

The Data Lifecycle

#harvarddatafest

Mercè Crosas
Chief Data Science and Technology Officer
Institute for Quantitative Social Science
Harvard University
[@mercecrosas](https://twitter.com/mercecrosas)

Data

Facts that can be analyzed or used in an effort to gain knowledge or make decisions; information.

Latin Datum “Something Given”

(The American Heritage Dictionary)

Museum of Fine Arts

The Art of the Ancient World



FRIDAY, JANUARY 13, 2017, 7PM

Giza Excavation 1902-1947



THE HARVARD UNIVERSITY BOSTON
MUSEUM OF FINE ARTS EXPEDITION

Archeological Finds as Data



Photo ID number: B5789_NS

Photo subjects: Object(s)
photograph: Site: Giza; view: G
7632, G 7510, street G 7000,
street G 7300, street G 7400,
street G 7500, avenue G 2

Description: Bronze objects: top row: [G 7632 A II] 25-2-822 (= MFA 25.2436, amulet of Horus), [G 7510 G] 25-2-916a (= MFA 25.2465, amulet of ichneumon), [street G 7500] 25-2-296 (= MFA 25.2211, Apis bull amulet), [street G 7300] 25-1-143 (= MFA 25.1591, Apis bull amulet), [avenue G 2] 25-1-1363 (= MFA 25.2066, Apis bull amulet), [street G 7500] 25-2-297 (= MFA 25.2212, amulet of Anubis); second row, left: [street G 7500] 25-1-1385 (= MFA 25.2082, ladle); second row, right: [G 7130-7140: G 7130 A] 24-12-513 (= MFA 24.2866, unidentified deity, fragment), [avenue G 2] 25-1-1364 (= MFA 25.2067, fragment of ring); third row, right: [street G 7000] 24-11-39 (= MFA 24.2271, fish on standard); bottom row: [street G 7000] 24-11-838 (= MFA 24.2638, cat on standard), [G 7510 D] 25-3-252 (= MFA 25.2672, flail), [street G 7400] 25-1-620 (= MFA 25.1766, beard), [street G 7000] 24-11-223 (= MFA 24.2371, amulet of Mehit and Inher), [avenue G 2] 25-1-1424 (= MFA 25.2103, uraeus), [street G 7400] 25-1-842 (= MFA 25.1853, uraeus)

Photo date: 10/21/1925

The Giza Archive and catalogue:

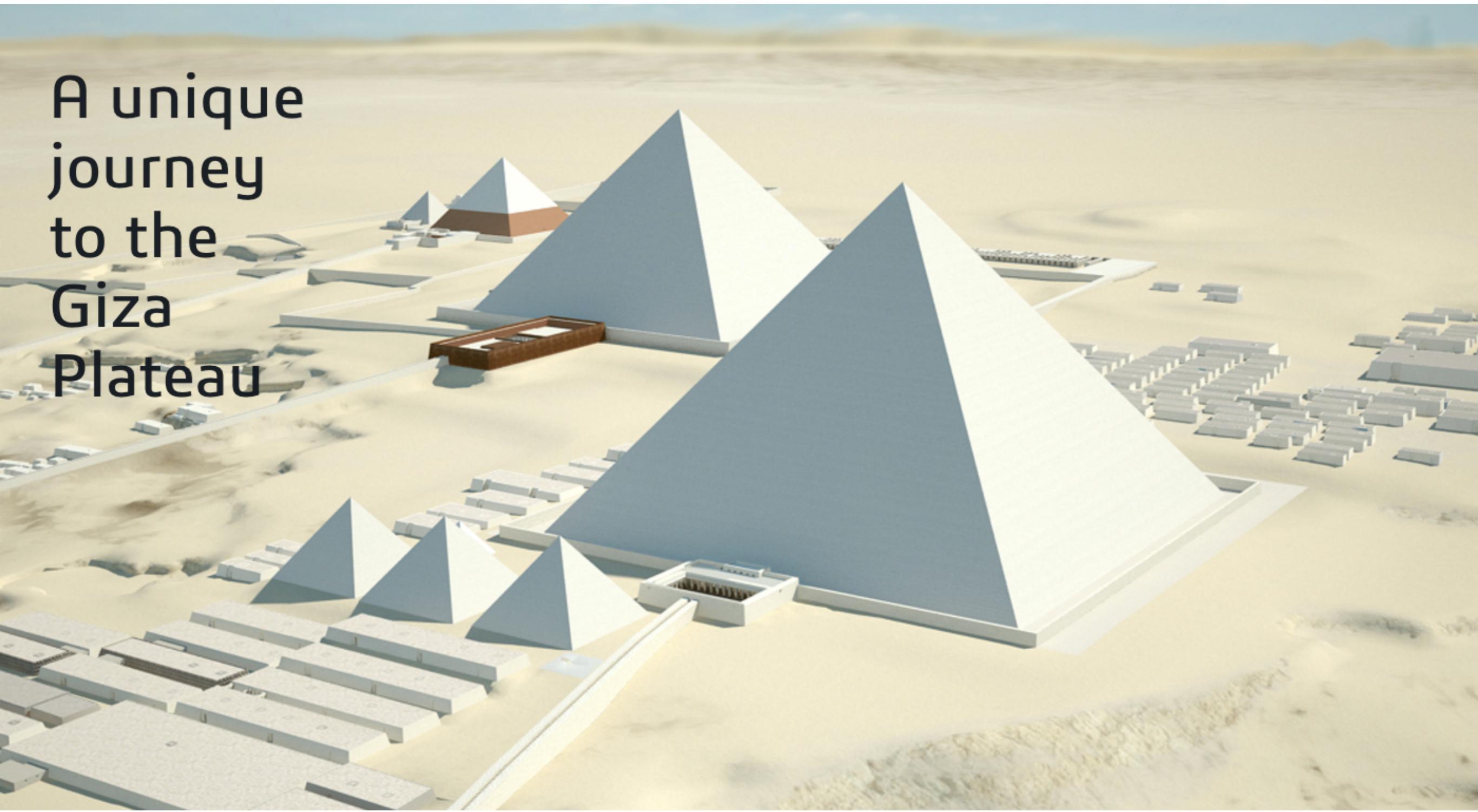
- 21,000 finds
- 3,000 photos
- 3000 tomb and monument records
- 3,000 diary pages
- 7,000 maps
- books, manuscripts
- metadata

[Click here to order a publication quality version of this image.](#)

[Click here to download a study copy of this image.](#)

3D Visualization to Re-Explore Giza

A unique
journey
to the
Giza
Plateau



Peter Der Manuelian, Harvard's Visualization Center (Geological Museum),
DARTH, and Dassault Systèmes in Paris

Books as Data

Harvard Cultural Observatory (Aiden, Michel, SEAS) + Google:

- 5 million digitized books (from 130 million total)
- From year 1500 to 2008
- Optical Character Recognition (OCR)
- Quantify unstructured text to n-gram count matrix

			1
1584	382341	926	1
1590	270386	581	1
1592	50762	129	1
1596	187852	721	1
1600	24755	75	1
1610	1732	16	1
1620	68848	292	1
1628	81026	228	1
1635	47119	130	1
1647	34008	110	1
1648	309670	724	1
1649	13677	59	1
1656	96789	473	3
1658	694132	978	2
1660	110823	231	1
1661	56308	187	1
1662	100701	200	1
1663	267812	574	3
1664	108710	222	2
1665	117171	236	1
1666	282982	638	3
1667	220639	439	2
1668	92431	181	1
1669	220639	529	1
		5'10_ADJ	1912
		6'10_ADJ	1964
		6'10_ADJ	1966
		6'10_ADJ	1970
		6'10_ADJ	1971
		6'10_ADJ	1972
		6'10_ADJ	1977
		6'10_ADJ	1978
		6'10_ADJ	1980
		6'10_ADJ	1981
		6'10_ADJ	1983
		6'10_ADJ	1984
		6'10_ADJ	1986
		6'10_ADJ	1987
		6'10_ADJ	1991
		6'10_ADJ	1992
		6'10_ADJ	1993
		6'10_ADJ	1994
		6'10_ADJ	1995
		6'10_ADJ	1996
		6'10_ADJ	1999
		6'10_ADJ	2002
		6'10_ADJ	2003
		6'10_ADJ	2004
		6'10_ADJ	2006
		6'10_ADJ	2008

