Title: TBA

Date: Mar 27, 2018 02:00 PM

URL: http://pirsa.org/18030113

Abstract:

Producing Knowledge in Academy and Industry

Andrew Mao

CTRL—labs

How did I end up here?

HARVARD
School of Engineering
and Applied Sciences

EconCS Group

Research at the intersection between computer science and economics

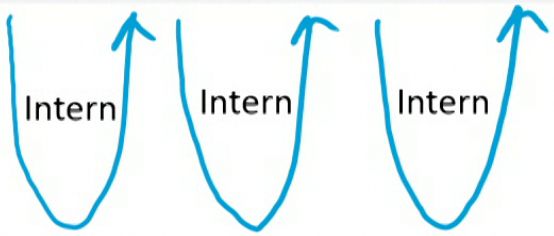
Microsoft®
Research

How did I end up here?

HARVARD
School of Engineering and Applied Sciences

EconCS Group

Research at the intersection between computer science and economics



Microsoft® Research

How did I end up here?

HARVARD
School of Engineering and Applied Sciences

EconCS Group

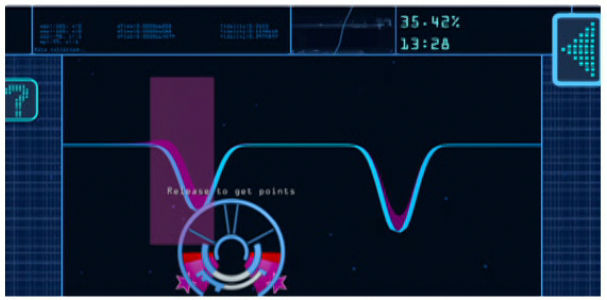
Research at the intersection between computer science and economics

Intern Intern Intern Postdoc

Microsoft® Research
Computational Social Science

Tenure-track

SCIENCE AT HOME



AARHUS BSS
SCHOOL OF BUSINESS AND SOCIAL SCIENCES
AARHUS UNIVERSITY

How did I end up here?

HARVARD
School of Engineering and Applied Sciences

EconCS Group

Research at the intersection between computer science and economics

Intern Intern Intern Postdoc

Microsoft® Research
Computational Social Science

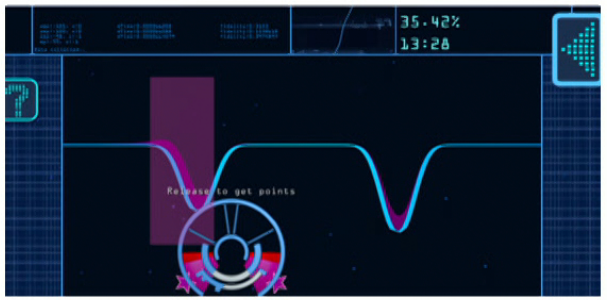
Tenure-track

SCIENCE AT HOME

AARHUS BSS
SCHOOL OF BUSINESS AND SOCIAL SCIENCES
AARHUS UNIVERSITY

CTRL—labs

Start-up



Two provocations…

Important areas of research, under the purview of companies, will not be done

It’s impossible to produce open-source software in academia

Companies have an advantage.

More (behavioral) data

More computational resources

Companies have an advantage.

LETTER

doi:10.1038/nature11421

A 61-million-person experiment in social influence and political mobilization

Robert M. Bond[1], Christopher J. Fariss[1], Jason J. Jones[2], Adam D. I. Kramer[3], Cameron Marlow[3], Jaime E. Settle[1] & James H. Fowler[1,4]

More (behavioral) data

More computational resources

Companies have an advantage.

LETTER

doi:10.1038/nature11421

A 61-million-person experiment in social influence and political mobilization

Robert M. Bond[1], Christopher J. Fariss[1], Jason J. Jones[2], Adam D. I. Kramer[3], Cameron Marlow[3], Jaime E. Set
& James H. Fowler[1,4]

Bigtable: A Distributed Storage System for Structured Data

Fay Chang, Jeffrey Dean, Sanjay Ghemawat, Wilson C. Hsieh, Deborah A. Wallach
Mike Burrows, Tushar Chandra, Andrew Fikes, Robert E. Gruber

{fay,jeff,sanjay,wilsonh,kerr,m3b,tushar,fikes,gruber}@google.com

Google, Inc.

More (behavioral) data

More computational resources

Companies can do unique research.

Data scientists at Facebook conduct large-scale, global, quantitative research to gain deeper insights into how people interact with each other and the world around them.

Our findings directly inform decisions to improve people's everyday experiences on Facebook, make it easier and more intuitive to use, and find ways to facilitate meaningful social interactions. Research efforts span a variety of disciplines, including computational social science, econometrics, operations research, market intelligence, survey science, and statistical computing. We employ a mixture of methods to accomplish our goals, including machine learning, field experiments, surveys, and information visualization. We also build scalable platforms for the collection, management, and analysis of data, and actively contribute our scientific findings to the academic research community

PNAS

Experimental evidence of massive-scale emotional contagion through social networks

Adam D. I. Kramer[a,1], Jamie E. Guillory[b], and Jeffrey T. Hancock[c,d]

[a]Core Data Science Team, Facebook, Inc., Menlo Park, CA 94025; [b]Center for Tobacco Control Research and Education, University of California, San Francisco, CA 94143; and Departments of [c]Communication and [d]Information Science, Cornell University, Ithaca, NY 14853

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved March 25, 2014 (received for review October 23, 2013)

Emotional states can be transferred to others via emotional contagion, leading people to experience the same emotions without their awareness. Emotional contagion is well established

demonstrated that (*i*) emotional contagion occurs via text-based computer-mediated communication (7); (*ii*) contagion of psychological and physiological qualities has been suggested based

Facebook Manipulated User News Feeds To Create Emotional Responses

Gregory S. McNeal, CONTRIBUTOR
I write about technology, law, and policy. FULL BIO

TWEET THIS

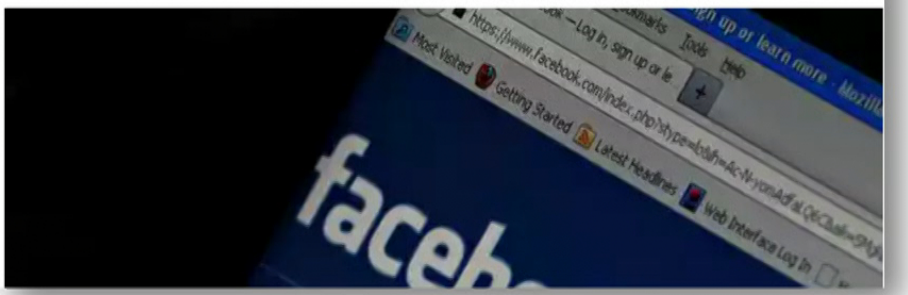
Facebook conducted a massive psychological experiment on 689,003 users, manipulating their news feeds to assess the effects on their emotions.

The short version is, Facebook has the ability to make you feel good or bad, just by tweaking what shows up in your news feed.

Facebook conducted a
experiment on 689,00
their news feeds to as
emotions. The d

Facebook fiasco: was Cornell's study of 'emotional contagion' an ethics breach?

A covert experiment to influence the emotions of more than 600,000 people. A major scientific journal behaving like a rabbit in the headlights. A university in a PR tailspin



Facebook Manipulated User News Feeds To Create Emotional Responses

Gregory S. McNeal, CONTRIBUTOR

I write about technology, law, and policy. FULL BIO

Opinions expressed by Forbes Contributors are their own.

TWEET THIS

Facebook conducted a massive psychological experiment on 689,003 users, manipulating their news feeds to assess the effects on their emotions.

The short version is, Facebook has the ability to make you feel good or bad, just by tweaking what shows up in your news feed.

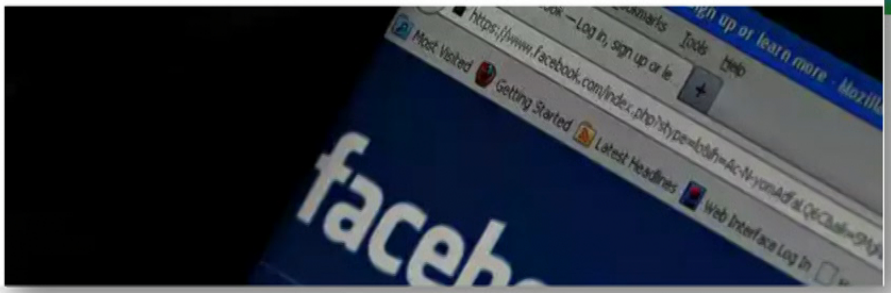
Facebook conducted a
experiment on 689,00
their news feeds to as
emotions. The d

Everything We Know About Facebook's Secret Mood Manipulation Experiment

It was probably legal. But was it ethical?

Facebook fiasco: was Cornell's study of 'emotional contagion' an ethics breach?

A covert experiment to influence the emotions of more than 600,000 people. A major scientific journal behaving like a rabbit in the headlights. A university in a PR tailspin




- certification of open source/ collaborative practices
- placement mechanisms for free riders

How Trump Consultants Exploited the Facebook Data of Millions

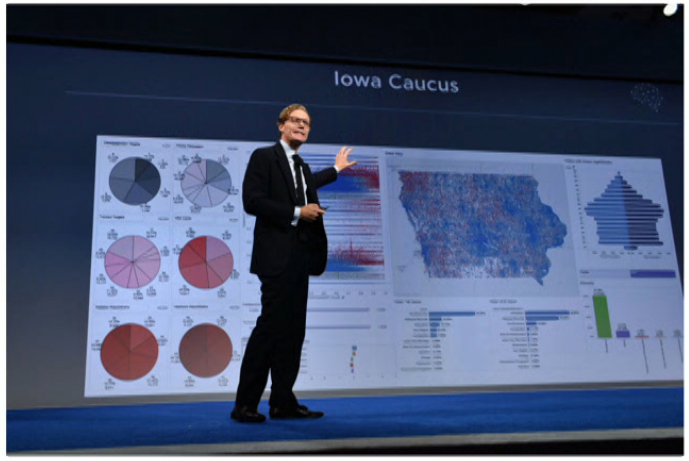
Leer en español

By MATTHEW ROSENBERG, NICHOLAS CONFESSORE and CAROLE CADWALLADR MARCH 17, 2018

Mark Zuckerberg Thinks We're Idiots.

by Jean-Louis Gassée

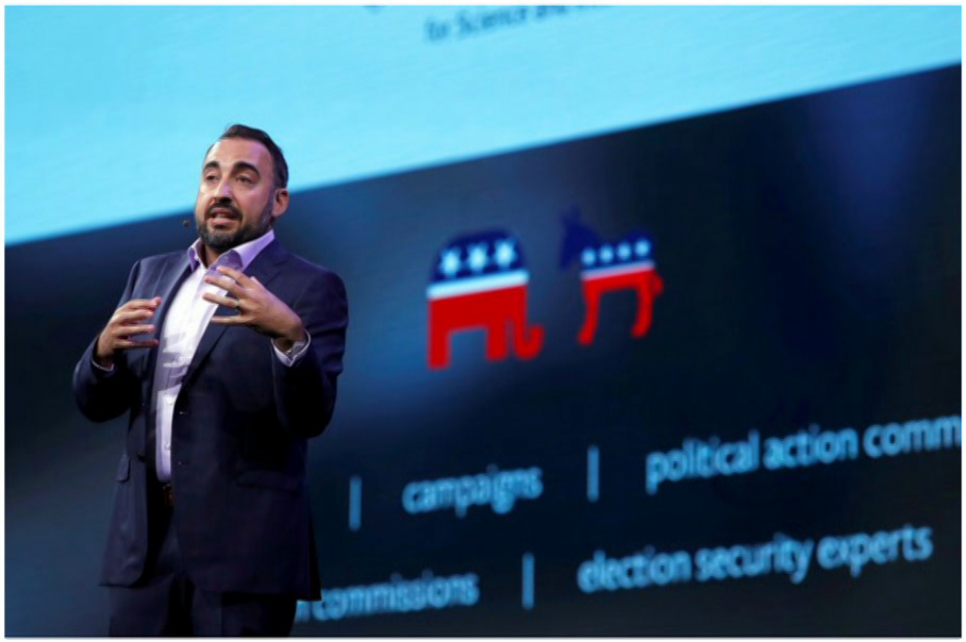
Surprise: Thanks to the Cambridge Analytica revelations, we're finding out that Facebook allowed a much broader and deeper prostitution of our private data than it had previously claimed. Facebook's disingenuous explanations call for more questions and even less trust.



