

The Big Picture: Big Data, Big Theory, and Big Challenges

William G. Kennedy, Ph.D.,
Captain, USN (Ret.)

17 May 2004

Think. Learn. Succeed.



Motivation

In 1990, Ed Wegman freed us from FlatLand with his paper: "Hyperdimensional Data Analysis Using Parallel Coordinates"

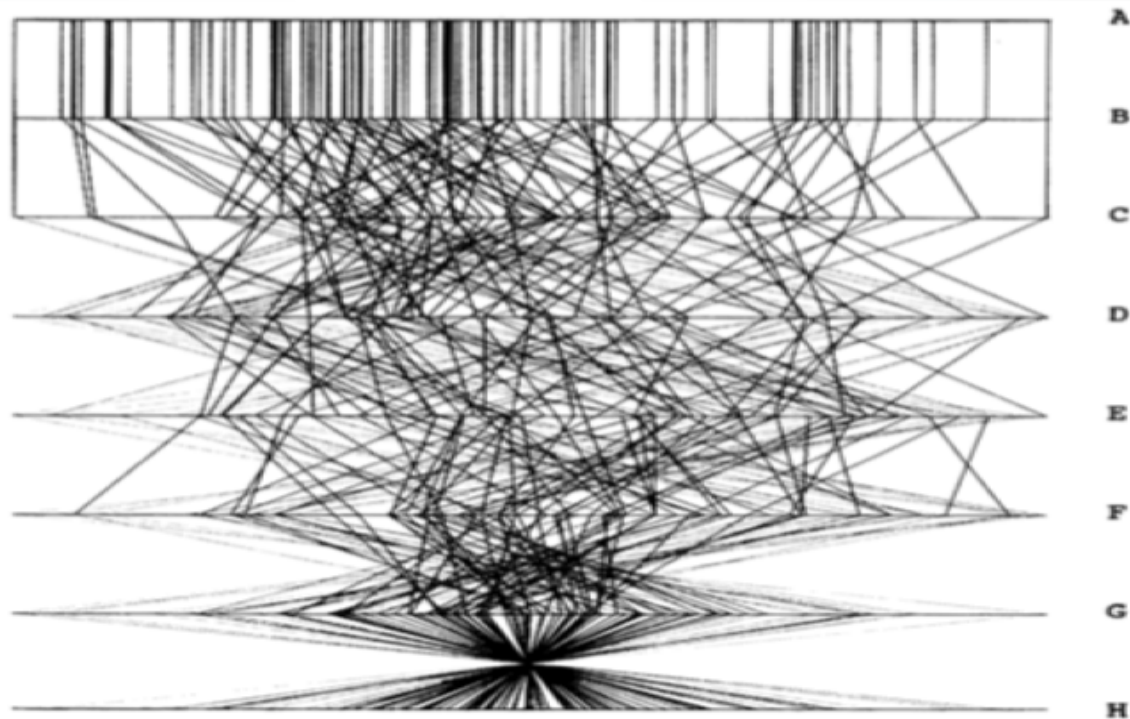


Figure 3. Parallel Coordinate Plot of Six-Dimensional Data Illustrating Correlations of $\rho = 1, .8, .2, 0, -.2, -.8, \text{ and } -1$.

