

Lecture 1: The Revolution is here!

D Donoho
Stats 285 Stanford

20210329

Outline

[The Computing Discontinuity](#)

- [Mobile is Eating the world](#)
- [Mobile Drives IT Revolution](#)
- [AWS is Eating the World](#)
- [New AWS Services are Proliferating](#)

[The Revolution in Computational Science](#)

[Case Study: Deep Learning](#)

- [The Sudden Emergence of Deep Learning](#)
- [Emergence of Prediction Challenges](#)
- [The Emergence of the Common Task Framework](#)
- [CTF Goes Mainstream](#)

[Digital Transformation of Science](#)

[Resistance](#)

- [Intellectual impoverishment](#)
- [Solution: The Great Enrichment](#)

[Painless Computational Experiments](#)

Disclaimer

This document contains images obtained by routine Google Images searches. Some of these images may perhaps be copyright. They are included here for educational noncommercial purposes and are considered to be covered by the doctrine of Fair Use. In any event they are easily available from Google Images.

It's not feasible to give full scholarly credit to the creators of these images. We hope they can be satisfied with the positive role they are playing in the educational process.

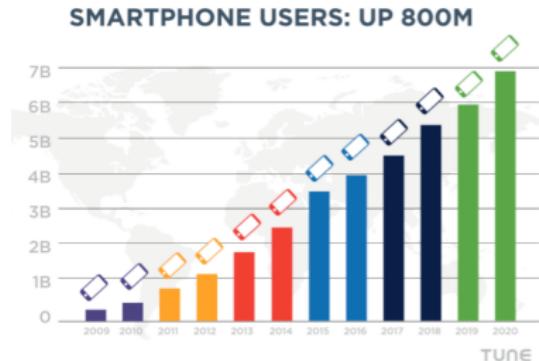
The Computing Discontinuity
The Revolution in Computational Science
Case Study: Deep Learning
Digital Transformation of Science
Resistance
Painless Computational Experiments

Mobile is Eating the world
Mobile Drives IT Revolution
AWS is Eating the World
New AWS Services are Proliferating

The Mobile Revolution

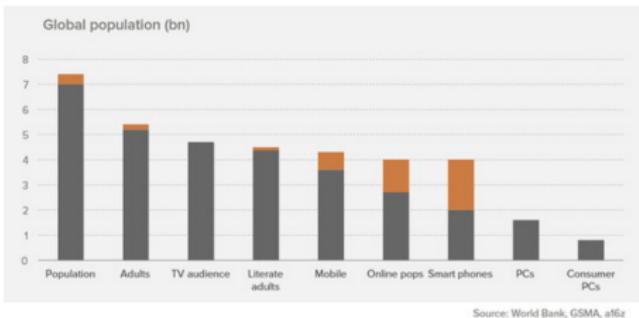


Smartphones are Spreading Everywhere



The world in 2020

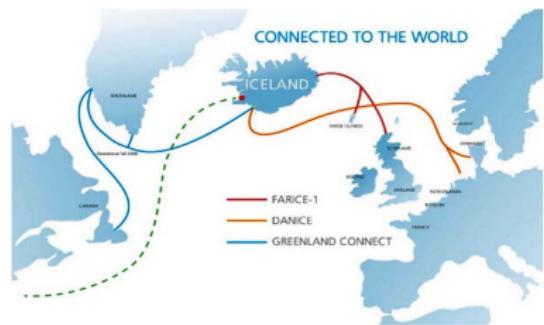
By 2020 80% of the adults on earth will have a smartphone



The Computing Discontinuity
The Revolution in Computational Science
Case Study: Deep Learning
Digital Transformation of Science
Resistance
Painless Computational Experiments

Mobile is Eating the world
Mobile Drives IT Revolution
AWS is Eating the World
New AWS Services are Proliferating

24/7 Deluge Spawns Global Computational Services



Cloud Paradigm

Cloud Paradigm:

- ▶ Billions of smart devices each drive queries to cloud servers
- ▶ Millions of business relying on cloud for all needs

Symbiosis of cloud and economy is *lasting* and *disruptive*.

AWS is Eating the world: Stock Market



TECH

TECH | MOBILE | SOCIAL MEDIA | ENTERPRISE | CYBERSECURITY | TECH GI

Amazon shares soar after massive earnings beat

- Amazon reported its third quarter results Thursday after the bell.
- It was a huge beat across the board.
- Amazon shares jumped over 7 percent in after hours trading.

Eugene Kim | @eugenekim222

Published 3:24 PM ET Thu, 26 Oct 2017 | Updated 6:55 PM ET Thu, 26 Oct 2017

CNBC

AWS is Eating the World, II

Amazon Web Services sales

Amazon will break out specific sales data for AWS on Thursday for the first time. Here's Robert W. Baird & Co. analyst Colin Sebastian estimates.

