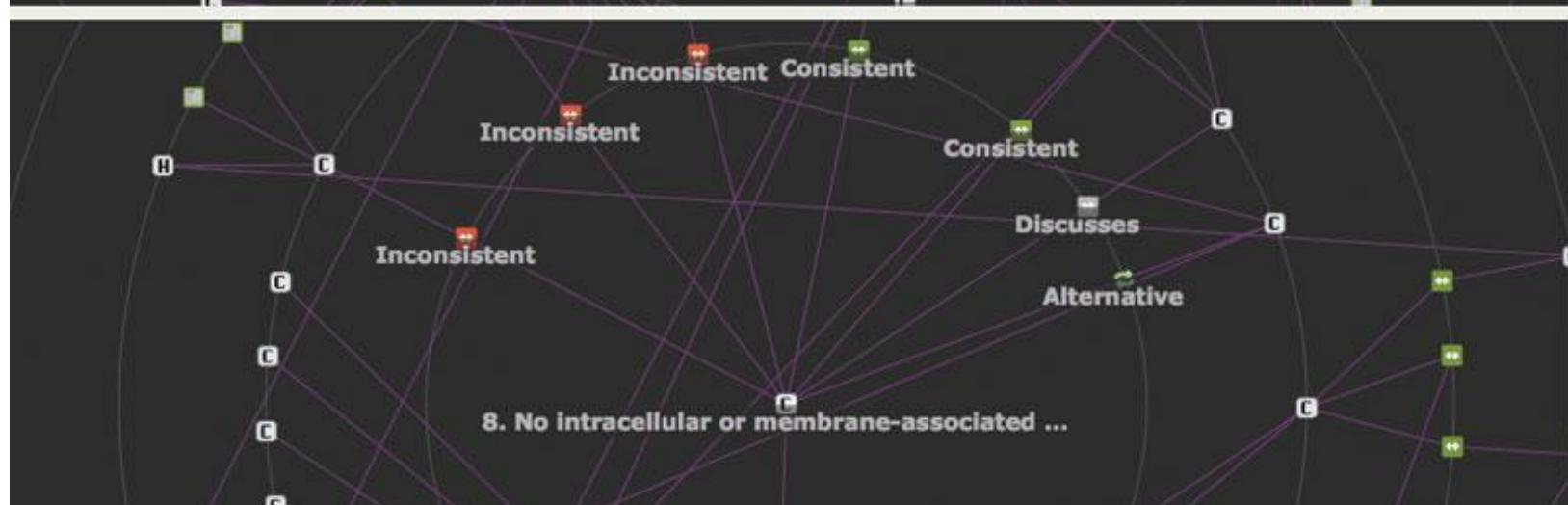
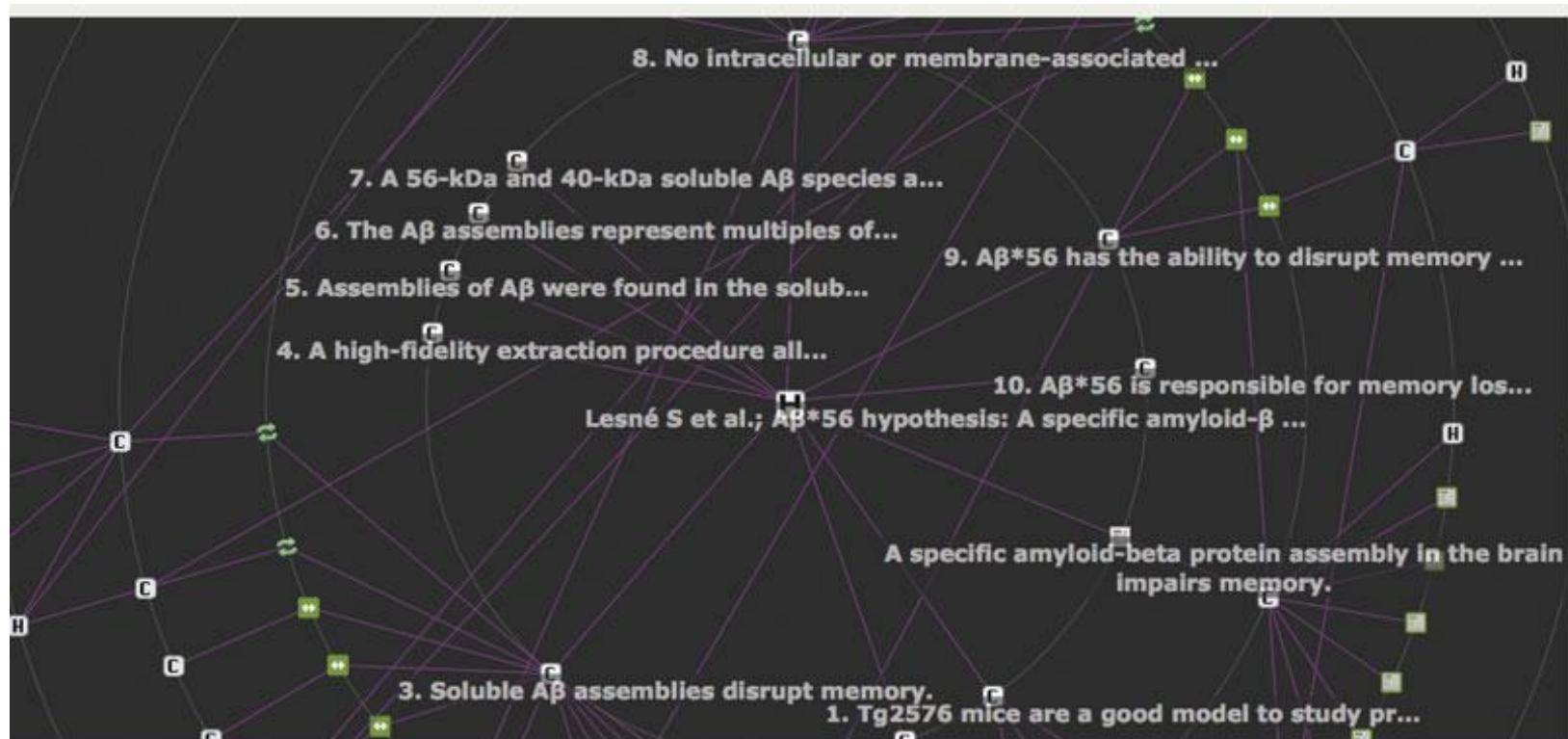
NATURE|Vol 463|28 January 2010

