

# Criminal Machine Learning - Biased Training Set?



(a) Three samples in criminal ID photo set  $S_c$ .



(b) Three samples in non-criminal ID photo set  $S_n$

# Criminal Machine Learning - Biased by Smile?



(a) -0.98



(b) -0.68



(c) -0.28



(d) -0.38



(e) 0.76



(f) 0.98



(g) 0.66



**Avoid falling for Bullshit.**

**Statistical Bullshit, the new norm.**

**Develop Critical Thinking &  
Data Science Skills.**

# Falling for Bullshit is way too easy.

# 50 COGNITIVE BIASES TO BE AWARE OF SO YOU CAN BE THE VERY BEST VERSION OF YOU

## Memory

## Social

## Learning

## Belief

## Money

## Politics

### Fundamental Attribution Error

We judge others on their personality or fundamental character, but we judge ourselves on the situation.



*Sally is late to class; she's lazy. You're late to class; it was a bad morning.*

### Self-Serving Bias

Our failures are situational, but our successes are our responsibility.



*You won that award due to hard work rather than help or luck. Meanwhile, you failed a test because you hadn't gotten enough sleep.*

### In-Group Favoritism

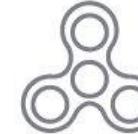
We favor people who are in our in-group as opposed to an out-group.



*Francis is in your church, so you like Francis more than Sally.*

### Bandwagon Effect

Ideas, fads, and beliefs grow as more people adopt them.



*Sally believes fidget spinners help her children. Francis does, too.*

### Groupthink

Due to a desire for conformity and harmony in the group, we make irrational decisions, often to minimize conflict.



*Sally wants to go get ice cream. Francis wants to shop for T-shirts. You suggest getting T-shirts with pictures of ice cream on them.*

### Halo Effect

If you see a person as having a positive trait, that positive impression will spill over into their other traits. (This also works for negative traits.)



*"Taylor could never be mean; she's so cute!"*

### Moral Luck

Better moral standing happens due to a positive outcome; worse moral standing happens due to a negative outcome.



*"X culture won X war because they were morally superior to the losers."*

### False Consensus

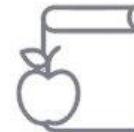
We believe more people agree with us than is actually the case.



*"Everybody thinks that!"*

### Curse of Knowledge

Once we know something, we assume everyone else knows it, too.



*Alice is a teacher and struggles to understand the perspective of her new students.*

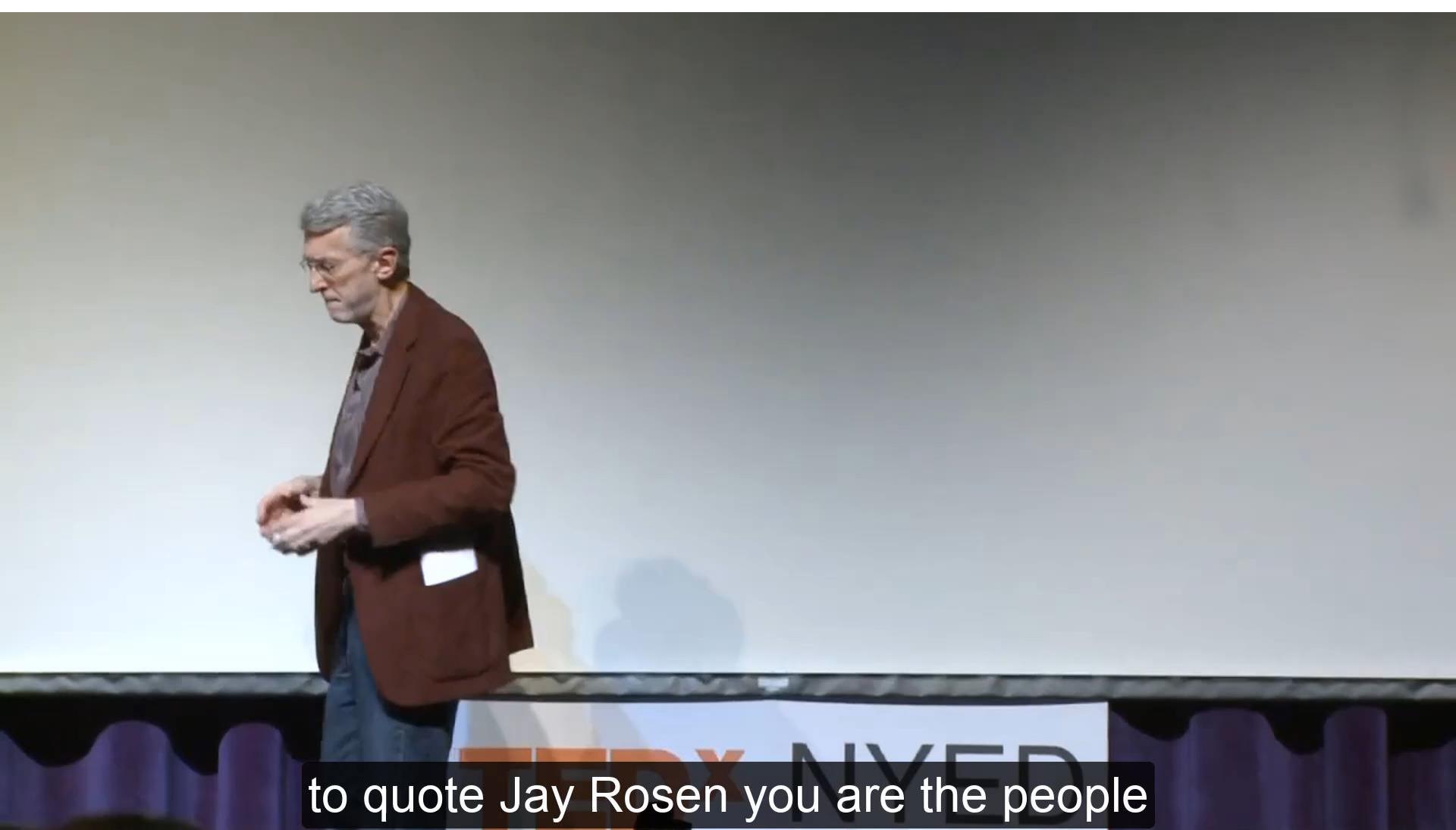
### Spotlight Effect

We overestimate how much people are paying attention to our behavior and appearance.

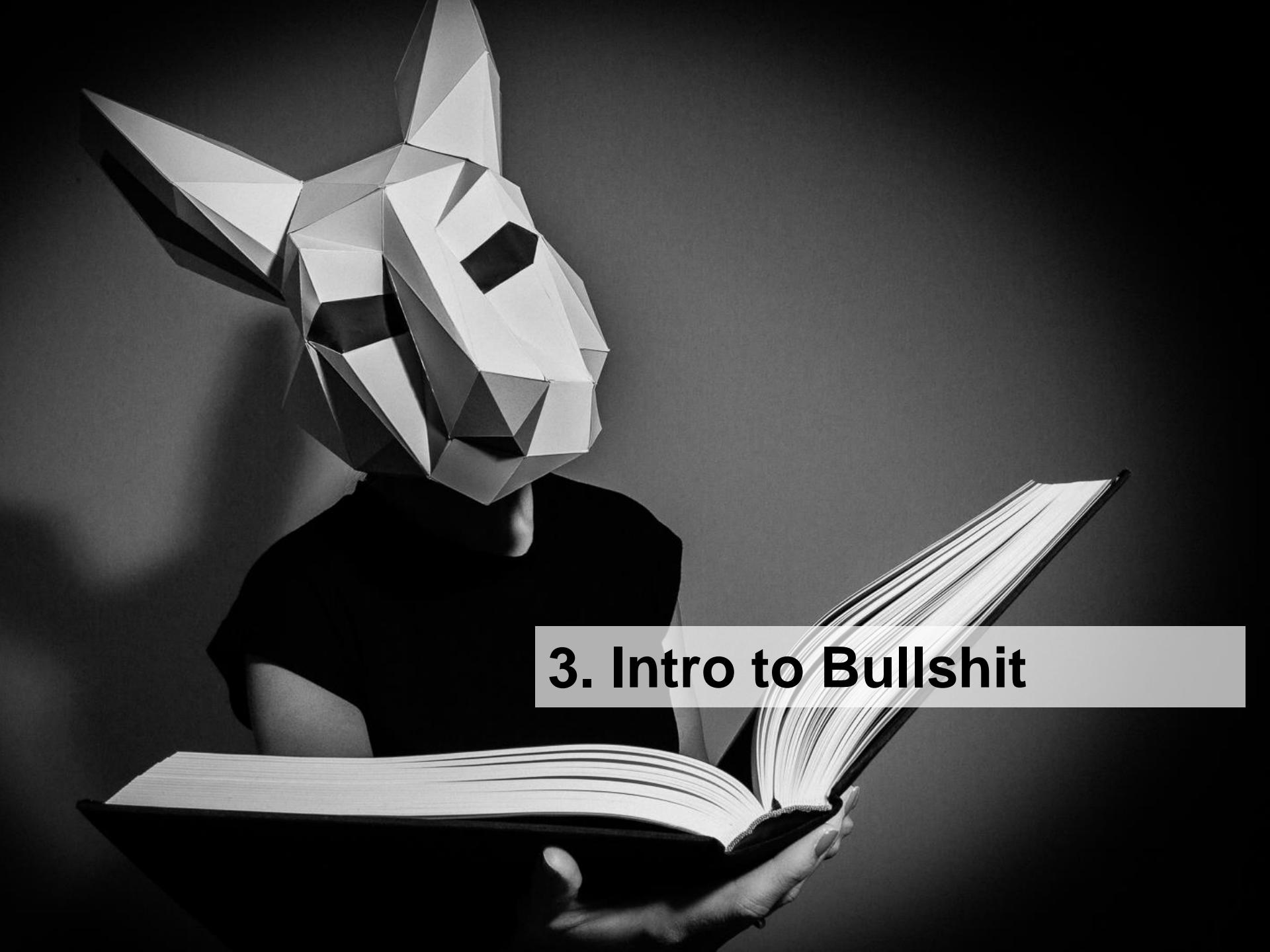


*Sally is worried everyone's going to notice how lame her ice cream T-shirt is.*

# Jeff Jarvis, Ted Talk über Bullshit

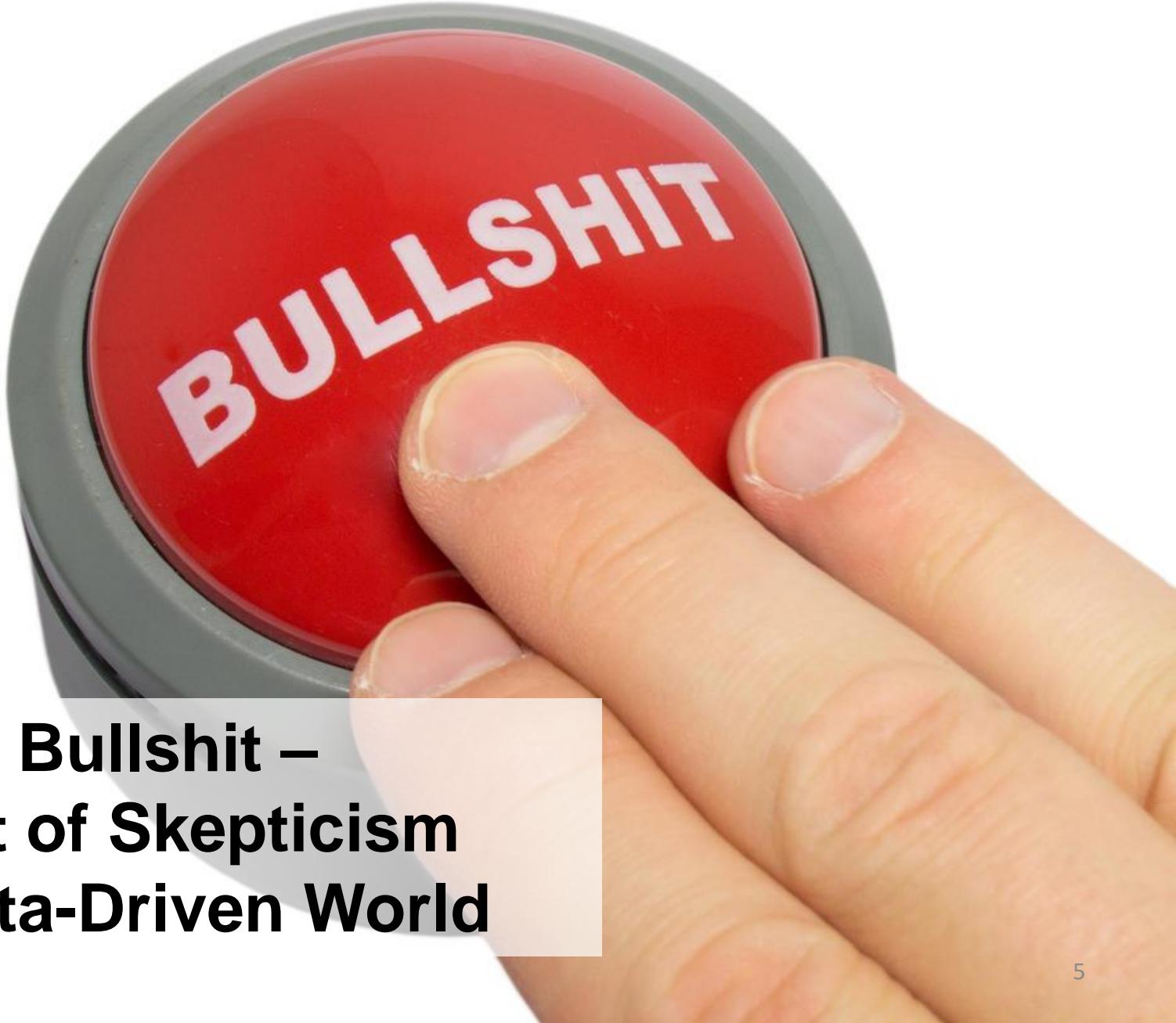


to quote Jay Rosen you are the people

A black and white photograph of a person from the chest up. They are wearing a highly detailed, geometric, low-poly paper mask of a donkey's head. The mask is white against a dark background. The person is holding an open book with both hands; the pages are visible and appear to be white. The lighting is dramatic, coming from the side to highlight the mask and the edges of the book.