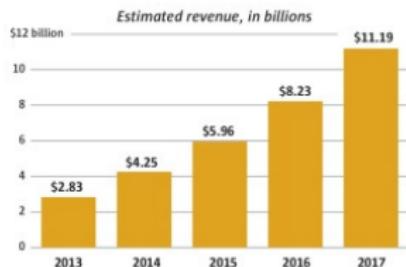
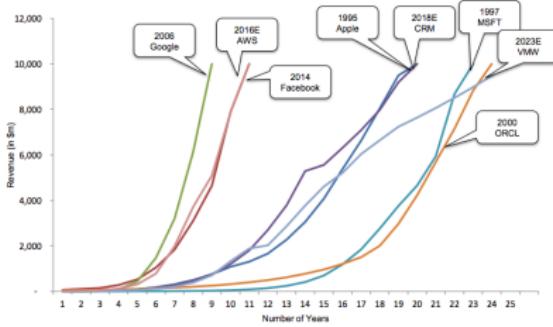
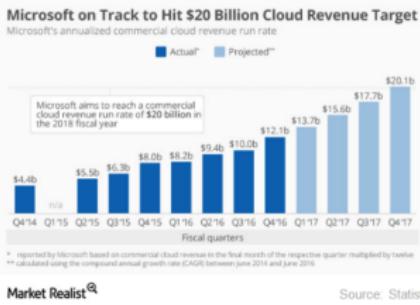


Figure 9: AWS is the Fastest-Growing Enterprise Technology Company Ever



Source: Deutsche Bank Estimates, Public Company Filings

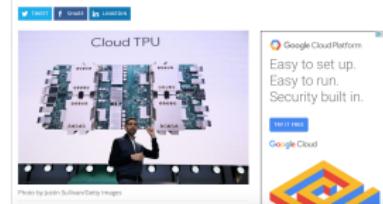
AWS is Eating the World: III



ALPHABET | GOOGLE | CLOUD COMPUTING
Google says it has tripled its big cloud deals over the last year

The number shows Google's cloud business is growing, but doesn't give us a good sense of how well it's competing.

BY REED HOLLOWAY | JUL 24, 2017 7:49PM EST



AWS Services Are Ubiquitous

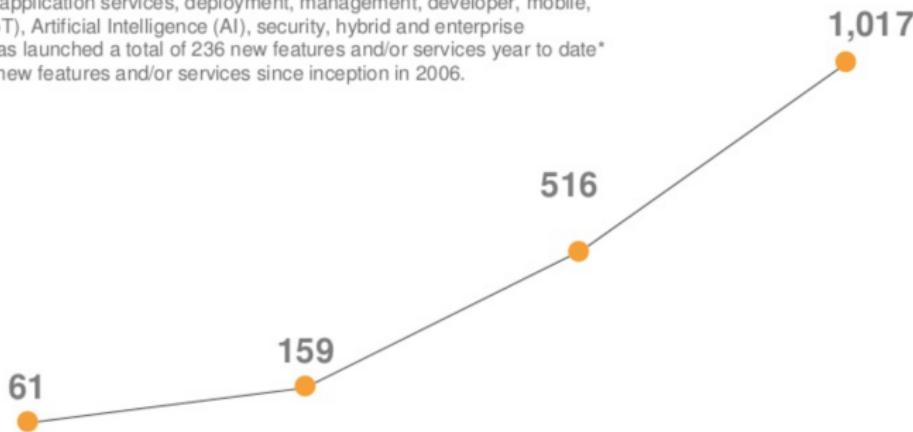
The AWS Platform



AWS Services are Proliferating

AWS Pace of Innovation

AWS has been continually expanding its services to support virtually any cloud workload, and it now has more than 90 services that range from compute, storage, networking, database, analytics, application services, deployment, management, developer, mobile, Internet of Things (IoT), Artificial Intelligence (AI), security, hybrid and enterprise applications. AWS has launched a total of 236 new features and/or services year to date* - for a total of 3,149 new features and/or services since inception in 2006.



Stack Paradigm I

Stack Paradigm:

- ▶ Organizations combine software components from other providers in a stack
- ▶ Massive new capabilities emerge by hybridizing components

Examples:

- ▶ Uber (next slide)
- ▶ Netflix relies on AWS
- ▶ Snap, Dropbox etc. small teams

Stack Paradigm II



Uber doesn't own their cars. They also don't directly employ their own drivers. So, one might ask, what do they own exactly as a core asset? The core application and ecosystem around the Uber experience is their primary asset and differentiator. But to deliver that experience, they apply rigorous focus.

