

> Dígame sus datos  
> Tell me your data

>> Hacks/Hackers Media Party Buenos Aires  
>> 26 August 2016

>> Christine Zhang  
>> @christinezhang

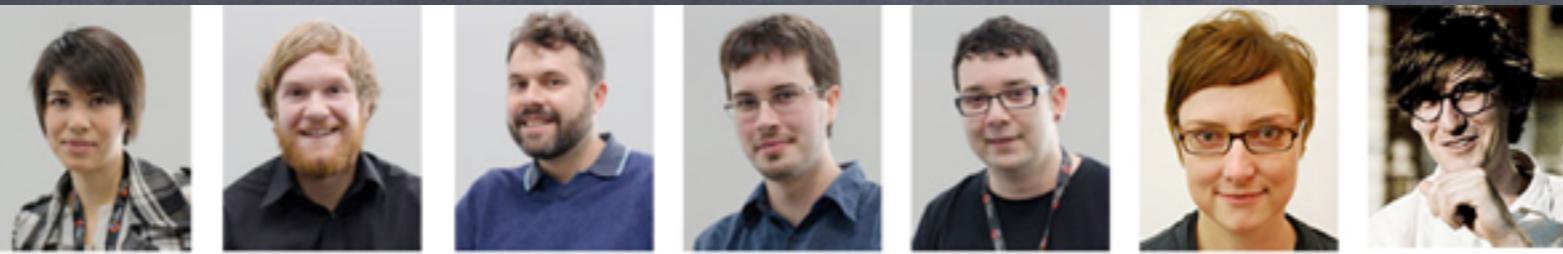
> Hello.



@Pizzeria Güerrin  
23 August 2016

Photo: Sandhya Kambhampati

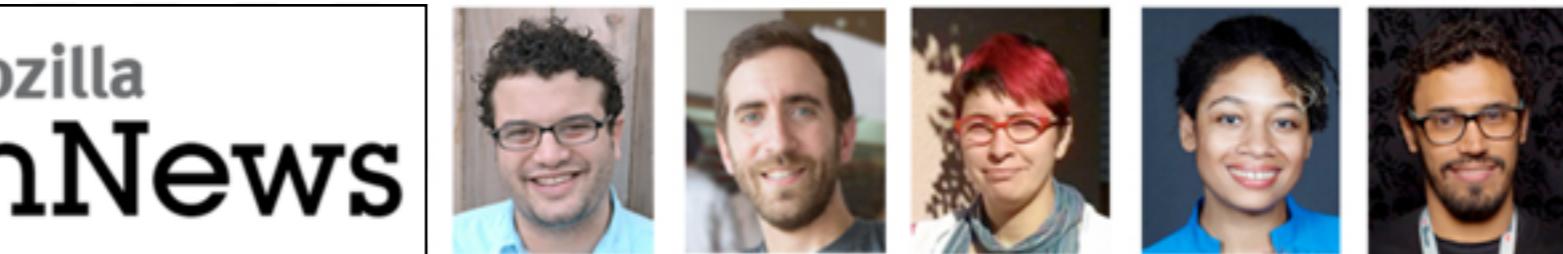
> @christinezhang



2012



2013



2014



2015



2016

>>

Knight-Mozilla OpenNews Fellows  
<https://opennews.org/what/fellowships/community/>

> @christinezhang



Christine Zhang  
@christinezhang

the fancy entrance @latimes



<https://twitter.com/christinezhang/status/712438206910697473>

> @christinezhang

# > Research paper

Introduction

Background

Data

Model

Methods

Results

Conclusion

*tell*

+

*show*

Adapted from p. 35 of *A Guide to Writing in Economics* by Paul Dudenhefer  
<http://lupus.econ.duke.edu/ecoteach/undergrad/manual.pdf>