Speed reading

Scientists are struggling to make sense of the expanding scientific literature.
Corie Lok asks whether computational tools can do the hard work for them.

SWAN (Semantic Web Applications in Neuromedicine)

<http://go.nature.com/nNz93Y>



@MarsCuriosity



<https://www.facebook.com/MarsCuriosity>



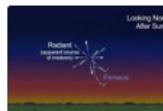
NASA's Curiosity Mars Rover
274,078 likes · 394,672 talking about this

1

Aerospace/Defense

Who's got six wheels, a laser and is now on the Red Planet? Me. I'm Curiosity, aka the Mars Science Laboratory rover.

About



Photos

"The trend has been in the form of increasing immediacy and increasing fidelity"

- Curiosity Rover** @MarsCuriosity 6 Aug
 No photo or it didn't happen? Well lookee here, I'm casting a shadow on the ground in Mars' Gale crater #MSL pic.twitter.com/cj1zFJty
[View photo](#)
- Curiosity Rover** @MarsCuriosity 6 Aug
 You asked for pics from my trip. Here you go! My 1st look (of many to come) of my new home... MARS! #MSL pic.twitter.com/894ouNJt
[View photo](#)
- Curiosity Rover** @MarsCuriosity 6 Aug
 I'm safely on the surface of Mars. GALE CRATER I AM IN YOU!!! #MSL
[Expand](#)
- Curiosity Rover** @MarsCuriosity 6 Aug
 Backshell separation. It's just you & me now, descent stage. Engage all retrorockets! #MSL
[Expand](#)
- Curiosity Rover** @MarsCuriosity 6 Aug
 Heatshield separation. Next up: Radar must lock on ground #MSL
[Expand](#)
- Curiosity Rover** @MarsCuriosity 6 Aug
 Parachute deployed! Velocity 900 mph. Altitude 7 miles. 4 minutes to Mars! #MSL
[Expand](#)
- Curiosity Rover** @MarsCuriosity 6 Aug
 Guided entry is begun. Here I go! #MSL
[Expand](#)
- Curiosity Rover** @MarsCuriosity 6 Aug
 Entering Mars' atmosphere. 7. Minutes. Of. Terror. Starts. NOW. #MSL
[Expand](#)

The new technology enables new science

"It has empowered anybody, anywhere with an internet connection to have access to first-rate data, tools, literature, and, if they have the expertise and the talent, to do first-grade science.."