At the practical level, when you look at the technology components of Uber's world-renowned app, they decided to rely on other core platforms and technologies to power many of the key elements.

Jeetu Patel, *Software is still eating the world*, TechCrunch, Jan 2016

Explosion of Computational Resources

Cloud Paradigm:

- ▶ Billions of smart devices each drive queries to cloud servers
- ▶ Millions of business relying on cloud for all needs

Symbiosis of cloud and economy is *lasting* and *disruptive*.

Cloud provides *any user same-day* delivery:

- ▶ Tens to hundreds of thousands of hours of CPU
- ▶ Pennies per CPU hour

Any user can consume *1 Million CPU hours* over a few days for a few \$10K's.

Massive Computational Power Will Transform *Science*

Traditionally:

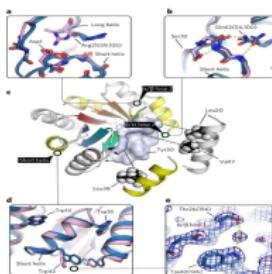
- ▶ Deduction (in math)
- ▶ Induction (in physical sciences)

Emerging new approach:

- ▶ Massive computational experiments

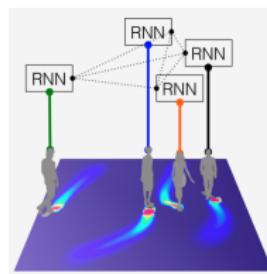
Massive Computations in Science

Traditionally computational fields



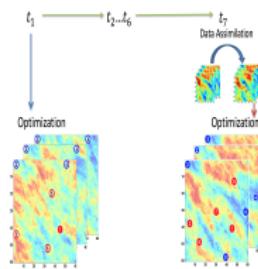
Protein Design

(Huang et al. 2016)



AI

(Alahi et al. 2016)

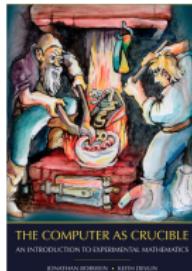
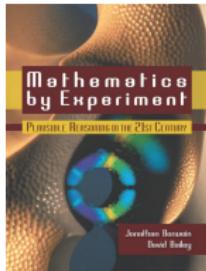


Oil Field Develop.

(Shirangi et al. 2015)

Massive Computations in Science

Traditionally **non-** computational field – Mathematics



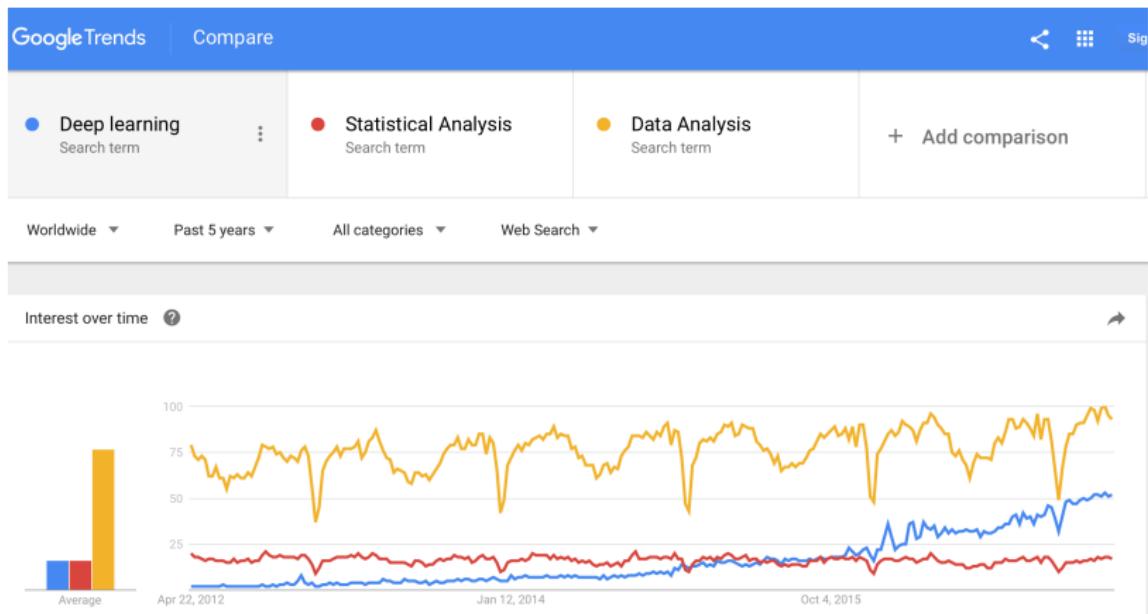
Borwein/Bailey

Borwein/Devlin

Individual Articles

The Computing Discontinuity
The Revolution in Computational Science
Case Study: Deep Learning
Digital Transformation of Science
Resistance
Painless Computational Experiments