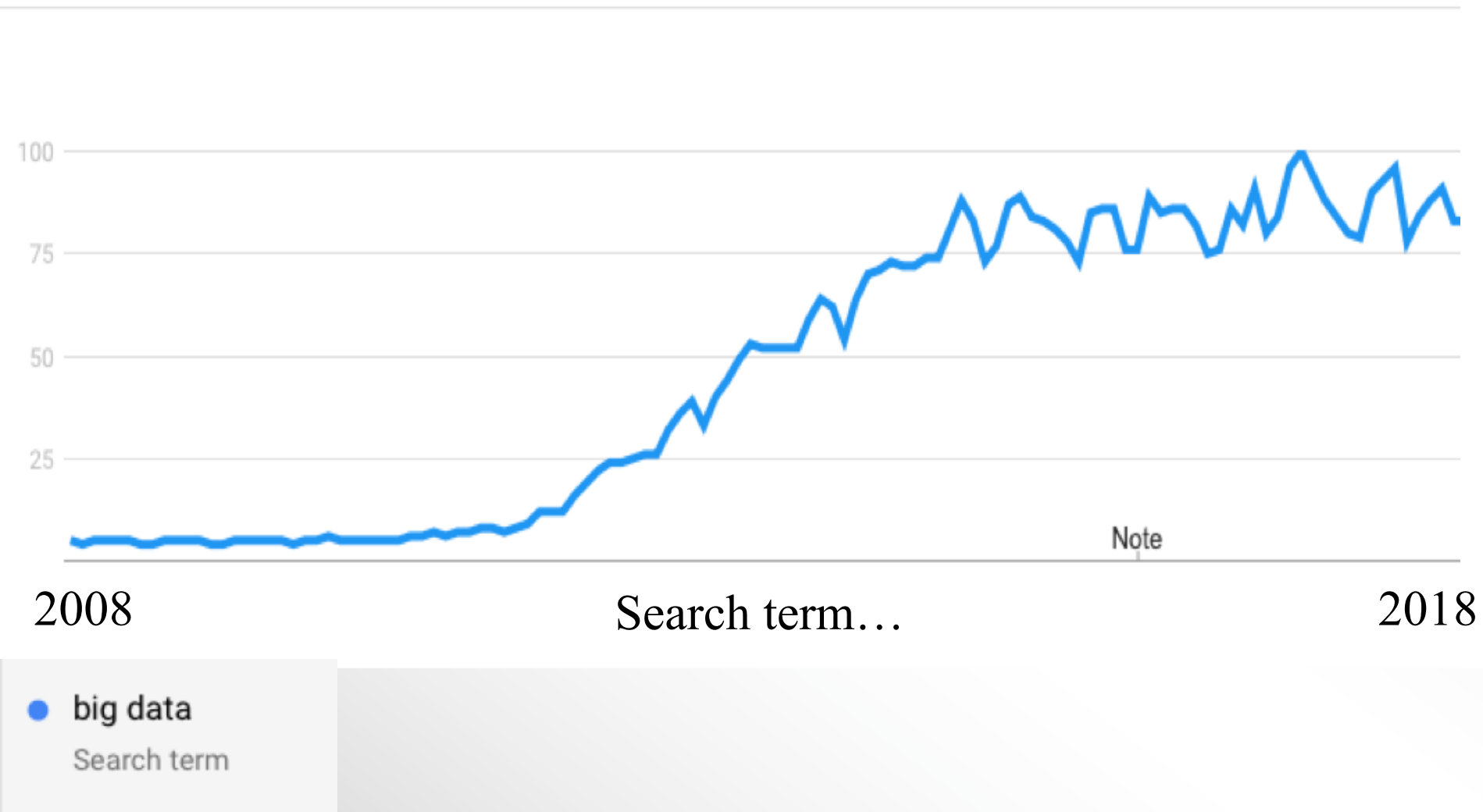
The Big Picture

Statistics

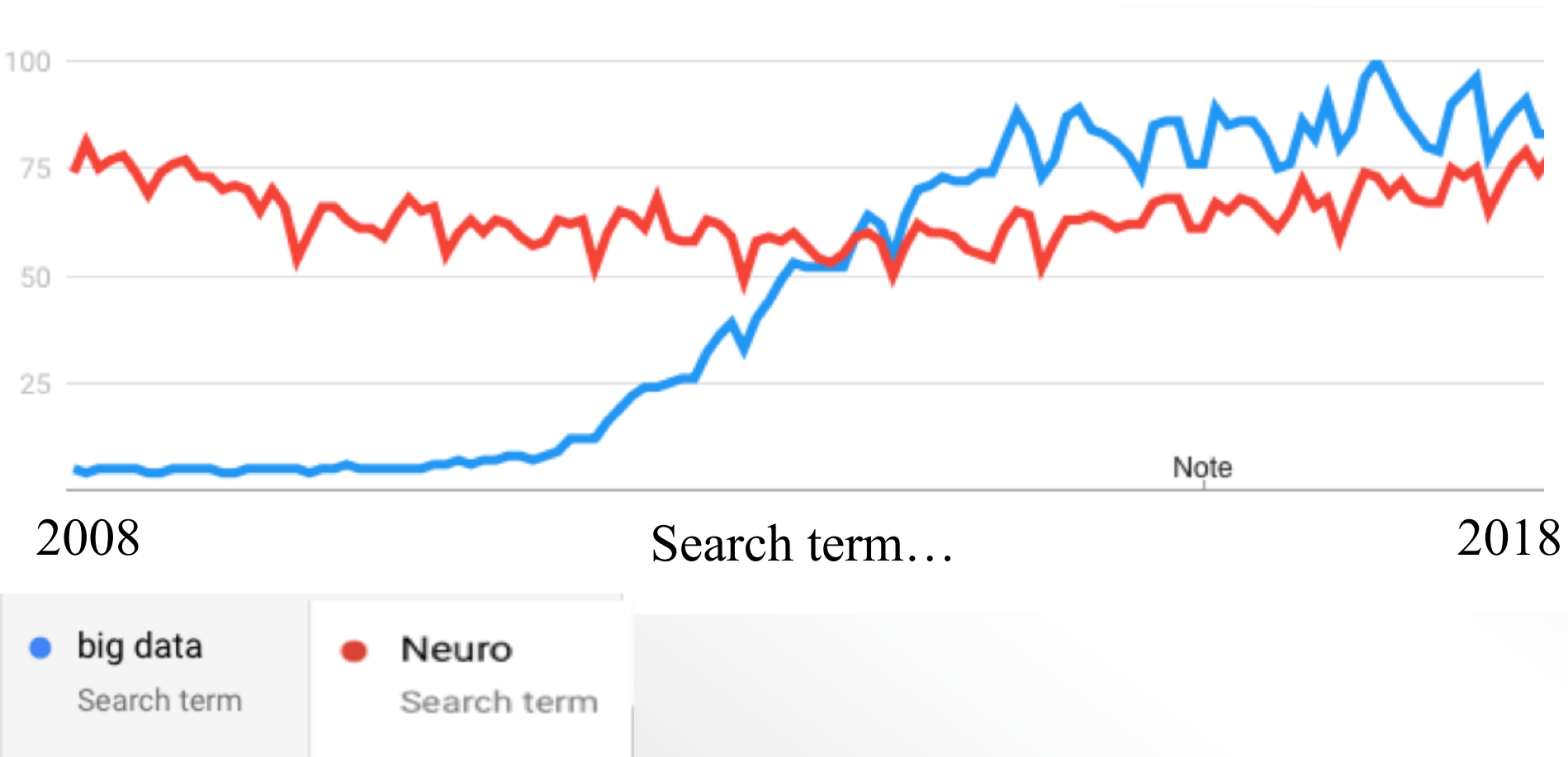
- Descriptive Statistics
- Inferential Statistics
- History
 - 1700's: demographic and economic data
 - 1800's: probability included
 - 1900's: analysis used in social and hard sciences
(“The Lady Tasting Tea” by Salsburg 2001)
 - 2000's: “Big Data”

“Big Data”

Big Data: Normalized Google Search



Big Data: Normalized Google Search



Big Data

Defined ...