ICE USES FACEBOOK DATA TO FIND AND TRACK IMMIGRANTS, INTERNAL EMAILS SHOW

Facebook Exit Hints at Dissent on Handling of Russian Trolls

By NICOLE PERLROTH, SHEERA FRENKEL and SCOTT SHANE MARCH 19, 2018



Alex Stamos, the chief information security officer for Facebook. He has urged more disclosure over Russian activity on Facebook. Steve Marcus/Reuters

RELATED COV

Mr. Stamos, who plans to leave Facebook by August, had advocated more disclosure around Russian interference of the platform and some restructuring to better address the issues, but was met with resistance by colleagues, said the current and former employees. In December, Mr. Stamos's day-to-day responsibilities were reassigned to others, they said.

Facebook is the largest socio-technical system of in the world.

But transparent research about its design and impact may never see the light of day.

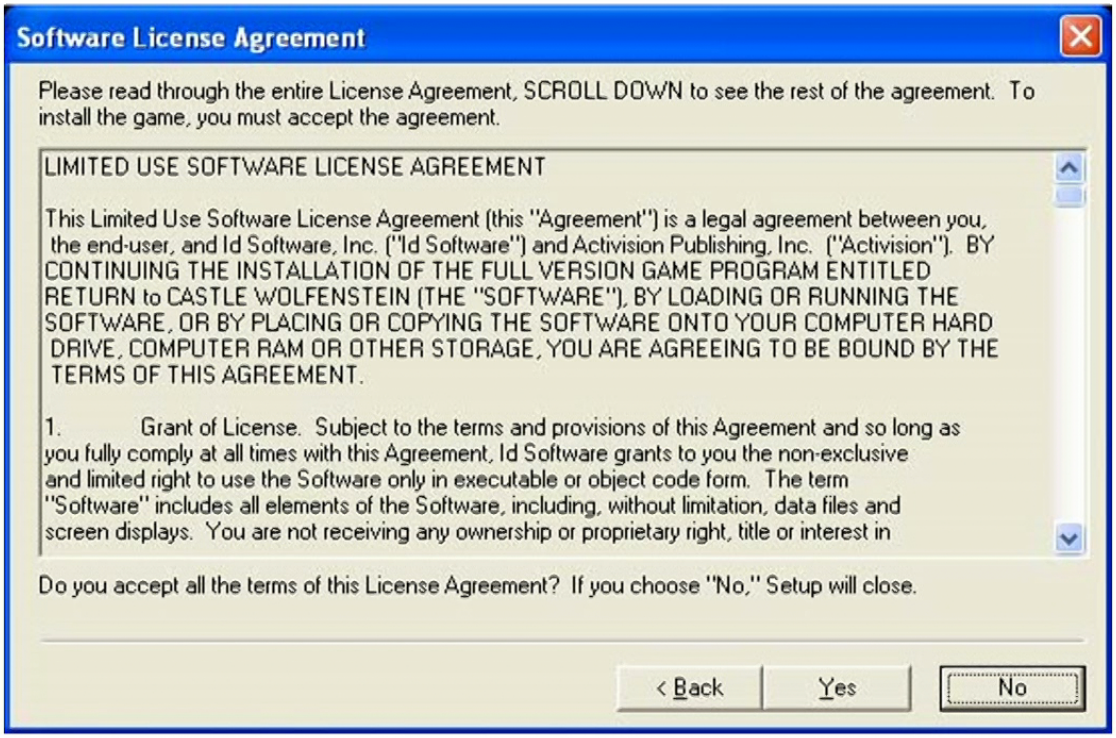
Another example: the sharing economy

Another example: the sharing economy


GiveWell

Open-source software in industry and academy

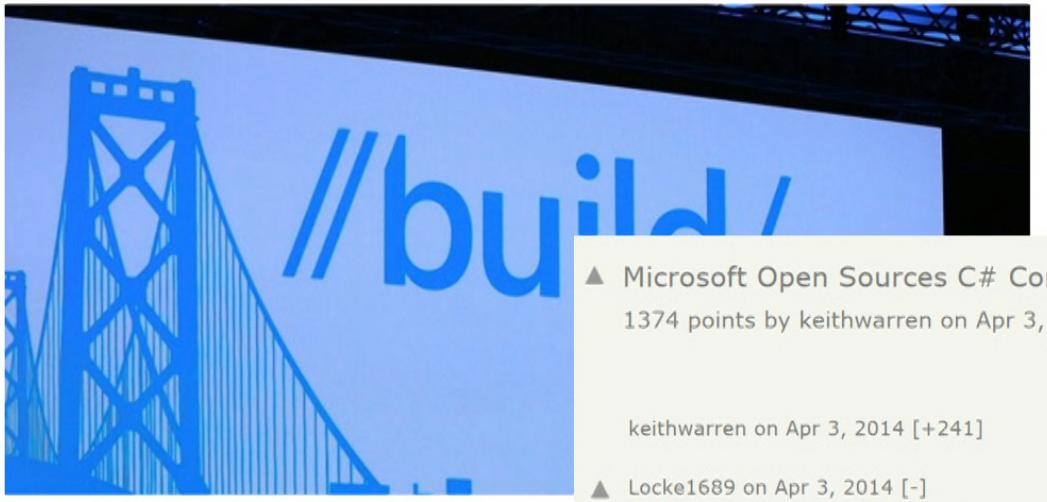
A long time ago…


Software License Agreement
Please read through the entire License Agreement, SCROLL DOWN to see the rest of the agreement. To install the game, you must accept the agreement.
LIMITED USE SOFTWARE LICENSE AGREEMENT
This Limited Use Software License Agreement (this "Agreement") is a legal agreement between you, the end-user, and Id Software, Inc. ("Id Software") and Activision Publishing, Inc. ("Activision"). BY CONTINUING THE INSTALLATION OF THE FULL VERSION GAME PROGRAM ENTITLED RETURN to CASTLE WOLFENSTEIN (THE "SOFTWARE"), BY LOADING OR RUNNING THE SOFTWARE, OR BY PLACING OR COPYING THE SOFTWARE ONTO YOUR COMPUTER HARD DRIVE, COMPUTER RAM OR OTHER STORAGE, YOU ARE AGREEING TO BE BOUND BY THE TERMS OF THIS AGREEMENT.
1. Grant of License. Subject to the terms and provisions of this Agreement and so long as you fully comply at all times with this Agreement, Id Software grants to you the non-exclusive and limited right to use the Software only in executable or object code form. The term "Software" includes all elements of the Software, including, without limitation, data files and screen displays. You are not receiving any ownership or proprietary right, title or interest in
Do you accept all the terms of this License Agreement? If you choose "No," Setup will close.
< Back
Yes
No


Ballmer: “Linux is a cancer that attaches itself in an intellectual property sense to everything it touches.”

Microsoft open sources .NET compiler platform "Roslyn" and announces open source .NET Foundation initiative

by EMIL PROTALINSKI — Apr 3, 2014 in MICROSOFT



7 2,507 SHARES

▲ Microsoft Open Sources C# Compiler (codeplex.com)

1374 points by keithwarren on Apr 3, 2014 | hide | past | web | favorite | 446 comments

keithwarren on Apr 3, 2014 [+241]

▲ Locke1689 on Apr 3, 2014 [-]

Everyone on Roslyn is really excited about this and we hope that it serves as a signal that big things are happening in .NET to make the entire platform more open and agile!

P.S. We're the Visual Basic compiler too :)

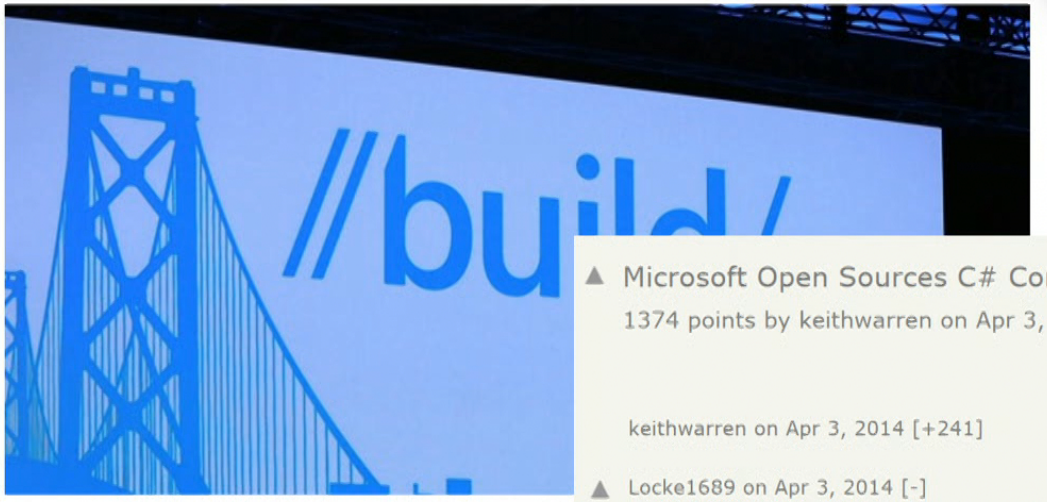
▲ stusmall on Apr 3, 2014 [-]

From an open source neckbeard to a MS employee, thanks!!!

Man, this is a crazy world. Cats and dogs living together. Open source MS projects. The falcon cannot hear the falconer

Microsoft open sources .NET compiler platform "Roslyn" and announces open source .NET Foundation initiative

by EMIL PROTALINSKI — Apr 3, 2014 in MICROSOFT



7 2,507 SHARES

Open source has won, and Microsoft has surrendered

Many Linux users are ticked off and anxious about Microsoft joining the Linux Foundation. They are missing the real significance of that move.

▲ Microsoft Open Sources C# Compiler (codeplex.com)

1374 points by keithwarren on Apr 3, 2014 | hide | past | web | favorite | 446 comments

keithwarren on Apr 3, 2014 [+241]

▲ Locke1689 on Apr 3, 2014 [-]

Everyone on Roslyn is really excited about this and we hope that it serves as a signal that big things are happening in .NET to make the entire platform more open and agile!