The Sudden Emergence of Deep Learning
The Emergence of the Common Task Framework
CTF Goes Mainstream



The Computing Discontinuity
The Revolution in Computational Science
Case Study: Deep Learning
Digital Transformation of Science
Resistance
Painless Computational Experiments

The Sudden Emergence of Deep Learning
The Emergence of the Common Task Framework
CTF Goes Mainstream

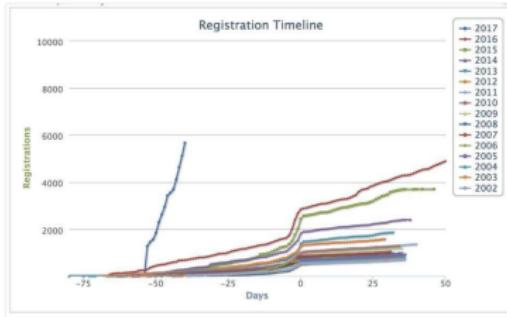


Alex Lebrun
@lxbrun

Follow



Deep learning hype in one picture
(NIPS conference registrations, 2002 through
2017) #nips2017



8:20 AM - 15 Sep 2017

758 Retweets 1,005 Likes



20 758 1.0K

Synchronies, 1

Over same timeframe – 2010-2014

- ▶ Instagram, Snapchat emerge to global prominence
- ▶ Deep Learning catapults to global attention

Coincides with emergence of

- ▶ Smartphone photography
- ▶ Cloud computing
- ▶ Cloud storage of selfie/smartphone photography

Synchronies, 2

"Six decades into the computer revolution, four decades since the invention of the microprocessor, and two decades into the rise of the modern Internet, all of the technology required to transform industries through software finally works and can be widely delivered at global scale."

Marc Andreessen - WSJ - 2011

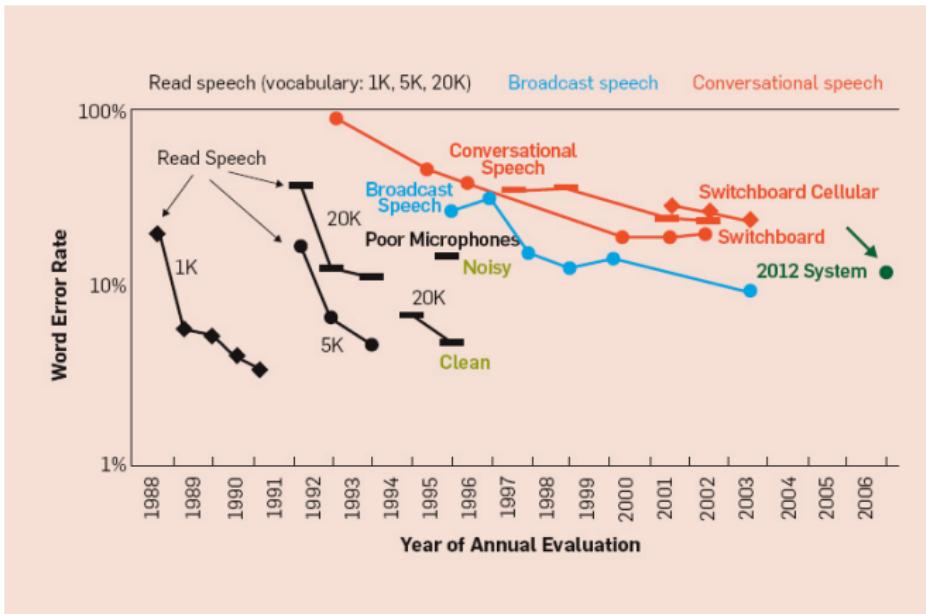
Common Task Framework (1980's)

Under CTF we have the following ingredients

- (a) A **publicly available training dataset** involving, for each observation, a list of (possibly many) feature measurements, and a class label for that observation.
- (b) A set of **enrolled competitors** whose **common task** is to **infer** a class **prediction rule from the training data**.
- (c) A **scoring referee**, to which competitors can submit their prediction rule. The referee runs the prediction rule against a testing dataset which is sequestered behind a Chinese wall. The referee objectively and automatically reports the score achieved by the submitted rule.

See Mark Liberman's description (Liberman, 2009).

CTF *Really* Works!