### **3. Intro to Bullshit**

# **Calling Bullshit – The Art of Skepticism In a Data-Driven World**



# Bullshit in Sign Language

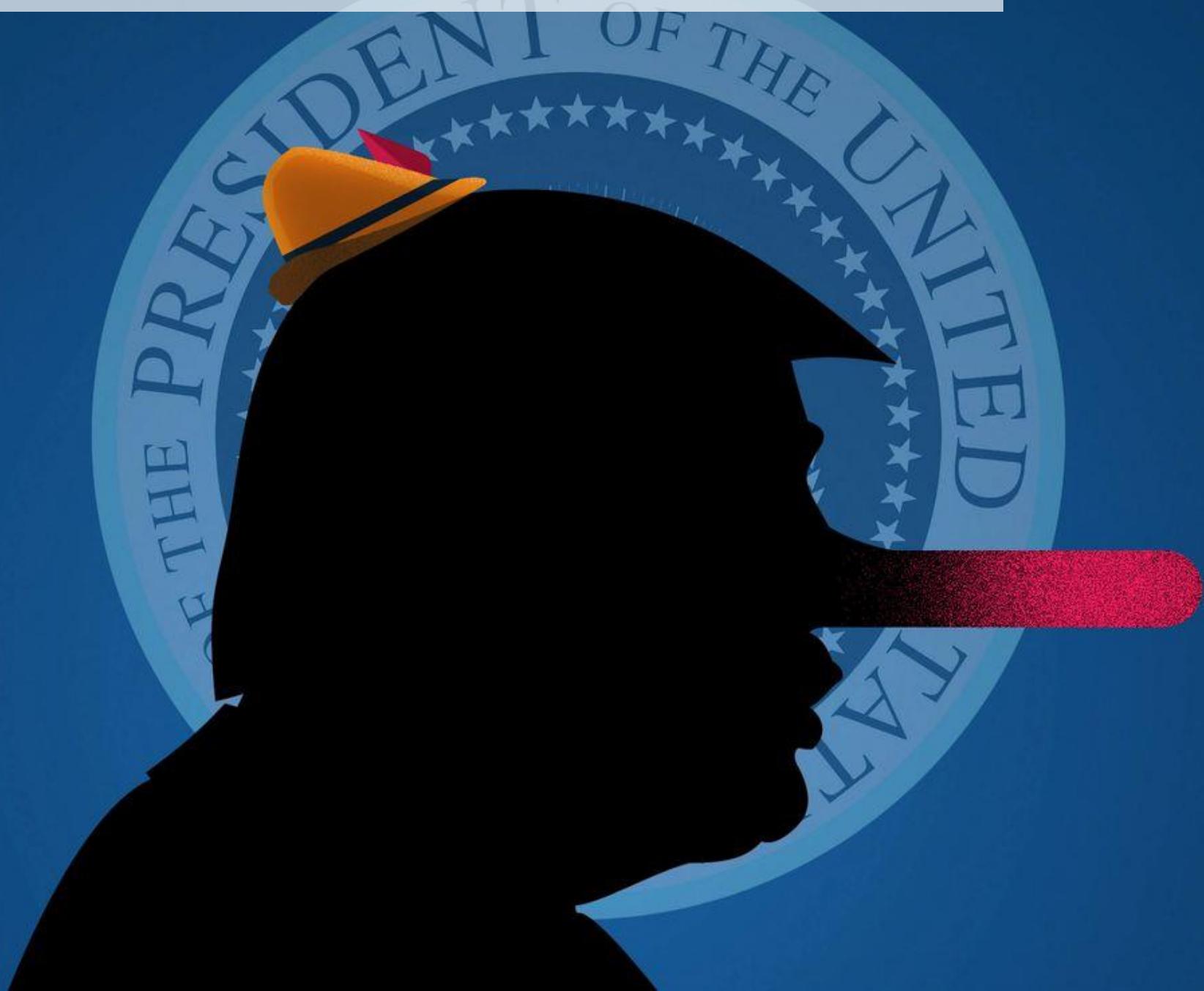
bullshit

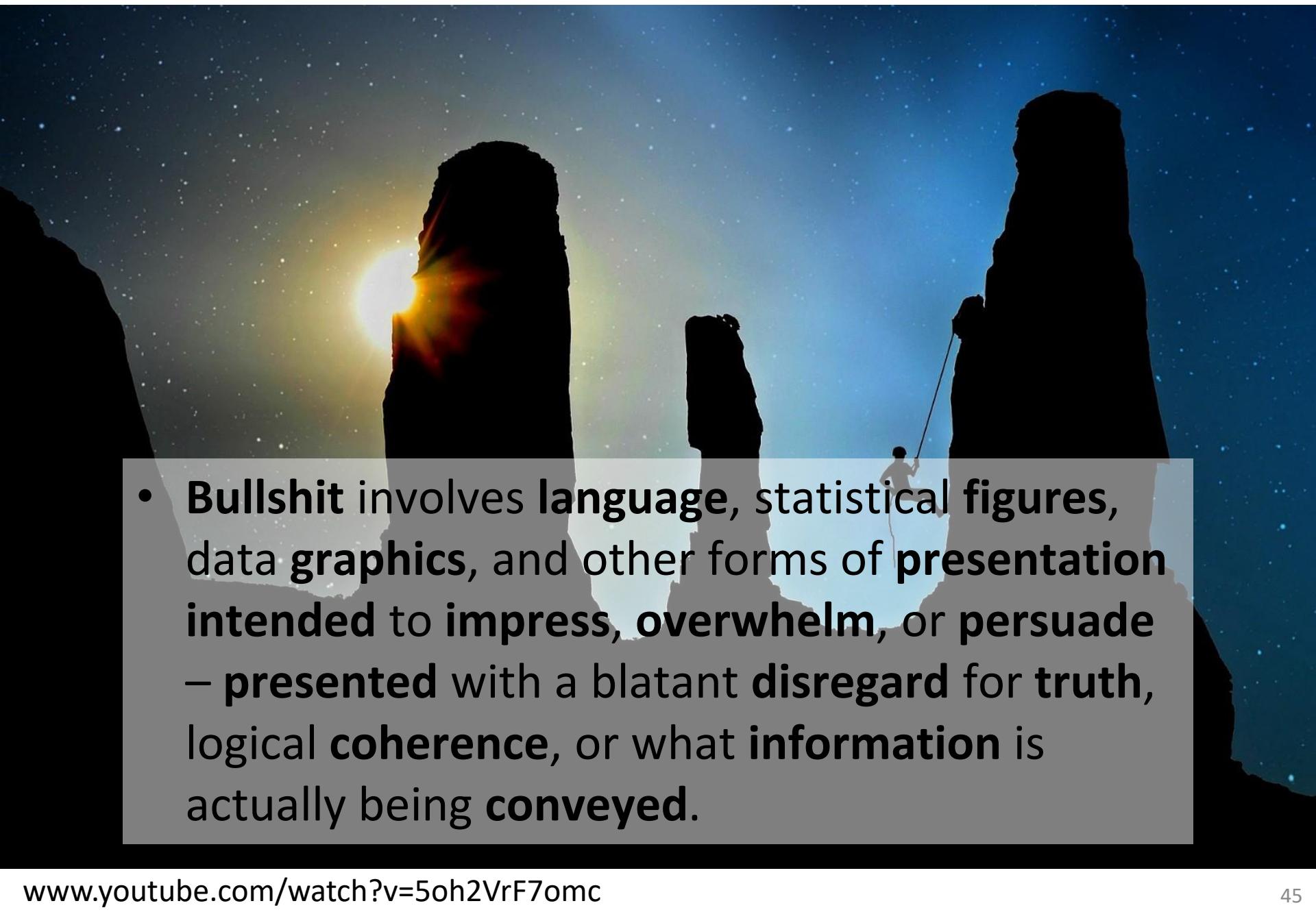


# Harry Frankfurt, Princeton University, 2005: On Bullshit

- 
- A **Liar** knows the **truth** and is trying to **convince** us of **something different**.
  - A **Bullshitter** either doesn't know the **truth**, or doesn't **care** – and is just trying to be **persuasive**.

Vox.com: The Bullshitter-in-Chief, 5/2017



- 
- A large, semi-transparent rectangular box covers the bottom half of the slide. Inside this box is a list of bullet points. In the background, there is a silhouette of a person climbing a tall, dark rock formation against a colorful, star-filled sky transitioning from yellow to blue.
- **Bullshit** involves **language**, statistical **figures**, data **graphics**, and other forms of **presentation** intended to **impress**, **overwhelm**, or **persuade** – **presented** with a blatant **disregard** for **truth**, logical **coherence**, or what **information** is actually being **conveyed**.

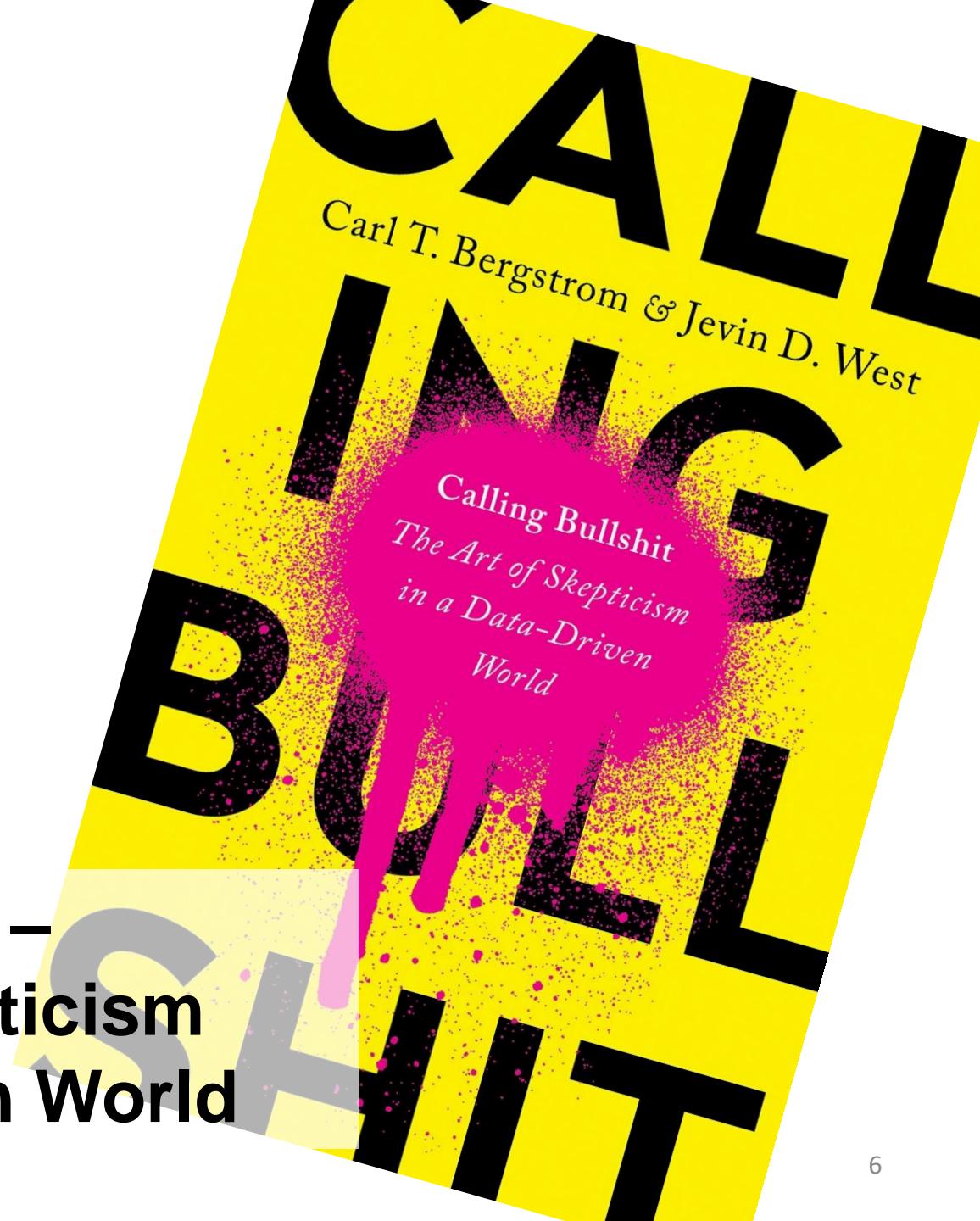
## Bergstrom, Calling Bullshit:

- If a **claim** seems **too good** — or **too bad** — to be **true**,
- it probably is.**

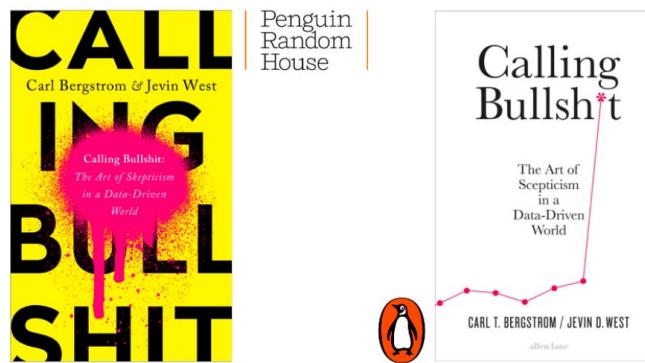
A photograph showing a group of Amish people, including adults and children, walking in a line on a dirt path. They are dressed in traditional Amish clothing: men in dark hats and vests over white shirts, and women in dark bonnets and dresses. The scene is set outdoors with trees and a fence in the background.