Frequency of Words with Time

Google Books Ngram Viewer

Graph these comma-separated phrases: case-insensitive

between and from the corpus with smoothing of



Historical events as Data

my ab	POSITION	Places Notes	Terrain Notes	ID	TYPE	NAME	DATE	STEP	VALUE	UNITS	TEXT	LAT1	LON1	LAT2	LON2	PATH	CERTAINTY
		Needs static event map	Needs static event map				4/6/1760		0								
							4/7/1760		0								
	18.3648, -76.8852	Frontier misplaced to NW			1 Conspiracy	Frontier	4/7/1760		1	5 Rebel conspirators		18.3648	-76.8852				1
	18.3542, -76.8990	Trinity misplaced to NW			1 Rebels	Trinity	4/7/1760		2	100 Rebel force		18.3542	-76.899				1
		Still glitchy; start					4/8/1760		0								

POSITION	Places Notes	Terrain Notes	ID	TYPE	NAME	DATE	S
	Needs static event map	Needs static event map					
		Needs static event map					
18.3648, -76.8852	Frontier misplaced to NW			1 Conspiracy	Frontier	4/7/1760	
18.3542, -76.8990	Trinity misplaced to NW			1 Rebels	Trinity	4/7/1760	
18.3542, -76.8990 to Ft. Haldane	Still glitchy; start should be from Trinity database fixed?			1 Rebels	Ft. Haldane	4/8/1760	
18.3542, -76.8990	Trinity misplaced to NW			1 Clash	Ft. Haldane	4/8/1760	
18.3542, -76.8990				1 Rebels	Trinity	4/8/1760	

18.3223, -76.8875 to Esher	Rebels should move from Ballard's Valley to Esher To extent possible, movement should follow roads (dotted line) To extent possible, movement should follow roads (dotted line)		1 Rebels	Esher	4/8/1760	6			18.3223	-76.8875	18.2809	-76.8675	-76.8675	-76.8675	-76.8882], [18.2709, -76.8851], [18.2704, -76.8822], [18.2703, -76.8773], [18.2722, -76.8736], [18.2783, -76.8703], [18.2809, -76.8639], [18.2809, -76.8906], [18.2885, -76.8921], [18.2908, -76.8964], [18.2919, -76.8995], [18.2952, -76.8995], [18.2941, -76.8976], [18.3032, -76.8937], [18.305, -76.8882]]	1
18.3223, -76.8875	Eliminate movement; Militia appear at single point, Ballard's Valley To extent possible, movement should follow roads (dotted line)		1 Clash	Esher	4/8/1760	7		5 whites killed	18.2809	-76.8675					1	
Esher to 18.3050, -76.8882	Starting pt is Esher To extent possible, movement should follow roads (dotted line)		2 Militia	Ballard's Valley	4/8/1760	8			18.3223	-76.8875					0	
1 Rebels	Whitehall	4/8/1760	9	400 Rebel force					18.2809	-76.8675	18.305	-76.8882	-76.8882	-76.8882	1	

Geospatial coordinates

Time

Jamaica Slave Revolt in Space and Time (1760 - 1761)

← PAUSE II → 1 of 17 Risings on multiple estates, including The Delve...

May 26, 1760

Mr. Macfarlane came to us in a great hurry and gave us a distressing message: that Negroes had broken out in rebellion against the white people not far from here, and they had gruesomely murdered many, and that all the white people in this parish were going to rifle practice...--Moravian diary, Mesopotamia estate, 26 May 1760

"They have been joined by the Negroes of the sev.l adjacent Estates & tis said they are 900 Strong, but this is not certain as we are receiv.g frequent Contradictory acc.ots, by an Express just now come in we hear they are got into Hanover." --Francis Treble, 2, June 1760

"Immediately upon receiving an Express from Westmoreland that a Rebellion was broke out there... a Quantity of Arms and Ammunition was dispatched for the use of the Militia in those

Places Map | Terrain Map

The Delve
Click for details
Trelawny Town
Rebel force: 200

Powered by Leaflet

Map Legend

JUN JUL AUG SEP

Animated maps to narrate and explore history data
by Vincent Brown (Harvard's History Department)

Polls as Data

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Trump	Clinton	Johnson	Other	Undecided	poll_slug	survey_hous	start_date	end_date	question_tex	sample_subp	observations	margin_of_e	mode	partisanship	partisan_affili
2	41	45	4	2	8	insights-west	Insights Wes	11/4/16	11/7/16	As you may k Likely Voters	940	3.2	Internet	Nonpartisan	None	
3	6	89	1	0	4	insights-west	Insights Wes	11/4/16	11/7/16	As you may k Likely Voters - Democrat		Internet	Nonpartisan	None		
4	82	7	3	2	6	insights-west	Insights Wes	11/4/16	11/7/16	As you may k Likely Voters - Republican		Internet	Nonpartisan	None		
5	38	43	8	4	7	insights-west	Insights Wes	11/4/16	11/7/16	As you may k Likely Voters - independent		Internet	Nonpartisan	None		
6	43	41	7	4	5	ibd-tipp-268	IBD/TIPP	11/4/16	11/7/16	Likely Voters	1107	3.1	Live Phone	Nonpartisan	None	
7	42	43		5	9	ibd-tipp-268	IBD/TIPP	11/4/16	11/7/16	Likely Voters	1107	3.1	Live Phone	Nonpartisan	None	
8	41	45	5	5	4	yougov-econ	YouGov/Econ	11/4/16	11/7/16							
9	6	90	1	2	2	yougov-econ	YouGov/Econ	11/4/16	11/7/16							
10	84	4	4	4	4	yougov-econ	YouGov/Econ	11/4/16	11/7/16							
11	44	31	9	9	8	yougov-econ	YouGov/Econ	11/4/16	11/7/16							
12	45	49		7	yougov-econ	YouGov/Econ		11/4/16	11/7/16							
13	6	91		2	yougov-econ	YouGov/Econ		11/4/16	11/7/16							
14	88	6		5	yougov-econ	YouGov/Econ		11/4/16	11/7/16							
15	49	38		13	yougov-econ	YouGov/Econ		11/4/16	11/7/16							
16	40	45	5	2	7	lucid-the-tim	Lucid/The Tim	11/4/16	11/6/16							
17	43	46		5	6	bloomberg-s	Bloomberg/S	11/4/16	11/6/16	If						
18	6	86		3	5	bloomberg-s	Bloomberg/S	11/4/16	11/6/16	If						
19	85	5		6	4	bloomberg-s	Bloomberg/S	11/4/16	11/6/16	If						
20	44	38		10	8	bloomberg-s	Bloomberg/S	11/4/16	11/6/16	If						
21	41	44	4	2	8	bloomberg-s	Bloomberg/S	11/4/16	11/6/16	If						
22	43	47	4	4	1	abc-post-267	ABC/Post	11/3/16	11/6/16							
23	46	49		3	2	abc-post-267	ABC/Post	11/3/16	11/6/16							
24	44	50	4	3	1	monmouth-l	Monmouth L	11/3/16	11/6/16	If						
25	44	50		5	1	monmouth-l	Monmouth L	11/3/16	11/6/16							
26	44	48	3	4	1	fox-26784	FOX	11/3/16	11/6/16							
27	4	91	2	3	0	fox-26784	FOX	11/3/16	11/6/16							
28	86	9	2	3	1	fox-26784	FOX	11/3/16	11/6/16							
29	42	36	9	9	5	fox-26784	FOX	11/3/16	11/6/16							
30	44	48		2	7	fox-26784	FOX	11/3/16	11/6/16							
31	4	90		1	4	fox-26784	FOX	11/3/16	11/6/16							
32	85	8		1	6	fox-26784	FOX	11/3/16	11/6/16							
33	43	37		3	18	fox-26784	FOX	11/3/16	11/6/16							
34	43	41	6	4	5	ibd-tipp-267	IBD/TIPP	11/3/16	11/6/16							
35	42	43		5	9	ibd-tipp-267	IBD/TIPP	11/3/16	11/6/16							
36	39	44		9	8	ipsos-reuters	Ipsos/Reuter	11/2/16	11/6/16	If						
37	7	84		5	5	ipsos-reuters	Ipsos/Reuter	11/2/16	11/6/16	If						
38	81	6		7	7	ipsos-reuters	Ipsos/Reuter	11/2/16	11/6/16	If the 2016 p Likely Voters - Republican		Internet	Nonpartisan	None		
39	24	29		28	19	ipsos-reuters	Ipsos/Reuter	11/2/16	11/6/16	If the 2016 p Likely Voters - independent		Internet	Nonpartisan	None		
40	39	42	6	6	7	ipsos-reuters	Ipsos/Reuter	11/2/16	11/6/16	If the 2016 p Likely Voters	2195	2.4	Internet	Nonpartisan	None	
41	7	81	4	3	4	ipsos-reuters	Ipsos/Reuter	11/2/16	11/6/16	If the 2016 p Likely Voters - Democrat		Internet	Nonpartisan	None		
42	81	6	5	3	5	ipsos-reuters	Ipsos/Reuter	11/2/16	11/6/16	If the 2016 p Likely Voters - Republican		Internet	Nonpartisan	None		
43	24	28	15	16	17	ipsos-reuters	Ipsos/Reuter	11/2/16	11/6/16	If the 2016 p Likely Voters - independent		Internet	Nonpartisan	None		
44	43	45	4	5	4	rasmussen-2	Rasmussen	11/2/16	11/6/16	If the 2016 p Likely Voters	1500	2.5	IVR/Online	Nonpartisan	None	
45	43	47		5	5	cbs-times-26	CBS/Times	11/2/16	11/6/16	Likely Voters	1426	3	Live Phone	Nonpartisan	None	
46	7	88		1	4	cbs-times-26	CBS/Times	11/2/16	11/6/16	Likely Voters - Democrat		Live Phone	Nonpartisan	None		
47	86	6		5	3	cbs-times-26	CBS/Times	11/2/16	11/6/16	Likely Voters - Republican		Live Phone	Nonpartisan	None		

