

# KDD 2014

Data Science for Social Good

20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining

Aug 24-27 New York City



# KDD Ignite Talks

In Partnership With

**DataKind**  
USING DATA IN THE SERVICE OF HUMANITY

Co Organizers: Paul Bradley, Brian d'Alessandro w/ Jake Porway.

Special Thanks to:

*Rayid Ghani, Prem Melville, Claudia Perlich, Ted Senator, Sofus Mackasssey,  
Misha Bilenko, Eric Bloedorn, Usama Fayyad, Johannes Gehrke, Ronny Kohavi, Ying Li, Gabor Melli, Jian Pei, Ramasamy Uthurusamy, Michael Zeller,  
Rajesh Parekh, Gregory Piatetsky, Mohammed Zaki*

# KDD 2014

## Data Science for Social Good

20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining

Aug 24-27

New York City



# Jake Porway, Executive Director and Founder



# KDD 2014

Data Science for Social Good

20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining  
Aug 24-27 New York City



[UK](#) | [Register](#) | [Log In](#)

Search



[TV SHOWS](#) ▾

[VIDEOS](#) ▾

[PHOTOS](#) ▾

[TV LISTINGS](#)

## THE NUMBERS GAME



[HOME](#)

[ABOUT](#)

[VIDEOS](#)

[PHOTOS](#)