not by size,

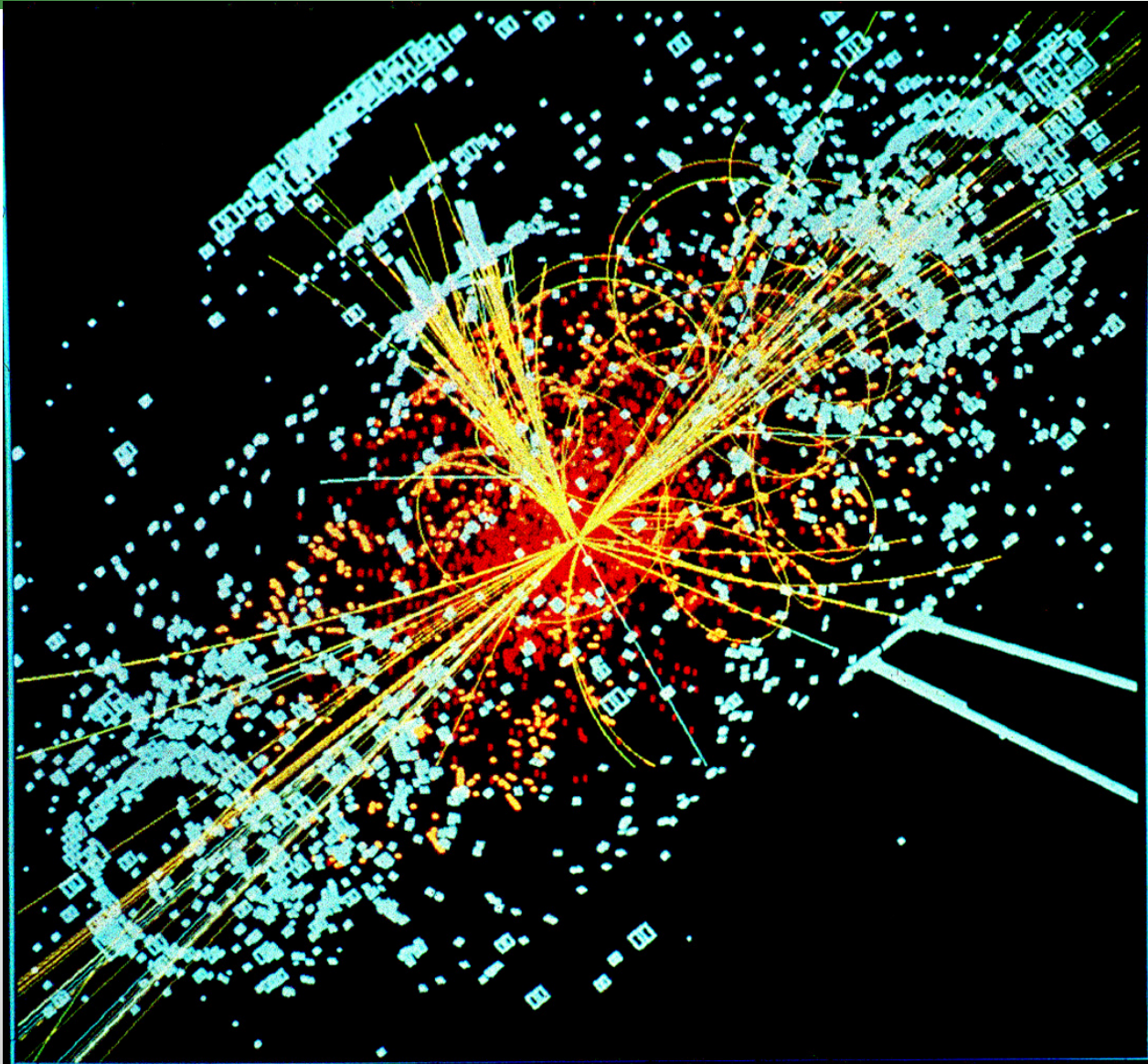
not by source or storage,

but by the challenge of processing it

Big Data

- Processing
 - In computer's memory?
 - Like eating an elephant?
 - “One bite at a time” (byte?)
 - i.e., sequentially?
 - As piranha, *in parallel*
 - On the cloud...
 - or volunteered “at home” systems

“Big Theory”



“Big Theory”

- GUT (“Grand Unified Theory”)
Model of particle physics in which, at high energy, forces of the Standard Model merge into one
- Unifying gravity with GUT is TOE
 (“Theory of Everything”)

“Big Theory”

“Grand Theory” (social science)

- C. Wright Mills (1959) “The Sociological Imagination”
- “Universal scheme to understand the unity of social structures” (Wikipedia)

“Big Theory”

General AI?