- 2016 Election Data:
- 100s scientific polls
 - Mixed models and sample frames
 - interviewer-administered telephone polls
 - interactive-voice-response polls
 - probability-based web polls
 - good quality non-probability based web polls
 - Inference and forecasting

(no comment)

 FiveThirtyEight
2016 Election Forecast

President
Updated Nov. 8, 2016

Senate
Updated Nov. 8, 2016

Analysis
Updated Nov. 9, 2016

We're forecasting the
election with three models

Polls-plus forecast
What polls, the economy and
historical data tell us about Nov. 8

Polls-only forecast
What polls alone tell us about Nov. 8

Now-cast
Who would win the election if it
were held today

 National overview

Updates
National polls

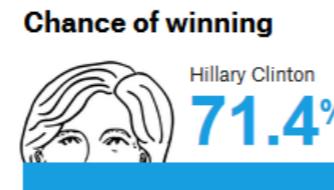
States to watch

Arizona
Colorado
Florida
Georgia
Iowa
Maine
Michigan
Minnesota
Nevada
New Hampshire
New Mexico
North Carolina
Ohio
Pennsylvania
Utah
Virginia
Wisconsin

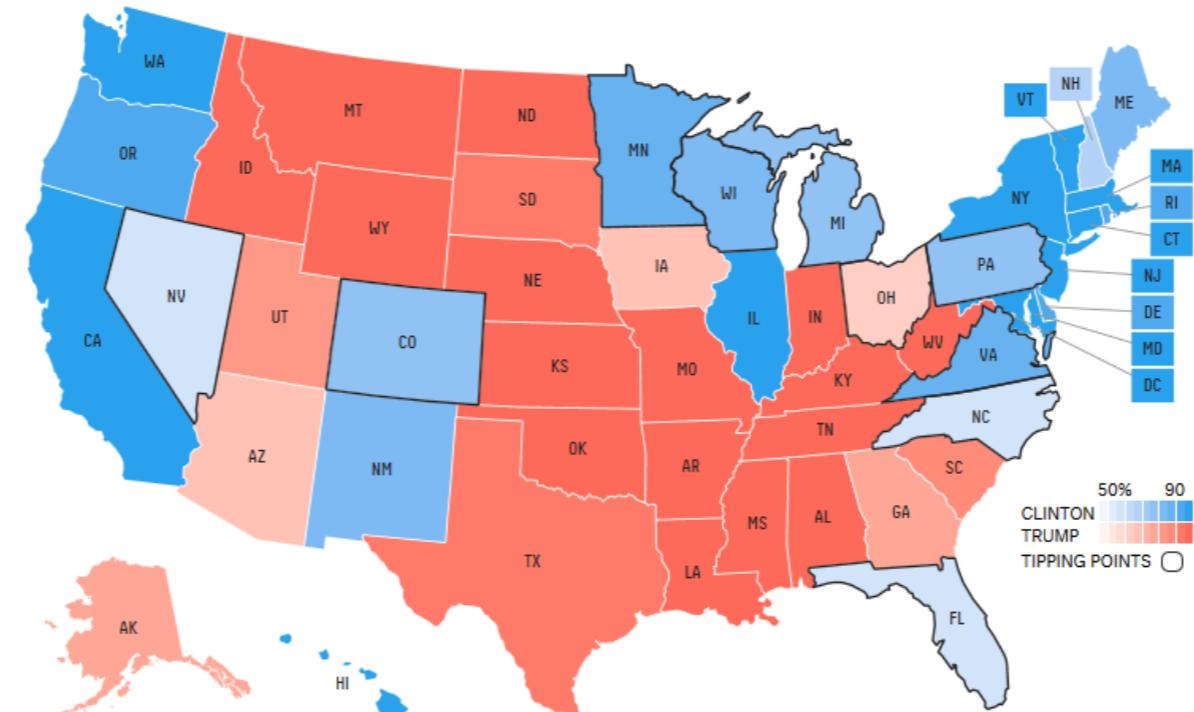
All states

Alabama
Alaska
Arizona
Arkansas
California
Colorado

Who will win the presidency?