# > (Data) journalism

**Results**  
(e.g., lede, nut graph, viz)

**show**

+

Data    Model    Methods

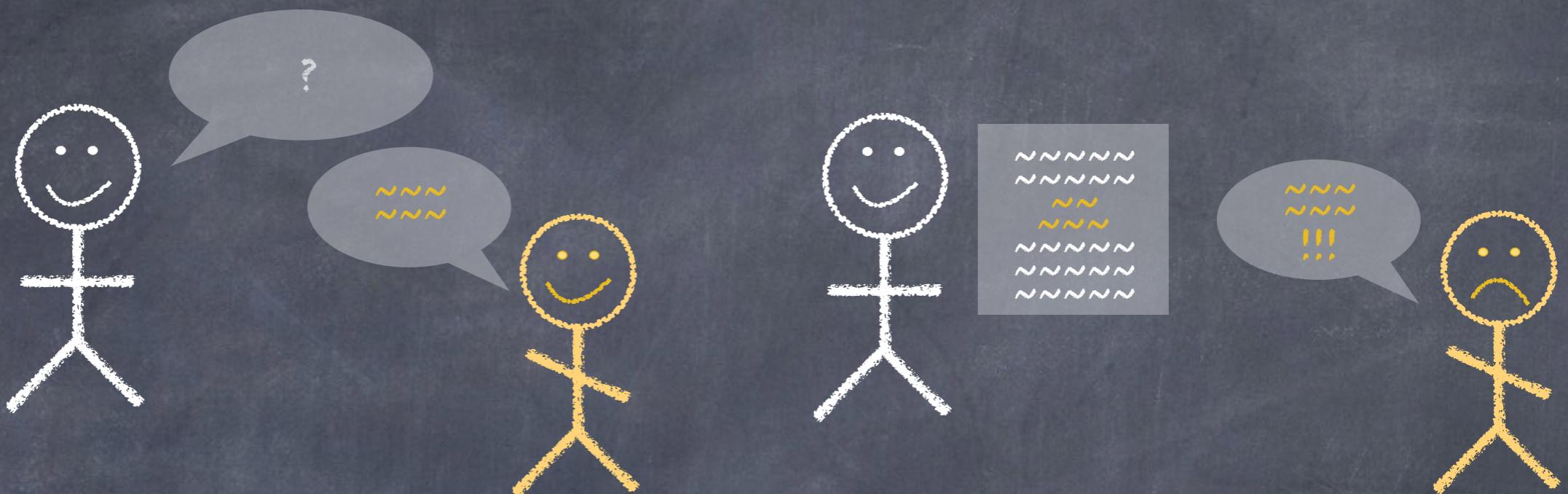
**(tell)**

Adapted from p. 35 of *A Guide to Writing in Economics* by Paul Dudenhefer  
<http://lupus.econ.duke.edu/ecoteach/undergrad/manual.pdf>

**Tell me your data**

# WHY?

# interviewing a person



## "interviewing" a dataset



data can't talk back!

openness, transparency,  
collaboration, etc.

**credibility**

# HOW?

# > The “nerd box”

**Los Angeles Times**

MOST POPULAR LOCAL SPORTS ORANGE COUNTY POLITICS ENTERTAINMENT OPINION 66°

ENTERTAINMENT / ENVELOPE / COMPANY TOWN

f t m

## As deaths rise on reality TV, film sets, many blame need to get 'dramatic footage'

Top, from left: The helicopter that crashed on the set of 1982's "Twilight Zone: The Movie"; the set of "Midnight Rider"; bottom left: Michael Huber, who was killed while dismantling a set; Michael Donatelli, who was killed in a helicopter crash in Acton; a still from the movie "Lone Ranger." (Handouts / Associated Press)

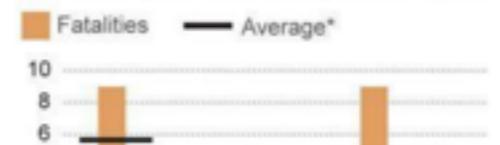
By **Richard Verrier** - Contact Reporter

But reality TV shows — especially ones that highlight danger — have added risks, industry safety experts say, because producers need dramatic footage to compete for viewers in an increasingly crowded field.