<http://dx.doi.org/10.1038/nature11165>

LETTER

doi:10.1038/nature11165

Possible tropical lakes on Titan from observations of dark terrain

Caitlin A. Griffith¹, Juan M. Lora¹, Jake Turner¹, Paulo F. Penteado², Robert H. Brown¹, Martin G. Tomasko¹, Lyn Doose¹ & Charles See¹

<http://dx.doi.org/10.1126/science.1175930>

A Population of Compact Elliptical Galaxies Detected with the Virtual Observatory

Igor Chilingarian,^{1,2,3*} Véronique Cayatte,⁴ Yves Revaz,⁵ Serguei Dodonov,⁶ Daniel Durand,⁷ Florence Durret,^{8,9} Alberto Micol,¹⁰ Eric Slezak¹¹

<http://dx.doi.org/10.1051/0004-6361/201117297>

A&A 532, L5 (2011)
DOI: [10.1051/0004-6361/201117297](https://doi.org/10.1051/0004-6361/201117297)
© ESO 2011

**Astronomy
&
Astrophysics**

LETTER TO THE EDITOR

Two very nearby ($d \sim 5$ pc) ultracool brown dwarfs detected by their large proper motions from WISE, 2MASS, and SDSS data*

R.-D. Scholz, G. Bihain, O. Schnurr, and J. Storm

<http://dx.doi.org/10.1111/j.1365-2966.2011.19085.x>

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY



Mon. Not. R. Astron. Soc. 418, 705–728 (2011)

doi:10.1111/j.1365-2966.2011.19085.x

A new grid of synthetic spectra for the analysis of [WC]-type central stars of planetary nebulae

Graziela R. Keller,^{1,2*} James E. Herald,² Luciana Bianchi,² Walter J. Maciel¹ and Ralph C. Bohlin³

¹Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, Cidade Universitária, São Paulo/SP, Brazil

²Department of Physics and Astronomy, The Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218, USA

³Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA

<http://dx.doi.org/10.1038/nature08608>

LETTERS

A single sub-kilometre Kuiper belt object from a stellar occultation in archival data

H. E. Schlichting^{1,2}, E. O. Ofek¹, M. Wenz³, R. Sari^{1,4}, A. Gal-Yam⁵, M. Livio⁶, E. Nelan⁶ & S. Zucker⁷

<http://dx.doi.org/10.1038/nature08082>

LETTERS

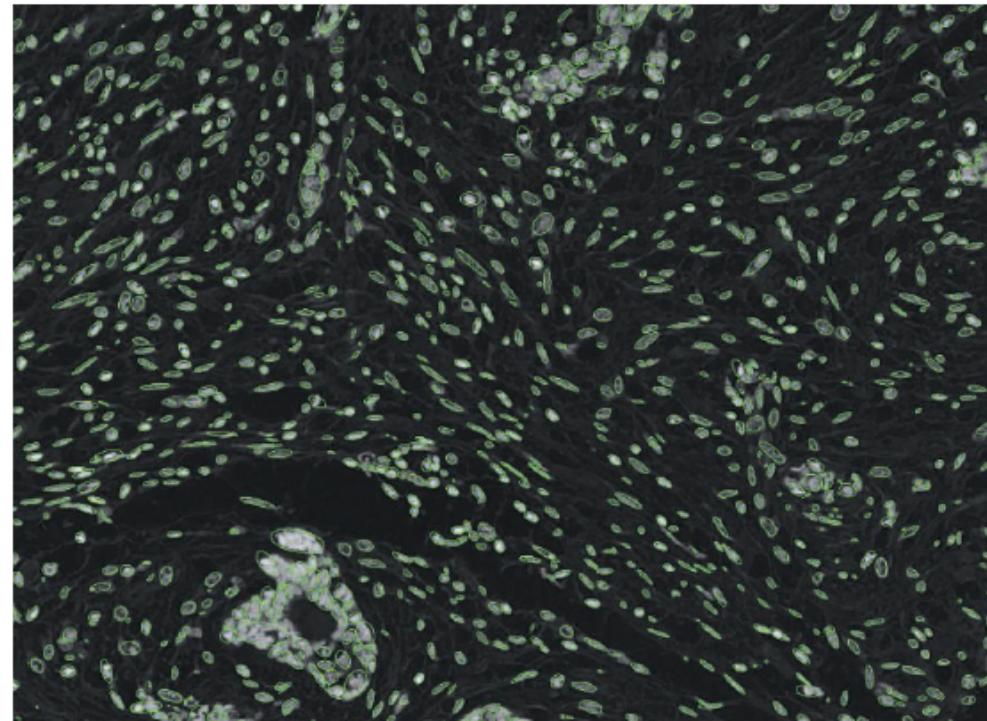
Type IIn supernovae at redshift $z \approx 2$ from archival data