Chance of winning



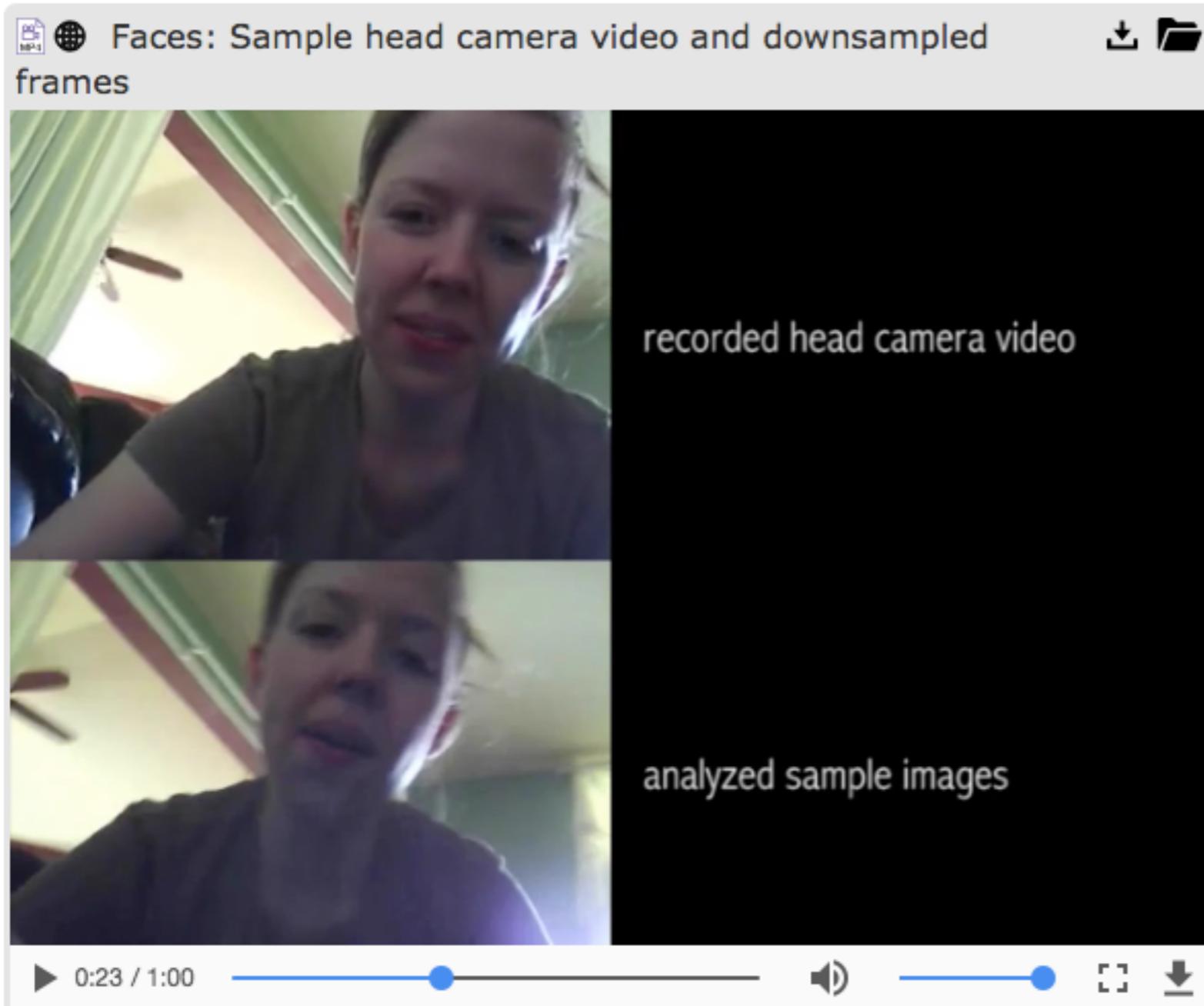
Electoral votes

 Hillary Clinton	302.2
 Donald Trump	235.0
 Evan McMullin	0.8
 Gary Johnson	0.0

Popular vote

 Hillary Clinton	48.5%
 Donald Trump	44.9%
 Gary Johnson	5.0%
 Other	1.6%

Video as Data

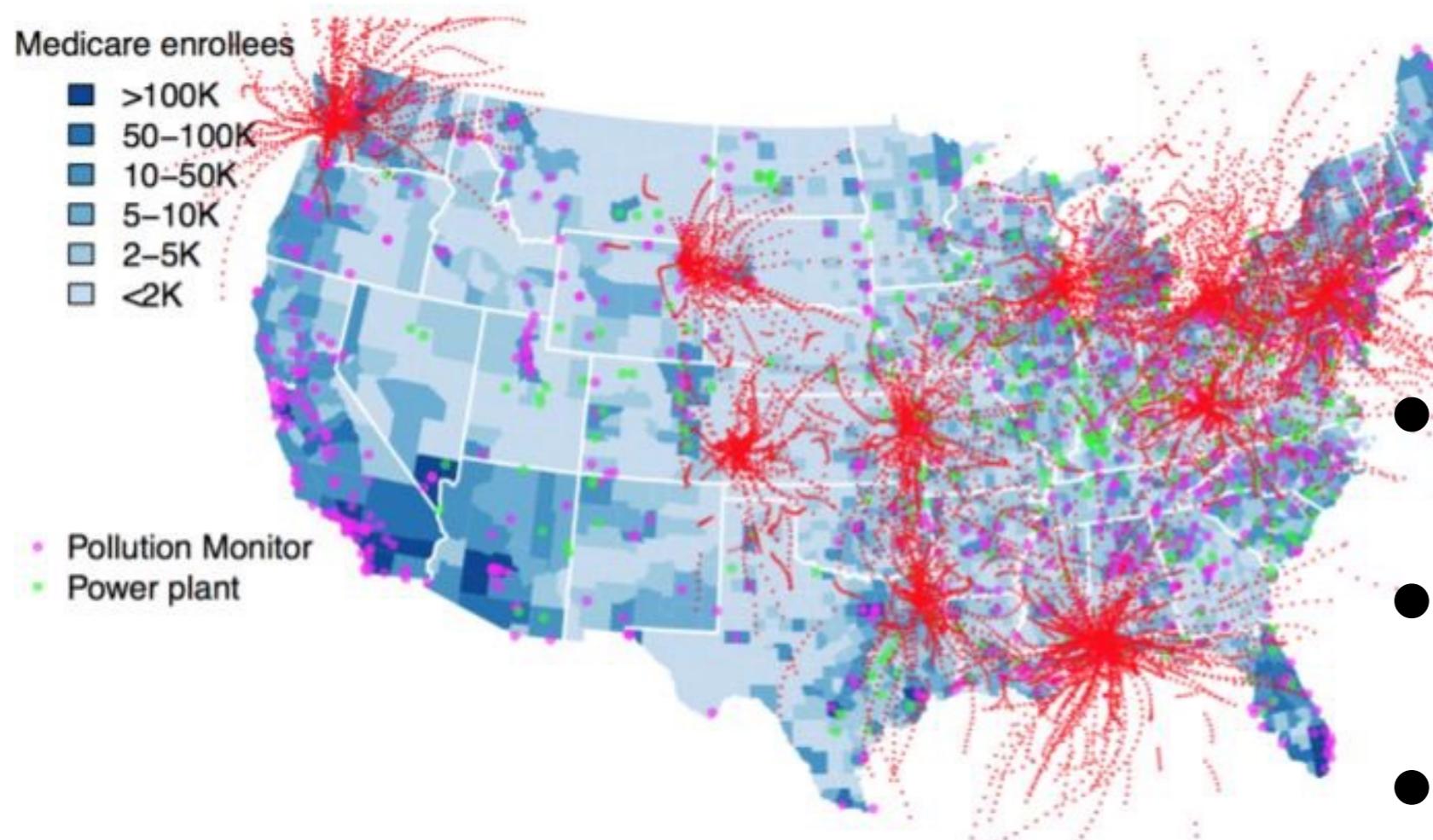


Video in developmental psychology:

- 143 hours of infant-perspective scenes, collected from 34 infants aged 1 month to 2 years
- Coding and annotations

Fausey, C.M., Smith, L.B. & Jayaraman, S. (2015). From faces to hands: Changing visual input in the first two years. Databrary

Pollution measurements as Data



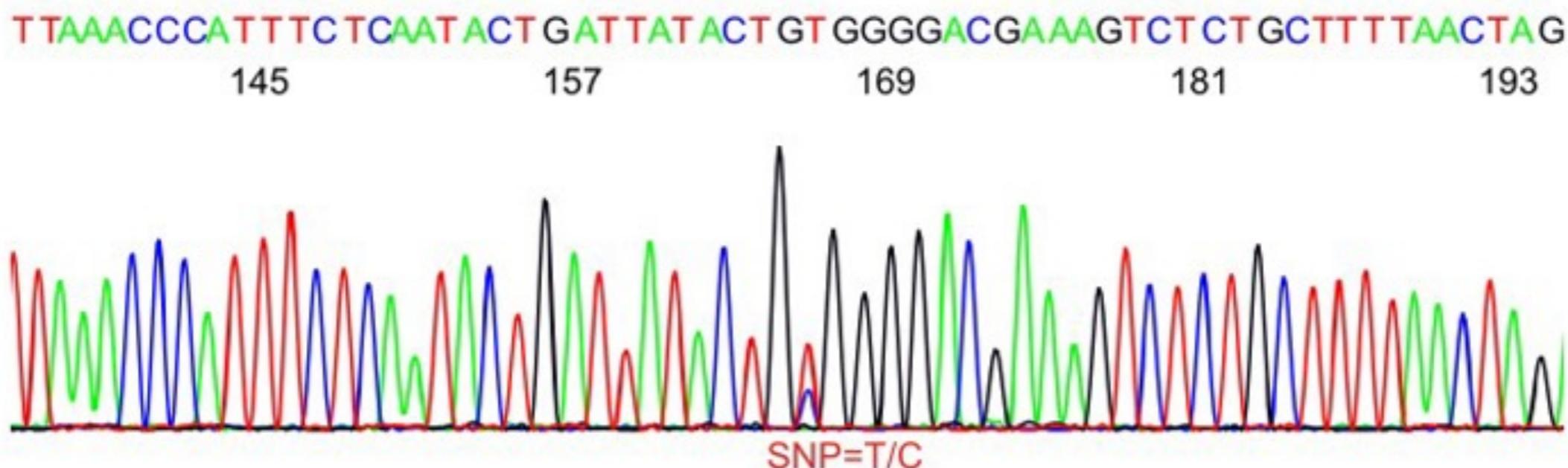
Combine data from diverse sources:

- 3000 air pollution monitors
- 1800 power plants SO₂ emissions
- claims data from 60 M Medicare beneficiaries
- air pollution trajectories based on weather models