## 4. Old School Bullshit

# Calling Bullshit – The Art of Skepticism In a Data-Driven World

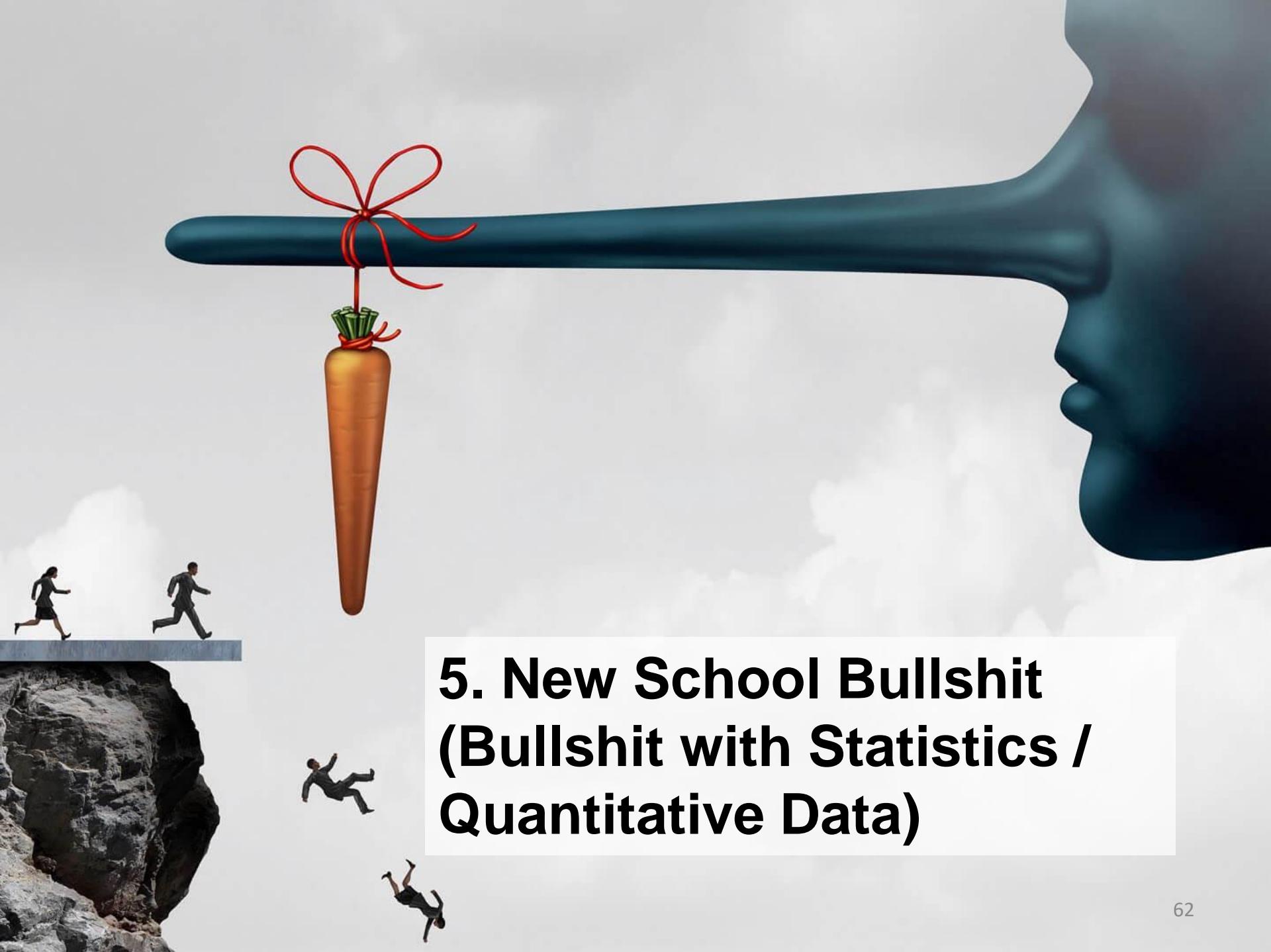


# Website **callingbullshit.org**



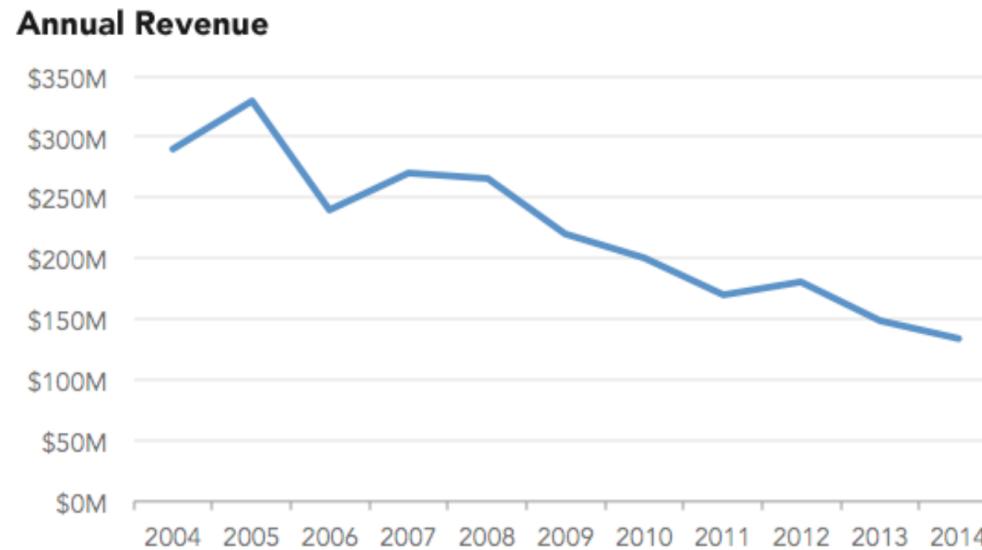
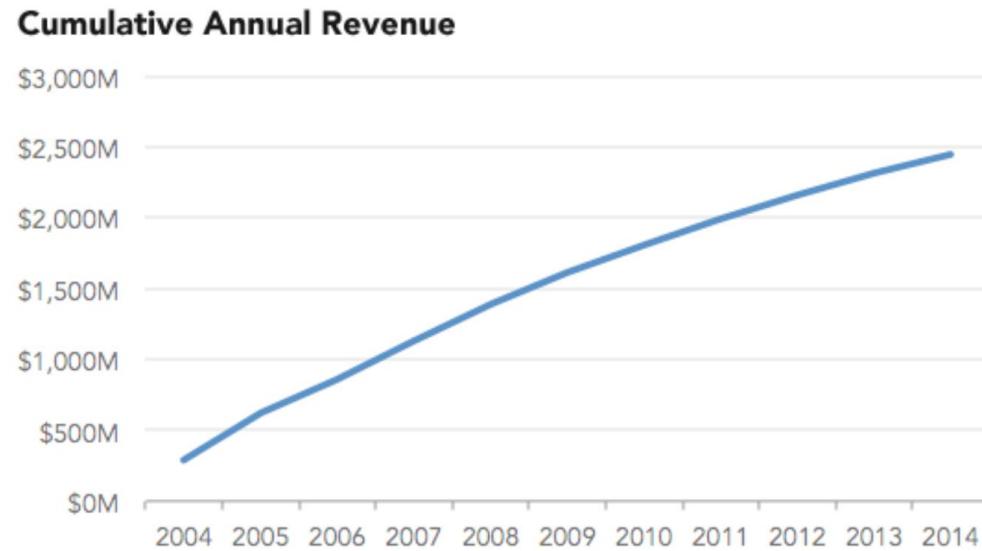
**Now available!** *Calling Bullshit: The Art of Skepticism in a Data-Driven World*, by Carl Bergstrom and Jevin West. [Available here.](#)

**The world is awash in bullshit.** Politicians are unconstrained by facts. Science is conducted by press release. Higher education rewards bullshit over analytic thought. Startup culture elevates bullshit to high art. Advertisers wink conspiratorially and invite us to join them in seeing through all the bullshit — and take advantage of our lowered guard to bombard us with bullshit of the second order. The majority of administrative activity, whether in private business or the public sphere, seems to be little more than a sophisticated

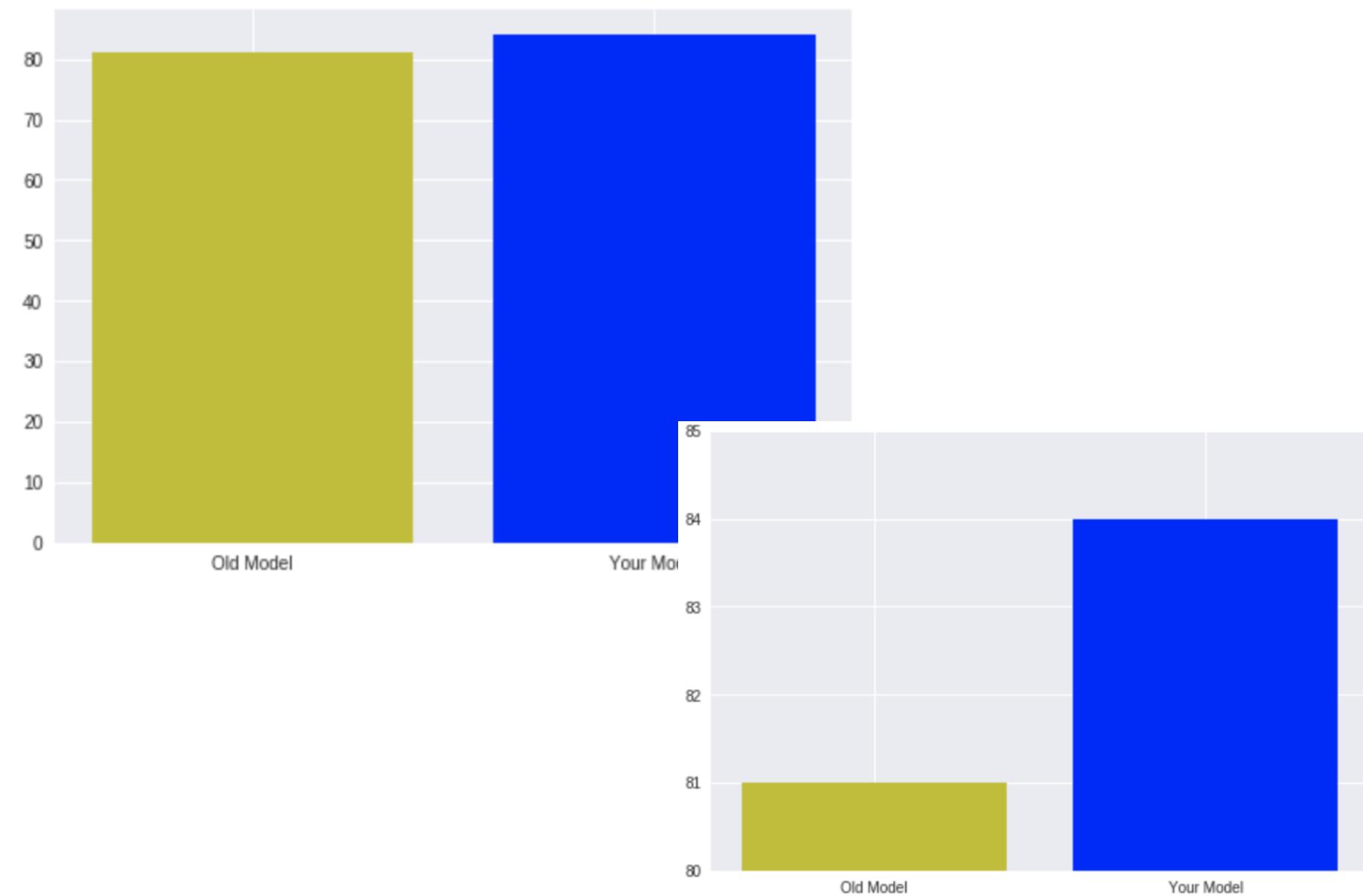


## **5. New School Bullshit (Bullshit with Statistics / Quantitative Data)**

# Bullshit: Cumulative graphs

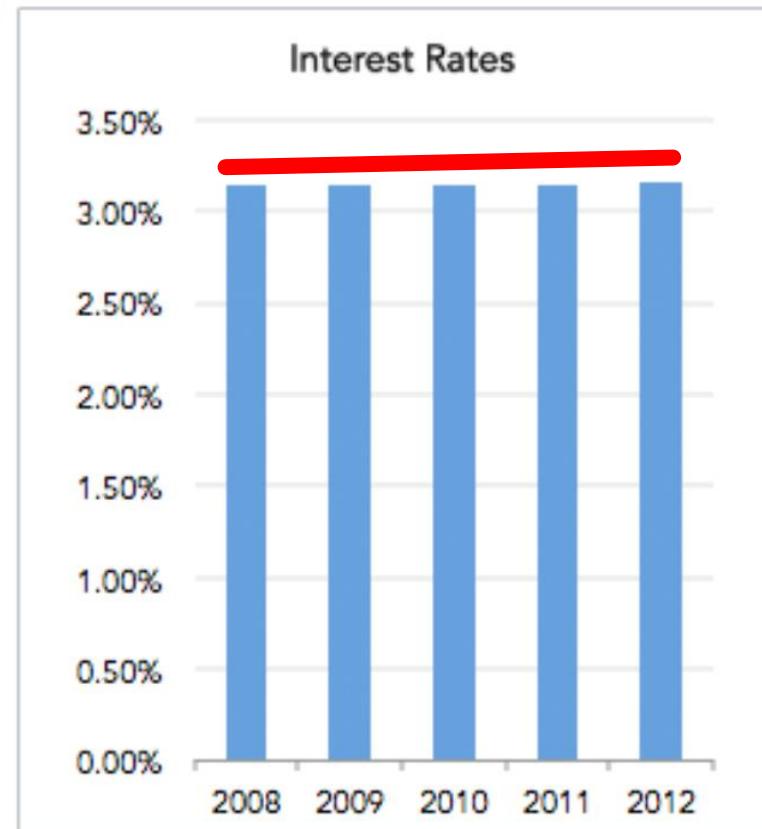
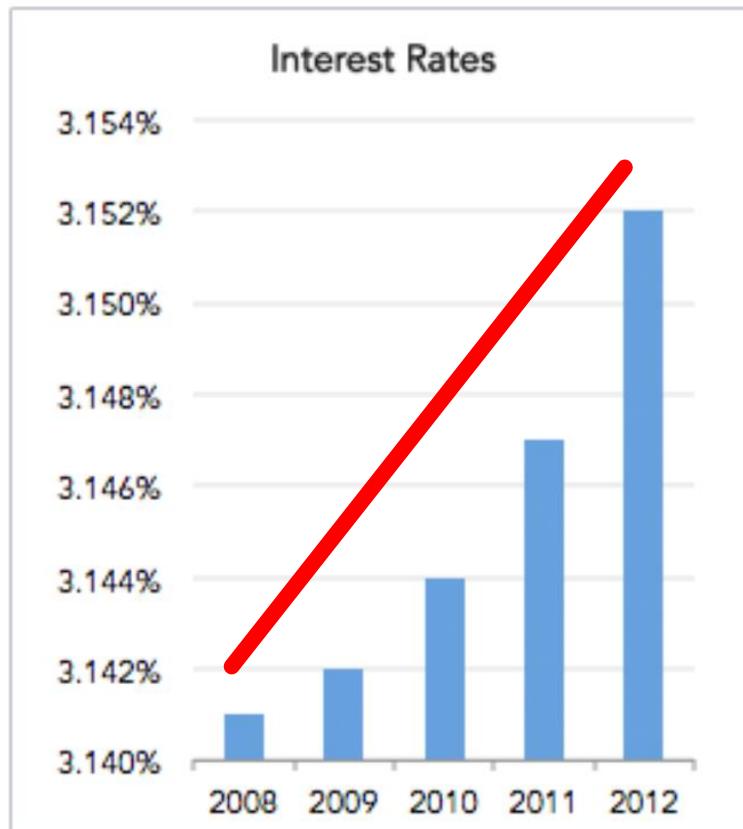