CTF Lifestyle – 1

1. Researchers set up local copies of Challenge
 - ▶ Data – Training, Test carved out of public dataset
 - ▶ Scoring – same as challenge scoring rule
2. Researcher's job: '*tuning models*'
 - ▶ Think up a family of model variations – '*tweak's*
 - ▶ Run a full '*experiment*' – suite of tweaks – '*grid*'
 - ▶ Score each tweak
 - ▶ Submit best-scoring result to central authority
3. Successful researchers perpetually motivated by
Game-ification: tweaking, scoring, winning.
4. Researchers who tweak more often, win more often!.
5. If easier to implement tweaks and faster to evaluate them,
more likely to win!.

The Computing Discontinuity
The Revolution in Computational Science
Case Study: Deep Learning
Digital Transformation of Science
Resistance
Painless Computational Experiments

The Sudden Emergence of Deep Learning
The Emergence of the Common Task Framework
CTF Goes Mainstream

CTF Lifestyle – 2



Sebastião Salgado Work

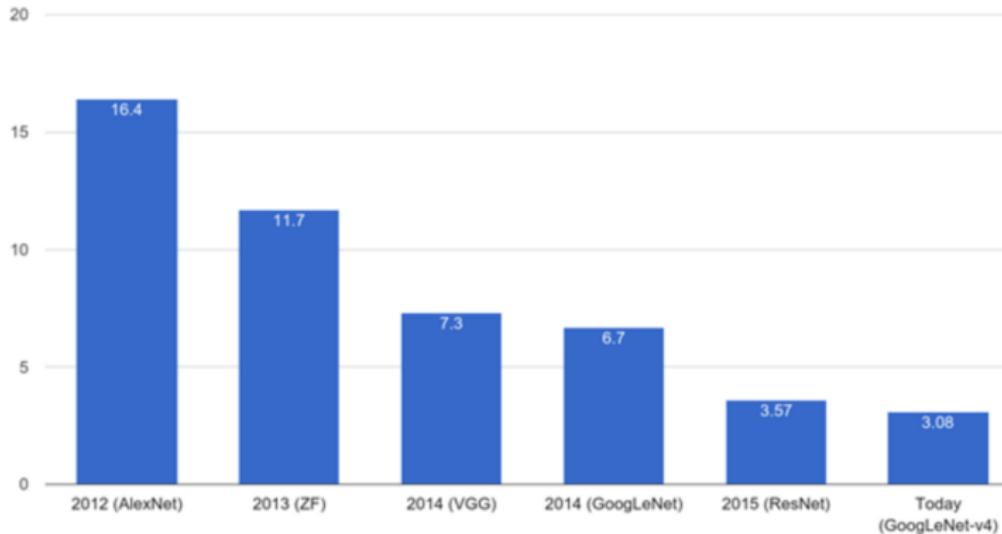
D Donoho Stats 285 Stanford

Lecture 1: The Revolution is here!

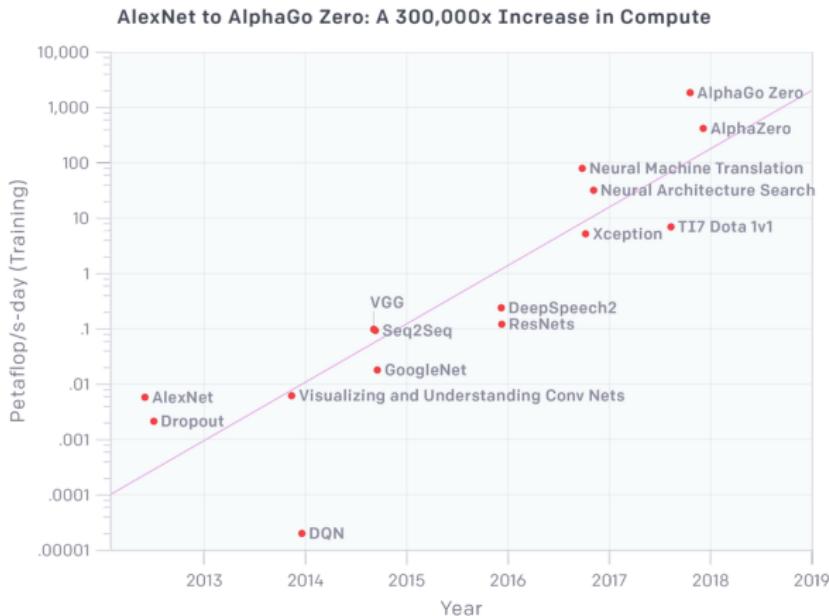
CTF Goes Mainstream

1. Netflix Challenge (2009)
\$1 Million Prize
2. Kaggle (2010)
1 Million'th competitor Sept. 2017
3. Fei-Fei Li masterminds ImageNet 2008-2010
4. Hinton's Deep Learning Team wins ImageNet 2012