Your Genome as Data

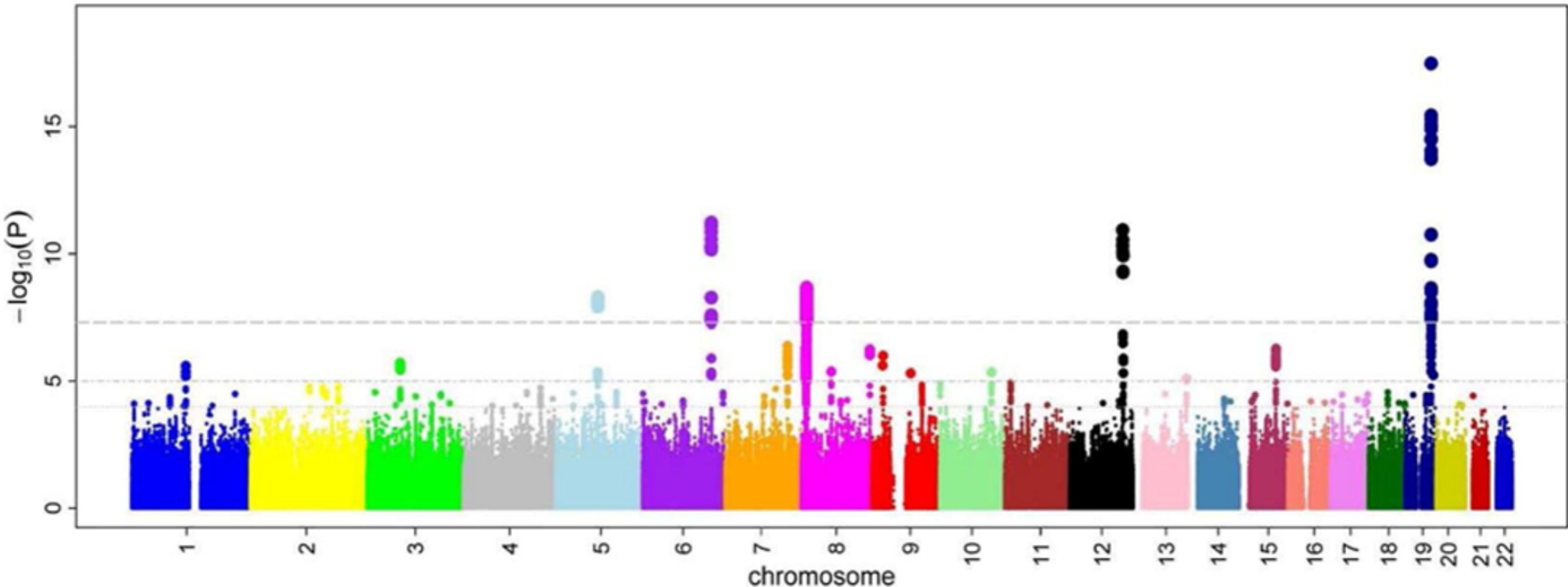


Measurement of Single-Nucleotide Polymorphisms (SNPs) with Sequencer instrument



Genome Wide Association Study

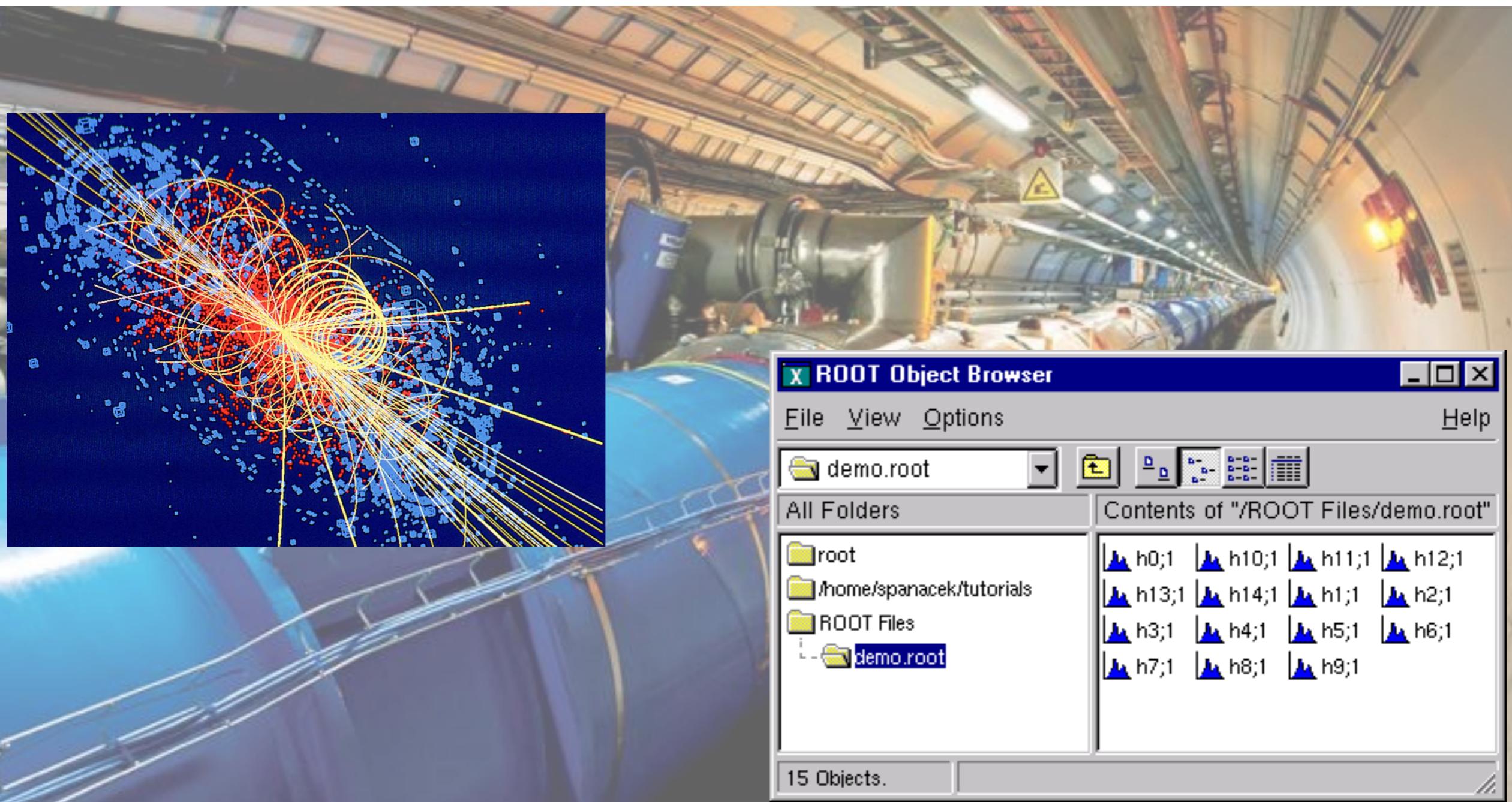
Comparing genomes from individuals with
(cases) and without disease (controls)



Genetic variants more often found in individuals with constrictions in small blood vessels
(GWAS,Wikipedia)