P.S. We're the Visual Basic compiler too :)

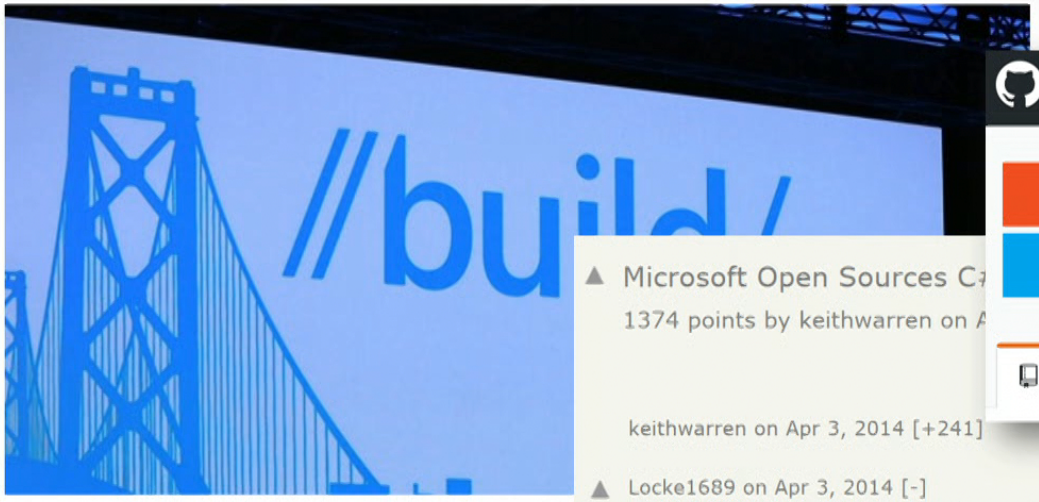
▲ stusmall on Apr 3, 2014 [-]

From an open source neckbeard to a MS employee, thanks!!!

Man, this is a crazy world. Cats and dogs living together. Open source MS projects. The falcon cannot hear the falconer

Microsoft open sources .NET compiler platform "Roslyn" and announces open source .NET Foundation initiative

by EMIL PROTALINSKI — Apr 3, 2014 in MICROSOFT



7 2,507 SHARES

Open source has won, and Microsoft has surrendered

Many Linux users are ticked off and anxious about Microsoft joining the Linux Foundation. They are missing the real significance of that move.

This organization Search Pull requests Issues Marketplace Explore

Microsoft

Open source, from Microsoft with love

Redmond, WA https://opensource.micro... opensource@microsoft.co...

Repositories 1,693 People 3,756

▲ Microsoft Open Sources C#

1374 points by keithwarren on A

keithwarren on Apr 3, 2014 [+241]

▲ Locke1689 on Apr 3, 2014 [-]

Everyone on Roslyn is really excited about this and we hope that it serves as a signal that big things are happening in .NET to make the entire platform more open and agile!

P.S. We're the Visual Basic compiler too :)

▲ stusmall on Apr 3, 2014 [-]

From an open source neckbeard to a MS employee, thanks!!!

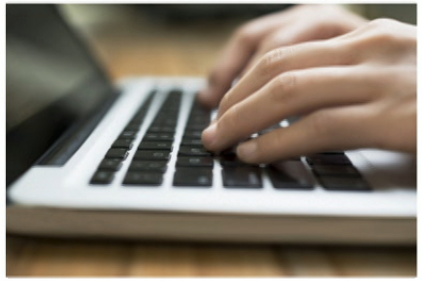
Man, this is a crazy world. Cats and dogs living together. Open source MS projects. The falcon cannot hear the falconer

Git After It: Why GitHub Won When So Many Other Big Companies Failed

MAR 13, 2015 / BY HANNAH FLEISHMAN

Tweet Share Like 0 Share

Yesterday, Google announced they would officially be pulling the plug on open source project hosting service, Google Code. What launched ten years ago as an alternative to SourceForge never really grew up to be *the* social coding platform Google envisioned. In other words, it wasn't GitHub.



"**Facebook sees open source as being good for business**: “It means we build better software, write better code, our engineers are able to work with more pride, and we’re able to retain the world’s best engineers because they know they can open-source their work.” Ultimately, because engineers can see for themselves the kinds of things Facebook is working on, **it makes it easier to attract the top talent**. “It’s not all altruism, there’s solid business sense behind this.”-

- James Pearce, Facebook

Meanwhile, in academia…

Academics aren’t software engineers

Why do many talented scientists write horrible software?

126

I am a software engineer and I have been working for years with people hired from academic background. Many times I've noticed that (even otherwise brilliant scientists) produce code of extremely low quality (unless their background was precisely Computer Science).

Since those people are very good in doing their research - and eventually obtain remarkable results - it seems they are clever enough to write *decent* code. Is it just that they don't think it's worth the effort? Plain arrogance? Lack of time?

Truly reusable software is hard.

The Rule of Three:

1. It is **three** times as difficult to build reusable components as single use components, and
2. a reusable component should be tried out in **three** different applications before it will be sufficiently general to accept into a reuse library.

Facts and Fallacies of Software Engineering

3. It's **three** times harder to document a software product well than just writing the code.

Software tools != citations.

Software tools != citations.

All hail ggplot2—The code powering all those excellent charts is 10 years old

By Dan Kopf | June 18, 2017

Happy birthday ggplot2. (Reuters/Suzanne Plunkett)

On June 10, 2007, the statistician and developer Hadley Wickham officially released ggplot2, a chart-making system for the statistical programming language R. It would alter the course of his life and the future of data visualization.

Software tools != citations.

On **Hadley Wickham:**

his motivation … was primarily to provide better ways of accomplishing routine tasks in R, an immensely useful contribution that sadly wasn't recognized in an academic setting.

All hail ggplot2—The code powering all those excellent charts is 10 years old

By Dan Kopf | June 18, 2017

Happy birthday ggplot2. (Reuters/Suzanne Plunkett)

On June 10, 2007, the statistician and developer Hadley Wickham officially released ggplot2, a chart-making system for the statistical programming language R. It would alter the course of his life and the future of data visualization.

Software tools != citations.

On **Hadley Wickham:**

his motivation … was primarily to provide better ways of accomplishing routine tasks in R, an immensely useful contribution that sadly wasn't recognized in an academic setting.

"There are definitely some academic statisticians who just don't understand why what I do is statistics, but basically I think they are all wrong . What I do is fundamentally statistics. **The fact that data science exists as a field is a colossal failure of statistics.** To me, that is what statistics is all about. It is gaining insight from data using modelling and visualization. Data munging and manipulation is hard and statistics has just said that's not our domain."

All hail ggplot2—The code powering all those excellent charts is 10 years old

By Dan Kopf | June 18, 2017

Happy birthday ggplot2. (Reuters/Suzanne Plunkett)

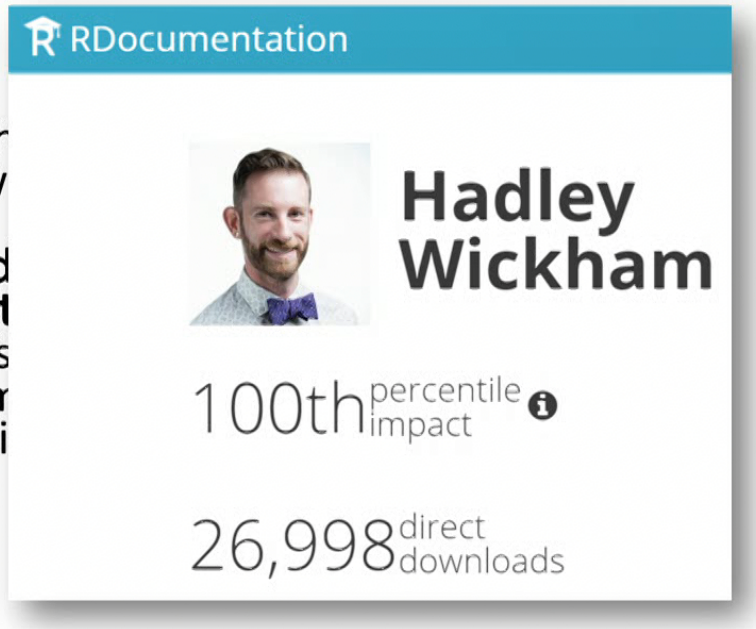
On June 10, 2007, the statistician and developer Hadley Wickham officially released ggplot2, a chart-making system for the statistical programming language R. It would alter the course of his life and the future of data visualization.

Software tools != citations.

On **Hadley Wickham:**

his motivation … was primarily to provide better ways of accomplishing routine tasks in R, an immensely useful contribution that sadly wasn't recognized in an academic setting.

"There are definitely sor
don't understand why w
think they are all wrong
statistics. **The fact that d**
colossal failure of statist
all about. It is gaining ins
and visualization. Data m
and statistics has just sai



All hail ggplot2—The code powering all those excellent charts is 10 years old

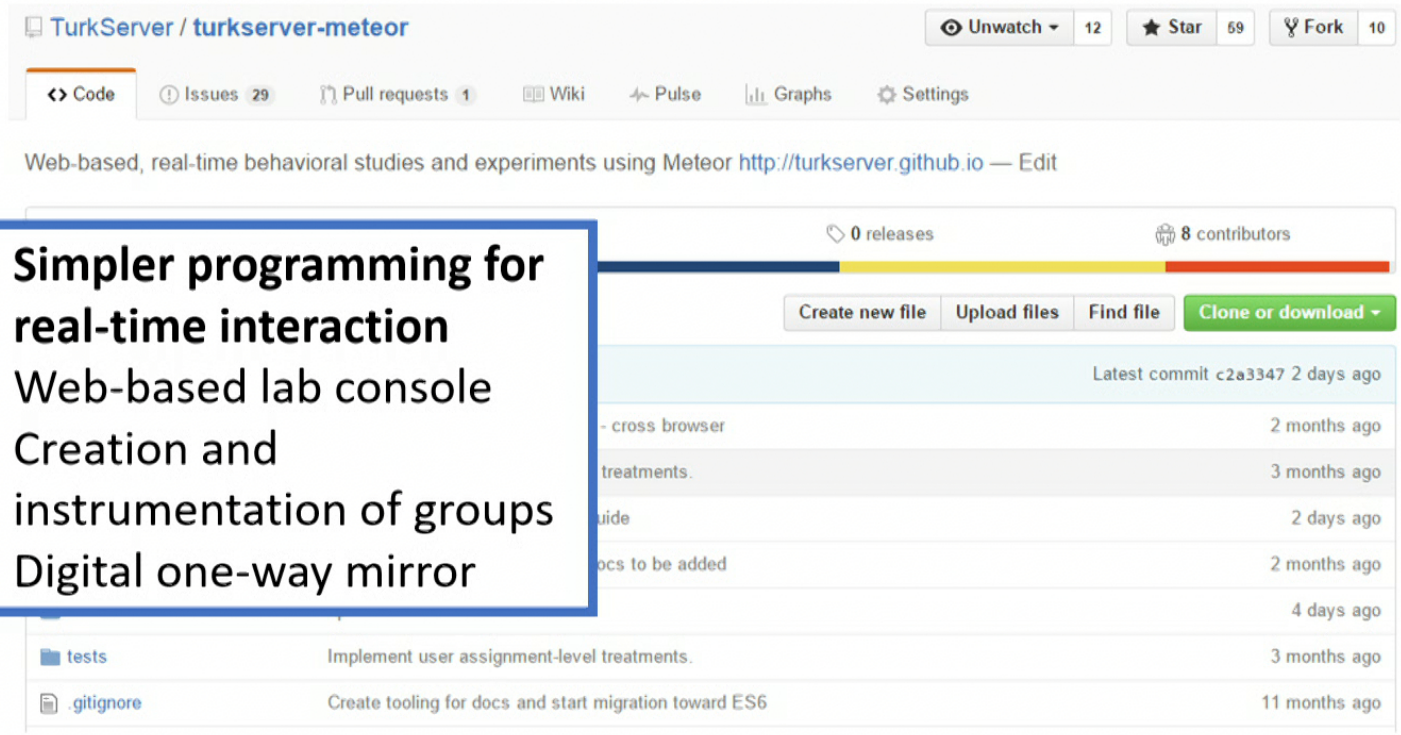
By Dan Kopf | June 18, 2017

Happy birthday ggplot2. (Reuters/Suzanne Plunkett)

On June 10, 2007, the statistician and developer Hadley Wickham officially released ggplot2, a chart-making system for the statistical programming language R. It would alter the course of his life and the future of data visualization.