### Analysis

#### Film industry fatalities

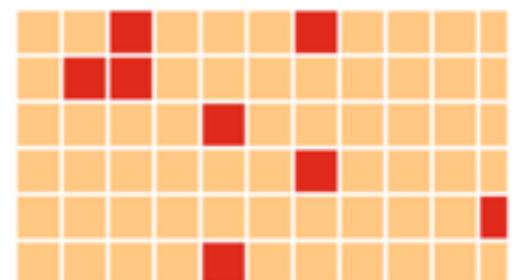
After a 15-year decline, fatalities in the motion picture industry have risen in the last five years even as all private sector fatalities fell.



#### How we examined the numbers

The cause of Monday's crash in Argentina — the deadliest yet in connection with a reality TV show — is under investigation. It occurred when two helicopters collided in midair during the filming of the series "Dropped" for French TV, in which participants are abandoned in the wilderness and must find their way back to civilization.

### Database



#### Database: Search film set accidents

"It's truly remarkable to me that production companies can use ultra-advanced technology to make spectacular films but too often they won't spend the modest resources necessary to make sure their workers are not injured or killed on the job," said David Michaels, assistant secretary of Labor for the Occupational Safety and

In this atmosphere, safety can take a back seat to ratings, said Angela Plasschaert, a Los Angeles-based risk management consultant who works with film producers and insurance companies to assess risks and develop safety programs.

"The attitude is, 'Do what you've got to do, but don't get in the way,'" she said of some of her clients.

The cause of Monday's crash in

Helicopters were also involved in two of the worst U.S. filming accidents in recent decades — a 2013 crash in Acton for a reality TV show that killed three people, and the 1982 "Twilight Zone: The Movie" crash near Santa Clarita that killed actor Vic Morrow and two children.

Safety investigators say Hollywood producers often fail to take proper precautions.

# > The “nerd box”

**Los Angeles Times**

MOST POPULAR LOCAL SPORTS ORANGE COUNTY POLITICS ENTERTAINMENT OPINION 68°

f t m

## Analysis How we examined the numbers

By Doug Smith and Ryan Menezes  
Contact Reporters

MARCH 10, 2015, 6:07 PM

Movie and television set fatalities have risen in the last five years, according to a Times analysis of U.S. government data. But wide variation from year to year makes it difficult to judge how the trend has changed from 1992 through 2014.

**READ MORE:** Filming accidents increased in last five years, data show

A procedure known as change-point analysis can help us see the changes. This analysis can identify the points at which the average number of fatalities changed over the 23 years of data. With the help of the R programming language and a package called “[strucchange](#)” we determined that there are five distinct segments over the time period. Fatalities on movie sets dropped during the middle part of the 1990s before rising in the early 2000s. But after a seven-year period of improvement (three years of which had zero fatalities), the average number of fatalities has more than doubled. Since 2010, an average of four people per year have died on movie sets.

### Film industry fatalities

After a 15-year decline, fatalities in the motion picture and video industry doubled over the last five years even as all private sector fatalities dropped.

Fatalities (orange bars) and Average\* (black line). Data points: 1992 (~4), 1994 (~3.5), 1996 (~2.5), 1998 (~1.5), 2000 (~4.5), 2002 (~3.5), 2004 (~1.5), 2006 (~3.5), 2008 (~1.5), 2010 (~3.5), 2011 (~7), 2014 (~3).

<http://www.latimes.com/entertainment/envelope/cotown/la-et-about-this-analysis-data-20150310-htmlstory.html>

**Film set accidents**

By JON SCHLEUSS, DOUG SMITH, RICHARD VERRIER  
MARCH 11, 2015

A breakdown of 205 production accidents from 1990 through 2014 involving 251 catastrophic injuries and 44 deaths. They are classified by the cause and type.