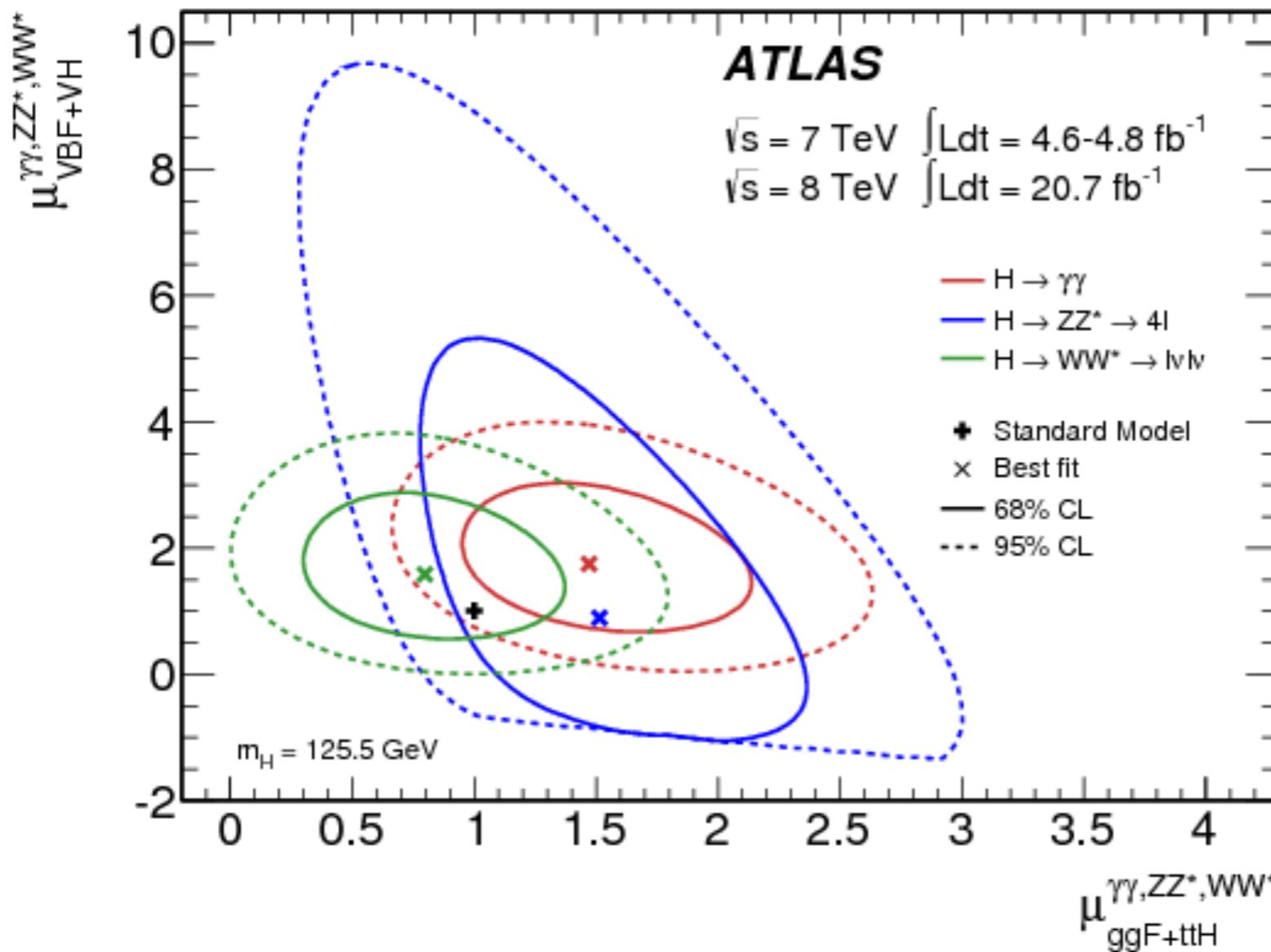
Particle Detections as Data

The Large Hadron Collider (LHC), CERN



Yes, there is a Higgs Boson!

(but is it the one we expected?)



Likelihoods as result
of comparing LHC
observations with
predictions from
particle physics
Standard Model

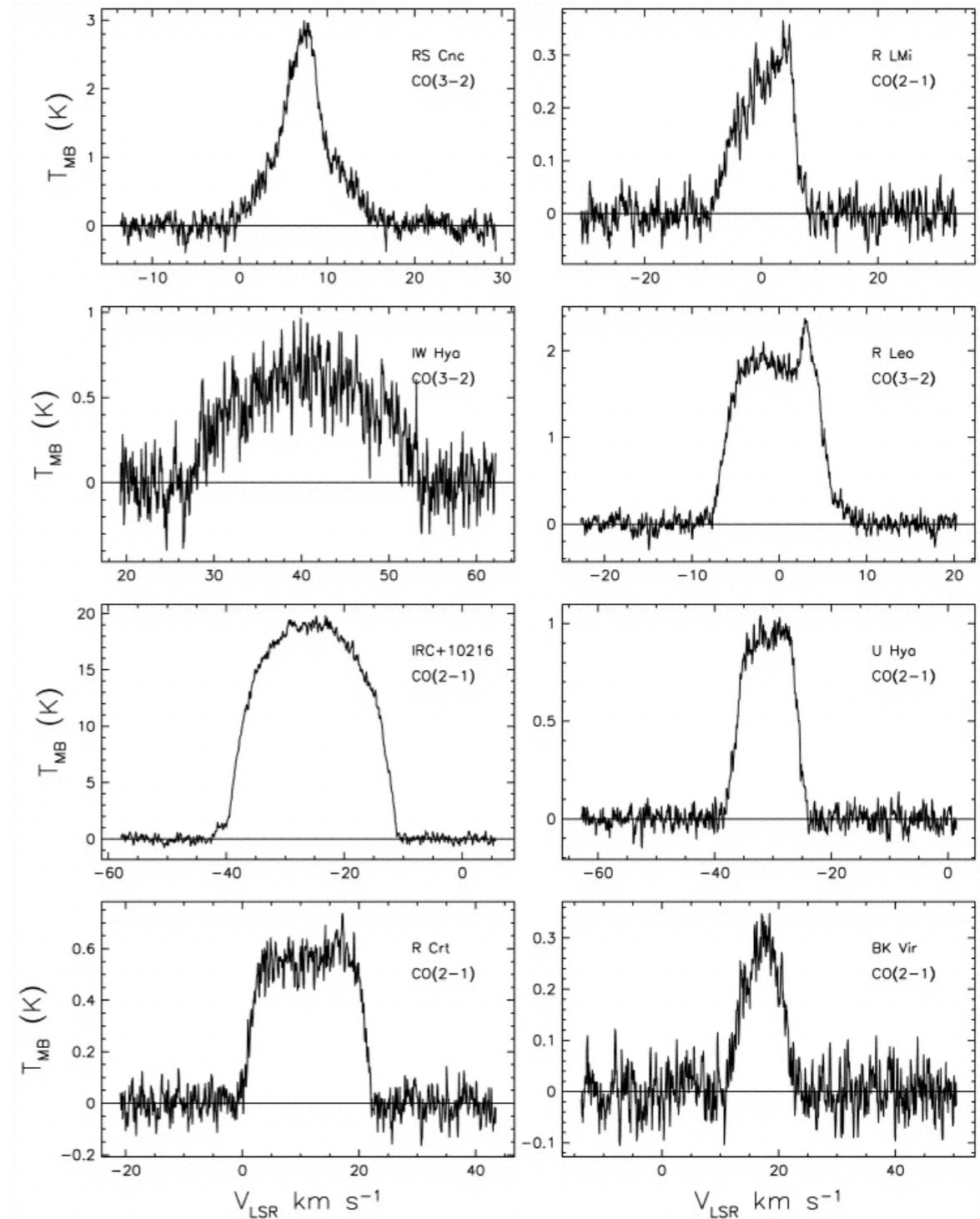
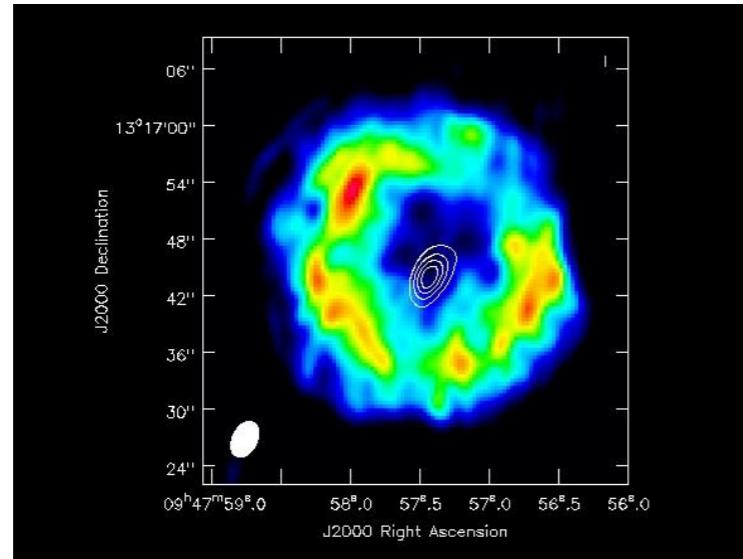
Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC (Atlas Collaboration, 2013, with 2923 authors).