An example from experience…

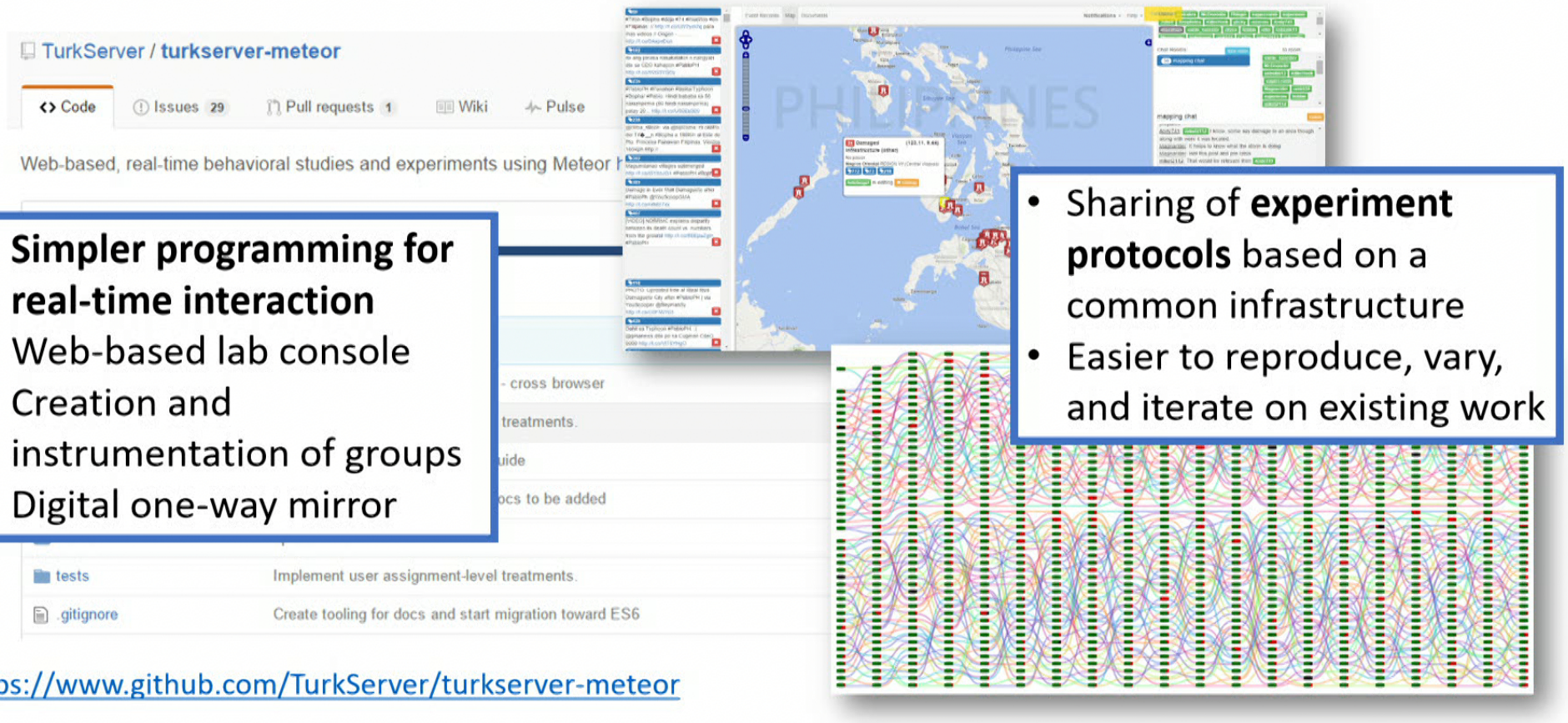
TurkServer: OSS platform + experiments



- **Simpler programming for real-time interaction**
- Web-based lab console
- Creation and instrumentation of groups
- Digital one-way mirror

https://www.github.com/TurkServer/turkserver-meteor

TurkServer: OSS platform + experiments

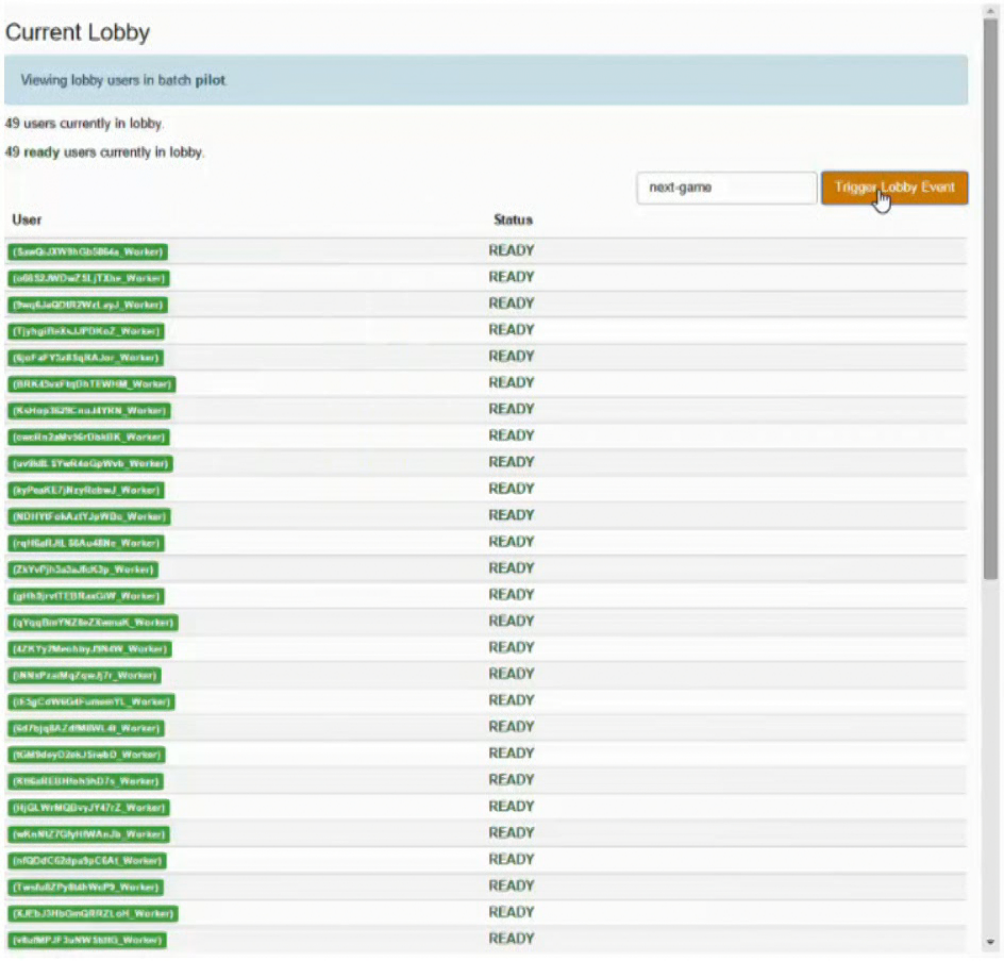


- **Simpler programming for real-time interaction**
- Web-based lab console
- Creation and instrumentation of groups
- Digital one-way mirror

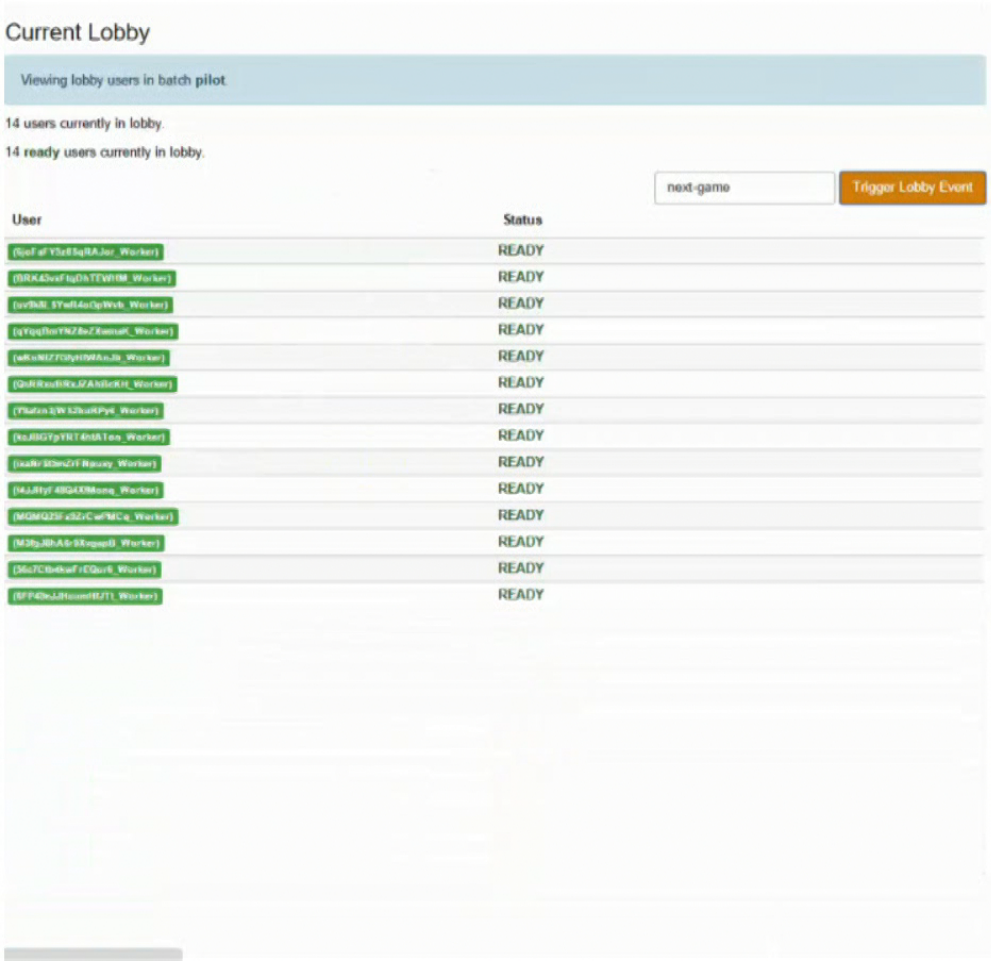
- Sharing of **experiment protocols** based on a common infrastructure
- Easier to reproduce, vary, and iterate on existing work