# Bullshit: Truncated Y-Axis

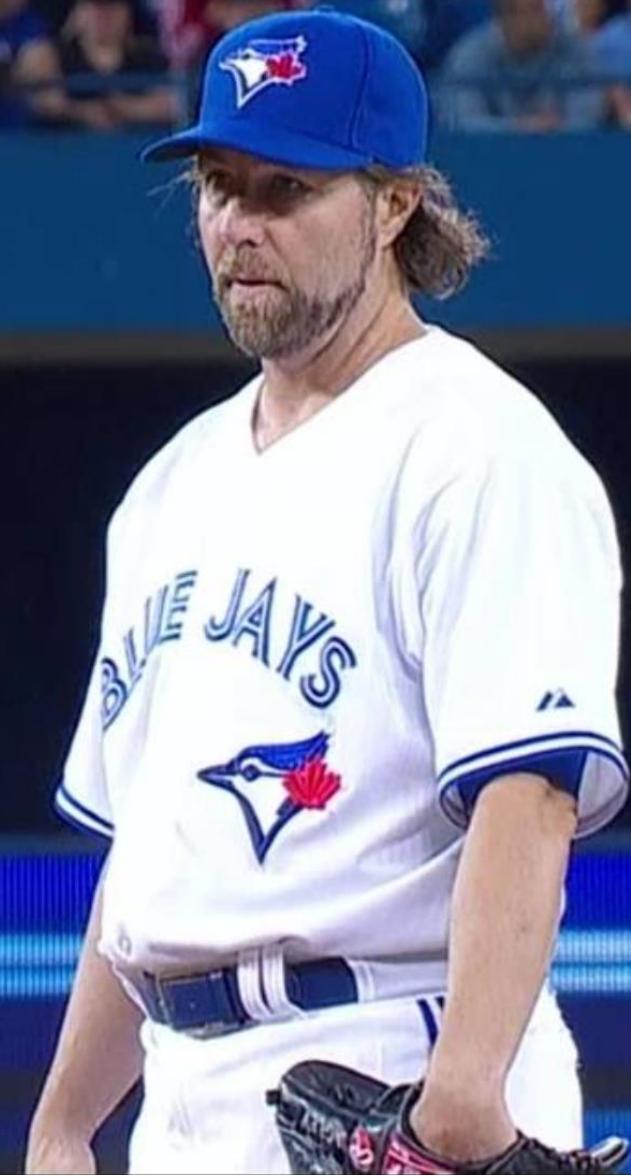
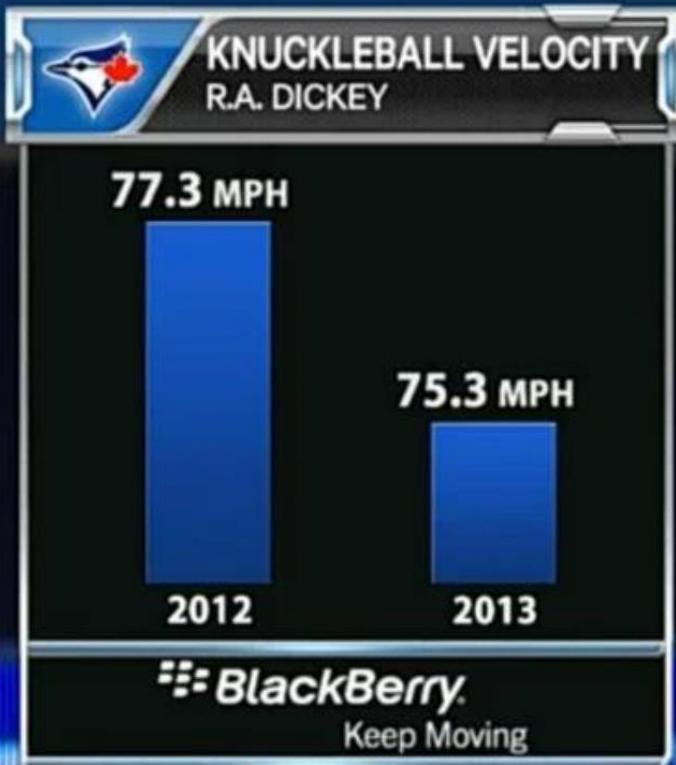


# Bullshit: Truncated Y-Axis

## Same Data, Different Y-Axis



# Bullshit: Truncated Y-Axis



# Michael Amberg

## Todays Learning Objectives:

1. **What is Bullshit and what is it not?**
2. **How do you identify Bullshit?**
3. **How to counteract Bullshit?**



# Causation or Correlation?

## Bergstrom, Calling Bullshit:

- Two **variables** are correlated when knowing the **value of one** gives you **information** about the likely **value** of the **other**.
- Two **states** (variables, events) are **causally related** when one state (...) **influence** the other via a **cause-and-effect process**.

# **Michael Amberg**

Todays Content:

- 1. Motivation**
- 2. Context Data Science**
- 3. Intro to Bullshit**
- 4. Old School Bullshit**
- 5. New School Bullshit**
- 6. Causation / Correlation**
- 7. Bullshit in Science**
- 8. Bullshit in AI / ML**
- 9. Summary**





## 7. Bullshit in Science



ULTRA!  
Top

## 8. Bullshit in AI / ML

# Inhalte des Moduls

## Vorlesung

Mittwoch, 11:30 Uhr über Zoom

Inhalte der Vorlesung:

- **Data Manipulation (Bullshit):** The Art of Skepticism in a Data Driven World
- **Data Handling:** Handling of Data in different scenarios
- **Data Privacy:** Protecting the Privacy & Integrity of Data
- **Data Driven Innovation:** Innovative Business & Applications
- **Project Management:** Managing Projects successfully
- **Business Systems:** Relevant Enterprise Software Systems
- **Emerging Technologies:** Implications of selected Technologies
- **Machine Learning I, II, III:** Overview of ML Techniques & Methods
- **ML & Robotics:** Current Developments & Implications

# Inhalte des Moduls

## Vorlesung

Mittwoch, 11:30 Uhr über Zoom

Inhalte der Vorlesung:

- **Data Manipulation (Bullshit)**
- **Data Handling**
- **Data Privacy**
- **Data Driven Innovation**
- **Project Management**
- **Business Systems**
- **Emerging Technologies**
- **Machine Learning I, II, III**
- **ML & Robotics**

## Projektarbeit

Donnerstag, 13:15 Uhr über Zoom

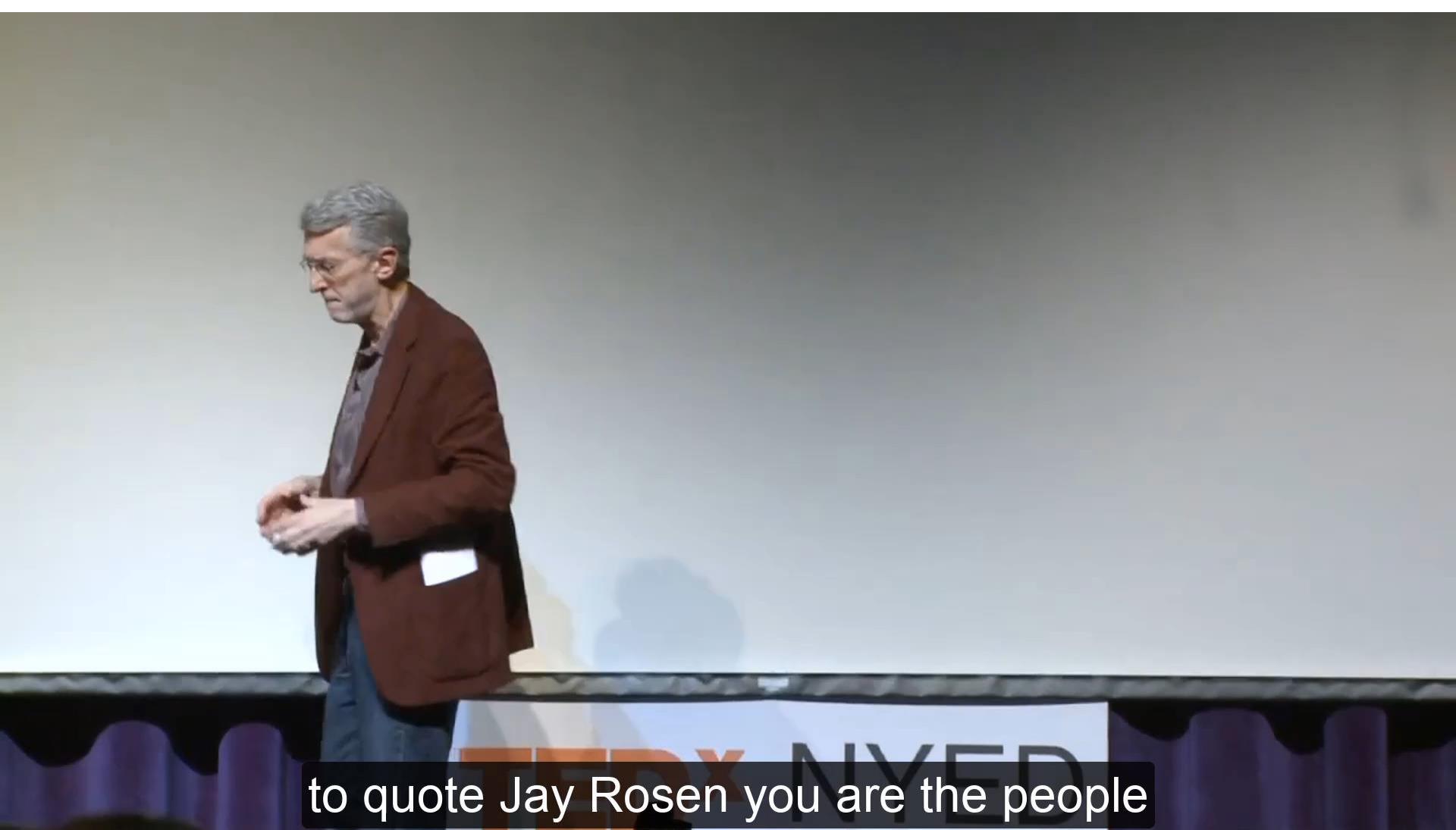
Drei toolgestützte Datenanalysen selbstständig unter Anleitung durchführen und in Form einer **wissenschaftlichen Arbeit toolgestützt** dokumentieren.

- **Selbstlernkurse** über StudOn
- **digitales Tutorium** über Zoom

salesforce



# Jeff Jarvis, Ted Talk über Bullshit



to quote Jay Rosen you are the people

# Correlation or Causation?

## Bergstrom, Calling Bullshit:

- Two **variables** are correlated when knowing the **value of one** gives you **information** about the likely **value** of the **other**.
- Two **variables** (states, events) are causally related when one variable (...) **influence** the other via a **cause-and-effect process**.



ULTRA!

*Top*

## 8. Bullshit in AI / ML

# Criminal Machine Learning - Biased Training Set?



(a) Three samples in criminal ID photo set  $S_c$ .