Acknowledgment Kyle Cranmer (NYU)

Radio Emissions as Data



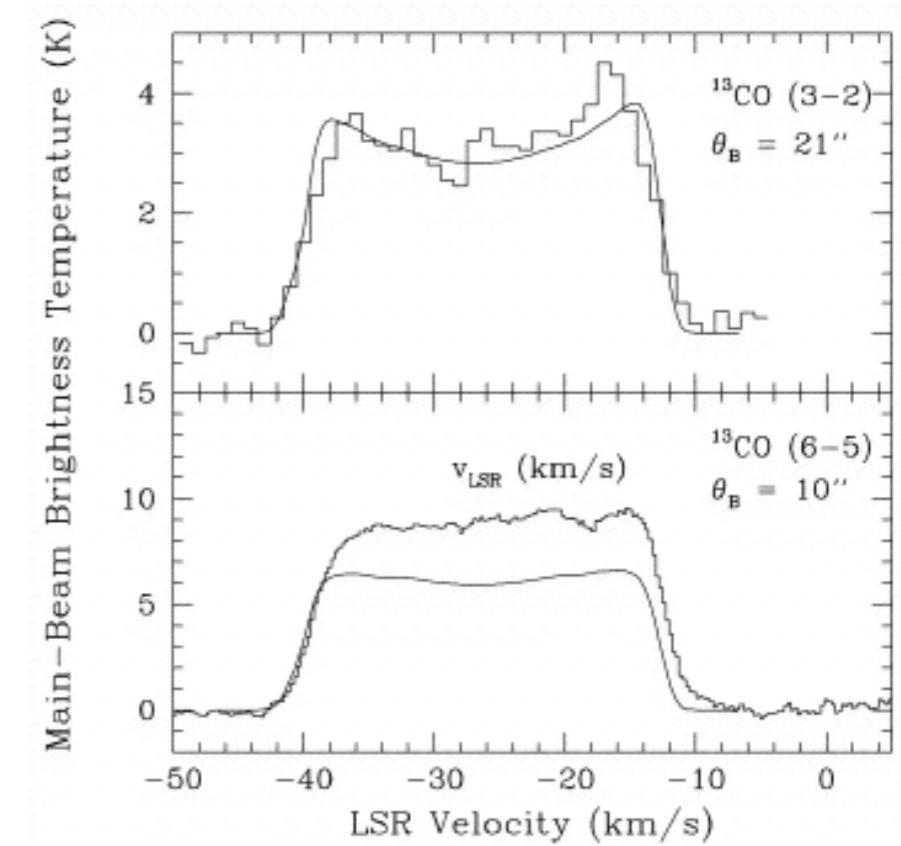
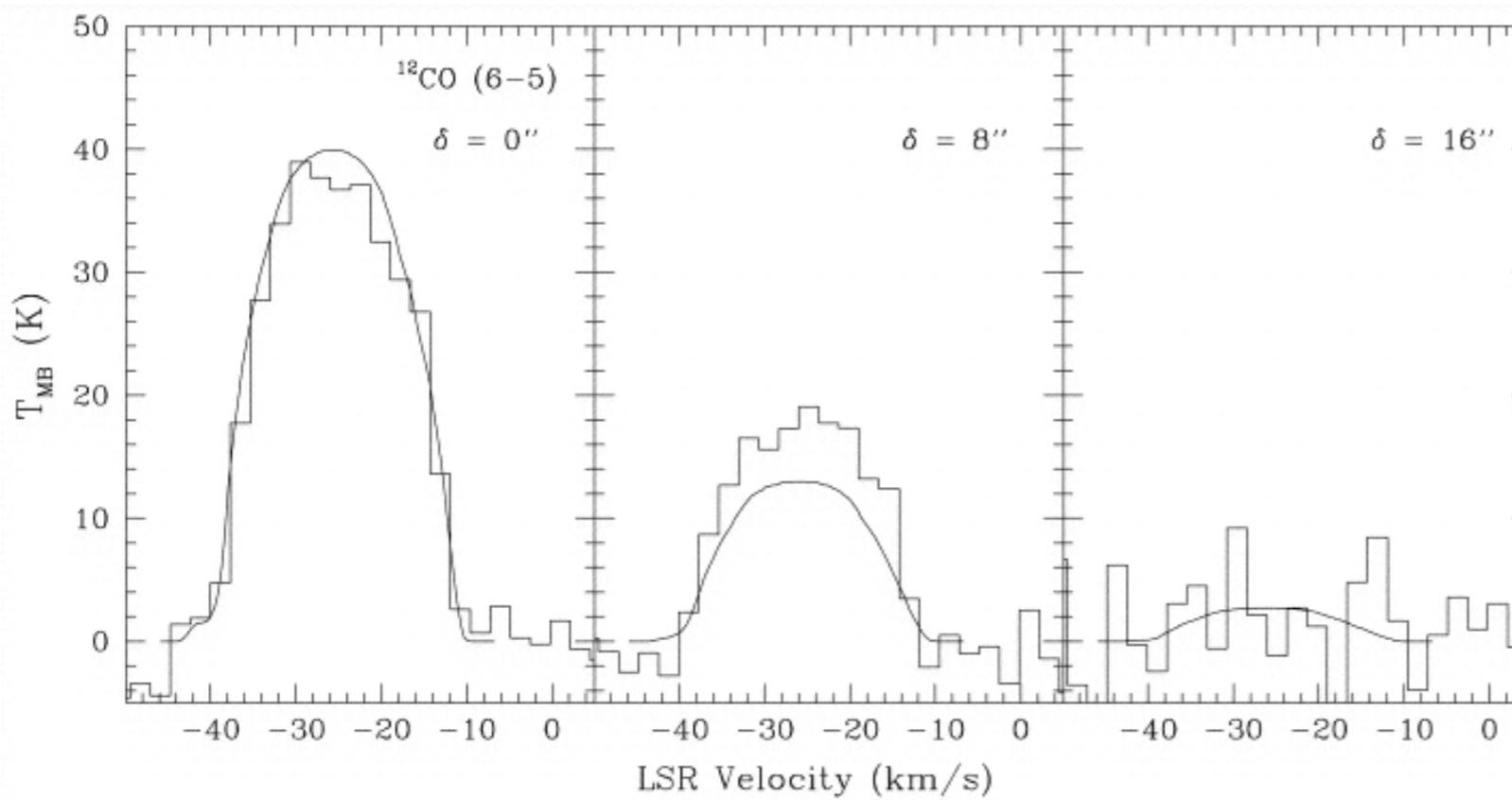
Caltech Submillimeter Observatory, California



Temperature, Mass, and Distance derived from Stellar Molecular Lines

$$\frac{dT}{dr} = (2 - 2\gamma) \left(1 + 0.5 \frac{d \ln v}{d \ln r} \right) \frac{T}{r} + \frac{\gamma - 1}{n_{\text{H}_2} k v} (H - C). \quad (11)$$

Monte Carlo simulations of radiative transfer compare observations with predictions from physical laws



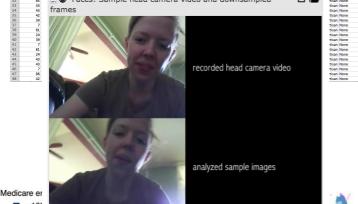
raw or primary
data



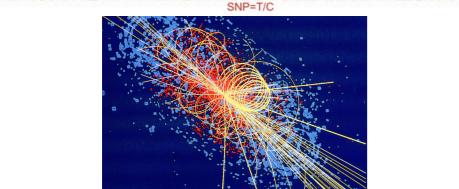
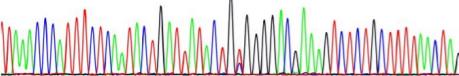
order, transform,
and analyze them

Humanities

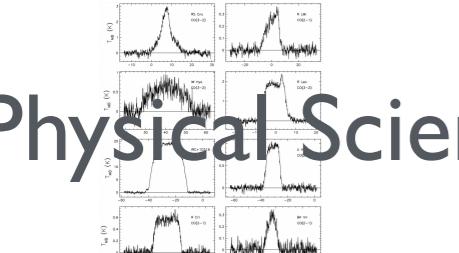
POSITION	Photo Note	Termin Notes	ID	TYPE	NAME	DATE
10.1442 - 26.1862	Frontier replicated	Needs static event map	1	Composite	Frontier	4/9/1700
10.1442 - 26.1862	Frontier replicated	Needs static event map	1	Robot	Frontier	4/9/1700
10.1442 - 26.1862	Still glitchy, want to fix it	Needs static event map	1	Robot	Ft. Harrison	4/9/1700
10.1442 - 26.1862	Still glitchy, want to fix it	Needs static event map	1	Robot	Tower	4/9/1700



Social Sciences



Health and
Life Sciences



Physical Sciences

Catalog
Classify
Visualize
Quantify
Summarize
Geo reference
Inference
Missing data
Forecast
Causal Inference
Coding
Annotations
Associations
Likelihoods
Compare with theory

gain knowledge,
make decisions

Learn about
the whole
from a part.

Tell a story.

Make a
prediction.

Ultimately
explain.