Unifying Theories of Cognition?

“Big Challenges”

“Big Challenges”

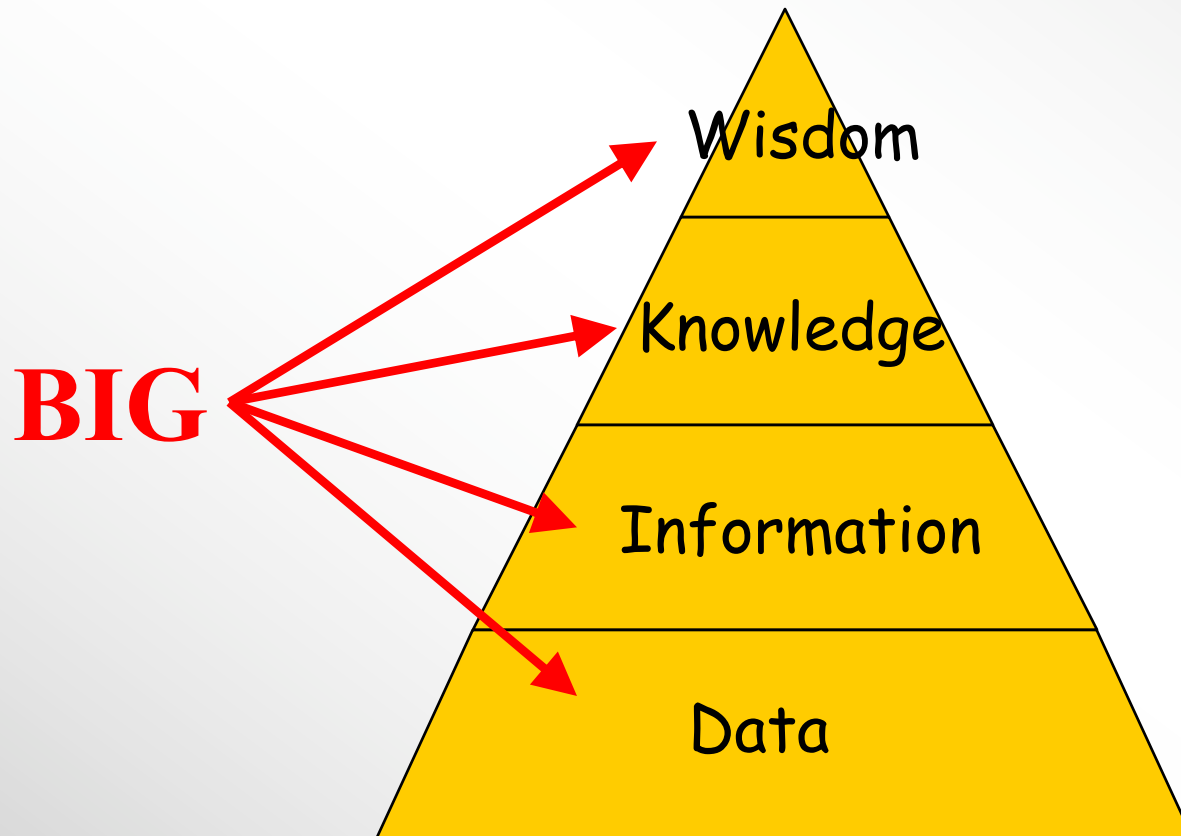
1. Is “Big Data” just more or is it different?

- Where are new concepts to join mean, mode,
- distributions, correlations, and theory testing?
- Hierarchies?
- Hierarchical hidden Markov models?
- (Ray Kurzweil, “How to Create a Mind”)

“Big Challenges”

1. Is “Big Data” just more or is it different?
- 2. Relevance of “Big Data” to DIKW?**
(Data, Information, Knowledge, Wisdom)

Processing “Big Data”



“Big Challenges”

1. Relevance of “Big Data” to DIKW?
2. Is “Big Data” just more or is it different?
3. **Does “Big Data” relieve us from understanding?**

“Big Challenges”

We all learned multiplication tables
up to 10×10 ,
 12×12 ,
 15×15 ?