(b) Three samples in non-criminal ID photo set  $S_n$

# Criminal Machine Learning - Biased by Smile?



(a) -0.98



(b) -0.68



(c) -0.28



(d) -0.38



(e) 0.76



(f) 0.98



(g) 0.66

- 
1. Motivation
  2. Context Data Science
  3. Basics of Bullshit
  4. Old School Bullshit
  5. New School Bullshit
  6. Bullshit in Science
  7. Correlation / Causation
  8. Bullshit in AI / ML
  9. Summary



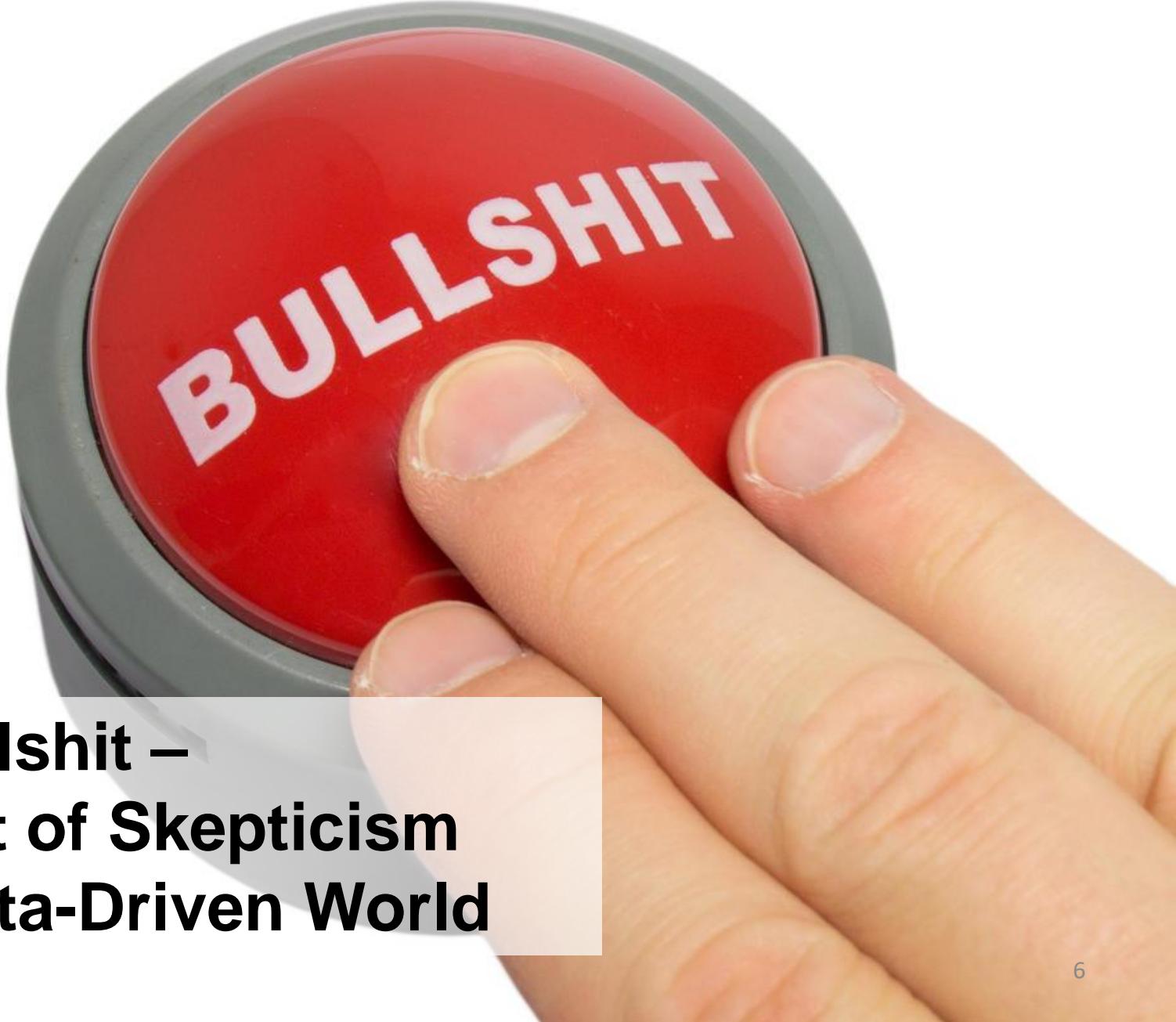
We are drowning in Bullshit.

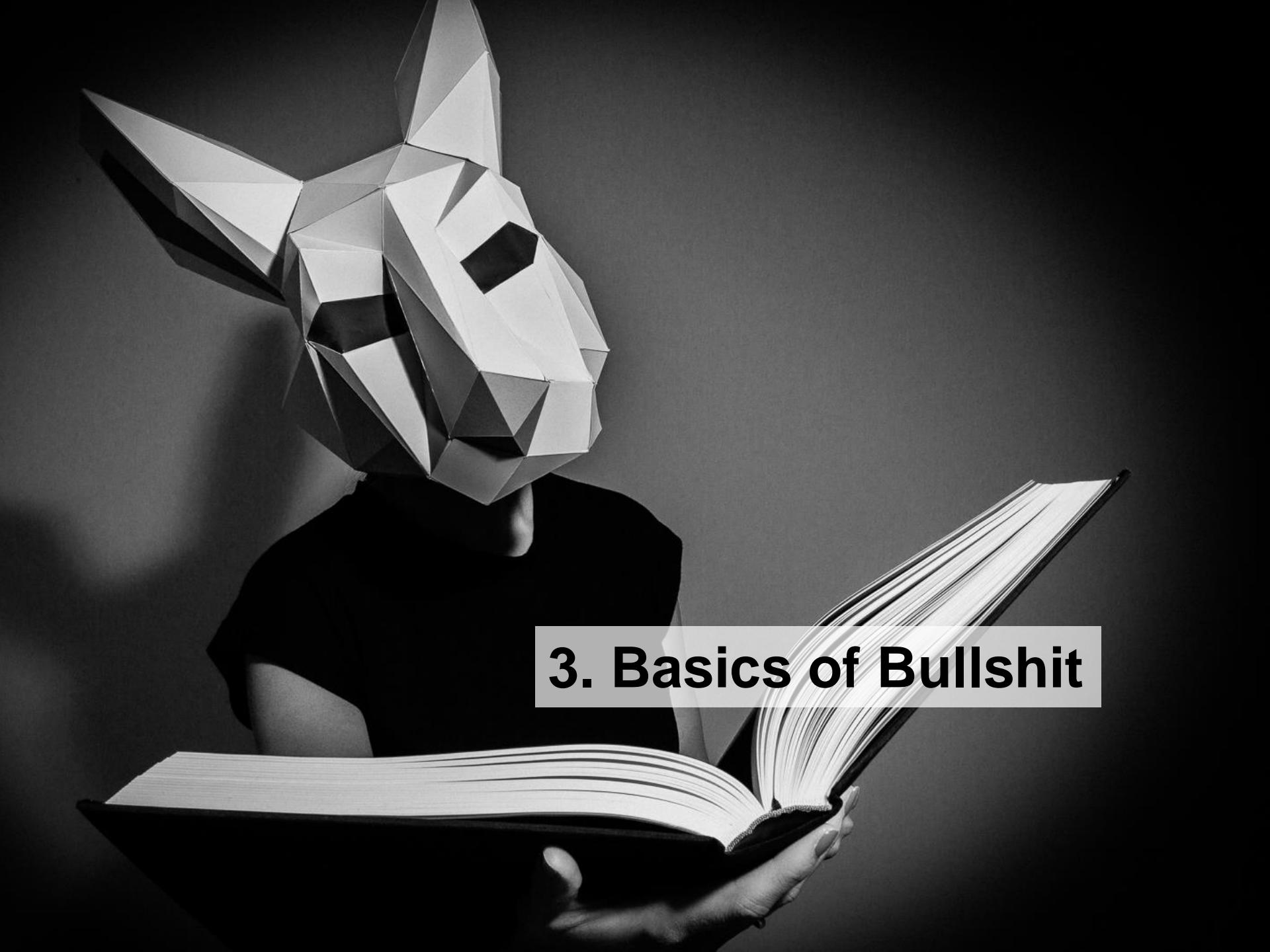
Statistical Bullshit is the new norm.

Avoid falling for Bullshit.

Society needs Critical Thinking & Data Science Skills more than ever.

# **On Bullshit – The Art of Skepticism in a Data-Driven World**



A black and white photograph of a person from the chest up. They are wearing a highly geometric, low-poly paper mask that looks like a donkey's head. The mask is white against a dark background. The person is holding an open book with both hands, and the pages are visible. The lighting is dramatic, coming from the side to highlight the mask and the edges of the book.

### **3. Basics of Bullshit**

# Bullshit in Sign Language

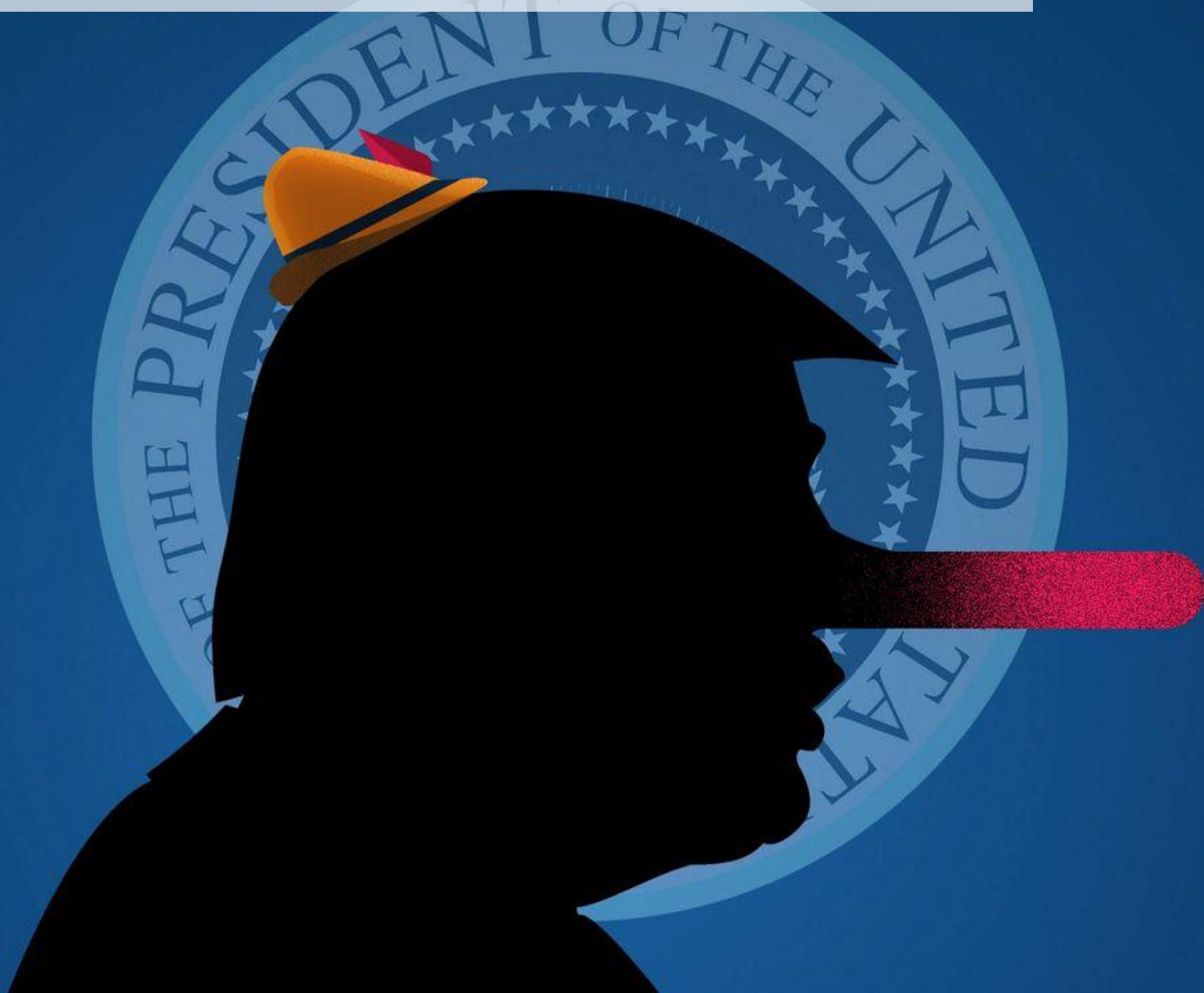
bullshit

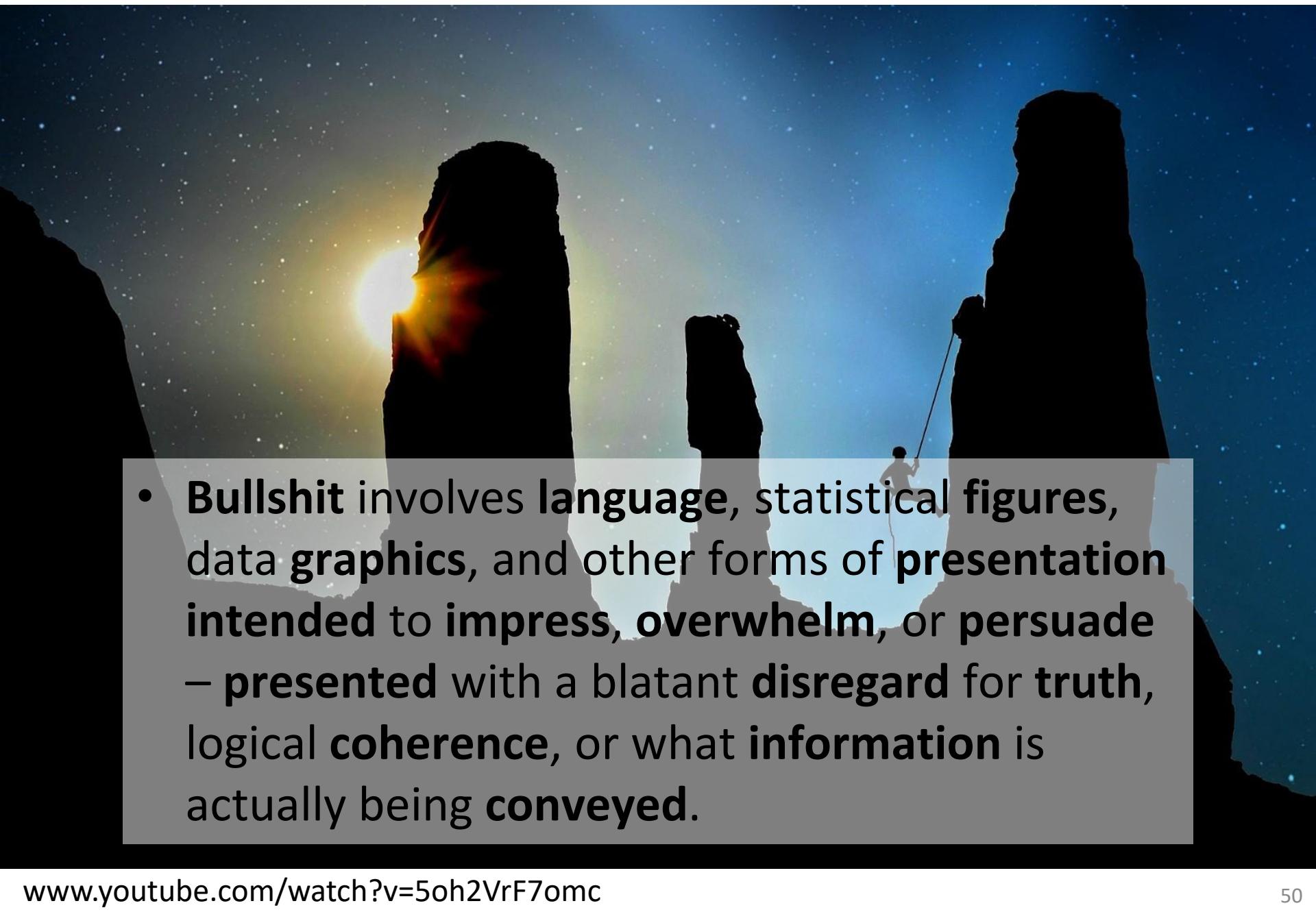


# Harry Frankfurt, Princeton University, 2005: On Bullshit

- 
- A **Liar** knows the **truth** and is trying to **convince** us of **something different**.
  - A **Bullshitter** either doesn't know the **truth**, or doesn't **care** – and is just trying to be **persuasive**.