https://www.github.com/TurkServer/turkserver-meteor

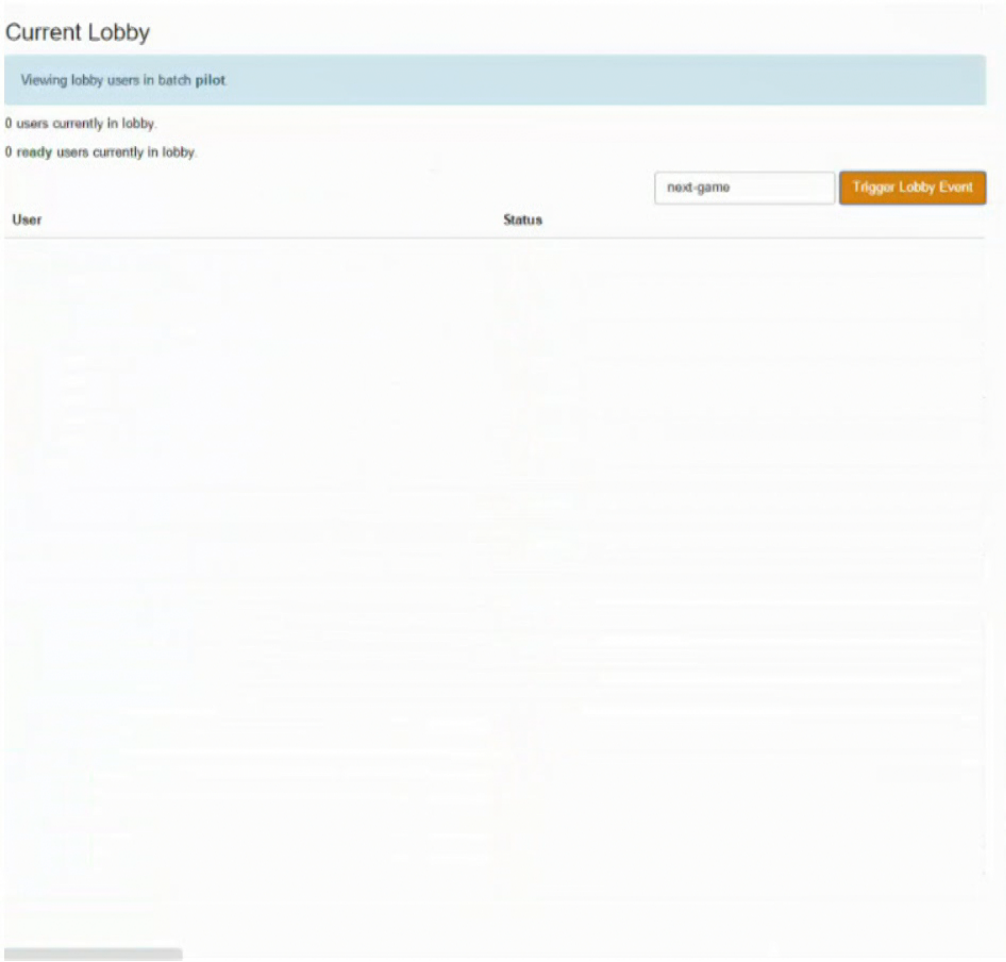
A virtual lab: web-based real-time view



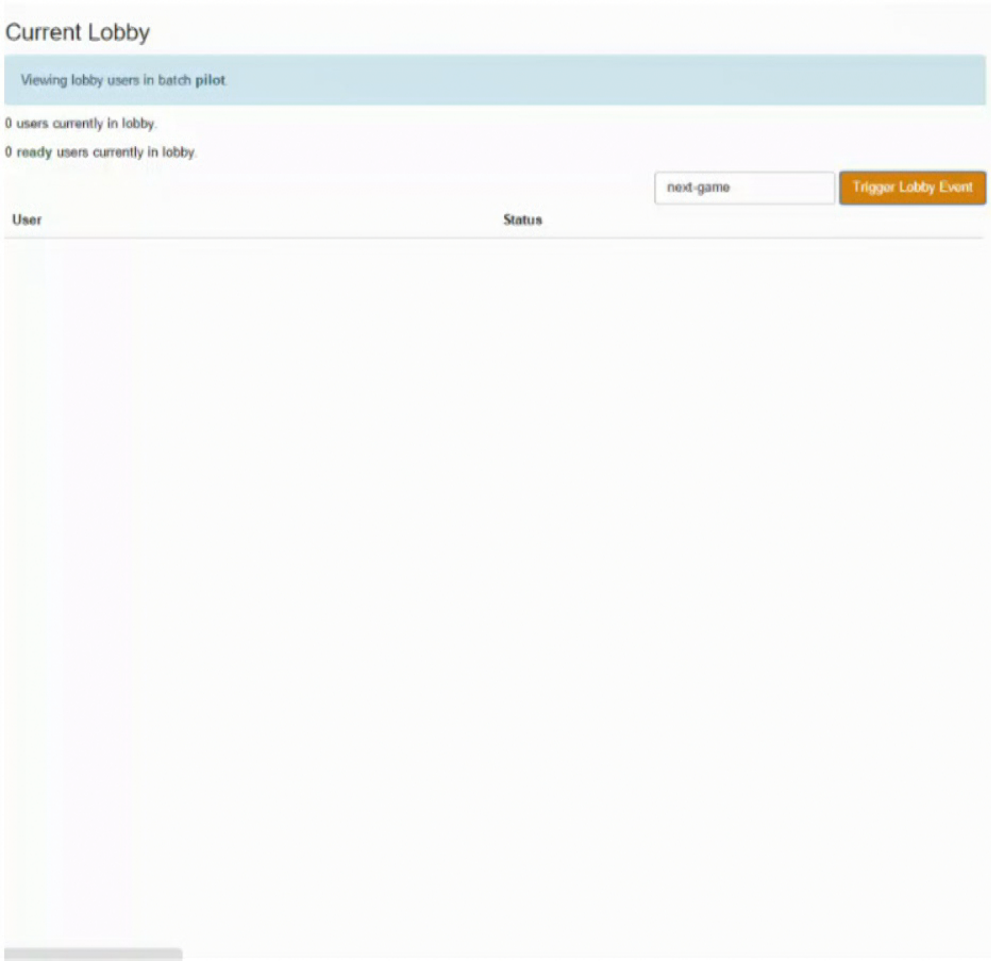
A virtual lab: web-based real-time view



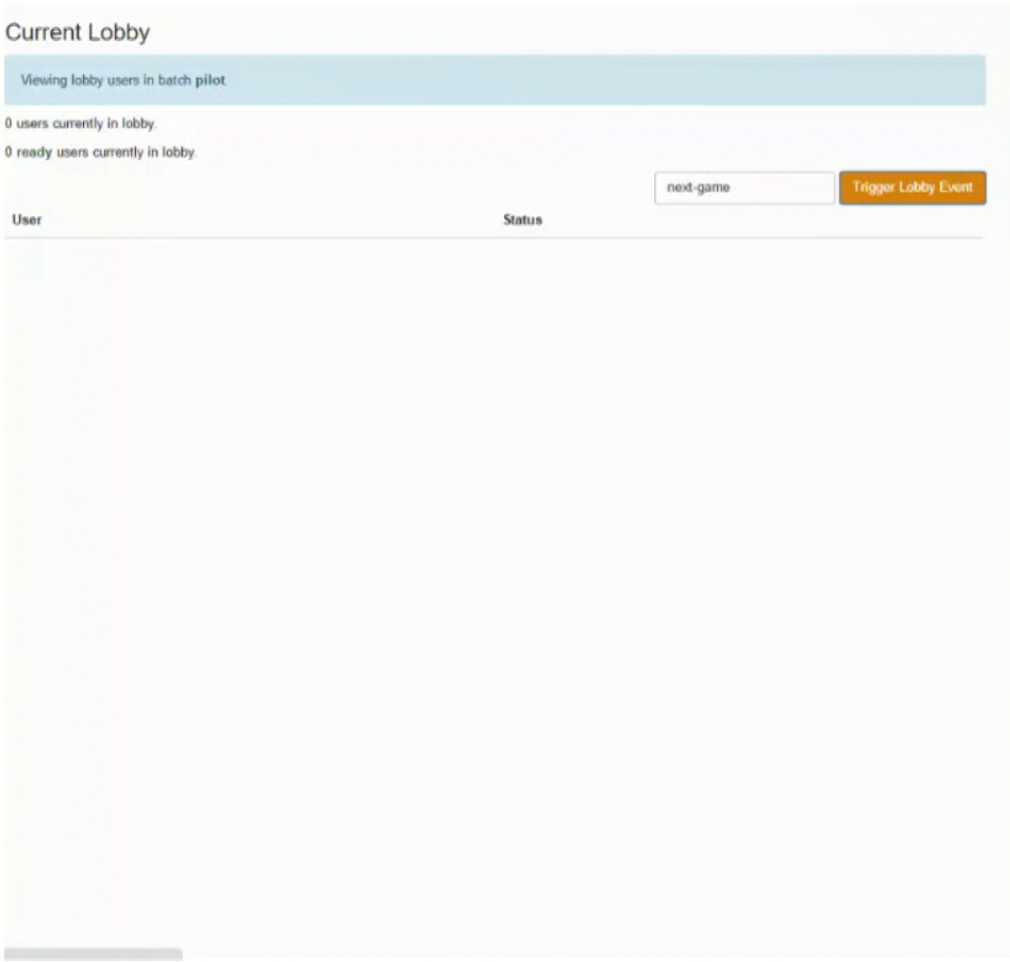
A virtual lab: web-based real-time view



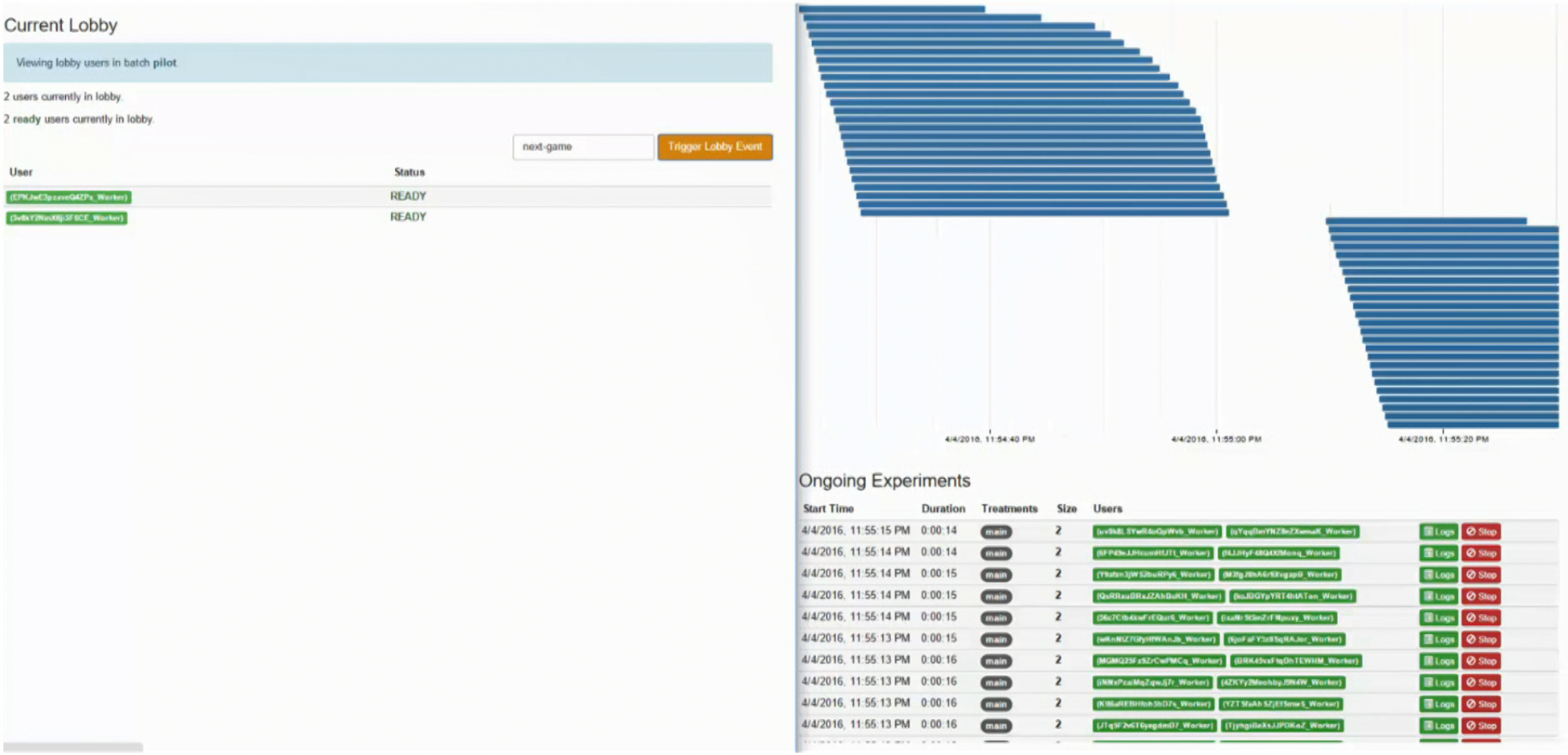
A virtual lab: web-based real-time view



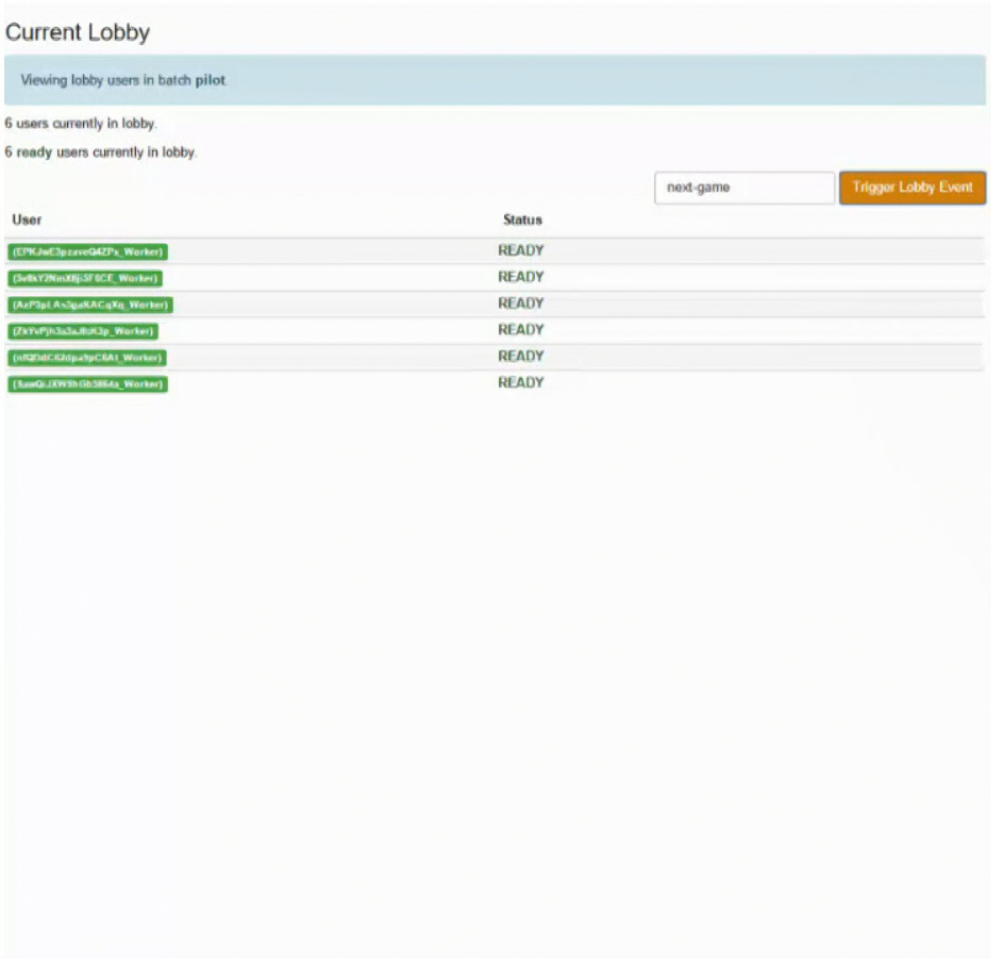
A virtual lab: web-based real-time view



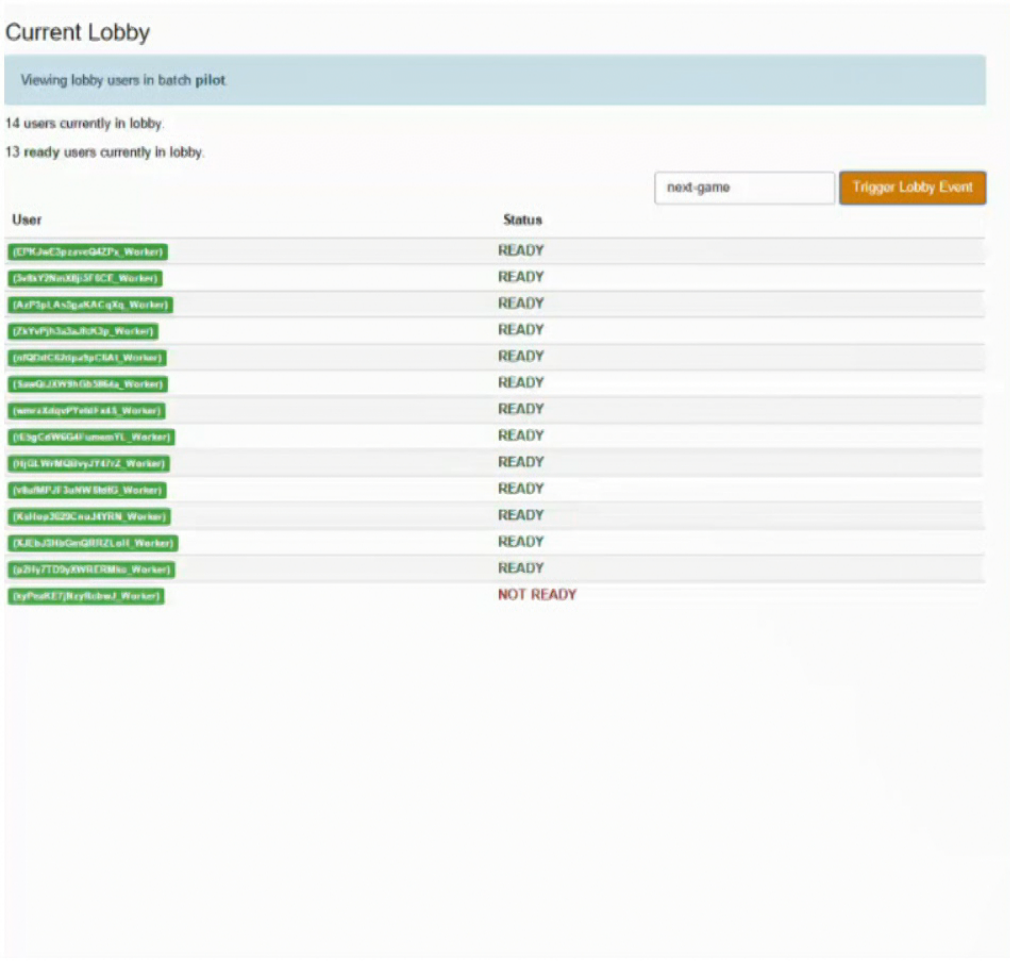
A virtual lab: web-based real-time view



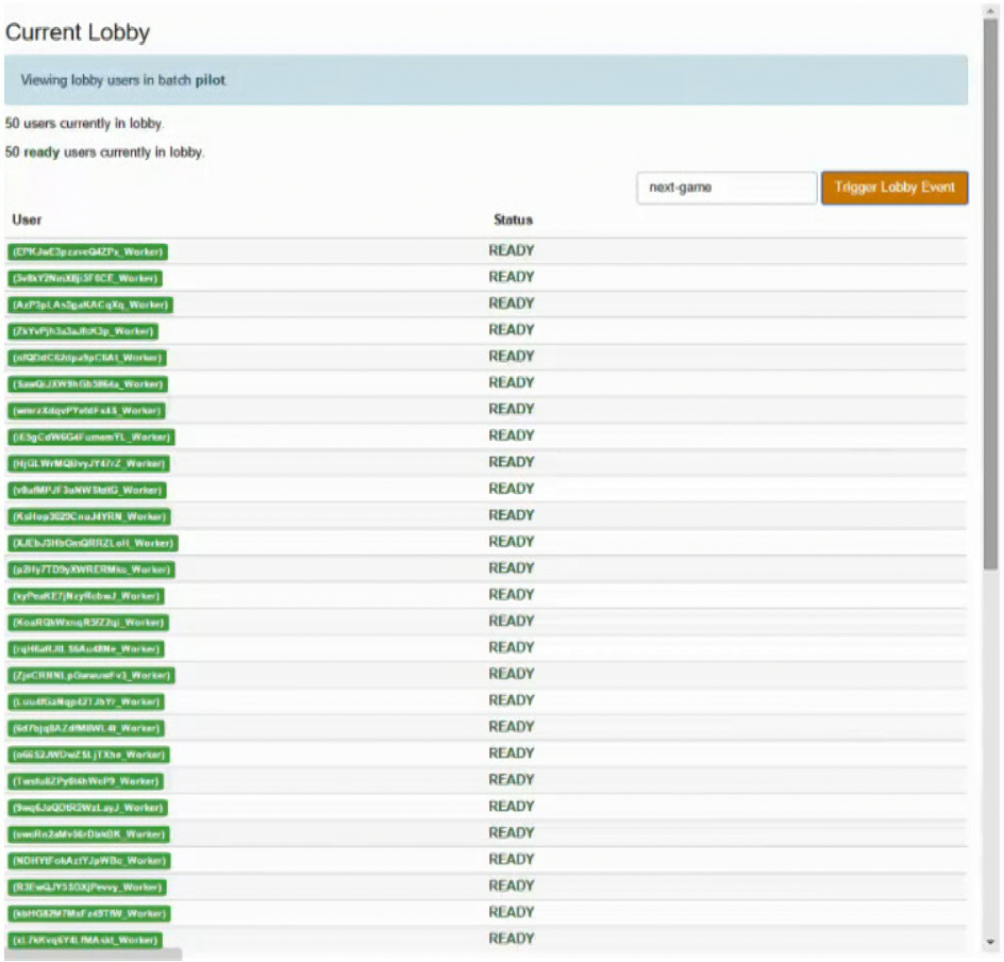
A virtual lab: web-based real-time view

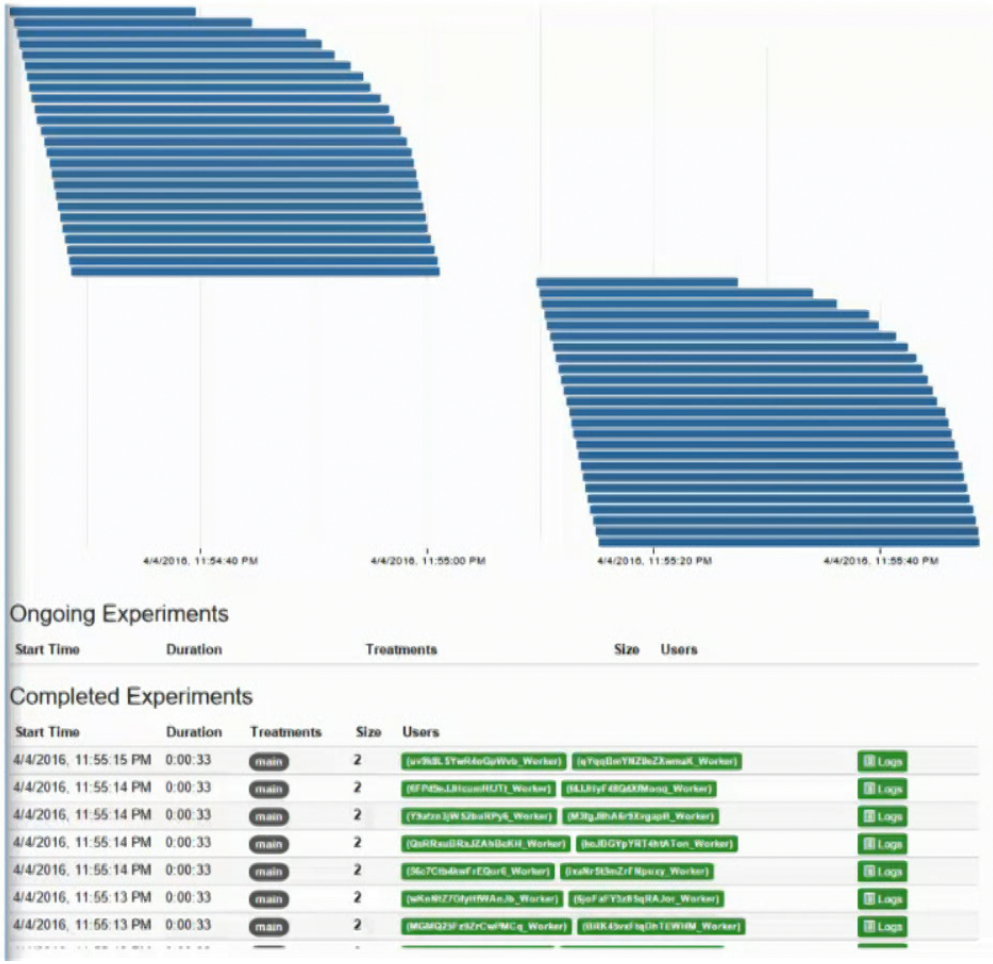


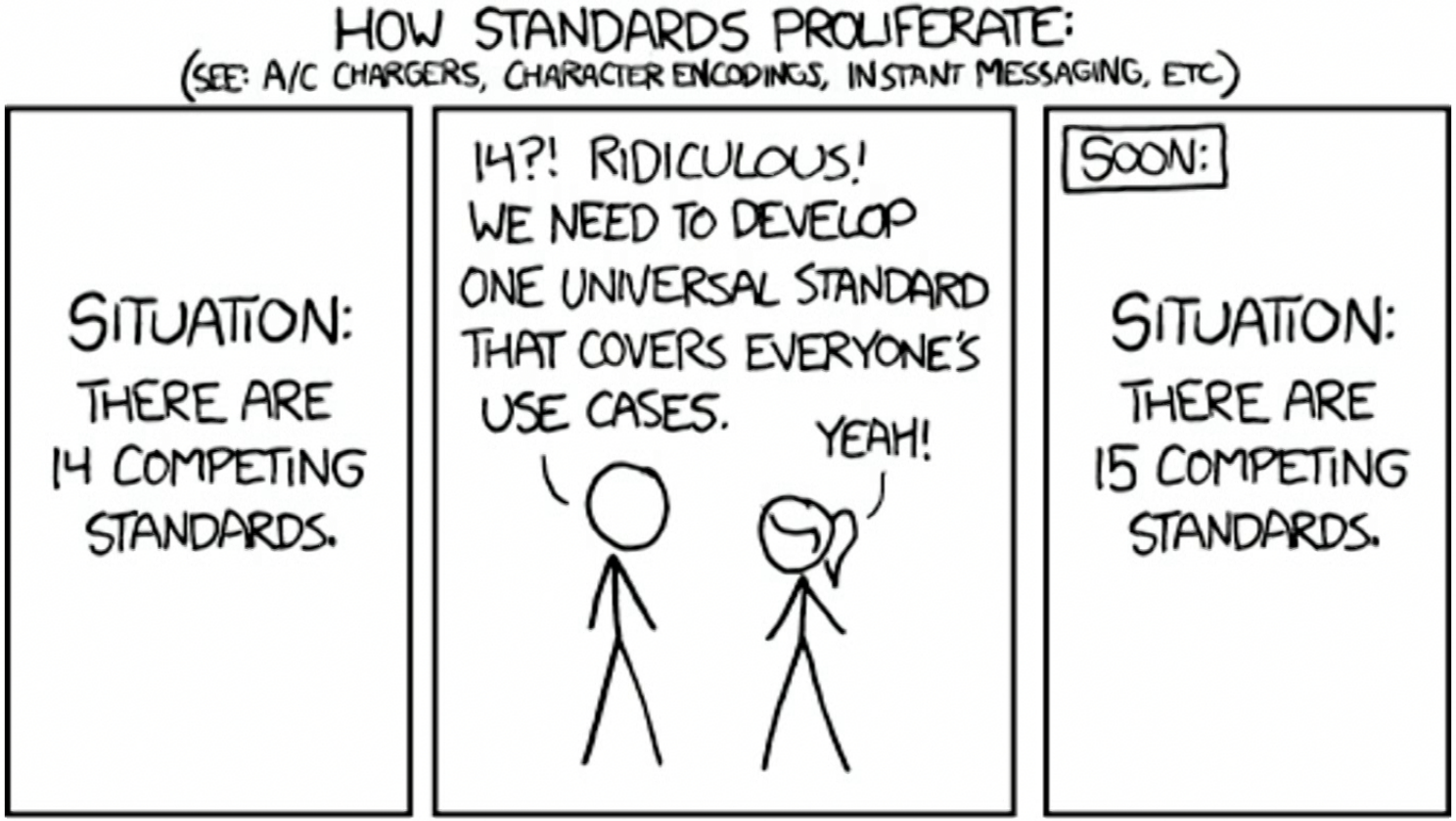
