



Content Notes

This presentation contains references to:

- ◇ Disordered eating
- ◇ Police violence
- ◇ Pregnancy loss





Hoarding, poisoned silos, and the madness of machines

The Death of Data



OUTR FRIEND
ETRGOT



Data Hoarding



© Erik Kessels/Gijs Van Den Berg/Caters News







Keep it

- ◇ Legal and fiduciary data
- ◇ Historical company info
- ◇ Old ads
- ◇ Archival material, like student transcripts and newspaper scans
- ◇ Data to show a trajectory
- ◇ Data you trained people to retain





Toss it

- ◇ Meeting invites
- ◇ Resolved items over a certain age
- ◇ Rejected paths and plans
- ◇ Data we don't want to see in the courtroom in 40 years, RJ Reynolds.
- ◇ Expired logging data, notifications, and anything we automate into a folder and mark as read.
- ◇ Ugly, duplicated, or unlabeled pictures.



@wiredferret | The Death of Data | Signal SFO 2017







BE HUMBLE.

WE CANNOT INTUIT
INNER STATE, EMOTIONS,
PRIVATE SUBJECTIVITY

—Carina C. Zona



RUBYCONF

SAN ANTONIO | 2015





Criminal Element





Banker Crime Spree!

Defaulting! Plummeting property values!
Criminal association. They're no angels. Where
are the police? Recidivism at an all-time high.
Back on the streets again! Revolving door
penal system. Financial district: are you safe?



The Death of Data





Hold on a minute!



I could use that!



Archives are not
backups







Sunk cost fallacy

“

It's practically free!



So is it really a sale if you didn't
intend to buy?












Profiling is about people,
not just their data points.



Anger, bargaining, and denial

We have feelings about data



To kill data is to
acknowledge the loss
of its potential.







@wiredferret | The Death of Data | Signal SFO 2017





Rex Rizzo
@RexRizzo



 Follow

Wired: "Machine learning will TAKE OVER THE WORLD!"

Amazon: "We see you bought a wallet. Would you like to buy ANOTHER WALLET?"

RETWEETS

8,300

LIKES

11,583



11:25 AM - 28 Jul 2016



BMI

Case study in poisoned data







The Poison Propagates





400m sprint

1896 and 2016 average height and weight of sprint athletes

1912 400m sprint
gold medallist
Charles Reidpath



Height

173cm



1896

Weight

60kg

178cm

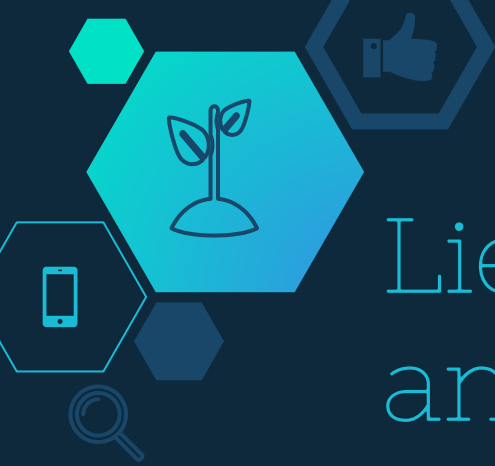


2016

77kg



Current Olympic
defending champion
Kirani James



Lies, Damned Lies, and Statistics





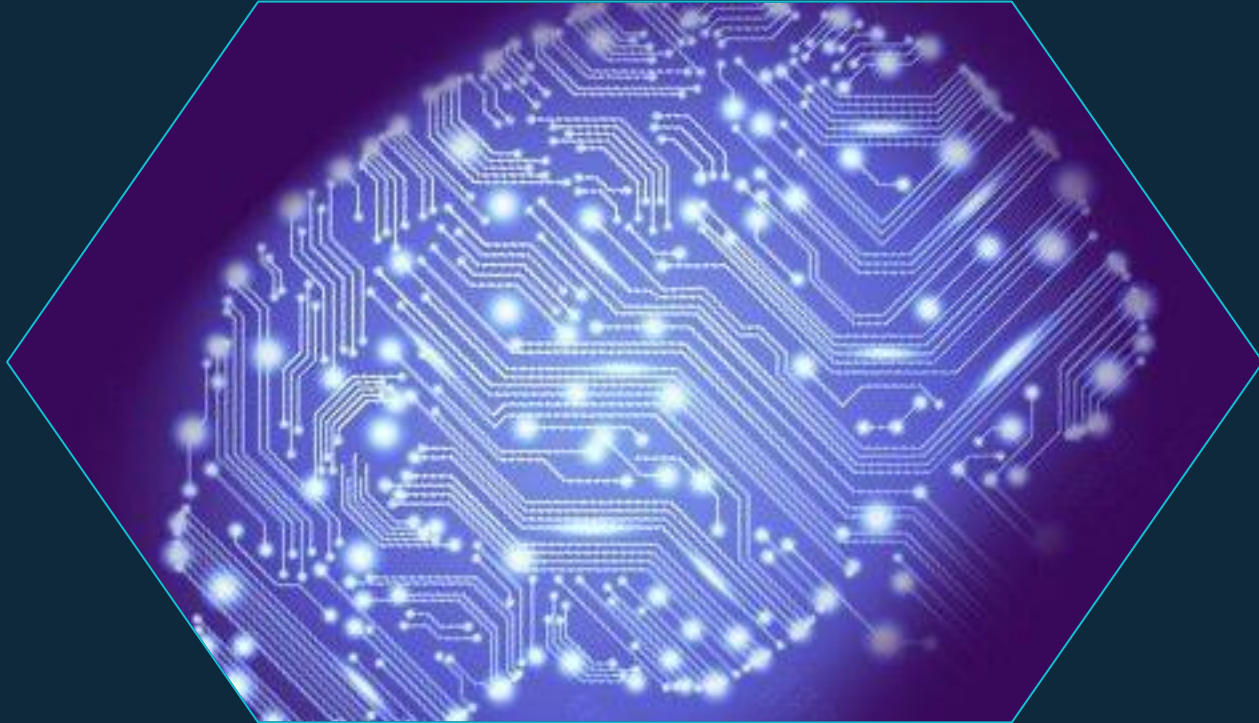
1,347 steps

10,000



11 hr 42 min asleep

Goal met





Old data is not neutral
“**cr**nk matter, it is the dead
hand of history, and it
can infect current data.

We are symbiotic
with big data



Useful Ideas

A decorative graphic on the left side of the slide consists of several hexagons of varying shades of blue and cyan. Some hexagons contain white icons: a lightbulb, a thumbs-up, a smartphone, a magnifying glass, and a gear. There is also a network-like icon with a central node and four surrounding nodes. A large, solid cyan hexagon is positioned in the center of this graphic cluster.

Hire an Archivist

A decorative graphic on the left side of the slide consists of several hexagons of varying shades of blue and cyan. Some hexagons contain icons: a lightbulb, a thumbs-up, a smartphone, a magnifying glass, and a gear. There is also a network-like icon with a central node and four surrounding nodes. A large, solid cyan hexagon is positioned in the center of this graphic cluster.

Constrain storage space

A decorative graphic on the left side of the slide consists of several hexagons of varying shades of blue and cyan. Some hexagons contain icons: a lightbulb, a thumbs-up, a smartphone, a magnifying glass, and a gear. There is also a network-like icon with a central node and four surrounding nodes. A large, solid cyan hexagon is positioned in the center of this graphic cluster.

Train Data Retention

AUTOMATE DELETION



- ◇ Collect carefully
- ◇ Ingest mindfully
- ◇ Delete boldly





Hoarding, poisoned silos, and the madness of machines

The Death of Data