Vox.com: „The Bullshitter-in-Chief“, 5/2017



- 
- **Bullshit** involves **language**, statistical **figures**, data **graphics**, and other forms of **presentation** intended to **impress**, **overwhelm**, or **persuade** – **presented** with a blatant **disregard** for **truth**, logical **coherence**, or what **information** is actually being **conveyed**.

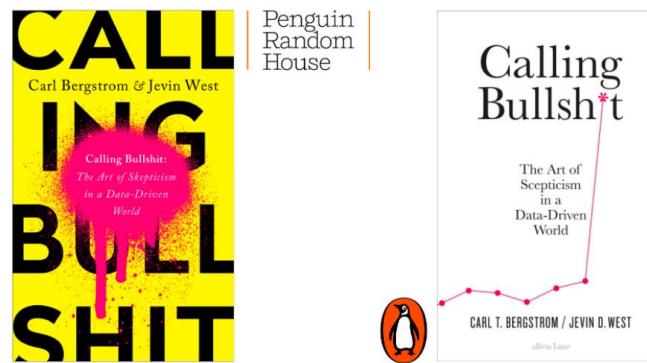
## Bergstrom, Calling Bullshit:

**“If a claim seems  
too good — or too bad —  
to be true,  
it probably is.”**



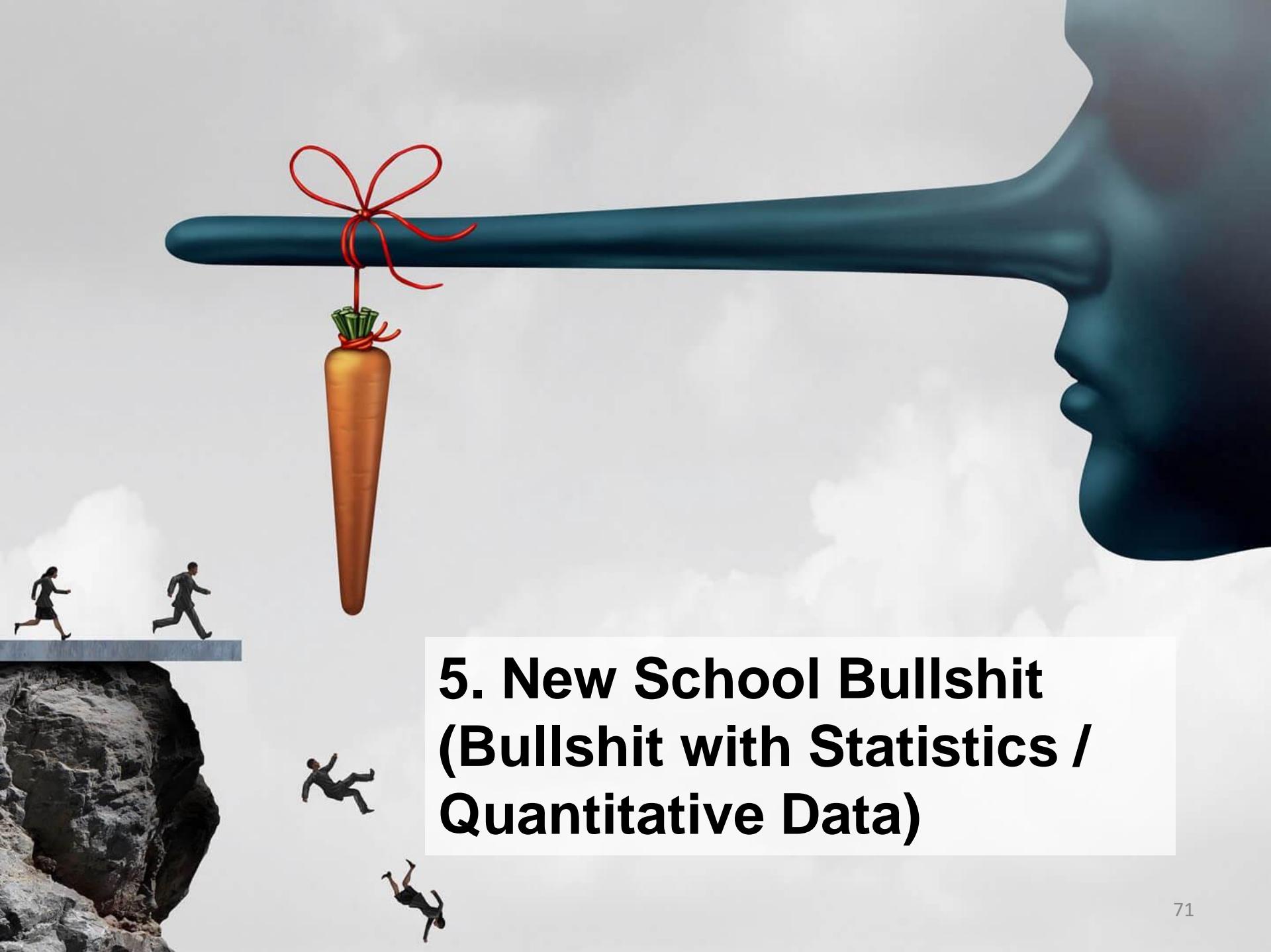
## 4. Old School Bullshit

# Website **callingbullshit.org**



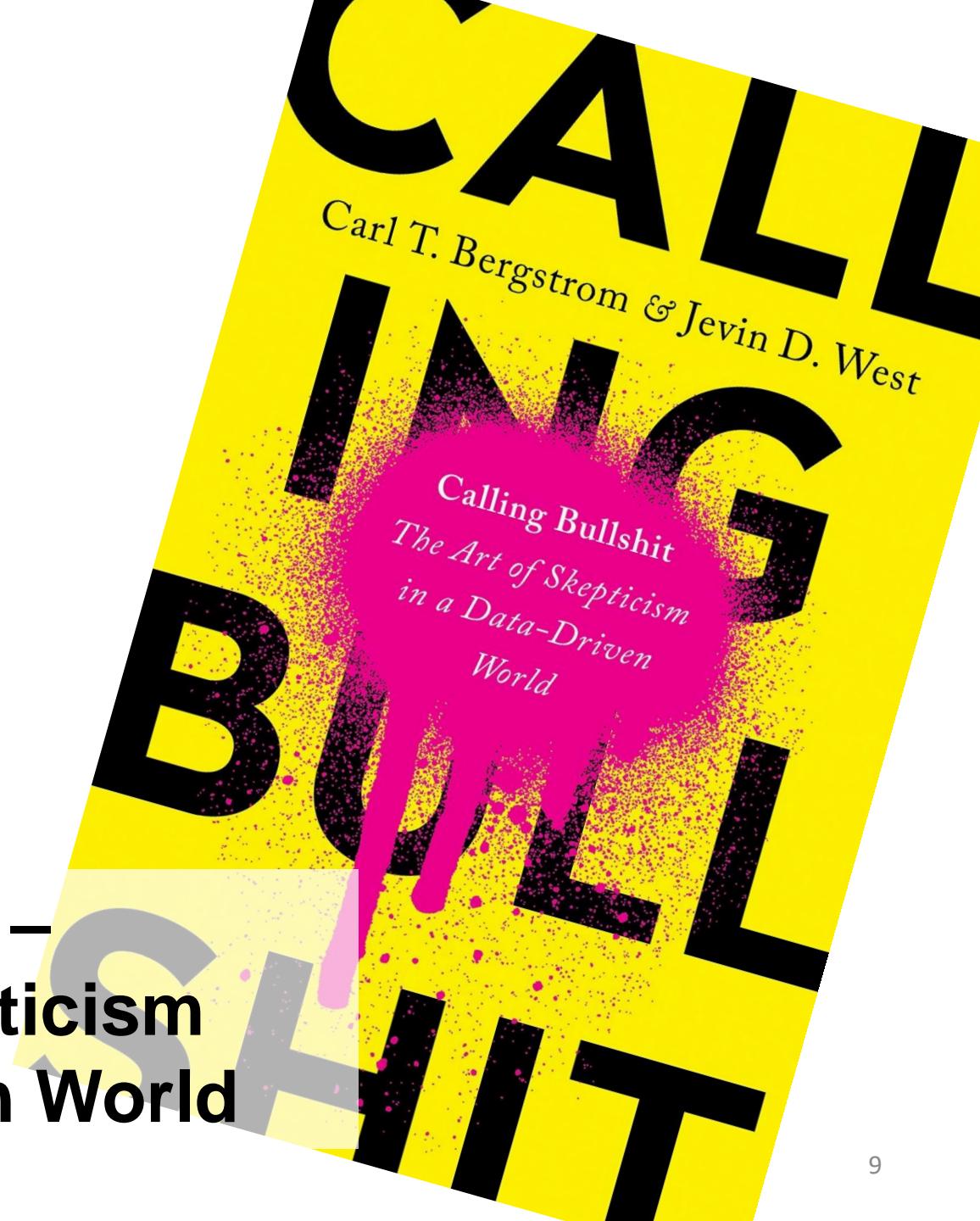
**Now available!** *Calling Bullshit: The Art of Skepticism in a Data-Driven World*, by Carl Bergstrom and Jevin West. [Available here.](#)

**The world is awash in bullshit.** Politicians are unconstrained by facts. Science is conducted by press release. Higher education rewards bullshit over analytic thought. Startup culture elevates bullshit to high art. Advertisers wink conspiratorially and invite us to join them in seeing through all the bullshit — and take advantage of our lowered guard to bombard us with bullshit of the second order. The majority of administrative activity, whether in private business or the public sphere, seems to be little more than a sophisticated

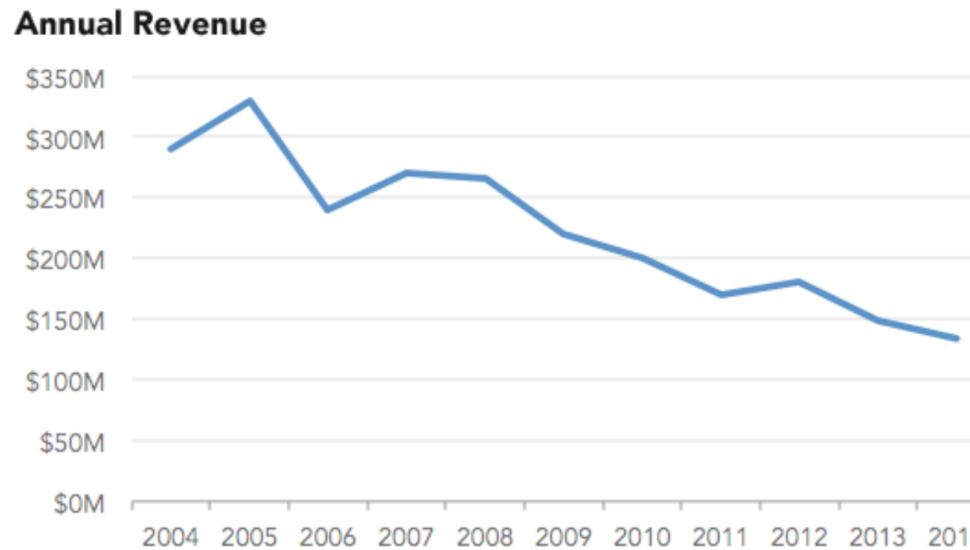
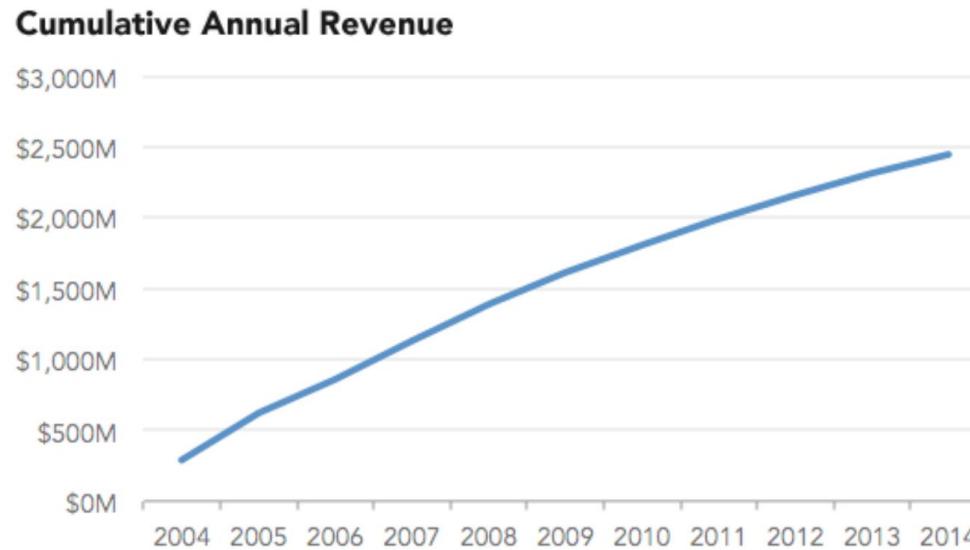