A virtual lab: web-based real-time view

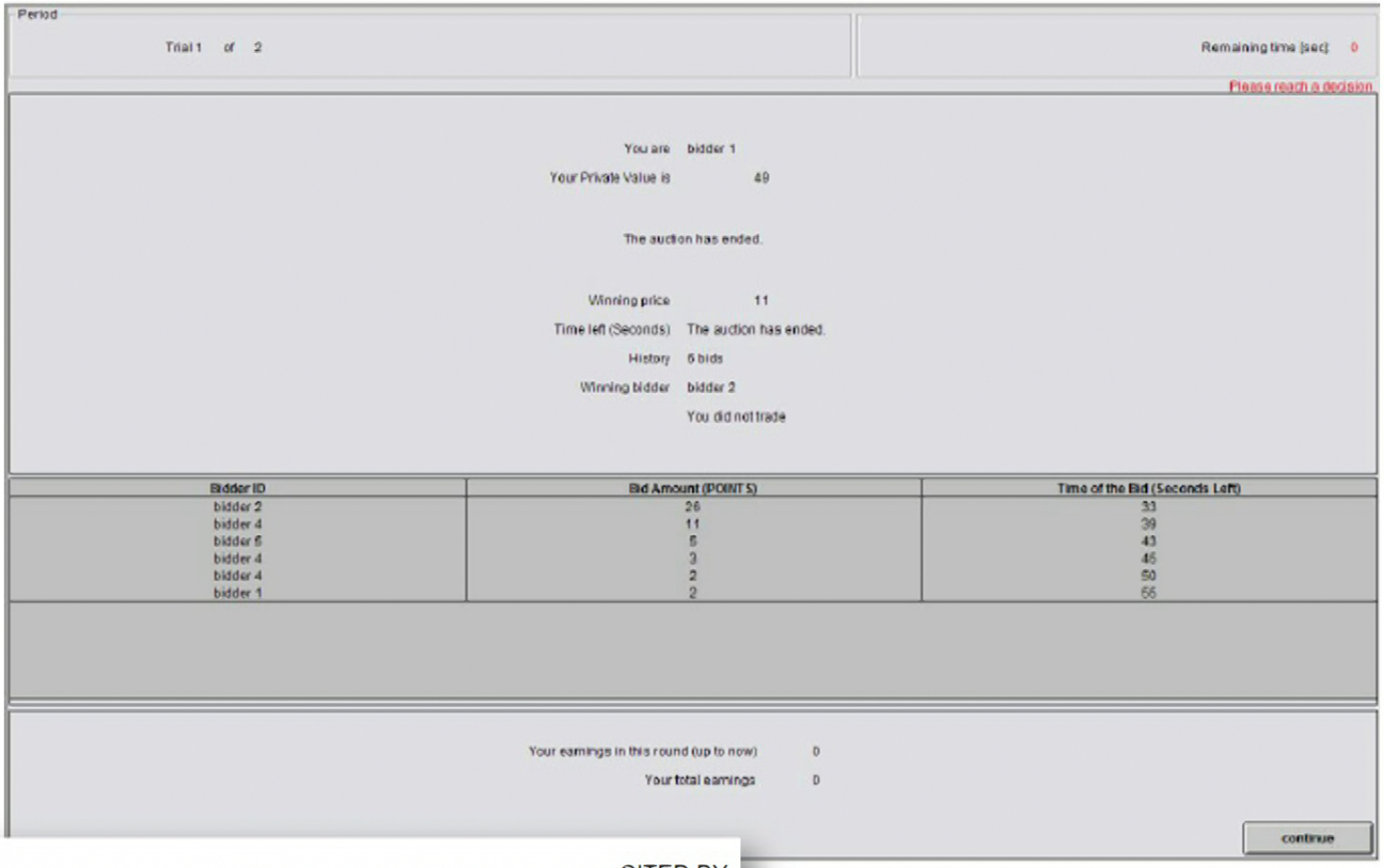


A virtual lab: web-based real-time view






HOW STANDARDS PROLIFERATE:
(SEE: A/C CHARGERS, CHARACTER ENCODINGS, INSTANT MESSAGING, ETC)
SITUATION:
THERE ARE
14 COMPETING
STANDARDS.
14?! RIDICULOUS!
WE NEED TO DEVELOP
ONE UNIVERSAL STANDARD
THAT COVERS EVERYONE'S
USE CASES.
YEAH!
SOON:
SITUATION:
THERE ARE
15 COMPETING
STANDARDS.


Period
Trial 1 of 2
Remaining time [sec]: 0
Please reach a decision
You are bidder 1
Your Private Value is 49
The auction has ended.
Winning price 11
Time left (Seconds) The auction has ended.
History 6 bids
Winning bidder bidder 2
You did not trade
Bidder ID
Bid Amount (POINTS)
Time of the Bid (Seconds Left)
bidder 2 26 33
bidder 4 11 39
bidder 5 5 43
bidder 4 3 45
bidder 4 2 50
bidder 1 2 55
Your earnings in this round (up to now) 0
Your total earnings 0
continue


TITLE
CITED BY
z-Tree: Zurich toolbox for ready-made economic experiments
8068
U Fischbacher
Experimental economics 10 (2), 171-178, 2007

So…

- Companies probably won't do essential research
- Academics probably won't build open-source tools

So…

- Companies probably won't do essential research
- Academics probably won't build open-source tools

Machine Learning: A Case Study

For better or for worse.

Standard Datasets

ImageNet Challenge

IMAGENET

- 1,000 object classes (categories).
- Images:
  - 1.2 M train
  - 100k test.