Reliance on searches for answers is not
understanding, not knowledge

Siri, Alexa, Wikipedia, Google may know, but
we don't.

**If we don't know it, we can't reason with it
and it can't “come to mind.”**

“Big Challenges”

1. Is “Big Data” just more or is it different?
2. Relevance of “Big Data” to DIKW”
3. Does “Big Data” relieve us from understanding?
4. **Is sufficiently big “Big Data” indistinguishable from magic?**

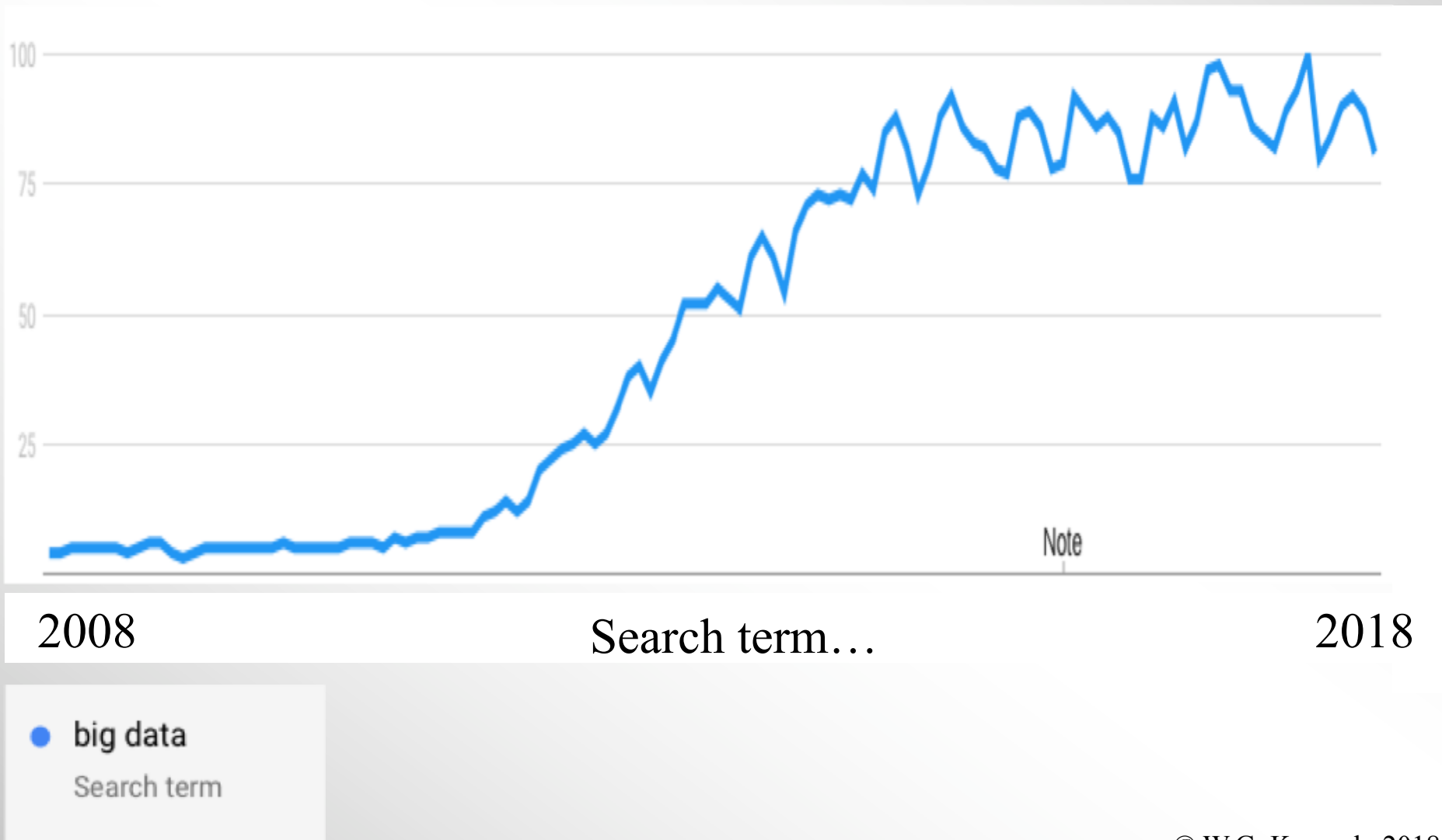
(apologies to Arthur C. Clark)