## **5. New School Bullshit (Bullshit with Statistics / Quantitative Data)**

# Calling Bullshit – The Art of Skepticism in a Data-Driven World

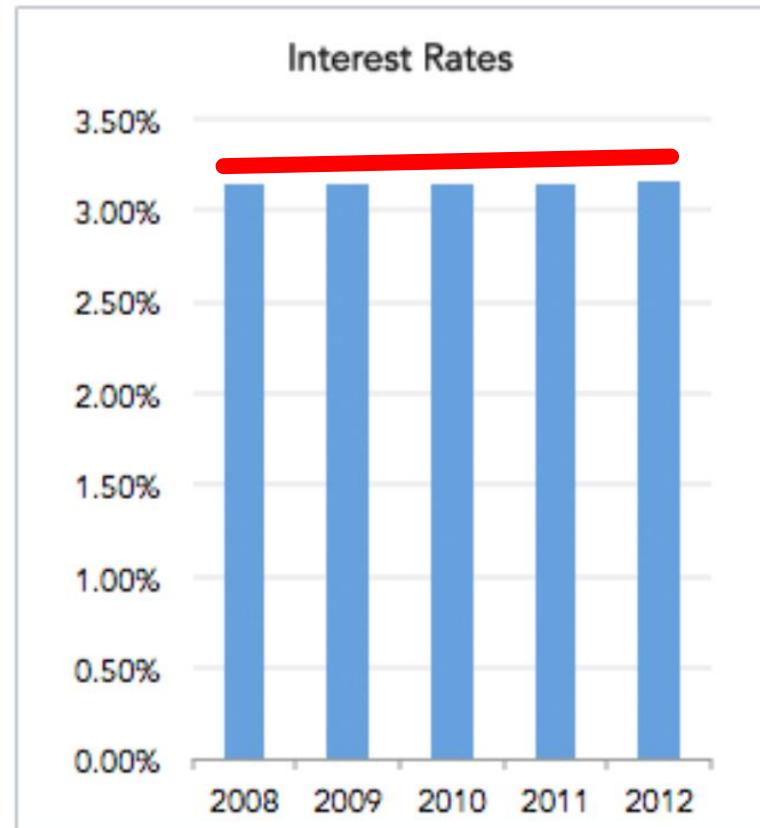


# Bullshit: Cumulative graphs

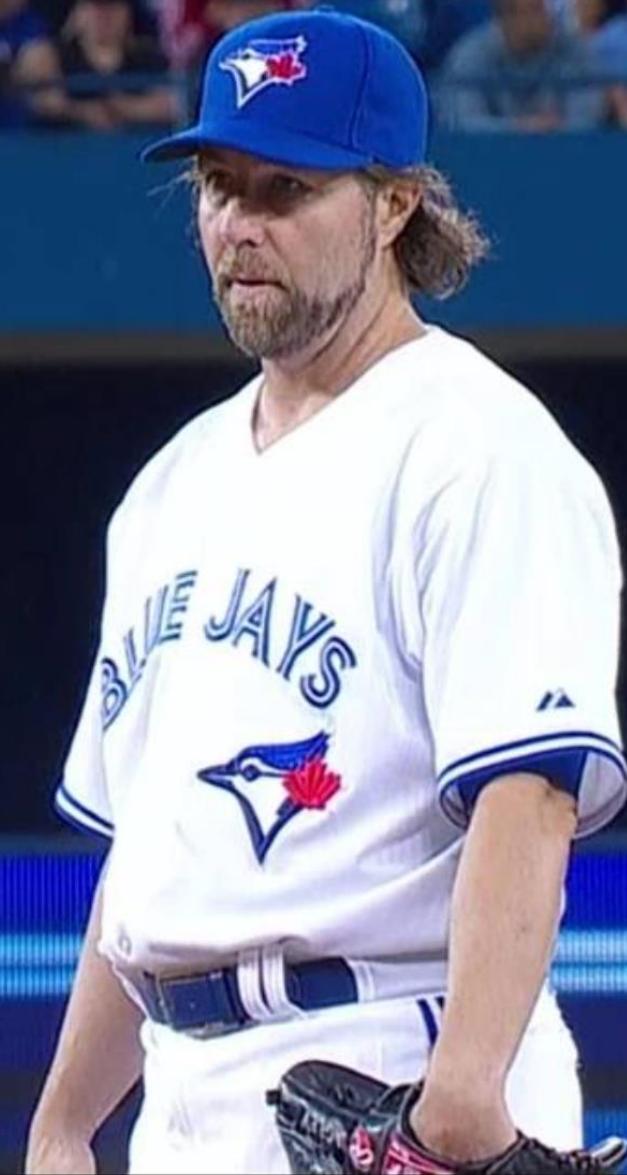
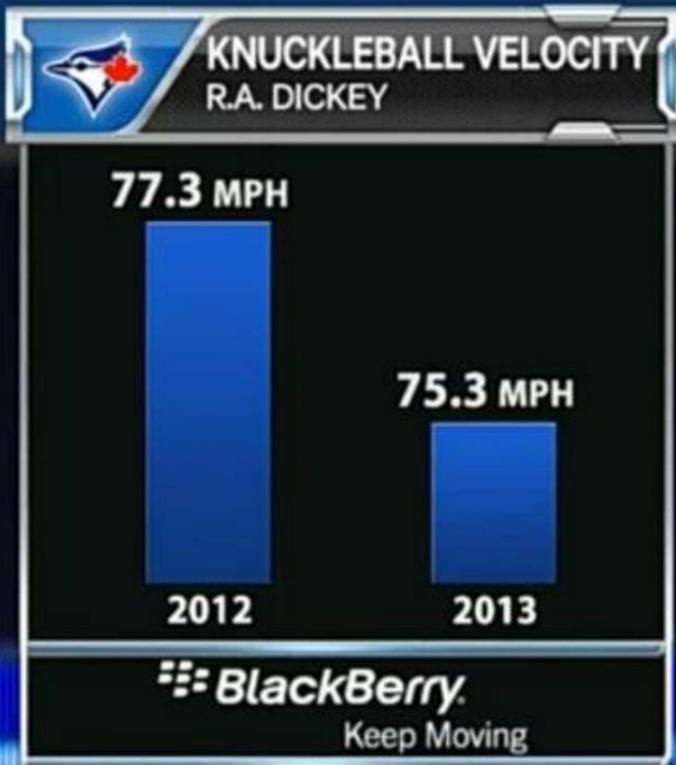


# Bullshit: Truncated Y-Axis

## Same Data, Different Y-Axis



# Bullshit: Truncated Y-Axis



# Michael Amberg

Todays Learning Objectives:

1. Better Understand Bullshit
2. Learn Techniques & Tips to Identify Bullshit





## 6. Bullshit in Science (with Scientific Research)

# Michael Amberg

## Todays Content:

- 1. Motivation**
- 2. Context Data Science**
- 3. Basics of Bullshit**
- 4. Old School Bullshit**
- 5. New School Bullshit**
- 6. Bullshit in Science**
- 7. Correlation / Causation**
- 8. Bullshit in AI / ML**
- 9. Summary**



# 1 Bullshit with Data

## - The Art of Skepticism in a Data-Driven World

### (1) Motivation



Critical Thinking!  
Think Better. Be Better. Become a Champion.

# 1 Bullshit with Data

## - The Art of Skepticism in a Data-Driven World (2) Context Data Science



# 1 Bullshit with Data

## - The Art of Skepticism in a Data-Driven World

### (3) Basics of Bullshit



## The Ultimate Cheatsheet for Critical Thinking

Want to exercise critical thinking skills? Ask these questions whenever you discover or discuss new information. These are broad and versatile questions that have limitless applications!

<b>Who</b>	<ul style="list-style-type: none"><li>... benefits from this?</li><li>... is this harmful to?</li><li>... makes decisions about this?</li><li>... is most directly affected?</li></ul>	<ul style="list-style-type: none"><li>... have you also heard discuss this?</li><li>... would be the best person to consult?</li><li>... will be the key people in this?</li><li>... deserves recognition for this?</li></ul>
<b>What</b>	<ul style="list-style-type: none"><li>... are the strengths/weaknesses?</li><li>... is another perspective?</li><li>... is another alternative?</li><li>... would be a counter-argument?</li></ul>	<ul style="list-style-type: none"><li>... is the best/worst case scenario?</li><li>... is most/least important?</li><li>... can we do to make a positive change?</li><li>... is getting in the way of our action?</li></ul>
<b>Where</b>	<ul style="list-style-type: none"><li>... would we see this in the real world?</li><li>... are there similar concepts/situations?</li><li>... is there the most need for this?</li><li>... in the world would this be a problem?</li></ul>	<ul style="list-style-type: none"><li>... can we get more information?</li><li>... do we go for help with this?</li><li>... will this idea take us?</li><li>... are the areas for improvement?</li></ul>
<b>When</b>	<ul style="list-style-type: none"><li>... is this acceptable/unacceptable?</li><li>... would this benefit our society?</li><li>... would this cause a problem?</li></ul>	<ul style="list-style-type: none"><li>... will we know we've succeeded?</li><li>... has this played a part in our history?</li><li>... can we expect this to change?</li></ul>



# 1 Bullshit with Data

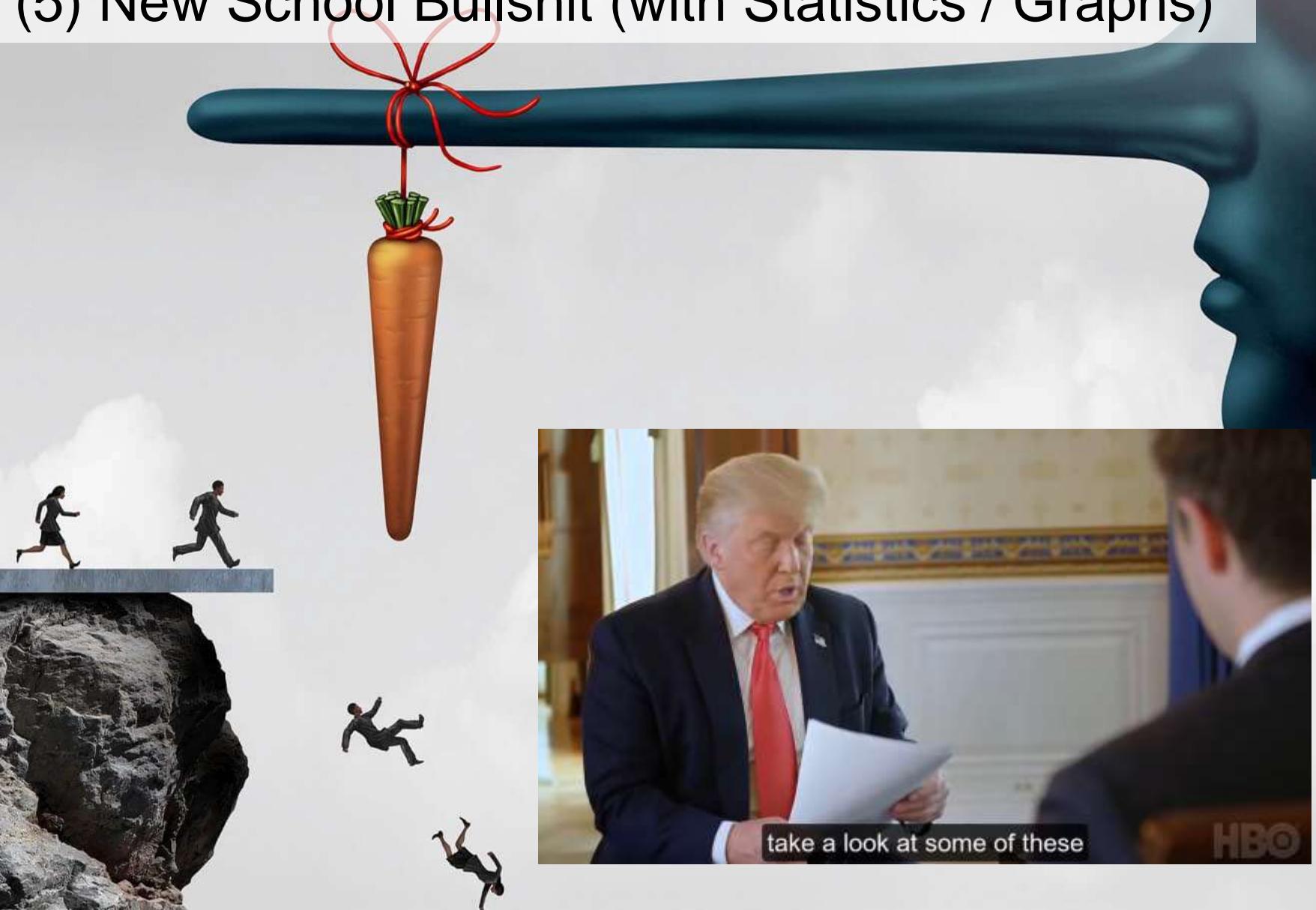
## - The Art of Skepticism in a Data-Driven World

### (4) Old School Bullshit



# 1 Bullshit with Data

- The Art of Skepticism in a Data-Driven World  
(5) New School Bullshit (with Statistics / Graphs)





# 1 Bullshit with Data

- The Art of Skepticism in a Data-Driven World

(6) Bullshit in Science



In patients with multiple risk factors for heart disease,

**Lipitor** reduces risk of heart attack by **36%**\*

If you have risk factors such as family history, high blood pressure, age, low HDL ('good' cholesterol) or smoking

\*That means in a large clinical study, 3% of patients taking a sugar pill or placebo had a heart attack compared to 2% of patients taking Lipitor.

→ Inventor of the Jumbo Artificial Heart and Lifesaver

DR. ROBERT JARVIS

**LIPITOR**  
atorvastatin calcium  
tarjevant

**FRESH CONCERN ABOUT THE EFFECTIVENESS OF MASKS**

TUCKER CARLSON tonight • #Tucker

see you in jail miss grant harsh words

The image consists of two side-by-side video stills. On the left, Tucker Carlson is shown from the chest up, wearing a dark suit and tie, speaking with a serious expression. Behind him is a blurred background of what appears to be a news studio. On the right, there is a close-up shot of a manufacturing process. A blue surgical mask is being processed by a machine with various metal components and tubes. The background is dark and industrial.