Jeff Cooke¹, Mark Sullivan², Elizabeth J. Barton¹, James S. Bullock¹, Ray G. Carlberg³, Avishay Gal-Yam⁴ & Erik Tollerud¹

“Even more important than the growth of data volume is the growth of data complexity”



<http://dx.doi.org/10.1126/science.331.6018.696>



NEWS

Is There an Astronomer in the House?

With biomedical researchers analyzing stars and astronomers tackling cancer, two unlikely collaborations creatively solve data problems

pathologists look for different biomarkers—specific proteins—in the patient’s cancerous tissue. For example, an overexpression of the

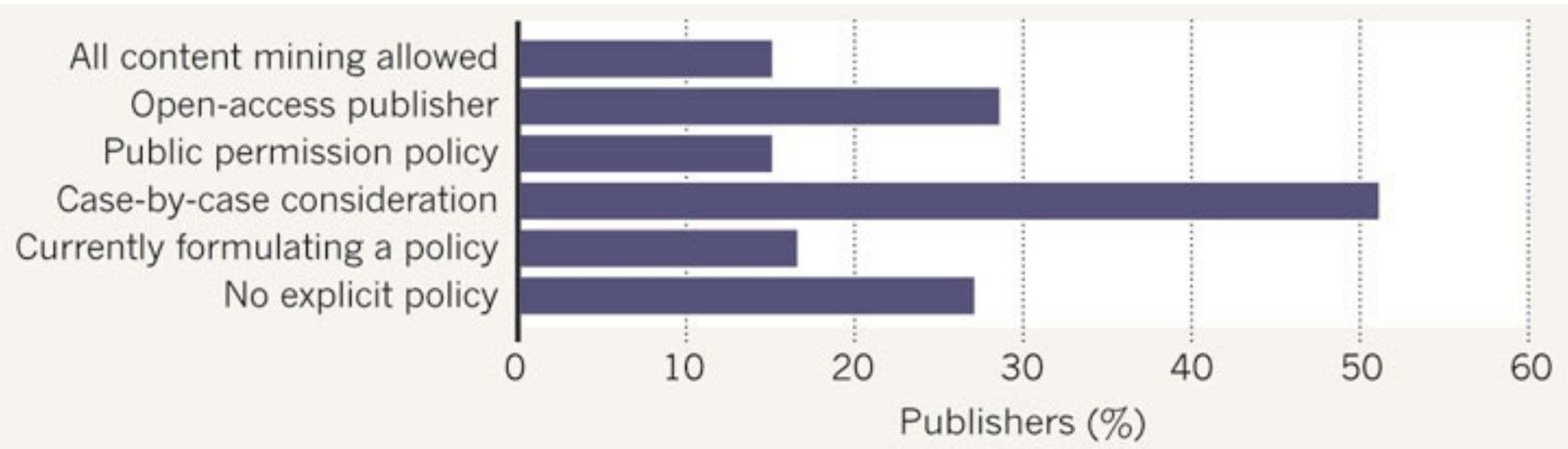
“Most universities still operate on a medieval model of how we teach”

Publishers, too:

PUBLISHING

Trouble at the text mine

Computers can rapidly scan through thousands of research papers to make useful connections, but work is being slowed by publishers' unease.



<http://dx.doi.org/10.1038/483134a>



The new technology enables new science

“Now, an output of a theory is no longer a formula, but a data set”

“There are now archives stuffed with data which will never be printed and probably never be seen by a human being.”

This presentation:
http://www.ppenteado.net/ast/pp_if_20121113.pdf