**Look up individual accidents**

Injury Cause of injury

Search < 1/1 > 250 per page

A table with 20 rows of accident data. Columns include Date, City, State, Company, Injury type, Injury source, Injuries, Fatalities, Fine, and Description. Key entries:  
- 8/27/14 Omaha NB Langley Productions Other Firearm 1 1 N/A Audio technician accidentally shot during film of show "Cops".  
- 8/8/14 Universal City CA FF7 Unknown Unknown 1 N/A Unknown  
- 7/28/14 Stanley NM Warner Horizon/A&E trauma Motor Vehicle 1 1 N/A Two crew members injured when camera equipment stuck a power line.  
- 4/2/14 Los Angeles CA Paramount Pictures Amputation Machine 1 N/A A carpenter's thumb was amputated.  
- 3/5/14 Los Angeles CA Red Studios Fracture Buildings/Structures 1 N/A Freelance photographer fell off a 19-foot 10-inch platform.  
- 2/23/14 Jesup GA Film Alman / Midnight Rider Multiple trauma Train 8 1 N/A Train collides into film crew.  
- 1/7/14 Los CA It's a Laugh Other Scissor Lift 1 N/A Employee

<http://spreadsheets.latimes.com/film-set-accidents/>  
> @christinezhang

# > GitHub + Jupyter

Los Angeles Times

MOST POPULAR LOCAL SPORTS ORANGE COUNTY POLITICS ENTERTAINMENT OPINION 65°

LOCAL / L.A. Now

f t e

## Parties, fights and trash fires: Signs of trouble seen weeks before arson at Westlake apartment



A body is lowered via cable from a charred building in the Westlake neighborhood. (Francine Orr / Los Angeles Times)

By Ruben Vives, Emily Alpert Reyes, Ben Welsh and Kate Mather

JUNE 15, 2016, 4:10 PM

**F**or weeks, Westlake neighbors complained about the green, two-story building along West 8th Street, empty except for the squatters who had moved inside.

<http://www.latimes.com/local/lanow/la-me-ln-westlake-fatal-fire-arrest-20160615-snap-story.html>

**p** Ben Welsh  @palewire Following

The quick analysis of vacant building complaints conducted by [@LATimesemily](#) now open-sourced here w/[@ProjectJupyter](#)



**datadesk/la-vacant-building-complaints-analysis**  
la-vacant-building-complaints-analysis - A quick analysis conducted for "Fire officials were concerned about Westlake building where 5 died in a blaze", a June 15, 2016, story by th...  
[github.com](https://github.com)

<https://twitter.com/palewire/status/743228155322736641>

> @christinezhang

# > GitHub + Jupyter

This screenshot shows a GitHub repository page for 'la-vacant-building-complaints-analysis'. The repository has 7 commits, 1 branch, and 0 releases. It includes files like .gitignore, README.md, combined\_complaints.csv, and la-vacant-building-complaints-analysis.ipynb. The README.md file describes a quick analysis of vacant building complaints.

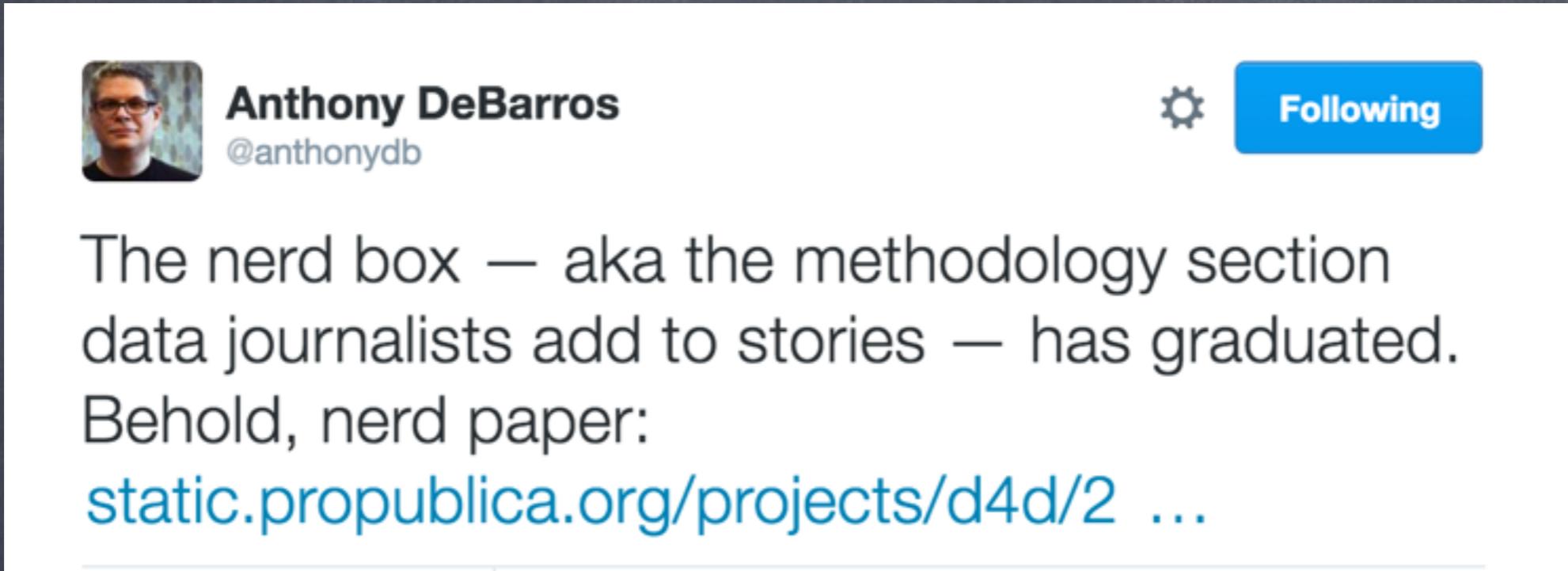
The right side of the screen displays a Jupyter notebook cell output titled 'L.A. vacant building complaints analysis' by Emily Alpert Reyes and Ben Welsh. The cell contains Python code using the agate library to print the schema of a CSV file. The schema is as follows:

column	data_type
CSR Number	Number
LADBS Inspection District	Number
Address House Number	Number
Address House Fraction Number	Date
Address Street Direction	Text
Address Street Name	Text
Address Street Suffix	Text

<https://github.com/datadesk/la-vacant-building-complaints-analysis>

> @christinezhang

# > The white paper / "nerd paper"



The nerd box — aka the methodology section data journalists add to stories — has graduated. Behold, nerd paper:  
[static.propublica.org/projects/d4d/](http://static.propublica.org/projects/d4d/) ...

<https://twitter.com/anthonydb/status/710434955340414976>

# > The white paper / "nerd paper"

Dollars for Doctors

## Now There's Proof: Docs Who Get Company Cash Tend to Prescribe More Brand-Name Meds

The more money doctors receive from drug and medical device companies, the more brand-name drugs they tend to prescribe, a new ProPublica analysis shows. Even a meal can make a difference.

by Charles Ornstein, Ryann Grochowski Jones and Mike Tigas, ProPublica, March 17, 2016, 5 a.m.



This story was co-published with [NPR](#), the [Boston Globe](#) and the [Tampa Bay Times](#).

Doctors have long disputed that the payments they receive from pharmaceutical companies have any relationship to how they prescribe drugs.

There's been little evidence to settle the matter — until now.

A ProPublica analysis has found for the first time that doctors who receive payments from the medical industry do indeed tend to prescribe drugs differently than their colleagues who don't. And the more money they receive, on average, the more brand-name medications they prescribe.

We matched records on payments from pharmaceutical and medical device makers in 2014 with corresponding data on doctors' medication choices in Medicare's prescription drug program. (You can [read our methodology here](#).)

Matching Industry Payments to Medicare Prescribing Patterns: An Analysis

Ryann Grochowski Jones and Charles Ornstein  
ProPublica

March 2016

### 1 Summary

This analysis examined Medicare's Part D prescription data as well as pharmaceutical and medical device company payment data to see if there was a relationship between industry payments and brand-name prescribing. We found that physicians in five common medical specialties who accepted at least one industry payment were more likely to prescribe high rates of brand-name drugs than physicians who did not receive any payments. Then we compared average prescribing rates among physicians divided in several ways: those who received payments and those who didn't; those who received large dollar amounts of payments and those who received smaller amounts; and those who received certain types of payments (meals, speaking) and those who didn't. In all cases, the group receiving larger payments had a higher brand-name prescribing rate on average. Additionally, the type of payment made a difference: those who received meals alone from companies had a higher rate of brand-name prescribing than physicians who received no payments, and those who received speaking payments had a higher rate than those who received other types of payments.