“Nullius in Verba”

(1665, the Royal Society motto and origin of Philosophical Transactions)



- **Replication***: Independent scientific experiments to validate findings
- **Reproducibility***: Calculation of quantitative results by others using original datasets and methods

(Stodden, Leisch, Peng, Implementing Reproducible Research, 2014)

* Replication and reproducibility definitions vary across disciplines

“Answering even a simple scientific question requires lots of choices that can shape the results”

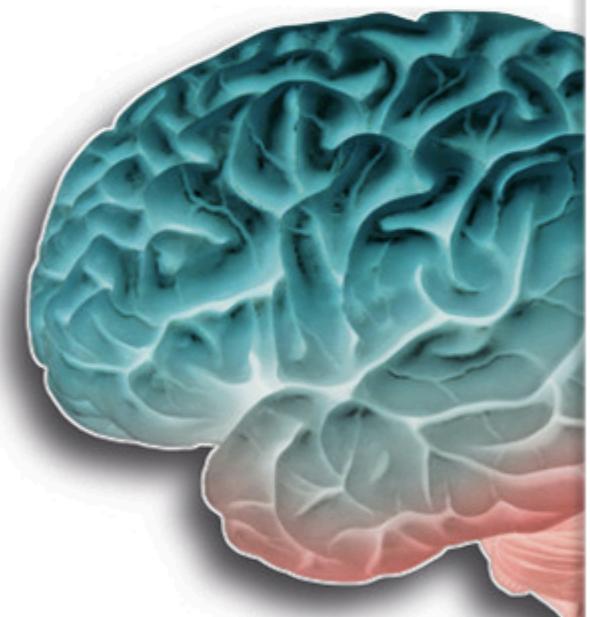
The Economist World politics Business & finance Econ

Problems with scientific research

How science goes wrong

Scientific research has changed the world. Now it needs to change itself.

Oct 19th 2013 | From the print edition



A SIMPLE idea underpins science: “trust, but verify”. Results must be challenged by experiment. That simple but powerful idea has built our knowledge. Since its birth in the 17th century, modern science has moved beyond recognition, and overwhelmingly for the better.

INSIGHTS



SCIENTIFIC INTEGRITY

Self-correction in science at work

Improve incentives to support research integrity

By Bruce Alberts,¹ Ralph J. Ciccone,² Stephen E. Fienberg,³ Alexander Kamb,⁴ Marcia McNutt,^{5*} Robert M. Nerem,⁶ Randy Schekman,⁷ Richard Shiffrin,⁸ Victoria Stodden,⁹ Subra Suresh,¹⁰ Maria T. Zuber,¹¹ Barbara Kline Pope,¹² Kathleen Hall Jamieson^{13,14}

Week after week, news outlets carry word of new scientific discoveries, but the media sometimes give suspect science equal play with substantive discoveries. Careful qualifications about what is known are lost in categorical headlines. Rare instances of misconduct or instances of irreproducibility are translated into concerns that science is broken. The Octo-

ber 2013 *Economist* headline proclaimed “Trouble at the lab: Scientists like to think of science as self-correcting. To an alarming degree, it is not” (*I*). Yet, that article is also rich with instances of science both policing itself, which is how the problems came to

The Economist’s attention in the first place, and addressing discovered lapses and irreproducibility concerns. In light of such

activities of science policing, to be leveled in any other result, as Popp of the very few the c are s fairly

(3). Instances in address flaws in of success, not f onstrate the unc nisms of science

Still, as in all writ large does n als. Although att Wakefield study between autism

FiveThirtyEight Science



THE SCIENTIFIC METHOD | 7:00 AM | AUG 19, 2015

Science Isn’t Broken

It's just a hell of a lot harder than we give it credit for.

By CHRISTIE ASCHWANDEN
Graphics by RITCHIE KING

If you follow the headlines, your confidence in science may have taken a hit lately. Peer review? More like self-review. An investigation in November uncovered a scam in which researchers were rubber-stamping their own work, circumventing peer review at five high-profile

“When possible, make data, methods, and code open to verify”

“Science/research might be imperfect, but is self-correcting”

“It’s not unreliable, but more challenging that we give it credit for”



SCIENTIFIC DATA



Altmetric: 442

Views: 29,187 Citations: 32

[More detail >](#)

Comment | [OPEN](#)

The FAIR Guiding Principles for scientific data management and stewardship

Mark D. Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E. Bourne, Jildau Bouwman, Anthony J. Brookes, Tim Clark, Mercè Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris T. Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, Alasdair J.G. Gray, Paul Groth, Carole Goble, Jeffrey S. Grethe, Jaap Heringa, Peter A.C. 't Hoen, Rob Hooft, Tobias Kuhn, Ruben Kok, Joost Kok, Scott J. Lusher, Maryann E. Martone, Albert Mons, Abel L. Packer, Bengt Persson, Philippe Rocca-Serra, Marco Roos, Rene van Schaik, Susanna-Assunta Sansone, Erik Schultes, Thierry Sengstag, Ted Slater, George Strawn, Morris A. Swertz, Mark Thompson, Johan van der Lei, Erik van Mulligen, Jan Velterop, Andra Waagmeester, Peter Wittenburg, Katherine Wolstencroft, Jun Zhao & Barend Mons - Show fewer authors

**Findable
Accessible
Interoperable
Reusable
(data, code,
methods)**

Harvard Data Resources

Harvard Library Services and Archives	Harvard Dataverse (IQSS, Library, HUIT)	Research Computing Council	Harvard Data Group (OVPR)	VPAL Research Group	
HMS Data Management group	Harvard Chan Bioinformatics Core	DARTH Arts and Humanities Computing	Center for Geographic Analysis	IQSS Data Science Services & Survey Research	HBS Research Computing Services

Services and Resources

Orchestra (HMS)	Odyssey (FAS)	RCE (IQSS)	HBS RC (HBS)
--------------------	------------------	---------------	-----------------

Computing & Storage Infrastructure

DataFest

Follow your sessions
using SCHED

Tuesday, January 17

Data Concepts

Project Planning a Data
Science Workflows

Wednesday, January 18

Processing and
Analyzing Data*

Dissemination

Plan how to
address your
research question

Address it!

Share your
results, data, code
with others



*Does not include handling sensitive data

Don't miss!

8:30am	Continental Breakfast	Computing Help Desk
10:00am	Coffee Break	Archive-a-thon signup
12:00pm	Lunch	Archive-a-thon signup
5:30pm	Reception	

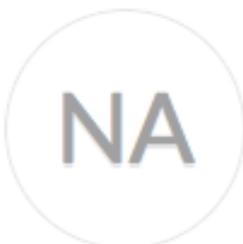
**If you leave DataFest thinking
“I need to learn more about X”,
we have already succeeded.
(however, we welcome your
feedback, using SCHEd)**

Harvard DataFest 2017

Schedule Speakers Attendees About Tickets Search



Troy Adair
Harvard Business School
Senior Director, Research Computing Services



Nicole Alexander



Sonia Barbosa



Lex Berman
Center for Geographic Analy
Web Services Mgr, GIS Specia



Judson Harward
Harvard University
Director, Arts&Humanities Research Computing



Amir Karger
Harvard Medical School
Associate Director of Research Computing



Radhika Khetani



Scott Lapinski
Harvard Medical School
Scholarly Communication & Research Librarian



Jeff Blossom
Harvard University



Daina Bouquin
Harvard-Smithsonian Center for Astrophysics
Head Librarian



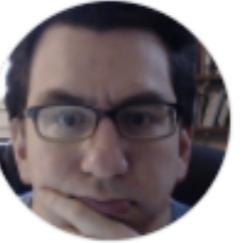
Alex Caracuzzo
Baker Library, Harvard Business School
Collections and Data Management Librarian



Kareem Carr



Ben Lewis



Derek Miller
Harvard University
Assistant Professor of English



Christopher Morse
Harvard University Arts & Humanities Research Computing
Senior Humanities Research Computing Specialist



David Osterbur
Harvard Medical School, County Library of Medicine
Director of Public and Access Services



Christine Choirat



Mercè Crosas
IQSS, Harvard University



Gustavo Durand



Bob Freeman
Harvard Business School
Dir, Research TechOps



Jessica Pierce



Caroline Shamu
Harvard Medical School
Assistant Professor



Diane Sredl
Maps, Media, Data, and Government Information, Lamont Library
Data Reference Librarian



Dustin Tingley
Harvard University
Professor of Government



Christopher Gandrud
IQSS



Simo Goshev
IQSS, Harvard University



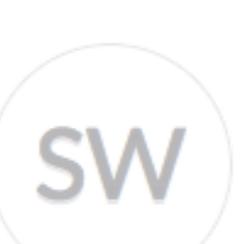
Wendy Guan



Chase Harrison



Hugh Truslow
Harvard University
Head, Social Sciences and Visualization



Steven Worthington



FAS Research Computing
Sr. Team Lead for HPC



The Institute for Quantitative Social Science at Harvard
Data Science Specialist