# 1 Bullshit with Data

- The Art of Skepticism in a Data-Driven World  
(7) Correlation vs Causation



# 1 Bullshit with Data

## - The Art of Skepticism in a Data-Driven World

### (8) Bullshit in AI / ML



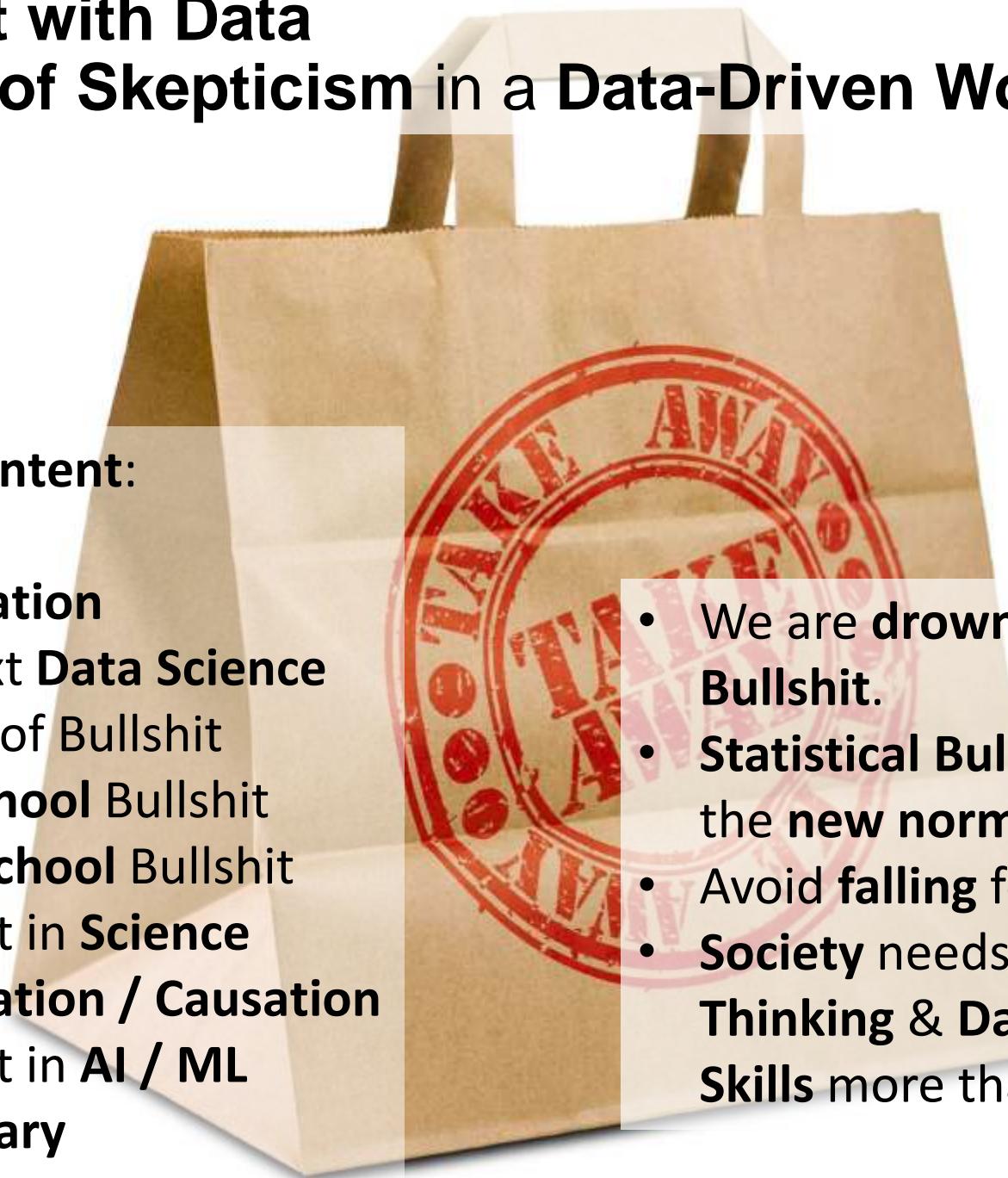
# 1 Bullshit with Data

## - The Art of Skepticism in a Data-Driven World

Todays Content:

- 1. Motivation**
- 2. Context Data Science**
- 3. Basics of Bullshit**
- 4. Old School Bullshit**
- 5. New School Bullshit**
- 6. Bullshit in Science**
- 7. Correlation / Causation**
- 8. Bullshit in AI / ML**
- 9. Summary**

- We are **drowning** in **Bullshit**.
- **Statistical Bullshit** is the **new norm**.
- Avoid **falling** for **Bullshit**.
- Society needs **Critical Thinking** & **Data Science Skills** more than ever.



# Which Statement about / in the context of Bullshit is Wrong?

Schwierigkeitsgrad	Art des Wissens	Auftragswissen (Vorlesung)	Anwendungswissen (Literatur)
Einfach			
Mittel			
Schwierig			

- a) A **Liar** knows the truth and is trying to convince us of something different.
- b) A **Bullshitter** either doesn't know the truth, or doesn't care – and is just trying to be persuasive.
- c) The amount of **energy** needed to refute **bullshit** is an order of magnitude bigger than to produce it.
- d) All.
- e) None.

In most cases, **Bullshit** in AI/ML is based on \_\_\_\_\_ instead of \_\_\_\_\_.



An der Wissensschwelle		Abfragewissen (Erinnerung)	Anwendungswissen (Literatur)
Schwierigkeitsgrad	Einfach	Mittel	Schwierig
schwierig		grün	
mittig		gelb	
einfach			rot

- a) biased Data / racist People
- b) biased People / biased Data
- c) racist People / biased Data
- d) biased Data / incompetent People
- e) incompetent People / biased Data

# Falling for Bullshit is way too easy.

# 50 COGNITIVE BIASES TO BE AWARE OF SO YOU CAN BE THE VERY BEST VERSION OF YOU

## Memory

## Social

## Learning

## Belief

## Money

## Politics

### Fundamental Attribution Error

We judge others on their personality or fundamental character, but we judge ourselves on the situation.



*Sally is late to class; she's lazy. You're late to class; it was a bad morning.*

### Self-Serving Bias

Our failures are situational, but our successes are our responsibility.



*You won that award due to hard work rather than help or luck. Meanwhile, you failed a test because you hadn't gotten enough sleep.*

### In-Group Favoritism

We favor people who are in our in-group as opposed to an out-group.



*Francis is in your church, so you like Francis more than Sally.*

### Bandwagon Effect

Ideas, fads, and beliefs grow as more people adopt them.



*Sally believes fidget spinners help her children. Francis does, too.*

### Groupthink

Due to a desire for conformity and harmony in the group, we make irrational decisions, often to minimize conflict.



*Sally wants to go get ice cream. Francis wants to shop for T-shirts. You suggest getting T-shirts with pictures of ice cream on them.*

### Halo Effect

If you see a person as having a positive trait, that positive impression will spill over into their other traits. (This also works for negative traits.)



*"Taylor could never be mean; she's so cute!"*

### Moral Luck

Better moral standing happens due to a positive outcome; worse moral standing happens due to a negative outcome.



*"X culture won X war because they were morally superior to the losers."*

### False Consensus

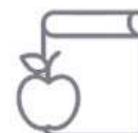
We believe more people agree with us than is actually the case.



*"Everybody thinks that!"*

### Curse of Knowledge

Once we know something, we assume everyone else knows it, too.



*Alice is a teacher and struggles to understand the perspective of her new students.*

### Spotlight Effect

We overestimate how much people are paying attention to our behavior and appearance.



*Sally is worried everyone's going to notice how lame her ice cream T-shirt is.*

# 1 Bullshit with Data

## - The Art of Skepticism in a Data-Driven World

### Content:

1. Motivation
2. Context Data Science
3. Basics of Bullshit
4. Old School Bullshit
5. New School Bullshit
6. Bullshit in Science
7. Correlation / Causation
8. Bullshit in AI / ML
9. Summary



# Anwendungen und Rahmenbedingungen von Data Science und Machine Learning in Unternehmen und in einer datengetriebenen Welt kennen und einschätzen lernen.



1. Bullshit with Data:
2. Personal Data Handling:
3. Protect Data:
4. Data Driven Innovation:
5. Project Management:
6. Business Systems:
7. Emerging Technologies:
8. Machine Learning I, II, III:
9. ML & Robotics:

**The Art of Skepticism in a Data Driven World**  
Handling of Data in Different Scenarios  
Protecting the Privacy & Integrity of Data  
Innovative Business & Applications  
Managing Projects successfully  
Relevant Enterprise Software Systems  
Implications of selected Technologies  
Overview of ML Techniques & Methods  
Current Developments & Implications

# Anwendungen und Rahmenbedingungen von Data Science und Machine Learning in Unternehmen und in einer datengetriebenen Welt kennen und einschätzen lernen.



1. Bullshit with Data:
2. Personal Data Handling:
3. Protect Data:  
4. Data Driven Innovation:
5. Project Management:
6. Business Systems:
7. Emerging Technologies:
8. Machine Learning I, II, III:
9. ML & Robotics:

The Art of Skepticism in a Data Driven World  
Handling of Data in Different Scenarios  
Protecting the Security & Integrity of Data  
Innovative Business & Applications  
Managing Projects successfully  
Relevant Enterprise Software Systems  
Implications of selected Technologies  
Overview of ML Techniques & Methods  
Current Developments & Implications

# Anwendungen und Rahmenbedingungen von Data Science und Machine Learning in Unternehmen und in einer datengetriebenen Welt kennen und einschätzen lernen.



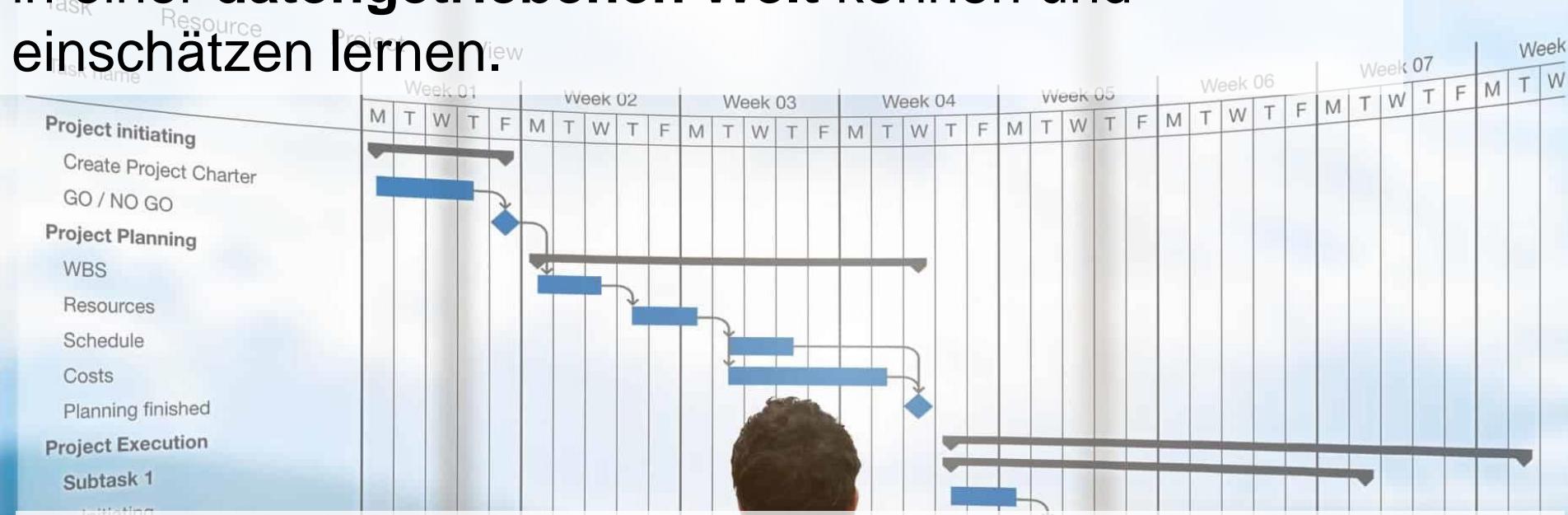
1. Bullshit with Data:
2. Personal Data Handling:
3. Protect Data:
4. Data Driven Innovation:  

---
5. Project Management:
6. Business Systems:
7. Emerging Technologies:
8. Machine Learning I, II, III:
9. ML & Robotics:

- The Art of Skepticism in a Data Driven World
- Handling of Data in Different Scenarios
- Protecting the Security & Integrity of Data
- Innovative Business & Applications  

---
- Managing Projects successfully
- Relevant Enterprise Software Systems
- Implications of selected Technologies
- Overview of ML Techniques & Methods
- Current Developments & Implications

# Anwendungen und Rahmenbedingungen von Data Science und Machine Learning in Unternehmen und in einer datengetriebenen Welt kennen und einschätzen lernen.



1. Bullshit with Data:
2. Personal Data Handling:
3. Protect Data:
4. Data Driven Innovation:
5. **Project Management:**
6. Business Systems:
7. Emerging Technologies:
8. Machine Learning I, II, III:
9. ML & Robotics:

**The Art of Skepticism in a Data Driven World**  
**Handling of Data in Different Scenarios**  
**Protecting the Security & Integrity of Data**  
**Innovative Business & Applications**  
**Managing Projects successfully**

---

**Relevant Enterprise Software Systems**  
**Implications of selected Technologies**  
**Overview of ML Techniques & Methods**  
**Current Developments & Implications**

# Anwendungen und Rahmenbedingungen von Data Science und Machine Learning in Unternehmen und in einer datengetriebenen Welt kennen und einschätzen lernen.



1. Bullshit with Data:
2. Personal Data Handling:
3. Protect Data:
4. Data Driven Innovation:
5. Project Management:
6. Business Systems:
7. Emerging Technologies:
8. Machine Learning I, II, III:
9. ML & Robotics:

The Art of Skepticism in a Data Driven World  
Handling of Data in Different Scenarios  
Protecting the Security & Integrity of Data  
Innovative Business & Applications  
Managing Projects successfully  
Relevant Enterprise Software Systems  
Business Implications of Technologies  
Overview of ML Techniques & Methods  
Current Developments & Implications