ImageNet Classification Error (Top 5)



Machine Learning is Transformed



Computing Supremacy is here, 1

Massively Multilingual Neural Machine Translation in the Wild: Findings and Challenges

Naveen Arivazhagan * Ankur Bapna * Orhan Firat *
Dmitry Lepikhin Melvin Johnson Maxim Krikun Mia Xu Chen Yuan Cao
George Foster Colin Cherry Wolfgang Macherey Zhifeng Chen Yonghui Wu
Google AI

Abstract

We introduce our efforts towards building a universal neural machine translation (NMT) system capable of translating between any language pair. We set a milestone towards this goal by building a single massively multilingual NMT model handling 103 languages trained on over 25 billion examples.

2016a; Ha et al., 2016c; Johnson et al., 2017) - using a single model to translate between multiple languages.

Multilingual NMT models are appealing for several reasons. Let's assume we are interested in mapping between N languages; a naive approach that translates between any language pair from the given N languages requires $O(N^2)$ individ-

Recent Lecture in my class: <http://stats285.stanford.edu>

Computing Supremacy is here, 2

Massively Multilingual Neural Machine Translation in the Wild: Findings and Challenges

Naveen Arivazhagan * Ankur Bapna * Orhan Firat *

Dmitry Lepikhin Melvin Johnson Maxim Krikun Mia Xu Chen Yuan Cao

George Foster Colin Cherry Wolfgang Macherey Zhifeng Chen Yonghui Wu

Google AI

Abstract

We introduce our efforts towards building a universal neural machine translation (NMT) system capable of translating between any language pair. We set a milestone towards this goal by building a single massively multilingual NMT model handling 103 languages trained on over 25 billion examples.

2016a; Ha et al., 2016c; Johnson et al., 2017) - using a single model to translate between multiple languages.

Multilingual NMT models are appealing for several reasons. Let's assume we are interested in mapping between N languages; a naive approach that translates between any language pair from the given N languages requires $O(N^2)$ individ-

- BIG RESULTS:** Translate Any/Any of 105 Languages
BIG MODEL: 80 **Billion** Parameters
BIG DATA: 25 **Billion** Examples

Computing Supremacy will change Science

Unprecedented:

- ▶ ... deliverables
- ▶ ... scales
- ▶ ... contributions
- ▶ .. collaborations

Paul Valéry's visions from 100 years ago...



WIKIPEDIA
The Free Encyclopedia

Main page
Contents
Featured content
Current events
Random article
Donate to Wikipedia
Wikipedia store

Interaction
Help
About Wikipedia
Community portal
Recent changes
Contact page

Tools
What links here
Related changes
Upload file
Special pages
Permanent link
Page information
Wikidata item
Cite this page

In other projects
Wikimedia Commons
Wikiquote
Wikisource
Print/export

Article Talk

Not logged in Talk Contributions Create account Log in

Read Edit View history

Search Wikipedia



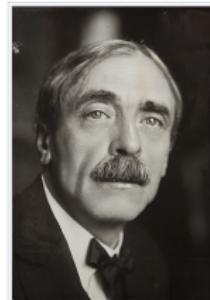
Paul Valéry

From Wikipedia, the free encyclopedia
(Redirected from Paul Valéry)

Ambroise Paul Toussaint Jules Valéry (French: [pol valéri]; 30 October 1871 – 20 July 1945) was a French poet, essayist, and philosopher. In addition to his poetry and fiction (drama and dialogues), his interests included aphorisms on art, history, letters, music, and current events. Valéry was nominated for the Nobel Prize in Literature in 12 different years.^[1]

Contents [hide]

- 1 Biography
- 2 Work
 - 2.1 The great silence
 - 2.2 *La Jeune Parque*
 - 2.3 Other works
 - 2.4 Technique
 - 2.5 Prose works
 - 2.6 The notebooks
- 3 In other literature
- 4 In popular culture
- 5 Selected works
- 6 See also
- 7 References
- 8 Further reading
- 9 External links



Paul Valéry



Biography [edit]

Valéry was born to a Corsican father and Genoese-Istrian mother in Sète, a town on the Mediterranean coast of the Hérault, but he was raised in Montpellier, a larger urban center close by. After a traditional Roman Catholic education, he studied law at university and then resided in Paris for most of the remainder of his life.