2

Deep Image: Scaling up Image Recognition

Ren Wu, Shengen Yan, Yi Shan, Qingqing Dang, Gang Sun

(Submitted on 13 Jan 2015 (v1), last revised 6 Jul 2015 (this version, v5))

IMAGENET **Large Scale Visual Recognition Challenge 2015 (ILSVRC2015)**

Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification

Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun

(Submitted on 6 Feb 2015)

Deep Image: Scaling up Image Recognition

Ren Wu, Shengen Yan, Yi Shan, Qingqing Dang, Gang Sun

(Submitted on 13 Jan 2015 (v1), last revised 6 Jul 2015 (this version, v5))

IMAGENET **Large Scale Visual Recognition Challenge 2015 (ILSVRC2015)**

Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift

Sergey Ioffe, Christian Szegedy

(Submitted on 11 Feb 2015 (v1), last revised 2 Mar 2015 (this version, v3))

Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification

Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun

(Submitted on 6 Feb 2015)

Deep Image: Scaling up Image Recognition

Ren Wu, Shengen Yan, Yi Shan, Qingqing Dang, Gang Sun

(Submitted on 13 Jan 2015 (v1), last revised 6 Jul 2015 (this version, v5))

IMAGENET **Large Scale Visual Recognition Challenge 2015 (ILSVRC2015)**

Batch Normalization: Accelerating D
Training by Reducing Internal Covari

Sergey Ioffe, Christian Szegedy

(Submitted on 11 Feb 2015 (v1), last revised 2 Mar 2015 (this version, v3

Delving Deep into Rectifiers: Surpassin
Level Performance on ImageNet Classi

Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun

(Submitted on 6 Feb 2015)

Deep Image: Scaling up Image Recogni

Ren Wu, Shengen Yan, Yi Shan, Qingqing Dang, Gang Sun

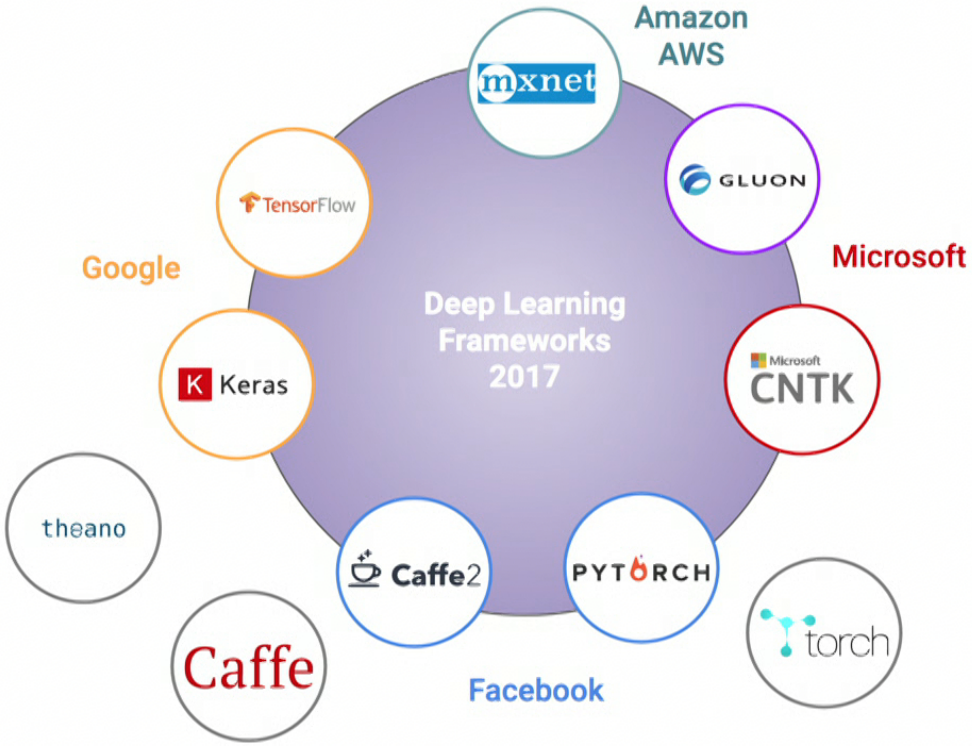
(Submitted on 13 Jan 2015 (v1), last revised 6 Jul 2015 (this version, v5))

Baidu banned from ILSVRC 2015

JUNE 4, 2015 BY SATYA MALLICK — 0 COMMENTS

IMAGENET **Large Scale Visual Recognition Challenge 2015 (ILSVRC2015)**

Open Source Everything



Industry/Academia revolving door

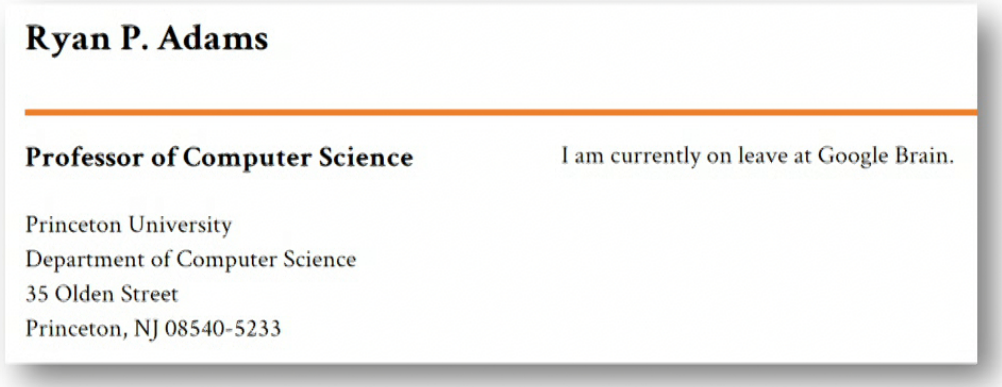
Twitter acquires machine learning startup Whetlab, will kill the service on July 15

The startup, a team of five, was only around a year old at the time of the acquisition.

Industry/Academia revolving door

Twitter acquires machine learning startup Whetlab, will kill the service on July 15

The startup, a team of five, was only around a year old at the time of the acquisition.



Ryan P. Adams

Professor of Computer Science

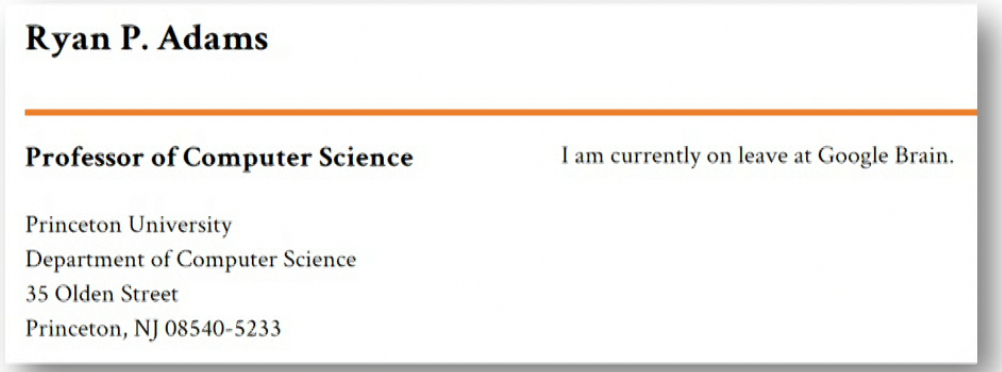
I am currently on leave at Google Brain.

Princeton University
Department of Computer Science
35 Olden Street
Princeton, NJ 08540-5233

Industry/Academia revolving door

Twitter acquires machine learning startup Whetlab, will kill the service on July 15

The startup, a team of five, was only around a year old at the time of the acquisition.



Ryan P. Adams

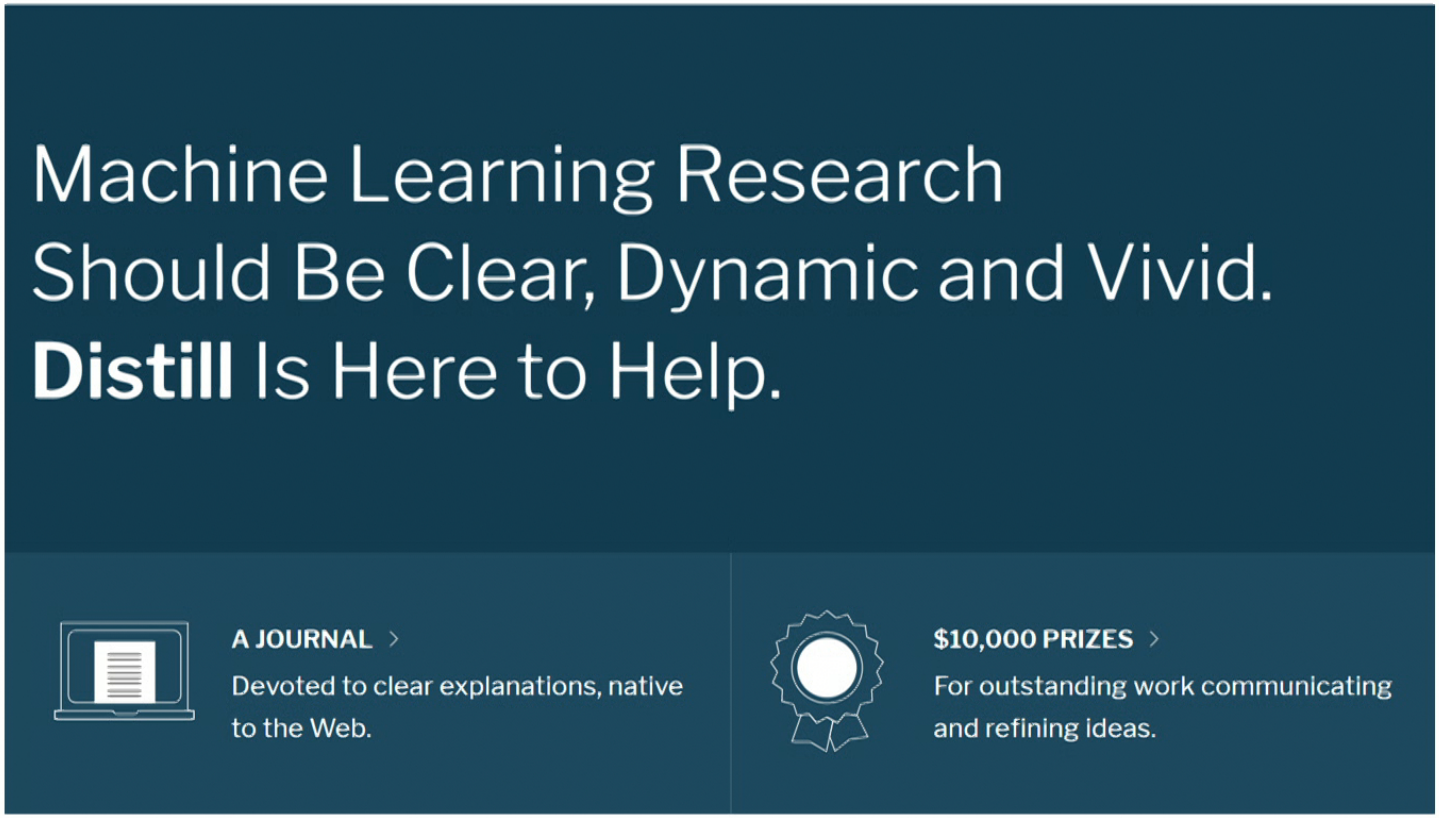
Professor of Computer Science

I am currently on leave at Google Brain.

Princeton University
Department of Computer Science
35 Olden Street
Princeton, NJ 08540-5233

Artificial Intelligence Teams Being Acquired For $2.5m/employee;

New Publication Methods



https://distill.pub/

Discussion

How do we produce open research on closed systems, especially in essential topics like computational social science?

How can academics build tools of lasting value? (Do they even need to?)