Big Picture

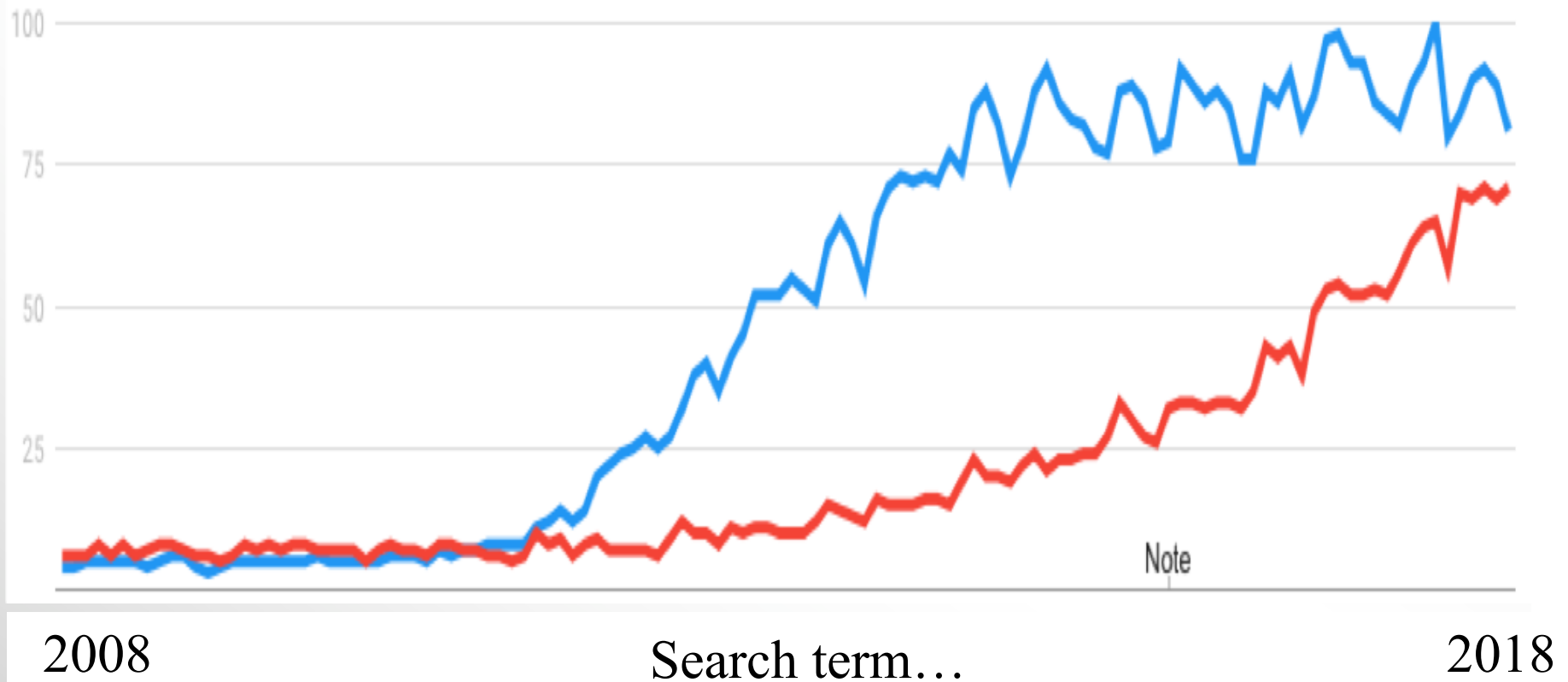
Last time we dealt with “big data,” it produced the field of statistics

What's next?