French literature

by category

French literary history

Medieval · Renaissance · 17th · 18th · 19th ·
20th century · Contemporary

French writers

Chronological list · Writers by category ·
Essayists · Novelists · Playwrights · Poets ·
Short story writers · Children's writers

Valéry's vision; a century later – original

Just as water, gas, and electricity are brought into our houses from far off to satisfy our needs in response to a minimal effort, so we shall be supplied with

visual or auditory images,
which will appear and disappear at a simple movement of the hand, hardly more than a sign.

after P. Valéry, Pièces sur L'Art, 1931; Le Conquête de l'ubiquité

Valéry's vision; a century later – updated, with apologies

Just as water, gas, and electricity are brought into our houses from far off to satisfy our needs in response to a minimal effort, so we shall be supplied with

results from massive computational experiments,
which will appear and disappear at a simple movement of the hand, hardly more than a sign.

after P. Valéry, Pièces sur L'Art, 1931; Le Conquête de l'ubiquité

Valéry's vision in Stats 285

Submitted to Harvard Data Science Review

1

Ambitious Data Science Can Be Painless*

Hatef Monajemi^{1,2}, Riccardo Murri³, Eric Jonas⁴, Percy Liang⁵, Victoria Stodden⁶ and David Donoho^{†,1}

January 28, 2019

Abstract—Modern data science research, at the cutting edge, can involve massive computational experimentation; an ambitious PhD in computational fields may conduct experiments consuming several million CPU hours. Traditional computing practices, in which researchers use laptops, PCs, or campus-resident resources with shared policies, are awkward or inadequate for experiments at the massive scale and varied scope that we now see in the most ambitious data science. On the other hand, modern cloud

I. INTRODUCTION

Tremendous increases in computing power in recent years are opening fundamentally new opportunities in science and engineering. Amazon, IBM, Microsoft and Google now make massive and versatile compute resources available on demand via their *cloud* infrastructure, making it in principle possible for a

Resistance – 1

We are at a university!

1. Q: *Where's the intellectual activity in tuning?*
2. Q: *I didn't come here to do hard manual labor!*
3. Q: *I didn't come here to compete as mindless drones!*

Resistance – 2

We are at a university!

1. Q: *Where's the intellectual activity in tuning?*
2. Q: *I didn't come here to do hard manual labor!*
3. Q: *I didn't come here to compete as mindless drones!*

What we see:



Sebastiao Salgado, *Work*

Resistance 3

We are at a university!

1. Q: *Where's the intellectual activity in tuning?*
2. Q: *I didn't come here to do hard manual labor!*
3. Q: *I didn't come here to compete as mindless drones!*

What we **imagine**:



Metaphor: Computers as Slavery

Traditionally, ‘using computers’ involves interactively running programs (Excel, Point-and-click)

Claerbout’s Dictum: “... dependence on an interactive program can be a form of slavery”

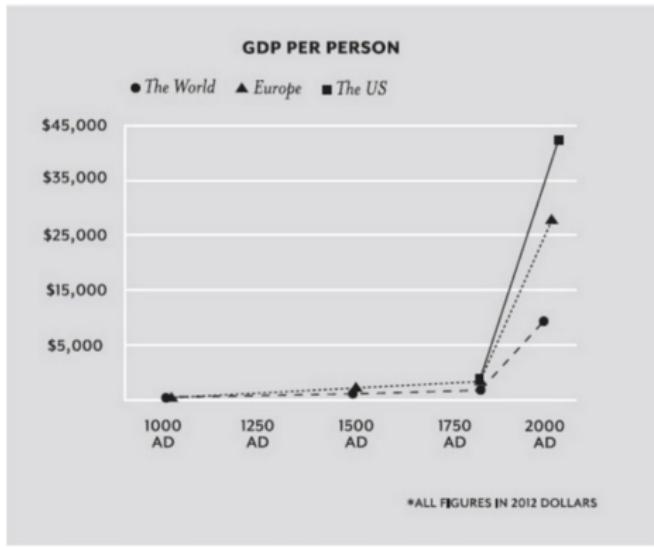
<http://sepwww.stanford.edu/sep/jon/reproducible.html>