[MORE](#) ▾



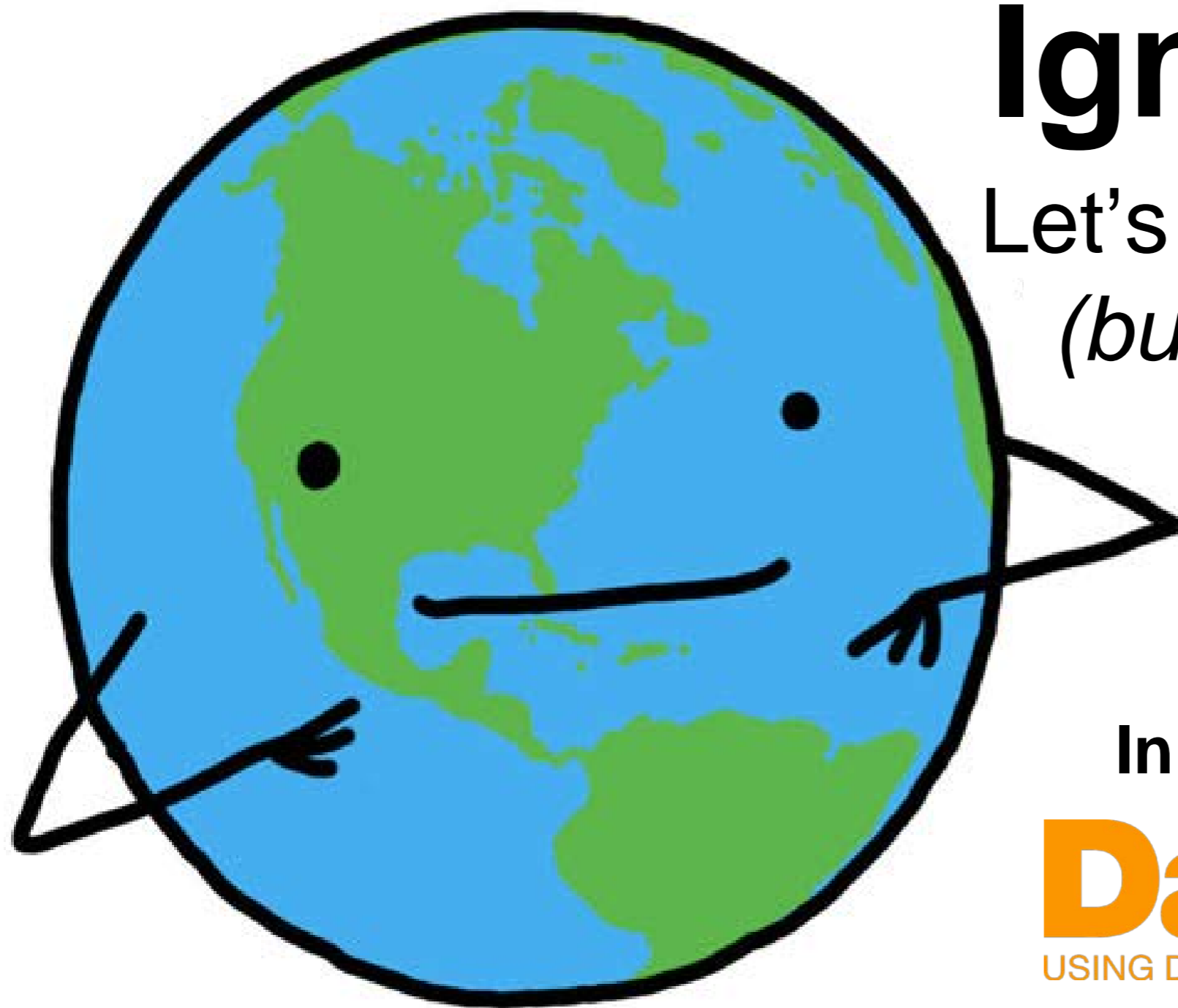
# KDD 2014

Data Science for Social Good

20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining

Aug 24-27

New York City



# Ignite Talks

Let's Save the World!  
*(but make it quick)*

In Partnership With

## DataKind

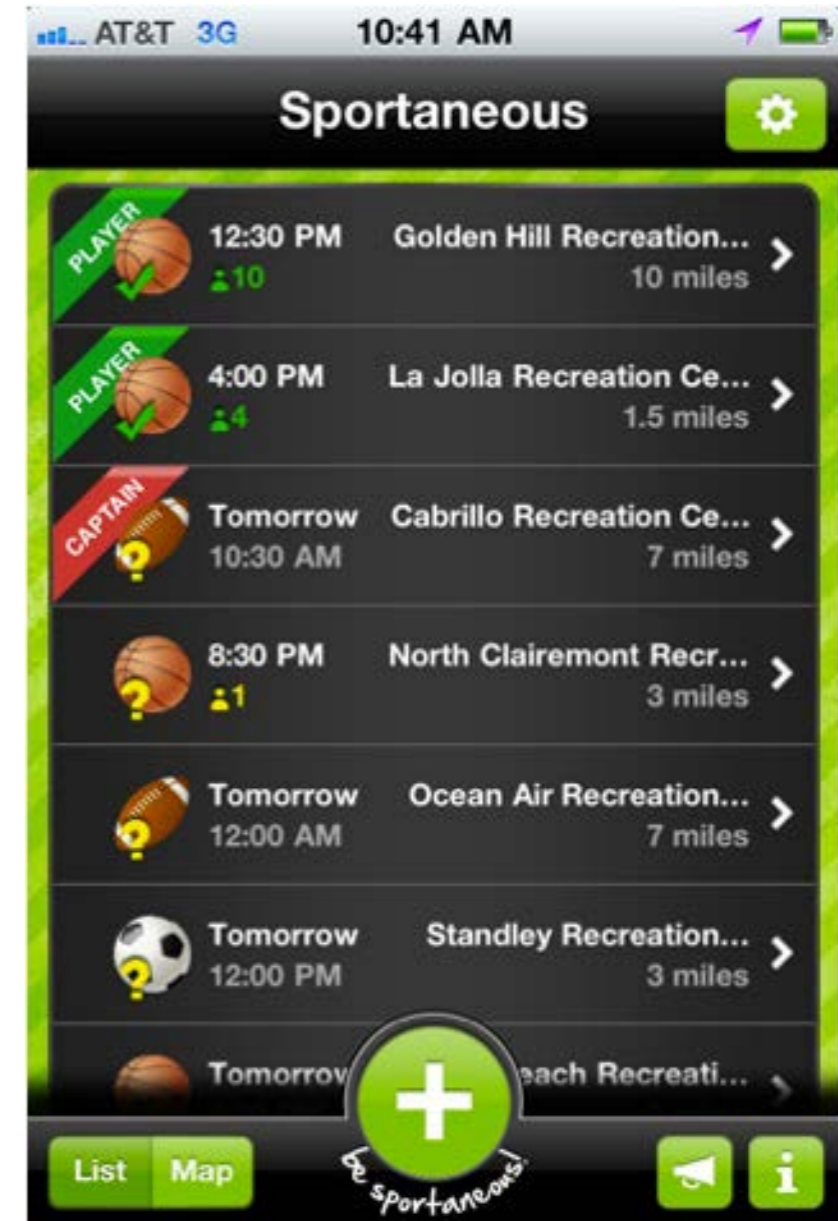
USING DATA IN THE SERVICE OF HUMANITY



# KDD 2014

Data Science for Social Good

20th ACM SIGKDD Conference on Knowledge Discovery and Data Mining  
Aug 24-27 New York City



<http://bestparking.com>

<http://salelocator.com>

<http://sportaneous.com>









# DataKind

brings organizations dedicated to solving the world's biggest challenges together with expert pro bono data scientists to improve the quality of, access to and understanding of data in the social sector.



DSI (ah Sottz)  
\$1gn \$tuff

DATA

datakind.org

BEST TREE INVENTORY  
LAT LON CENSUS

try MGT

ACCESS

CI

2005-present

- 2004-present
- COMBINE INTO ONE CSV
  - SCRUB NAMES
  - LOW LAT

- 1998-2008
- TREE MGMT SYSTEM
- ACCESS TO CSV
  - SCRUB NAMES
  - GEOTAG (partly NYC Parks)

Brooklyn only 1997-2008

What is Risk?

1. Prob. of tree falling?
2. # trees falling?



$$E(\#) = (\#) \cdot \#T_{\text{fall}}$$

TREE SPECIES CHARACTERISTICS

- USDA Plant Database
- US Tree Inventory
- NYC Parks website, Approved Species of Florida, IFAS, etc.