Big Picture

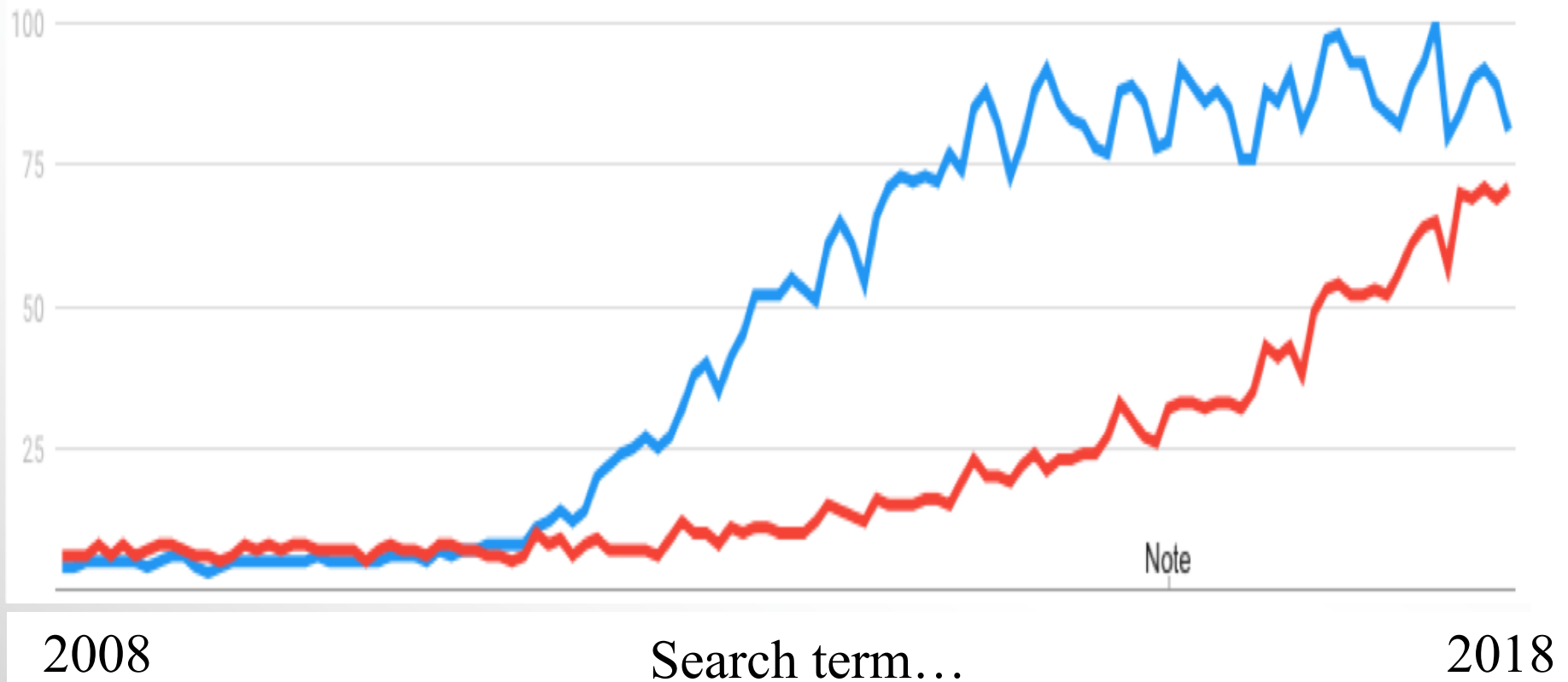


Big Picture



● big data
Search term

Big Picture



● big data
Search term

● data science
Search term



Thank you.

William G. Kennedy, Ph.D.,
Captain, USN (Ret.)
wkennedy@gmu.edu