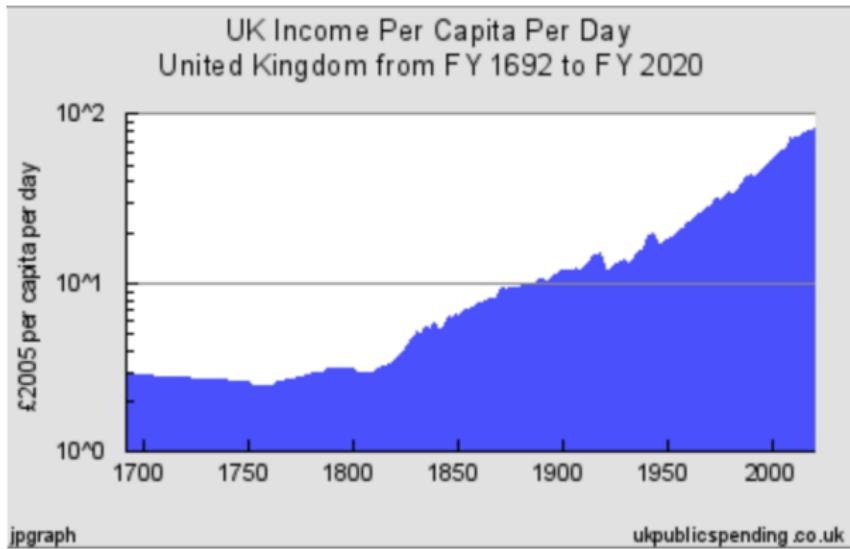
Photo: Jon Claerbout Cartoon: <http://fritsAhlefeldt.com>



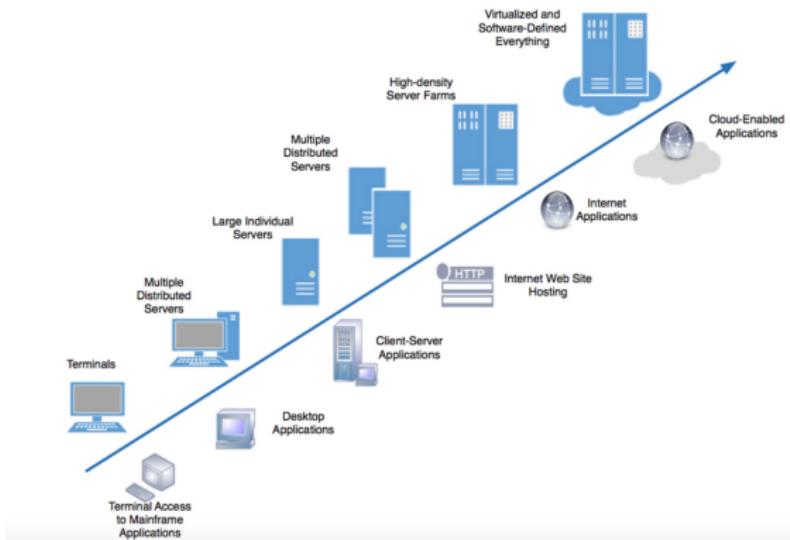
Digression: The Great Enrichment (Deidre McKloskey) 1



Digression: The Great Enrichment (Deidre McKloskey) 2



The Great IT Enrichment – 1



The Great IT Enrichment - 2

Our vision.

The intellectual poverty of the old interactive 'Excel'-era paradigm was real, but will be transcended.

New and better and more powerful abstractions will lift us out of the mud and out of slavery.

Coming Soon to a Scientific field near you

In the near future,

- ▶ Scientific research will be transformed
 - ▶ *1 million CPU Hours* behind research papers and theses
 - ▶ *Widespread acceptance* of empirical/simulation evidence
- ▶ 1 million-hour hurdle manageable through *new frameworks*.
- ▶ Frameworks offer Convenient and Efficient
 - ▶ ... definition of experiments
 - ▶ ... management of jobs
 - ▶ ... gathering of results
 - ▶ ... analysis and presentation
- ▶ Output:
 - ▶ Better science
 - ▶ Better math

Course Focus: Frameworks for Massive Experiments, 1

- ▶ Traditional issues
 - ▶ Experiments implicitly defined by executing unorganized code
 - ▶ Hard to understand what the baseline is, what variations are
 - ▶ Code dependencies unclear
 - ▶ Ordeal to get all the jobs to run, maybe gave up early
 - ▶ Tedious to harvest all the data, maybe missing some data
 - ▶ Confusing manual compilation and reporting
- ▶ Modern Frameworks
 - ▶ Systematic structure to coding
 - ▶ Base experiment clearly defined
 - ▶ Tweaks clearly defined
 - ▶ Code dependencies explicit
 - ▶ Grid of Jobs run systematically
 - ▶ Automatic transparent access of (cluster, AWS,...)
 - ▶ Data Harvested automatically to central data repository
 - ▶ Data analyzed automatically using defined tools

Course Focus: Frameworks for Massive Experiments, 2

- ▶ Example Frameworks
 - ▶ By individual research teams:
 - ▶ ClusterJob/ElastiCluster – Hatef Monajemi/Riccardo Murri
 - ▶ CodaLab – Percy Liang
 - ▶ Alpha – Vardan Papyan and XY Han
 - ▶ By startups:
 - ▶ Kedro
 - ▶ Ray

Global Economy → Computing → Science