### 2 Introduction

While it has long been theorized that pharmaceutical company payments to doctors influence their prescribing, empirical research on the topic has been scant. ProPublica launched its [Dollars for Docs](#) tool in 2010 to make it easier for consumers to look up whether their physicians received payments from certain drug companies that disclosed this information as part of lawsuit settlements. In 2013, ProPublica introduced [Prescriber Checkup](#), a tool that allows users to compare their physicians' prescribing patterns in the Medicare Part D

<https://www.propublica.org/article/doctors-who-take-company-cash-tend-to-prescribe-more-brand-name-drugs>  
<https://static.propublica.org/projects/d4d/20160317-matching-industry-payments.pdf?22>

> @christinezhang

"Journalists are not  
motivated by heaven;  
they are motivated by  
glory"

— Joe Germuska  
Chief Nerd, Knight Lab  
@JoeGermuska



Open Data

La Nacion

10. Open data award. Using freedom of information and/or other levers to make crucial databases open and accessible for re-use and for creating data-based stories. 1,000€ Prize

**SEARCH**

Enter a search term...

[Search](#)**AWARDS****AWARDS**[IRE AWARDS](#)[PHILIP MEYER AWARDS](#)[GOLDEN PADLOCK](#)[Subscribe to the IRE Radio Podcast](#)[Login | Register](#)[AWARDS » GOLDEN PADLOCK](#)

## The Golden Padlock Award

Investigative Reporters and Editors in 2013 launched a new award - dubbed the Golden Padlock - recognizing the most secretive publicly-funded agency or person in the United States.

"Thwarting the public's right to know has become a mission statement inside many government bureaucracies across the country," said Robert Cribb, chair of the Golden Padlock committee. "This award recognizes tireless commitment to that principle and honors the most creative ways in which it is executed every day by determined civil servants."

Governments at all levels – from local to federal – are eligible for the award. The "honor" is handed out annually at IRE's national conference.



### HOW TO APPLY

- Email your nomination to [goldenpadlock@ire.org](mailto:goldenpadlock@ire.org).
- Submissions should include the name of the agency or individual along with reasons and/or media coverage detailing the intransigence.

# > Take my survey

The NiemanLab website features a prominent image of wooden blocks spelling "OPEN DATA". Below the image is a call-to-action text: "How does your newsroom share the data behind your data journalism with the public? Fill out this survey!" The NiemanLab logo is at the top left, and a navigation bar includes "Fellowships", "Reports", "Lab", "Storyboard", and "Nieman Foundation at Harvard". A "SUBSCRIBE" button is also visible.

NiemanLab

ABOUT SUBSCRIBE

OPEN DATA

How does your newsroom share the data behind your data journalism with the public? Fill out this survey!

The survey interface shows a blue header with the question "News orgs: How do you share your data with the public?". Below the header, it says "Section 1 of 5". The main content area contains the question text and a descriptive paragraph about data stories in journalism. It includes fields for "Name" and "Organization \*".

← News orgs: How do you share your data with the public?

Section 1 of 5

News orgs: How do you share your data with the public?

There are a lot of data stories in journalism. Sometimes news organizations make their underlying data and methodology available to the public, sometimes they don't. We (L.A. Times Knight-Mozilla Fellow Christine Zhang <https://twitter.com/christinezhang> & Stanford Journalism Lecturer Dan Nguyen <https://twitter.com/dancow>) would like to explore what drives these decisions at the Computation + Journalism conference <http://journalism.stanford.edu/cj2016/>. Help us out by filling out this short questionnaire.

Thanks!

NAME

Organization \*

>> <http://bit.ly/2bEgaM5>

> @christinezhang

> Thank you

>> Find my slides

<https://github.com/underthecurve/tell-me-your-data>

>> Take my survey

<http://bit.ly/2bEgaM5>

>> Tell me your data

christine.zhang@latimes.com || @christinezhang

+ wherever there is good food & drink

>> <http://bit.ly/2bEgaM5>